



universität  
wien

# DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„I think therefore I uh...m?

A corpus-based study of syntactic and functional  
differences between *uh* and *uhm* in British English“

verfasst von / submitted by

Iris Paur

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magistra der Philosophie (Mag. phil.)

Wien, 2018 / Vienna, 2018

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on  
the student record sheet:

A 190 344 347

Studienrichtung lt. Studienblatt /  
degree programme as it appears on  
the student record sheet:

Lehramtsstudium UF Englisch UF Französisch

Betreut von / Supervisor:

Univ.-Prof. Mag. Dr. Gunther Kaltenböck, Privatdoz.



### **Declaration of Authenticity**

I confirm to have conceived and written this diploma thesis in English all by myself. Quotations from other authors are all clearly marked and acknowledged in the bibliographical references, either in the footnotes or within the text. Any ideas borrowed and/or passages paraphrased from the works of other authors are truthfully acknowledged and identified in the citations.

Signature: \_\_\_\_\_



## **Acknowledgements**

I am sincerely grateful to the many people who have directly and indirectly helped me to complete this thesis.

First of all, I would like to thank my supervisor Univ.-Prof. Mag. Dr. Gunther Kaltenböck, whose encouraging feedback and contagious passion for linguistics have guided and motivated me throughout this final phase of my studies.

I would also like to thank my family, who has always supported and encouraged me in the pursuit of my ambitions. This includes not only my parents and my sister, but also my beloved grandparents, all of whom have always been there for me when I needed them.

Finally, I am eternally grateful to my partner Philipp and my dear friend Lisa, who have carried me through this last phase of my studies by patiently supporting and encouraging me in every possible way.



# Table of Contents

<b>1. Introduction .....</b>	<b>1</b>
<b>2. The linguistic categorization of UH/M .....</b>	<b>3</b>
2.1. Are speakers in control of UH/M?.....	5
2.2. UH/M as disfluencies, pauses, and hesitation markers.....	6
2.2.1. Fluency, disfluencies, and UH/M.....	7
2.2.2. Pauses, hesitation, and UH/M .....	9
2.3. UH/M as English words.....	12
2.3.1. Should UH/M be classified as words? .....	12
2.3.2. UH/M as interjections .....	15
2.3.3. UH/M as pragmatic markers .....	17
<b>3. Are UH/M symptoms or signals? .....</b>	<b>19</b>
3.1. Defining <i>symptom</i> and <i>signal</i> .....	20
3.2. The symptom hypothesis .....	21
3.3. The signal hypothesis .....	22
3.4. UH/M as multifunctional phenomena.....	23
<b>4. The functions of UH/M in conversation .....</b>	<b>25</b>
4.1. The planning function of UH/M .....	25
4.1.1. Complexity, abstractness, and difficulty: task-related determinants of UH/M .....	26
4.1.2. Genre, register, and linguistic conventions as determinants of UH/M .....	29
4.1.3. Syntactic evidence for UH/M's planning function.....	32
4.1.4. UH/M as markers of information retrieval.....	34
4.1.5. UH/M as a marker of lexical retrieval.....	35
4.1.6. Interim conclusion: the planning function of UH/M.....	37
4.2. UH/M and the turn-management system .....	37
4.2.1. UH/M and turn-keeping .....	38
4.2.2. UH/M, turn-taking and turn-accepting.....	39
4.2.3. UH/M and turn-yielding.....	42
4.2.4. Interim conclusion: UH/M and the turn-management system .....	43
4.3. UH/M as correction markers.....	44
4.4. The role of UH/M in speech comprehension.....	46
4.4.1. Do listeners notice UH/M?.....	47
4.4.2. Do UH/M affect speech comprehension? .....	49
4.5. Interim conclusion: The functions of UH/M in conversation .....	53

<b>5. Towards a theory of UH/M.....</b>	<b>54</b>
5.1. Karl Bühler, Roman Jakobson, and UH/M.....	55
5.2. UH/M and the three metafunctions of Systemic Functional Linguistics .....	60
<b>6. Is there a difference between <i>uh</i> and <i>uhm</i>?.....</b>	<b>64</b>
<b>7. Studying functional differences between <i>uh</i> and <i>uhm</i>: an outline of research design .....</b>	<b>67</b>
7.1. Research questions.....	67
7.2. Data acquisition .....	69
7.3. Levels of analysis .....	71
7.3.1. Turns .....	73
7.3.2. The turn-construction unit.....	75
7.3.3. The clause level.....	80
7.3.4. The phrase level .....	83
7.3.5. The word level.....	85
7.3.6. <i>Uh</i> and <i>uhm</i> in self-interruption and repetitions.....	87
7.4. Method.....	88
<b>8. Results .....</b>	<b>89</b>
8.1. <i>Uh</i> , <i>uhm</i> and genre.....	89
8.2. <i>Uh</i> and <i>uhm</i> in turns and TCUs .....	97
8.3. The positions of <i>uh</i> and <i>uhm</i> in clauses .....	103
8.4. The positions of <i>uh</i> and <i>uhm</i> in phrases.....	106
8.5. The lexical/ pausal environment of <i>uh</i> and <i>uhm</i> .....	108
8.6. <i>Uh</i> and <i>uhm</i> in repetitions and self-interruptions.....	112
<b>9. Discussion .....</b>	<b>114</b>
9.1. Does genre affect the production of <i>uh</i> and <i>uhm</i> ? .....	115
9.2. Are there syntactic differences between <i>uh</i> and <i>uhm</i> ?.....	119
9.3. Are there functional differences between <i>uh</i> and <i>uhm</i> ?.....	122
9.3.1. <i>Uh</i> and <i>uhm</i> as markers of delay .....	122
9.3.2. <i>Uh</i> and <i>uhm</i> as signals of thinking and planning .....	123
9.3.3. <i>Uh</i> and <i>uhm</i> as turn-management devices.....	129
9.3.4. <i>Uh</i> and <i>uhm</i> as editing phrases.....	133
9.3.5. Interim conclusion: Are there any functional differences between <i>uh</i> and <i>uhm</i> ? .....	134
9.4. Are <i>uh</i> and/or <i>uhm</i> idiosyncratic phenomena?.....	135
<b>10. Conclusion.....</b>	<b>140</b>
<b>11. References .....</b>	<b>146</b>
<b>12. Appendices .....</b>	<b>155</b>



## Appendices

Appendix 1: ICE-GB corpus data studied in the analysis of <i>uh</i> (segmented into TCUs).....	155
Appendix 2: ICE-GB corpus data studied in the analysis of <i>uhm</i> (segmented into TCUs).....	167
Appendix 3: The position of <i>uh</i> and <i>uhm</i> in clauses .....	186
Appendix 4: Analysis of the precise syntactic positions of phrase-internal <i>uhs</i> and <i>uhms</i> .....	187
Appendix 5: The lexical/pausal environment of <i>uh</i> and <i>uhm</i> .....	188
Appendix 6: <i>Uh</i> and <i>uhm</i> in self-interruptions.....	188
Appendix 7: Individual speaker's production rates of <i>uh</i> and <i>uhm</i> per 1,000 words.....	189
Appendix 8: Abstract (English).....	190
Appendix 9: Abstract (German).....	191

## List of Tables

Table 1: Production rates of <i>uh</i> and <i>uhm</i> in individual corpus texts and genres.....	92
Table 2: Positions of <i>uh</i> and <i>uhm</i> in turns and TCUs.....	97
Table 3: Positions of <i>uh</i> and <i>uhm</i> in turns and TCUs (summary) .....	99
Table 4: Positions of <i>uh</i> and <i>uhm</i> in clauses .....	104
Table 5: Positions of <i>uh</i> and <i>uhm</i> in phrases.....	107
Table 6: Most frequent collocates of <i>uh</i> and <i>uhm</i> .....	110
Table 7: Examples of <i>uhs</i> and <i>uhms</i> occurring in the context of repetitions and self-interruptions .....	112
Table 8: Proportions of <i>uhs</i> and <i>ums</i> appearing in the context of repetitions and self-interruptions .....	113

## List of Figures

Figure 1: The anatomy of self-repairs .....	45
Figure 2: Bühler's <i>Organon model</i> .....	56
Figure 3: Roman Jakobson's model of communication.....	58
Figure 4: Possible positions of <i>uh</i> and <i>uhm</i> in TCUs.....	80
Figure 5: Possible syntactic positions of <i>uh</i> and <i>uhm</i> in clauses.....	82
Figure 6: Example of <i>uh</i> at a noun phrase boundary.....	83
Figure 7: Example of <i>uhm</i> at a verb phrase boundary.....	84
Figure 8: Example of <i>uhm</i> inside a preposition phrase .....	84
Figure 9: Example of <i>uh</i> in a noun phrase conjoin .....	85

Figure 10: Frequencies of <i>uh</i> and <i>uhm</i> per 1,000 words in three different genres.....	90
Figure 11: Distribution of <i>uh</i> and <i>uhm</i> across different genres (in percentages) .....	91
Figure 12: Ranking of <i>uh</i> frequencies per 1,000 words per corpus text (sorted by frequency).....	93
Figure 13: Ranking of <i>uhm</i> frequencies per 1,000 words per corpus text (sorted by frequency).....	94
Figure 14: Individual speakers' production rates of <i>uh</i> and <i>uhm</i> per 1,000 words in different genres.....	96
Figure 15: Positions of <i>uh</i> and <i>uhm</i> in turns and TCUs .....	98
Figure 16: The position of <i>uh</i> and <i>uhm</i> in turns (in percentages) .....	100
Figure 17: The distribution of uhs and uhms across different clausal positions .....	104
Figure 18: Positions of <i>uh</i> and <i>uhm</i> in phrases (in percentages).....	107
Figure 19: The lexical/phrasal environment of <i>uh</i> and <i>uhm</i> .....	109

## List of abbreviations

AUX	auxiliary verb
AJP	adjective phrase
CCID	conjunction, connective, interjection and/or discourse marker
CJ	conjoin
CP	non-transition relevance place completion point
DT	determiner
DU	discourse unit
ICECUP	The International Corpus of English Corpus Utility Program
ICE-GB	The British Component of the International Corpus of English
MOD	modal verb
MVB	main verb
NP	noun phrase
NPHD	noun phrase head
PoA	point of abortion
P	preposition
PP	preposition phrase
SFL	Systemic Functional Linguistics
TCU	turn-construction unit
TRP	transition relevance place
UH/M	<i>uh</i> and <i>uhm</i>
VP	verb phrase

## Transcription Conventions

S1, S2, etc.	Speaker 1, Speaker 2, etc.
//	TRP
/	non-TRP CP
[ ]	aborted TCU
=	point of interruption
<1> [...] </1>	overlap
<,>	short silent pause
<,,>	long silent pause
:	post-completion point expansion (e.g. tag questions)



## 1. Introduction

There are few phenomena in quotidian speech that are as common and yet as condemned by public opinion as the frequent uttering of *uh* and *uhm* – often also transcribed *er* and *erm* in British publications – between the words and phrases that construct our messages in spoken discourse. Despite their ubiquity in spontaneous speech (Tottie 2011: 173; Shriberg 1994: 1), *uh* and *uhm* – henceforth referred to as UH/M – are often deemed “undesirable and unnecessary” (Fox Tree 2001: 320); even worse, speakers producing high rates of UH/M are often judged to be “inarticulate”, “ill-prepared” or “disfluent” (Christenfeld 1995: 173). Consequently, speakers who want to make a good impression on their listeners devote a lot of effort to the elimination of UH/M from their speech, to the point that rhetoricians offer whole courses dedicated to teaching their students how to avoid them in public speaking (Fox Tree 2001: 320; Christenfeld 1995: 172; Clark and Fox Tree 2002: 98-99). Due to their negative reputation, UH/M have received only scarce academic attention until the 1950s, when the gradual introduction of spoken corpora and the increased interest in the descriptive analysis of spoken discourse have caused more and more researchers to investigate the reasons for and determinants of their production (Gilquin and De Cock 2011: 6-8; Rochester 1973: 52). Contrary to popular opinion, the subsequent mass of publications dedicated to this topic produced a significant body of evidence suggesting that UH/M are much more than the random, disruptive production of sounds in between messages they had been commonly assumed to be; instead, scholars have started to realize that UH/M do in fact fulfill numerous functions that are relevant to the accomplishment of smooth and effortless communication, including, amongst others, turn-keeping, turn-taking, self-correction, and the signaling of planning (Kjellmer 2003: 171; Swerts, Wichmann and Beun 1996: 1033; Rühlemann, Bagoutdinov and O'Donnell 2011: 87; see also chapter 4 of this paper).

Despite the large body of corpus-based and experimental evidence suggesting the communicative value of UH/M for the management of spoken discourse, there are still many questions concerning the precise nature and functions of UH/M that have not yet been answered. For instance, while most scholars now seem to agree that the uttering of UH/M *does* fulfill specific purposes such as signaling planning or turn-keeping, no agreement has yet been reached on whether they should be perceived as symptoms of a speaker's cognitive processes or as deliberate signals to the listener. As a consequence, and even though it is now generally assumed that UH/M are more than mere “verbal blunders” (Erard 2004) or “vocal hiccups” (Croucher 2004: 38), researchers remain uncertain about the linguistic classification

of UH/M, suggestions ranging from their classification as pausal phenomena to their categorization as conventional words of the English language. Finally, recent research has raised the question of whether or not the different realizations of UH/M within the same language – the vocalic *uh* on the one hand and the vocalic-nasal *uhm* on the other hand – fulfill different functions in spoken discourse. Interestingly, while the former two questions – concerning the functional and linguistic nature of UH/M – are the subject of a large body of publications, the latter – concerning the potential functional differences between *uh* and *uhm* – is only seldom addressed in existing literature. In an attempt to contribute to the closing of this research gap, the present thesis aims at investigating the question of whether or not *uh* and *uhm* fulfill different functions in conversation. To this end, the following paper will not only comprise a theoretical investigation of the linguistic characteristics and functions that have been commonly associated with UH/M, but also a corpus-based analysis of British English that seeks to explore potential differences between *uh* and *uhm* via the investigation of their syntactic position as well as their distribution across different genres. This analysis will mainly focus on the following three research questions:

- (1) How does genre affect the frequency of *uh* and *uhm*?
- (2) Are there observable differences between *uh* and *uhm* with regard to their preferred syntactic positions, and if yes, what are they?
- (3) How do *uh* and *uhm* differ from each other with regard to the functions they fulfill in spoken discourse?

Since one cannot possibly explore the potential functional differences between *uh* and *uhm* without being aware of the linguistic and functional nature of UH/M in general, the first part of this thesis will provide a comprehensive summary of existing literature on this subject. This theoretical part will be divided into five different chapters. The first one – Chapter 2 – will introduce the reader to the problem of terminology that characterizes the study of UH/M by discussing their conception as disfluencies, pauses, and conventional English words. Afterwards, the Chapter 3 will examine whether UH/M should be perceived as a symptom of a speaker's cognitive processes or rather as a signal to the listener. It is in this chapter that a multi-dimensional notion of UH/M – according to which they can act both as a symptom and a signal at the same time – will first be proposed. This multifunctional view of UH/M will then act as the basis for Chapter 4, which is dedicated to a discussion of the concrete functions fulfilled by UH/M, focusing particularly on their ability to signal a speaker's cognitive processes (cf. section 4.1), their roles in the turn-management system (cf. section 4.2), their

function as editing phrases in self-correction (cf. section 4.3), and their role in speech comprehension (cf. section 4.4). The findings obtained in these three chapters will be summarized in Chapter 5, in which it will be attempted to construct a theoretical framework of UH/M by analyzing the applicability of three theoretical linguistic models – Karl Bühler’s *Organon model*, Roman Jakobson’s model of communication, and Halliday’s model of the three metafunctions of language. This enterprise is deemed particularly important because, at least to my knowledge, there has never before been an attempt to embed UH/M in linguistic models, and I therefore aspire to provide some relevant insights into the mechanisms underlying the production of UH/M. Finally, after this detailed, but generic analysis of UH/M, the fifth and final section of the theoretical part of this thesis (Chapter 6) will be dedicated to a review of the scarcely existing literature on the question of whether or not there are functional differences between *uh* and *uhm*.

Having provided an extensive theoretical background, the second part of this paper is dedicated to the study of syntactic and/or functional differences between *uh* and *uhm* in British English. For the purpose of answering the three research questions presented above, the practical part of this thesis will be divided into four parts. The first part (Chapter 7) will outline the research design underlying the present study; this comprises not only the presentation of the studied corpus, but also the definition of the precise methodology and levels of analysis. After having presented the research design, Chapter 8 offers a detailed description of the obtained results on all defined levels of analysis. These results serve as a basis for Chapter 9, which will discuss the most pertinent findings with regard to the three research questions on the one hand and concerning potential implications for the study of *uh* and *uhm* in general on the other hand. Finally, chapter 10 concludes this thesis by summarizing its main points as well as suggesting possible directions for future research.

## **2. The linguistic categorization of UH/M**

As previously discussed in the introductory section of this paper, the study of UH/M is characterized by a strong heterogeneity of terminology; this seems to be largely due to the fact that scholars have not yet come to agree on the linguistic status and functions of UH/M. While UH/M used to be commonly perceived as mere “nuisance, [...] a kind of debris lying the way of an ordered exposition” (Kjellmer 2003: 170) and have thus been assigned highly depreciative names such as “verbal blunders” (cf. Erard 2008), “vocal hiccups” (Croucher 2004: 38) or even “verbal viruses” (cf. Berkley 2002), the extensive research in

psycholinguistics and corpus linguistics taking place since the second half of the twentieth century (Rochester 1973: 52) has provided evidence for a broad variety of interactional functions fulfilled by the uttering of UH/M (see chapter 4). As a consequence, UH/M have gained a reputation as valuable elements of spoken discourse that perform a broad variety of functions, to the point that some have come to ascribe them functional terms such as “planners” (cf. Tottie 2011: 193) or “speech management phenomena” (Allwood et al. 1990: 1).

While it is highly interesting to observe the vast range of individualistic terminological approaches to this field of study, it would be wrong to conclude that there are no observable tendencies. In fact, the majority of existing publications seem to rely on one of the three widely established terms *disfluencies* (cf., e.g., Gilquin and De Cock 2013), *hesitation markers* (cf., e.g., Gilquin 2008) and *filled pauses* (cf., e.g., Kjellmer 2003), all of which are generally thought to be less pejorative than the examples presented in the above paragraph. However, even though these three terms are highly frequent in the relevant literature, they are far from being mutually accepted; in fact, all of them have been criticized for still being too pejorative, too restrictive, and/or too inaccurate on numerous occasions. For instance, while the term *disfluency* seems inappropriate because “[...] the negative prosody of the prefix *dys-* tacitly portrays pauses, along with a range of similar phenomena, as part of a pathological speech condition” (Rühlemann et al. 2011: 59-60; see also chapter 2.2.1), the expression *hesitation marker* has been criticized for excluding those functions of UH/M that are not associated with the marking of hesitation (Kjellmer 2003: 171; see also chapter 4). Finally, the label *filled pause* has become increasingly unpopular for its implication that UH/M are pausal phenomena rather than words, thus contradicting recent findings on the lexical and communicative status of UH/M (cf. section 2.3).

Even from this very brief introduction, it becomes apparent that the issue of terminology seems to center around the problem of having to find a neutrally connotated term that still reflects its subject’s functions and linguistic nature; however, this enterprise is very challenging because the precise functions and linguistic nature of UH/M have not been reliably identified by now. From the many publications that have attempted to solve this problem, one gains the impression that apart from the analysis of the precise functions fulfilled by UH/M in conversation (which will be provided in chapter 4), the discussion on terminology and linguistic status seems to center around the following three questions: firstly, whether or not speakers can control their uttering of UH/M, secondly, in how far UH/M



contradicts or complements notions of fluency, and, finally, whether or not UH/M should be perceived as conventional English words. The present chapter seeks to answer all of these three questions, with each of them being addressed in a separate section.

## **2.1. Are speakers in control of UH/M?**

The question of whether or not speakers can control their uttering of UH/M has been subject to much debate in recent years and in fact still remains unanswered. While some (and especially earlier) publications seem to assume a certain degree of “speaker passivity” (de Leeuw 2007: 88), a growing body of evidence suggests that the production of UH/M can at least partially be controlled by the speaker (Clark and Fox Tree 2002: 98-99). In general terms, most evidence supporting this view can be ascribed to the frequent observation that a speaker’s production rate of UH/M is not constant but varies across different genres, contexts and registers (Clark and Fox Tree 2002: 98-99). For instance, Schachter et al. (1991: 365) have shown that lecturers tend to produce significantly more instances of UH/M in informal interviews than in their more formal university lectures; similar results were obtained by Duez (1982: 11) and Clark and Fox Tree (2002: 98-99), who found that proficient public speakers such as politicians or television presenters usually produce very few, if any, instances of UH/M when on duty. Of course, the lower frequency of UH/M in non-private genres such as lectures or political speeches could also be linked to the speaker’s thorough preparation rather than his/her conscious control of UH/M; however, it must be equally pointed out that non-proficient speakers, too, were found to generally produce most instances of UH/M in situations in which they did not feel pressured to monitor their speaking style, such as in casual conversations or when talking to themselves in isolation (cf. Broen and Siegel 1972: 219). Along with the mere existence of rhetoric courses that successfully teach their students to avoid UH/M in public speaking contexts, the above observations support the view that speakers seem to be able to at least partially control their uttering of UH/M (Clark and Fox Tree 2002: 98-99).

Interestingly, it has been suggested in several studies that speaker’s do not only adapt their production of UH/M to the respective speaking situation, but also to their audience. In other words, it has been proposed that a speaker’s wish to control his/her production of UH/M is not only influenced by register and genre, but also by the extent to which he/she pays regard to the needs of his/her audience. This hypothesis has been supported, for instance, by Christenfeld and Creager (1996: 45), who have found that the drunker people get, the fewer

instances of UH/M they generally produce; this observation has been ascribed to a certain loss of communicative skills caused by the consumption of alcohol. This view has been summarized by Clark and Fox Tree (2002: 99), who claim that “[...] [w]hen people get drunk, they presumably find it harder to monitor for upcoming pauses, or care less about helping their addressees, or both”. In addition to this finding, it has been suggested that the more the focus of a communicative situation shifts away from the transmission of content to one or more listener(s) and towards other, more style-related foci, such as speech rhythm, the fewer instances of UH/M are produced. This hypothesis has been supported by the finding that speakers produced fewer instances of UH/M when asked to the rhythm of a metronome; moreover, it also accounts for the finding that proficient public speakers produce hardly any instances of UH/M, for these speakers have to cater for the communicative needs of a broad mass of different people and can thus not adapt their speech for individual listeners (cf. Clark and Fox Tree 2002: 98-99).

Summing up, it seems that speakers are indeed able to control their uttering of UH/M, even if they might not be aware of their doing so. As is suggested by the abovementioned studies, it appears that speakers analyze the requirements, foci and objectives of their specific speaking situation and regulate their production of UH/M accordingly. This finding seems to suggest not only that UH/M do indeed fulfill certain communicative functions, but also that speakers are at least subconsciously aware of them; this, in turn, seems to refute the view of UH/M as mere disruptive noises that are void of meaning on the one hand, and, on the other hand, support their conception as words. However, the findings on controllability are only one of the many reasons for which it seems more apt to consider UH/M as words of the English language. This will be illustrated in the following two sections, where section 2.2 will investigate the extent to which UH/M contradicts or complements notions of fluencies, and section 2.3 will discuss the categorization of UH/M as words of the English language.

## **2.2. UH/M as disfluencies, pauses, and hesitation markers**

As was already briefly addressed in the introductory section of this chapter, one of the most common ways UH/M have been referred to in literature is by the name of *disfluencies*. However, this term has frequently been criticized for being too pejorative; as Rühlemann, Bagoutdinov and O'Donnell (2011: 59-60) point out, “the negative prosody of the prefix *dys-* tacitly portrays [silent and filled] pauses, along with a range of similar phenomena, as part of a pathological speech condition”. Still, due to a lack of neutrally connotated alternatives, the term has continued to be broadly used and has subsequently reached widespread acceptance.

In the course of the present section, the view of UH/M as disfluencies will be discussed by providing a brief definition of the terms *fluency* and *disfluency* (cf. section 2.2.1) before moving on to applying these definitions to related perceptions of UH/M, including UH/M as *filled pauses* or *hesitation markers* (cf. section 2.2.2).

### 2.2.1. Fluency, disfluencies, and UH/M

Before being able to discuss the conception of UH/M as disfluencies, it is essential to define the terms *fluency* and *disfluency* first. In the course of time, many scholars have proposed different definitions of fluency, and while it is impossible to cover the discussion associated with the many different conceptions of the term in all its complexity in the context of this paper, it is still interesting to consider some of the most recent developments in this field of research<sup>1</sup>.

In recent years, the definition of *fluency* has experienced a significant shift, moving away from the idea of fluency as *ideal delivery* and towards a more realistic approach that centers around the concept of efficient communication. Especially earlier publications often seem to associate the term *fluency* with a speaker's ability to construct grammatically and lexically flawless utterances that are "akin to artificial spoken dialogue in plays, films, scripted lectures or radio announcements" (Tottie 2016: 116) and that do not expose the listener to pauses, repetitions, UH/Ms or other typical characteristics of spoken discourse (Tottie 2016: 116). However, this former conception of fluency as ideal delivery has gradually come into question, and more recent publications on the topic now often assume that *fluency* does not concern a given message's linguistic accuracy, but rather a speaker's capacity to coherently and efficiently communicate it in a way that allows his/her listener to understand it effortlessly (Hasselgren 2002: 148). One example for this "modern" approach towards fluency can be found, for instance, in Gilquin (2008: 142-143), where the term is defined as a speaker's ability to present his/her message in a way that "[...] makes it possible for the listener to concentrate on what should be central to an utterance, namely its content". In line with this approach, Lennon (1990: 391-392) suggests that "fluency reflects the speaker's ability to focus the listener's attention on his or her message by presenting a finished product rather than inviting the listener to focus on the working of the production mechanisms". Summing up these definitions, it seems that the term *fluency* no longer necessarily designates

---

<sup>1</sup> For a more comprehensive summary of the topic, see especially Gilquin and De Cock (2013) and Gilquin (2008).

the idea of a grammatically accurate delivery that is free from signs of hesitations, planning, or uncertainty, but is rather thought to refer to one's ability to attain the ideal of effortless and efficient communication by facilitating comprehension processes for one's listeners.

One reason for the shift in the definition of fluency described above is the decreasing stigmatization of pauses, UH/Ms and other characteristics of spoken discourse that have formerly been dismissed as mere deviations from ideal delivery. It has been implied above that traditional notions of the term have commonly defined *fluency* over the absence of *disfluencies*, i.e. "phenomena that interrupt the flow of speech and do not add propositional content to an utterance [...] such as long pauses, repeated words or phrases, restarted sentences, and the fillers *uh* and *um*" (Fox Tree 1995: 709). However, the stigmatization of these phenomena as undesirable deviations from a presumed linguistic ideal has increasingly come into question. This is mainly because a growing body of evidence suggests that UH/M and other so-called disfluencies do not only constitute a frequent characteristic of spoken discourse, but also fulfill several functions that serve the process of speech production and turn-management<sup>2</sup> (Stenström and Svartvik 1994: 242; Tottie 2014: 26). In addition, it has also been pointed out in several publications on the topic of fluency that it is neither possible nor meaningful to avoid the production of (breathing) pauses, repetitions or UH/Ms in spoken language (O'Connell and Kowal 2005: 557); this is due to two main reasons. On the one hand, it has been suggested that the absence of such phenomena hinders efficient communication because listeners find it difficult to understand longer stretches of ideal delivery (O'Connell and Kowal 2005: 557-558). On the other hand, it is assumed that the creative processes involved in speech production cannot always occur simultaneously to the speaking process itself; as a consequence, speakers need to pause, rephrase, and hesitate in order to think, plan, or choose the right words for conveying their messages (Chafe 1980: 170; O'Connell and Kowal 2005: 556-557). Following this line of argument, it seems that UH/M and other phenomena that have formerly been referred to as *disfluencies* do not contradict, but even complement modern notions of fluency, because they allow for a more listener-friendly transmission of content (Tottie 2014: 26).

Considering the above findings, it seems inappropriate to use the term *disfluency* to designate UH/M. Following the more modern notion of *fluency*, which refers to a speaker's ability to transmit meaning in a the most listener-friendly way possible (see above), it appears that UH/M are not a threat to fluency; in fact, they might even contribute to the development of

---

<sup>2</sup> See chapter 4 for a detailed discussion of the various functions fulfilled by UH/M in spoken discourse.

fluent communication by helping speakers to efficiently handle the cognitive challenges involved in on-line speech production on the one hand and, on the other hand, “giv[ing] listeners time to take in the message that the speaker wishes to transmit in a way that more elegant delivery often does not” (Tottie 2014: 26). However, this applies only to contexts in which UH/M is moderately used, for an excessive production of UH/M might “invit[e] the listener to focus on the working of the production mechanisms” (Lennon 1990: 391-392) rather than content and thus inhibit the development of efficient, fluent communication (Tottie 2014: 26). In conclusion, the above discussion can be summarized by citing Tottie (2014: 26), who asserts that “[...] [u]nless it is overused or used in the wrong context, [...] *uh* and *um* serve the function of making speech more fluent and less disrupted. They may deserve to be called markers of fluency rather than signs of disfluency”.

Due to the increased awareness of the inappropriateness of the term *disfluency*, many have chosen to use other terms to designate UH/M instead, some of the most popular alternatives being the terms *filled pauses*, *fillers*, and *hesitation markers*. The following section briefly discusses each of these terms with regard to their appropriateness for the study of UH/M.

### **2.2.2. Pauses, hesitation, and UH/M**

Despite the large body of evidence refuting the conception of UH/M as disfluencies, the term continues to be used on a regular basis, mainly due to a lack of appropriate alternatives. Up to this point, scholars refusing to use the term *disfluency* while also opposing the perception of UH/M as conventional English words have proposed the terms *filled pauses*, *fillers*, and *hesitation markers*; however, neither of these options seems to be a satisfactory solution, since all of them fail to comprise the many different functions of UH/M that have been suggested by various studies (see chapter 4). The present section aims at illustrating the main points of criticism associated with each of the three abovementioned terms.

If all of the above terms have been criticized for various reasons to be outlined below, the conception of UH/M as *filled pauses* is particularly problematic because it imposes the assumption that UH/M are to be equated with silent pauses (Clark and Fox Tree 2002: 75). The categorization of UH/M as pausal phenomena is inappropriate insofar as a substantial body of evidence now suggests that UH/M and (silent) pauses are “very different creatures” (Christenfeld 1994: 193). This discrepancy between silent and “filled” pauses has been illustrated by findings suggesting that the frequency of UH/M and silent pauses seem to be

determined by different factors<sup>3</sup> (cf. Goldman-Eisler 1968, cited in Christenfeld 1994: 193; Mahl 2014[1987]: 277). Moreover, it has also been proposed on numerous occasions that the uttering of *uh* and *uhm* in lieu of a silent pause is by no means random; instead, UH/M have been argued to fulfill specific communicative functions that cannot be performed by silent pauses, e.g. turn-holding or the explicit signaling of delay in periods of excessive pausing (MacLay and Osgood 1959: 41-42). As a consequence of these findings, the term *filled pause* has been increasingly criticized for not sufficiently conveying UH/M's functional value on the one hand and belittling them as pauses on the other hand (Clark and Fox Tree 2002: 75); for this reason, many researchers have ceased to use the term in their publications.

In an attempt to overcome the problems associated with the classification of UH/M as *filled pauses*, certain scholars have proposed to use the more neutral term *filler* instead; however, this term, too, has been criticized on numerous occasions for inaccurately reflecting the linguistic nature and functions of UH/M. In that sense, Tottie (2011: 193) argues that “[f]iller’ is a rather negative and uninformative default term that is dependent on the ideal of fluency and that says nothing about the discourse functions of [UH/M] [...]”. Moreover, it should be pointed out that *filler* seems to be nothing more than a mitigated version of its predecessor *filled pause*, especially because it, too, presents UH/M as a vocalization that *fills* periods of emptiness. This becomes especially obvious when considering dictionary entries, which define the term *filler* as something that is “[...] used to cover or *fill in a space* between two parts of a structure [...] [my emphasis]” (Merriam Webster: online) on a general level, or, with regard to its linguistic meaning, as “a sound, word, or phrase [...] used to *fill pauses* in speaking [my emphasis]” (Merriam Webster: online). In other words, the use of the term *filler* constitutes a questionable approach to the issue of terminology, for the mere avoidance of the word *pause* provides only a very superficial solution to the problem.

With the terms *disfluency*, *filled pause*, and *filler* having increasingly become subject to criticism, researchers were confronted with the need to find a term that reflects the view of UH/M as non-words without denying their functions in spoken discourse; as a consequence, the term *hesitation marker* has become increasingly popular. One reason for this trend

---

<sup>3</sup> Even though the precise differences between the determinants of silent and filled pauses have not been reliably identified up to this point, it has been suggested by several scholars that the production of these two phenomena are influenced by different factors. For instance, Goldman-Eisler 1968, cited in Christenfeld 1994: 193) found that the manipulation of the abstractness of a given speaking task produced an effect on the production rates of silent pauses, but not of UH/M. Moreover, Mahl (2014[1987]: 277) has suggested a link between silent pauses and emotional factors that could not be observed in the case of UH/M.

towards describing UH/M as hesitation phenomena rather than pauses is probably the fact that the term *hesitation*, usually referring to a momentary delay or a short period of indecision (cf. Merriam Webster: online), seems to be a convenient umbrella term for the various planning functions that have been associated with the uttering of UH/M. According to this view, then, UH/M are inserted in periods of hesitation for the sake of turn-keeping or the representation of planning problems; in that sense, it sets itself apart from the pausal conception by assuming that UH/M transmit meaning to at least some extent.

Even though it seems significantly more appropriate than many of its alternatives, the term *hesitation marker*, too, has been criticized on several occasions. For instance, Kjellmer (2003: 189) has pointed out that the name *hesitation marker* is misleading because there is now a vast body of research suggesting that UH/M are multi-functional phenomena that cannot be reduced to the marking of hesitation. In addition, the term *hesitation*, too, has often been associated with those concepts of pausing and (dis-)fluency that have been argued to be inappropriate to the study of UH/M (cf. section 2.2.1). This can be observed, for instance, when considering Lounsbury's (1954: 98) statement that "[h]esitations *which interrupt the continuous flow of speech* are anything from very brief pauses to expended periods of halting, often filled with 'hemming and hawing' [my emphasis]" seems to suggest not only that UH/M are merely phonetic phenomena that carry hardly any meaning, but also that that they are disruptive disfluencies that contradict the notion of fluent interaction.

Even though the term *hesitation marker* has been criticized for associating UH/M with pausal phenomena and disfluencies (see above), it is equally important to point out that not all researchers evoke this connection deliberately. This becomes obvious when considering Gilquin (2008: 120), who equates UH/M with pausal phenomena and disfluencies on the one hand, but, on the other hand, also stresses that the signaling of hesitation "[...] is crucial as a conversational strategy" because it allows speakers to signal an upcoming delay without having to fear a turn-taking attempt by their interlocutors. This approach towards the term *hesitation marker* seems to relativize the above discussion insofar as it adds a more positive tone to the general conception of *hesitation*. Still, it can be concluded that even though the term *hesitation marker* constitutes the lesser of terminological evils in the conception of UH/M as non-lexical phenomena, its rather negative connotation as well as its misleading simplicity render it only marginally appropriate to the study of UH/M. Since all systems of classifying UH/M as non-lexical phenomena seem to be inappropriate because they fail to

account for many of their characteristics, some scholars have gradually come to consider their classification as English words instead; this will be the subject of the following section.

### **2.3. UH/M as English words**

With research providing a growing body of evidence supporting the conception of UH/M as markers of fluency rather than disfluencies (Tottie 2014: 26), some have raised the question of whether they should be classified as words of the English language (Clark and Fox Tree 2002: 73). This suggestion has caused a considerable outcry in the research community, and many have criticized the lexical approach for various reasons to be mentioned below. Still, there now seems to be a general trend towards recognizing UH/M as words; this trend is observable even amongst those researchers that had initially been highly critical of the idea (O’Connell and Kowal 2005: 573). The following sections outline popular arguments both for and against the lexical conception of UH/M before discussing the categorization of UH/M as interjections on the one hand (cf. section 2.3.2) and as pragmatic markers on the other hand (cf. section 2.3.3).

#### **2.3.1. Should UH/M be classified as words?**

In recent years, many notable researchers have come to advocate the classification of UH/M as words of the English language, pointing out that they share many characteristics with other “conventional English words” (Clark and Fox Tree 2002: 73). In that sense, some have argued that the relatively recent inclusion of *uh* and *uhm* in dictionaries seems to suggest that UH/M are verbal phenomena that carry a fixed meaning (Tottie 2011: 175-176; Gilquin and De Cock 2013: 13). The idea that speakers of English understand UH/M as words transmitting a specific message can also be observed with regard to their occasional use in written language, where *uh* and *uhm* are used to highlight hesitation (Tottie 2016: 99). In addition to these functional characteristics, Clark and Fox Tree (2002: 104) have argued that the frequent cliticization of UH/M, i.e. their intonational attachment onto preceding words such as *and* or *but* to create constructions such as *anduh* or *butuh*, seems to further support their conception as words. This argumentation is based on the assumption that *uh* and *uhm* could not be adapted to the prosody of a given utterance if they were not words (Clark and Fox Tree 2002: 104). Finally, it has been pointed out in several publications that the phonetic realization of *uh* and *uhm* are language-specific, which has led many to argue that UH/M should be regarded as (translatable) words rather than mere noises (de Leeuw 2007: 89); this argument will be further explored in the following paragraph.



One of the most popular arguments presented in favor of the lexical conception of UH/M is that *uh* and *uhm* are not in fact unique to English but are rather language-specific realizations of a phenomenon that can be found in most other languages as well; as a consequence, it has been argued that they should be classified as translatable words that form part of a language's lexicon. Indeed, cross-linguistic studies have revealed that markers of hesitation and planning are not universally realized as *uh* or *uhm* but rather in a way that renders them conveniently accessible for speakers of the respective language. In that sense, speakers of German do not usually utter *uh* and *uhm*, but *äh* and *ähm*, while French speakers produce *eu*, *eu**h* or *oe* and speakers of Swedish use *eh*, *äh*, *m*, *oh*, and *a* (Clark and Fox Tree 2002: 91). The hypothesis that UH/M form part of a language's lexicon can be further supported by considering that there are significant differences between their realizations in Indo-European languages and their counterparts in languages of different descents; this can be seen, for instance, when comparing the above examples to the Japanese hesitation markers *eto*, *ano*, *kono* and *sono* (Clark and Fox Tree 2002: 92). What is more, it should be pointed out that even within the large category of Indo-European languages, hesitation markers are not restricted to vocalic or vocalic-nasal realizations such as *uh* and *uhm*; instead, other words such as demonstratives (e.g. the Spanish *este*) or discourse markers might also be used to fulfill functions that are very similar to those accomplished by UH/M (Tottie 2011: 176). Finally, the view that UH/M form part of a language's lexicon is further supported by Clark and Fox Tree's (2002: 93) finding that native speakers seem to notice deviations from the usual realizations of *uh* and *uhm* (or the respective language's counterparts), asserting that "[s]peakers of English as a second language often import the fillers from their first language [...] and that is one reason they continue to be heard as non-native speakers". Together with the above findings, this observation suggests that UH/M cannot be directly transferred from one language to another, which in turn supports the view that UH/M form part of the respective language's lexicon.

Despite the above arguments, there is also a considerable body of literature contradicting the view that UH/M should be categorized as English words. One of the most popular arguments advanced in this context is that *listeners* do not seem to perceive UH/M as words of the English language. This has been proposed, for instance, by Kjellmer (2003: 190-191), who has pointed out that listeners often fail to reproduce instances of UH/M if asked to repeat a sentence they had just heard before; similarly, Lickley and Bard (1996: 1876) have provided evidence that even proficient transcribers find it difficult to locate each and every occurrence of UH/M. Based on these findings, it has been argued that UH/M are not processed in the

same way as conventional words but rather seem to “[...] operate below the level of consciousness and can therefore be an unobtrusive and effective instrument in facilitating spoken language”. In addition to this listener-based perspective, Corley and Stewart (2008: 589) have criticized the lexical conception of UH/M by pointing out that “there is little evidence that they are intentionally produced”. However, all of these arguments can be refuted for the simple reason that “even bona fide words like *well* or *you know* used in similar functions are rarely produced intentionally, and [that] they would probably not be included if someone was asked to repeat an utterance containing one of them” (Tottie 2011: 176). Consequently, it could be argued that *uh* and *uhm* should only be perceived as non-lexical elements if the lexical status of other expressions such as *well* or *you know*, which have been found to share many characteristics with UH/M especially with regard to their production and functions, was also questioned (cf. Tottie 2016: 116, Clark and Fox Tree 2002: 79).

As can be inferred from the previous paragraph, one of the main issues related to the discussion on the lexical status of UH/M seems to be the understanding of the term *conventional word* itself. It has been stated above that some have argued against the categorization of UH/M as words because they do not generally contribute to the primary meaning of a message (cf. Kjellmer 2003: 190-191); however, as has equally been discussed in the present section, this argumentation raises the question of why pragmatic markers and interjections such as *well*, *you know* and *I mean* have been granted lexical status despite not influencing the primary meaning of the message under construction (Tottie 2011: 176). One possible solution to this problem has been proposed by Clark and Fox Tree (2002: 105-107), who have classified UH/M – together with interjections and pragmatic markers – as *collateral signals* that allow a speaker to comment on his/her performance or cognitive/emotional state without adding to the primary message under construction. According to this view, UH/M deserve to be equated with the lexical categories of interjections or pragmatic markers because they carry meaning on the ongoing speech production process and thus act as valuable sources of information for the listener (cf. also chapter 4). This approach is highly interesting insofar as it allows for a lexical perception of UH/M that is able to account for the counterarguments presented in the above paragraph. Naturally, the successful categorization of UH/M as words would require significantly more theoretical and empirical research; however, for the purpose of this paper, it is considered sufficient information that there are good reasons to believe that UH/M are words of the English language that can be equated with other collateral word types such as interjections and pragmatic markers. For this reason,

the following two sections will discuss the classification of UH/M as either of these two word types.

### 2.3.2. UH/M as interjections

It has already been mentioned above that those scholars assuming a lexical status of UH/M have frequently counted them towards the category of interjections. One of the first to propose this classification was probably James (1972; cited in Clark and Fox Tree 2002: 76), who described *uh* as interjections that are applied in order to comment on one's current performance (Clark and Fox Tree 2002: 76). This approach was then further elaborated by Clark and Fox Tree (2002), who have argued that *uh* and *uhm* should be regarded as interjections that signal the initiation of a short or long delay respectively (Clark and Fox Tree 2002: 79). The present section discusses the categorization of UH/M as interjections by exploring whether or not they conform to the defining characteristics of this word class.

It has been pointed out in several publications on the topic that interjections are extremely difficult to define (cf., e.g., Heine et al. 2013: 170-171). However, a consideration of relevant literature reveals that scholars seem to generally agree that the words classified as interjections generally share the following seven characteristics:

- 1) They are associated with spoken rather than written discourse (O'Connell and Kowal 2005: 568);
- 2) They can appear anywhere in an utterance and are thus "syntactically independent from their environment", even though they seem to prefer utterance-initial positions (Heine et al. 2013: 172);
- 3) They are usually prosodically accentuated, not necessarily through deliberate pausing, but through other emphatic prosodic means such as increases in loudness, speed of delivery or intonation (Heine et al. 2013: 172; O'Connell, Kowal and Ageneau 2005: 166);
- 4) They may form an utterance on their own (Clark and Fox Tree 2002: 76);
- 5) They often indicate "a change in the emotional or cognitive state of the speaker" (Heine et al. 2013: 171)<sup>4</sup>;

---

<sup>4</sup> With regard to point five, it needs to be further pointed out that "index[ing] a change in the emotional or cognitive state of the speaker" (Heine et al. 2013: 171) is not the only function of interjections; instead, they have been found to also be able to fulfill phatic functions, which help a speaker to express his/her attitude towards another speaker's ongoing discourse (e.g. by uttering *mmm* or *yeah*) as well as conative functions,

- 6) They are multifunctional, i.e. their basic meaning(s) can be shaped by context to fulfill other functions or convey other messages as well (Clark and Fox Tree 2002: 77); and, finally,
- 7) They are defined via conventional practices, i.e. the settings they are used in, rather than paraphrases (Clark and Fox Tree 2002: 76-77)

At first glance, the above list of defining characteristics of interjections seems to a large extent applicable to *uh* and *uhm*. On a very general level, UH/M conform to points one and seven of the above list because they have been found to be linked to spoken rather than written discourse on the one hand (O’Connell and Kowal 2005: 568) and, on the other hand, are usually defined via the situations they are used in rather than by paraphrases (cf. Merriam Webster online dictionary, where *uh* is defined as being “*used to express hesitation [my emphasis]*”). Furthermore, as concerns their purposes in spoken discourse, UH/M have been shown to fulfill many different functions that all center around the basic meaning of “I am thinking” (Fischer 2006: 432; Tottie 2016: 115; see chapter 4), thus conforming to point six of the above list of defining characteristics of interjections. Finally, UH/M have been argued to conform to certain syntactic and prosodic characteristics that are typical of English interjections; this includes their ability to appear in any syntactic position (cf. point 2; Heine et al. 2013: 172), as well as the fact that “they are normally delivered with a parenthetical intonation, a monotone pitch that allows them to be segregated from the melody of the surrounding construction” (cf. point 3; Clark and Fox Tree 2002: 104).

Despite the above evidence, some scholars – and particularly O’Connell and Kowal (2005) – have pointed out several problems associated with the classification of UH/M as interjections. According to them, UH/M cannot be categorized as interjections because “[c]haracteristically, interjections do reflect emotion, whereas *uh* and *um* do not” (O’Connell and Kowal 2005: 568). However, this argument can be refuted by considering that research on the nature of interjections has suggested that the latter do not necessarily have to indicate emotional changes but can also fulfill functions such as focusing an audience’s attention towards a current speaker, expressing one’s cognitive status or managing spoken discourse (Heine et al. 2013: 171-172). Moreover, O’Connell and Kowal (2005: 572) have presented the irrefutable argument that UH/M, unlike most other interjections, cannot stand alone to form an

---

which are important for focusing the listener’s attention or influencing a listener’s behavior (e.g. by uttering *pst!* in order to calm an audience and focus their attention on oneself) (Heine et al. 2013: 172).

independent turn, and if they do, they are usually not “categorizable as a turn, [but] must be categorized as an unsuccessful interruption”<sup>5</sup> (O’Connell and Kowal 2005: 572). Indeed, given that the ability to form an autonomous turn on its own has very often been argued to constitute a defining characteristic of interjections (cf., for instance, Clark and Fox Tree 2002: 76; O’Connell and Kowal 2005: 572), the fact that UH/M are only seldom found to be uttered independently from a turn or a turn-taking attempt seems to be a major obstacle in their classification as interjections (O’Connell and Kowal 2005: 572). Consequently, this approach at linguistic categorization, despite having been supported by many valid arguments, does not seem to be an entirely satisfactory solution, and scholars have thus started to look for other potential ways to classify UH/M. One of the most popular such ways is the categorization of UH/M as pragmatic markers, which will be the subject of the following section.

### 2.3.3. UH/M as pragmatic markers

Given the abovementioned criticism of the classification of UH/M as interjections, some have proposed to count them towards the group of *pragmatic markers* or *discourse markers* instead. The two terms *pragmatic marker* and *discourse marker* have been used synonymously by some (e.g. Alami 2015: 1), while others have argued the latter to constitute a subcategory of the former (e.g. Fraser 1996: 187); in the course of this paper, I wish to follow Tottie (2016: 98) in using the expression “*pragmatic marker* as an umbrella term, reserving the term *discourse marker* for text-structuring functions” (cf. also Fraser 1996: 169 for similar practice). Generally speaking, pragmatic markers have been defined as elements that “have pragmatic meaning (e.g. interpersonal or textual meaning) and do not contribute to the content” (Archer, Aijmer and Wichmann 2012: 74). In other words, pragmatic markers “usually consist of lexical items whose literal meaning has been bleached” in a way that they no longer contribute to their surrounding message’s propositional content (Tottie 2016: 99, cf. also Fraser 1996: 169) and may thus be perceived as “separated and distinct” (Fraser 1996: 169) extra-clausal elements that are also syntactically independent from the rest of the utterance (Alami 2015: 7). Moreover, it has been pointed out on numerous accounts that pragmatic markers are typically “[...] highly variable and context-bound in their meaning” (Kaltenböck, Keizer and Lohmann 2016: 10). These and other defining features of pragmatic markers presented in relevant literature have been summarized by Alami (2015: 7), who

---

<sup>5</sup> Note that this observation is also confirmed in the data obtained in the empirical study conducted for the purpose of the present paper, cf. section 8.2.

concludes that pragmatic markers seem to be roughly characterized by the following nine properties:

1. They are almost used in all languages.
2. They are syntactically independent.
3. They are syntactically flexible, i.e. they may appear at the beginning, in the middle or at the end of an utterance. This flexibility contributes to their enormous usefulness and high frequency in discourse.
4. They do not affect the propositional meaning of [an] utterance.
5. They make no contribution to the informational content of discourse.
6. They deal with the pragmatic aspects of discourse
7. They are meaningful but non-truth conditional [i.e. “its truth or falsity is not relevant to the truth or falsity of the sentence use implicating it” (Richard 2005: online)]
8. They are multifunctional,
9. They are short, consisting of one to three syllables.

(Alami 2015: 7)

Using this comprehensive list as a starting point for the present discussion, it seems that there are good reasons to categorize UH/M as pragmatic, for they appear to fulfill most, if not all, of the points listed above. For instance, *uh* and *uhm* conform to point nine by consisting of no more than one or two phonemes respectively; moreover, they seem to fulfill points four, five, and seven because they do not influence the “propositional meaning” or truth of their surrounding utterance. In addition, as has been argued in chapter 2.3.1, many languages have been found to dispose of items that are phonologically and/or functionally similar to the English UH/M; thus, they also conform to point one. Furthermore, with regard to point 3, it has been found that UH/M are indeed “syntactically flexible” insofar as they may appear almost anywhere in an utterance, even though they display certain tendencies to appear at or near the onsets of turns or syntactic units (Kjellmer 2003: 174; see also chapter 4.1.3). Moreover, in line with point six of the above list, UH/M have been suggested to help interlocutors to negotiate meaning on the ongoing processes of speech production and interaction without having to disrupt the transmission of primary content, thus fulfilling a substantial function in interpersonal communication (cf., for instance, de Leeuw 2007: 110-111). Finally, UH/M have been argued to be multifunctional insofar as their meaning is highly context-dependent (cf. chapter 4.2), even though all potential meanings seem to center around the core meaning of “I am thinking” (Fischer 2006: 432; Tottie 2016: 115). For these reasons, it seems that Tottie (2016: 115) does seem to have a point in claiming that UH/M “should be characterized as pragmatic markers used by speakers to further plan the rest of their turn, with the core meaning ‘I’m thinking’”. Lastly, it should be pointed out that while one of the major drawbacks of a classification as interjections was the fact that UH/M cannot be usually found to construct an autonomous turn on their own, this problem does not arise in the case of

pragmatic markers, for the latter, by definition, usually serve “text-structuring function” (Tottie 2016: 98) and are thus, too, mostly embedded in turns.

Summing up the above discussion, it has been shown that it seems plausible to consider UH/M as words of the English language, even if “[t]he legitimation of *uh* and *uhm* as words is far from accomplished” (O’Connell and Kowal 2004: 469). Concerning the precise lexical categorization of UH/M, it has been found that there is substantial evidence in support of both their classification as interjections and as pragmatic markers; however, the fact that UH/M do not share the ability of interjections to construct independent turns leads me to prefer the latter (cf. also Tottie 2016). Nevertheless, it should equally be pointed out that there is not necessarily one definite, “correct” answer to the question of lexical classification, for, as Heine et al. (2013: 180) point out:

[...] linguistic categories are generally discontinuous and gradient rather than discrete. [...] There are various semantic, syntactic and morphological overlaps, with the effect that boundaries are notoriously fuzzy and some information units can be related to more than one category. For example [...], hesitation markers such as *uh* and *um*, which we loosely classified as interjections, have for equally good reasons been treated as discourse markers like *you know* and *like* [...]

In other words, the fact that linguistic categories are not usually discrete and clear-cut suggests that the perception of UH/M as interjections by no means excludes the possibility of their use as pragmatic markers, and vice versa. For this reason, it seems that one should depart from the enterprise of trying to identify the one “correct” linguistic category UH/M should be associated with, and instead focus on other questions such as the functions they fulfill. Therefore, the examination of linguistic classification will be closed at this point so as to move on to more functionally-based interrogations, such as whether UH/M serve the speaker, the listener, or both at the same time; and it is precisely this question which shall be addressed in the following chapter.

### **3. Are UH/M symptoms or signals?**

It has already been stated on numerous occasions throughout this paper that no undisputed theory on the functions fulfilled by UH/M in conversation exists up to this point. Even if the majority of scholars seem to agree nowadays that UH/M serve a linguistic purpose, the precise nature of this purpose has, for a long time, been subject to vivid academic discussion. While some have argued that UH/M merely constitute a by-product of the speaker’s cognitive processes associated with speech production (an approach which is commonly referred to as *the symptom hypothesis*; cf. de Leeuw 2007: 96; see also chapter 3.1), others have proposed

that they act as communicative signals to the listener that convey specific messages (*the signal hypothesis*; cf. de Leeuw 2007: 87; see also chapter 3.2). Finally, a third and very recent category has proposed a reconciliatory approach according to which UH/M serve both the speaker and the listener in reaching their common goal of efficient communication. The present chapter seeks to outline all three of these approaches, with the ultimate goal of providing the reader with a profound understanding of the multifunctional conception of UH/M that will be highly relevant to the further course of this paper.

In order to provide a comprehensive overview on the different approaches towards symptom-versus-signal debate, the present chapter will be structured into four main parts: firstly, an attempt at providing a working definition of the terms *symptom* and *signal* (cf. section 3.1), secondly, a discussion of the symptom hypothesis (cf. section 3.2), thirdly, an investigation of the arguments presented for and against the signal hypothesis (cf. section 3.3), and, finally, an insight into the multifunctional conception of UH/M, according to which they are both symptoms of a speaker's cognitive processes *and* a signal to the listener (cf. section 3.4).

### **3.1. Defining *symptom* and *signal***

In recent years, the symptom-versus-signal debate seems to have reached an impasse that is largely caused by the fact that researchers do not seem to agree on the definitions and implications of its two central terms *symptom* and *signal*. This becomes particularly obvious when considering the different approaches to the classification of UH/M as *symptoms*. While many adapt a *functional* approach towards the term by using it to designate UH/M's ability to signal a speaker's cognitive processes (cf. de Leeuw 2007: 10-11), others define it as an “*automatic, or involuntary, consequence of one or another process in speaking [my emphasis]*” (Clark and Fox Tree 2002: 75-76), thus suggesting that UH/M are mere by-products of the speech production process that do not serve any communicative purpose. According to this latter approach, UH/M – if classified as symptoms – cannot be controlled by the speaker; as a consequence, its advocates have interpreted evidence on the controllability of UH/M as presented in section 2.1 as contradicting the symptomatic perception of UH/M (Clark and Fox Tree 2002: 75). If UH/M cannot be classified as symptoms because of evidence supporting its controllability, the logical consequence would be that they are *signals*, which have been associated with a certain extent of intentionality and functionality (de Leeuw 2007: 88). However, the perception of UH/M as signals to one's listener(s) raises new issues, for Broen and Siegel's (1972: 219) aforementioned finding that speakers tend to utter UH/M



even when talking to themselves in isolation seems to contradict the assumption that UH/M are always directed towards an audience.

The problems associated with the words *symptom* and *signal* outlined above illustrate the necessity of providing precise working definitions of the two terms. For the purpose of this paper, it was decided to follow those scholars who have defined the terms *symptom* and *signal* as functional dimensions rather than categories related to linguistic production; this approach is convenient not only because it allows for a discussion that is focused on the *functions* of UH/M rather than their linguistic production, but partly also because it is in line with other theoretical models of the linguistic signs and especially with Karl Bühler's *Organon model* (cf. section 5.1) that will be applied to construct a model of UH/M as multifunctional phenomena in chapter 5 of this paper. Moreover, the perception of *symptoms* and *signal* as functional dimensions allows the construction of a multidimensional model of UH/M in which they may serve both a symptomatic, i.e. expressive function *and* a signaling function; this view of UH/M will be discussed in detail in section 3.4. Having provided a working definition of the terms *symptom* and *signal*, the present chapter will now move towards the discussion of the different approaches associated with the symptom-versus-signal debate.

### **3.2. The symptom hypothesis**

Especially in earlier publications, the so-called symptom hypothesis constitutes one of the most popular approaches to the study of UH/M, and it is still widely recognized today (de Leeuw: 2007: 86). Already its most central term, *symptom*, suggests that the approach perceives UH/M as “something that indicates the existence of something else” (Merriam Webster: online); adapted to the study of UH/M, this means that advocates of the symptom hypothesis regard UH/M as representations of the cognitive and creative processes involved in speech production (de Leeuw 2007: 86-87). Indeed, there is a vast body of evidence suggesting that UH/M can at least partially be linked to a speaker's cognitive efforts related to several levels of the speech production process; the present section constitutes only a brief overview on the most relevant findings, while a comprehensive discussion of related evidence will be provided in section 4.1 of this paper.

Generally speaking, the majority of evidence centers around the finding that UH/M seem to be particularly frequent in contexts that are thought to impose an increased cognitive load on the speaker. For instance, it has been shown that the frequency of UH/M has been shown to increase with the complexity and abstractness of a given speaking task (Reynolds and Paivio

1968: 172; Lay and Paivio 1969, cited in Rochester 1973: 65; see also section 4.1.1). Moreover, there is now also considerable evidence that UH/M tend to occur at the boundaries of syntactic units such as clauses or phrases rather than inside them, leading to the hypothesis that the uttering of UH/M is closely linked to syntactic planning and especially to the range of syntactic options available at a certain point (cf. Hawkins 1971: 277; Schachter et al. 1991: 362, Corley and Stewart 2008: 590; see also section 4.1.3). Finally, the hypothesis that UH/M are representations of a speaker's cognitive processes is further supported by the finding that they are often associated with points of lexical access and information retrieval (Maclay and Osgood 1959: 32-33; Rühlemann, Bagoutdinov and O'Donnell 2011: 82, cf. also section 4.1.5). In a different line of argument, the symptom hypothesis has also been supported by arguments refuting the idea that UH/M function solely as signals to a speaker's audience. This is particularly based on the finding that UH/M are also frequently produced in non-dialogic genres such as university lectures, which do not usually allow for spontaneous interruptions and should thus not be expected to include frequent signals of turn-keeping and turn-taking intentions (Schachter et al. 1991: 365). Consequently, it has been argued that the production of UH/M cannot exclusively be linked to turn-management, but also appears to be indicative of a speaker's cognitive processes (de Leeuw 2007: 86).

In summary, the above findings seem to support the assumption that UH/M act as representations of a speaker's cognitive processes involved in speech production; as a consequence, it appears impossible to entirely refute the symptom hypothesis. Still, in recent years, it has become fashionable to assume that UH/M are not mere symptoms, but (also) function as signals to a speaker's audience; this approach will be briefly explored in the following section.

### **3.3. The signal hypothesis**

In the course of the past years, it has been suggested on numerous occasions that UH/M cannot be reduced to the mere indication of a speaker's cognitive processes but instead seem to fulfill pragmatic and communicative functions that shape a listener's understanding and interpretation of a given contribution. Subsequently, many linguists have gradually started to criticize the perception of UH/M as a mere expression of a speaker's planning for the sake of gaining time, instead focusing on a notion of UH/M that treats them as signals to the listener that may communicate messages such as "I am thinking" (Tottie 2016: 115) or "I'm still in control – don't interrupt me!" (Maclay and Osgood 1959: 41).

One of the most frequent arguments advanced in support of the signal hypothesis is that there is now substantial evidence that the uttering of UH/M is often related to a speaker's attempt at turn-holding. This has been hypothesized, among others, by Maclay and Osgood (1959: 42), who assume that speakers produce instances of UH/M in order to "keep control of the conversational 'ball'" during pauses in speech production; empirical evidence in support of this hypothesis stems, for example, from Künzel (1997: 58; cited in de Leeuw 2007: 86), who observed that speakers usually produced higher rates of UH/M when talking on the telephone than in face-to-face conversations. This observation leads him to the hypothesis that speakers utter UH/M more frequently when lacking visual means to signal the intended continuation of the turn, thus establishing a direct link between turn-keeping and the uttering of UH/M. Paradoxically, however, UH/M has not only been found to be a signal of turn-holding but also of turn-taking, and, perhaps even more surprisingly, turn-yielding (Kjellmer 2003: 185-186). This has been proposed, amongst others, by Fischer (2006: 432) and Clark and Fox Tree (2002: 90), who have suggested that depending on its position inside a turn, UH/M may signal either one's intention to react to something that has been said (i.e. turn-taking, or, turn-accepting; cf. section 4.2.2) or, alternatively, that one has run out of things to say. Apart from these turn-management functions, Kjellmer (2003: 187) has also proposed that UH/M may act as a signal of "an important, semantically, heavy element in the delivery that is about to follow". The idea that UH/M could signal all of the above seems paradox at first, but has been confirmed by many (cf., for instance, Clark and Fox Tree 2002: 73; Kjellmer 2003: 185-186); thus, it seems that UH/M do indeed serve as signals to the listener, but that their meaning is highly context-dependent and can only be decoded through a consideration of their immediate situational and conversational context.

### **3.4. UH/M as multifunctional phenomena**

Given the vast body of evidence supporting both the symptom and the signal hypotheses outlined above, it seems impossible to refute either of them; therefore, it has been suggested that UH/M may in fact be *both*, a symptom of a speaker's cognitive processes *and* a signal to his/her listener. According to this approach, UH/M are multifunctional phenomena that serve both the speaker and the recipient at the same time; this is the case, for example, when they signal a speaker's intention to hold a turn by expressing that he/she is currently planning the rest of his/her utterance. This reconciliatory approach towards the *symptom-versus-signal hypothesis* has gained a substantial number of supporters in recent years. For instance, Kjellmer (2003: 189) points out that UH/M fulfill a broad variety of different functions

(including, for instance, signaling hesitation, turn-taking or the marking of self-correction) which are “not always easy to distinguish, and [...] are sometimes carried out simultaneously”; similarly, Tottie (2014: 25) suggests that UH/M have not always been, but are now, multifunctional phenomena, stating that “[i]t is likely that *uh* and *um* originated in situations of cognitive load, where speakers needed time to pause to think and plan, but that they [...] have now also acquired pragmatic meanings”. What all of these statements have in common is their perception of UH/M as multidimensional linguistic signs which serve both the speaker *and* the listener; this conception of UH/M will be illustrated with an example in the following paragraph.

To provide a brief example of the multidimensional notion of UH/M, consider, for instance, the following passage taken from *The International Corpus of English: the British component*, hereafter known as ICE-GB (Aarts, Nelson and Wallis 2006: S1A-015:55<sup>6</sup>), in which speaker 1 (a young man hereafter referred to as S1) has just told speaker 2 (a female friend henceforth referred to as S2) that he and his brother had planned on having portraits taken that were inspired by photographs of the British authors W.H. Auden and Christopher Isherwood (the relevant occurrence of UH/M is printed in bold):

(1) S2: Who would be who

S1: Uhm oh I don't know <,> but **uhm**<,,> I suppose we're not really like either of them

In this example, S1 could potentially have finished his turn after the first part of his utterance, *uhm oh I don't know*, but has apparently decided that he wants or needs to elaborate further on this answer. Assuming that the presence of a short and a long pause (marked by <,> and <,,> respectively) immediately preceding/following his uttering of *but uhm* indicates some kind of cognitive process, we can argue that *uhm* is *symptomatic* of the speaker's planning of upcoming speech; however, given that this planning process coincides with a *transition relevance place*, i.e. a point in speech that lends itself to speaker transition (see section 7.3.2 for a more detailed explanation of the transition relevance place), the speaker also faces the need to signal his intention of keeping the turn. He accomplishes this by uttering *uhm* during his pause, and we can thus conclude that UH/M, in this specific example, can be seen both as symptom of a planning process *and* as signal related to turn-keeping.

---

<sup>6</sup> Note that in this paper, references to the *International Corpus of English: Great Britain* (ICE-GB) will be made in the format [corpus text reference: line]. In the present example, for instance, I have cited corpus text S1A-015, line 55.

Summing up, the present chapter has shown that the substantial body of research supporting both the symptom hypothesis and the signal hypothesis makes it impossible to refute either of them; consequently, it has been proposed that UH/M are multifunctional phenomena that serve both the speaker and their listener(s) in attaining their shared goal of efficient communication. Having proposed this multifunctional notion of UH/M, this paper will now turn towards the examination of the precise functions that have been found to be accomplished by the insertion of UH/M into spoken discourse; this will be the topic of the next chapter.

#### **4. The functions of UH/M in conversation**

In the course of the previous chapter, it has been suggested that UH/M are multifunctional lexical phenomena that serve both the speaker and the listener in their attempt to attain their shared goal of efficient communication by fulfilling certain functions in spoken discourse. The question that remains unanswered, however, is *which* precise functions are fulfilled by UH/M; this will be the subject of this present chapter. The upcoming analysis will comprise four main parts. While the first three sections will each represent one of the three main functions that have been linked to the uttering of UH/M – planning, turn-management, and, self repair – the fourth and last one will be dedicated to the discussion of UH/M’s role in the speech comprehension process. The findings obtained in the course of this chapter will serve as a basis for the construction of a theoretical model of UH/M that attempts to demystify the linguistic nature of UH/M (see chapter 5). As will be seen in the course of this chapter, the meaning of UH/M in conversation seems to generally center around the basic message “I am thinking” (cf. Tottie 2016: 115); thus, the first section will be dedicated to the exploration of what is commonly referred to as “the planning function” of UH/M.

##### **4.1. The planning function of UH/M**

As stated above, the first function of UH/M to be discussed in the course of this chapter is the planning function, which is inferred in many seminal publications on the topic. In recent years, many researchers have suggested that the frequency of UH/M in conversation tends to increase with a speaker’s “cognitive load” (Corley and Stewart 2008: 590), an umbrella term that comprises aspects such as task complexity, a given topic’s level of abstractness, the difficulty of a posed question and/or the number of conceptual and linguistic options a speaker is facing at a given point in speech. Indeed, as will be shown below, there is substantial evidence suggesting that the presence of UH/M in conversation can be related to

the cognitive aspects of linguistic interaction (cf. Goldman-Eisler 1958a: 67), including, for instance, information retrieval, the conceptual planning of content, syntactic realization and lexical access. The present chapter aims at discussing this evidence in the course of six sections. For this purpose, sections 4.1.1 and 4.1.2 will investigate the determinants of UH/M that have been suggested to reveal their link to planning, while section 4.1.3 will discuss the syntactic evidence for the planning function of UH/M. Finally, sections 4.1.4 and 4.1.5 will explore the different manifestations of the planning function (with regard to information retrieval on the one hand and lexical access on the other hand) before the findings of all subchapters will be summarized in section 4.1.6.

#### **4.1.1. Complexity, abstractness, and difficulty: task-related determinants of UH/M**

As was briefly addressed above, the assumption that UH/M might represent a speaker's cognitive processes is to a large extent based on the observation that their frequencies tend to increase proportionally with the "cognitive load" (Corley and Stewart 2008: 590) imposed by contextual and task-related aspects of a given speaking environment; these aspects include, for instance, task complexity, topic abstractness, and the perceived difficulty of a given activity. This relationship has been noted, for instance, by Reynolds and Paivio (1968), who have shown that speakers tend to produce more instances of UH/M when defining abstract nouns such as *fun*, *crime* or *mind* (which was assumed to be the more challenging task) than when performing the same task with concrete words such as *hotel*, *mother* or *library* (Reynolds and Paivio 1968: 172). A similar finding was obtained by Siegman and Pope (1966: 241), who prompted subjects to spontaneously narrate stories based on given pictures. Some of these pictures were assumed to be more ambiguous than others, because they had brought forth a wide variety of different responses (i.e. different stories or themes) in previous experiments of the same kind (Siegman and Pope 1966: 241). In the following analysis of the different stories, it was found that those pictures that had been categorized as highly ambiguous tended to elicit considerably higher frequencies of UH/M than less ambiguous ones. Given that in both of the above experiments UH/M were found to appear most frequently in those situations that had been presumed to be more difficult, it seems that the production of UH/M is determined by task complexity and abstractness. In a different line of argument, Gilquin (2008: 141) has established a link between task difficulty and the uttering of UH/M by showing that learners of English as a Foreign Language tend to produce significantly higher frequencies of UH/M than native speakers. Based on the assumption that the process of speech production is considerably more difficult for non-native speakers, who

have to plan their utterances more consciously than native speakers, this finding seems to provide further evidence for a relationship between the production of UH/M and cognitive load.

Even if the above findings seem to clearly suggest a correlation between UH/M and task complexity or topic abstractness, it should be pointed out that not all publications on the subject confirm this link. For instance, when manipulating task abstractness by asking subjects to either describe or summarize cartoons (the latter task of which was deemed significantly more abstract because it required a more profound understanding of the cartoon as well as a greater degree of conciseness), Frieda Goldman-Eisler (1961: 25) found that the more abstract task did not produce the expected rise of UH/M despite observing an effect for silent pauses. This finding could also be reproduced in a later experiment, in which she compared the rates of UH/M produced in cartoon descriptions and cartoon interpretations (Goldman-Eisler 1968, cited in Christenfeld 1994: 193). Similarly, Greene, Lindsey and Hawn's (1989: 128) manipulation of complexity via the increase of social goals to be attained in a given communicative situation - where subjects in the "single goal condition" had to report on a fictional employee's job performance to a third party and those in the "multiple goal condition" had to give direct feedback without hurting their interlocutor's feelings (cf. Greene, Lindsey and Hawn 1990: 123) – did not elicit any increase in UH/M production rates. These three studies seem to contradict the previously outlined experiments by suggesting that task complexity does not affect the production of UH/M in spoken discourse.

One possible explanation for the discrepancy between the above findings lies in the ambiguity of the terms *complexity* and *abstractness*. This has been pointed out by Christenfeld (1994: 193), who has argued that it is almost impossible to make objective judgments on a specific task's "abstractness", "complexity", or "difficulty", because it seems likely that the perception of these characteristics is dependent on highly individual factors such as personal disposition, experience, or talent and might thus differ from speaker to speaker (Christenfeld 1994: 193). Moreover, it is equally important to point out that the various tasks differed in more respects than just their complexity or abstractness; for instance, it seems likely that subjects in Greene, Lindsey and Hawn's study (1989) were not only influenced by the number of social goals, but also by the differences in setting and interpersonal context (see section 4.1.2 for a discussion of the effects of setting and genre on the production of UH/M). Thus, even though the experiments presented in the previous two paragraphs do offer some interesting insights into potential determinants of UH/M, it seems that Christenfeld (1994: 193) still has a point in

concluding that neither of them do in fact prove “the presence or absence of a causal relationship” between complexity, abstractness, and UH/M.

In an attempt to overcome the problems associated with the terms *complexity* and *abstractness* outlined above, Christenfeld (1994) has proposed to investigate the relationship between UH/M and cognitive load via the more observable process of decision-making. This approach is based on the assumption that the number of options available at a given point in discourse produces a direct effect on the cognitive load imposed upon the speaker, for the latter has to consider and evaluate more potential options. The main advantage of Christenfeld’s proposal is that unlike the abstract notions of *task complexity* or *abstractness*, the number of options available to a speaker can be objectively manipulated. This has been done, for instance, in Christenfeld’s (1994) study on the effects of options on the production of UH/M, in which he studied the relationship between task complexity and UH/M by asking his subjects to verbally describe their ways through mazes that presented them with differing numbers of alternative paths. This research design was based on the assumption that if options were indeed a determinant of UH/M, speakers could be expected to produce higher rates of UH/M in the more complex mazes than in the simpler ones. Indeed, it was shown that the subjects’ production rates of UH/M increased proportionally with the complexity of the mazes they had to describe (Christenfeld 1994: 194); consequently, Christenfeld’s experiment not only supports the hypothesis that UH/M are a direct consequence of a speaker’s increased cognitive load, but also suggests that they may act as expressions of conceptual planning and decision-making.

Interestingly, however, not all findings obtained in Christenfeld’s (1994) experiment seemed to confirm the hypothesis that the production of UH/M is determined by the number of options available to a speaker. In a follow-up of his maze experiment, Christenfeld (1994: 198) found that subjects produced significantly more instances of UH/M when they were asked to navigate through the simplest of all mazes (which the participants knew not to include any options) by just using the four words *up*, *down*, *left* and *right* – a task which the author assumed to be “one of the easiest possible speech tasks” (Christenfeld 1994: 198) – than when they were allowed linguistic freedom in their descriptions (Christenfeld 1994: 198). This observation led Christenfeld (1994: 198) to the conclusion that the production of UH/M is not only determined by the number of available options *present*, but also by those *absent*; in other words, according to Christenfeld (1994: 198), UH/M seems to be not only



determined by the number of linguistic options available, but also by the extent to which a given speaking task allows for or hinders the development of fluent speech.

The hypothesis that linguistic restriction might trigger the production of UH/M also offers a possible explanation for Goldman-Eisler's (1961; 1968) aforementioned failure to prove a causal relation between task complexity and UH/M. This has been suggested by Christenfeld (1994: 198), who states that:

Describing a cartoon may be simpler than interpreting it, but it may also be a more fragmented task. The description is more likely to consist of a series of unconnected observations without a unifying theme. If filled pauses are produced when the speech apparatus is not flowing smoothly, then the lack of unity in a description task might cancel the benefits of its simplicity. (Christenfeld 1994: 198)

As can be seen from this quote, Christenfeld uses his unexpected finding on the increased frequency of UH/M in the "simplest", lexically restricted speaking task to hypothesize that the uttering of UH/M is not only dependent on the number of lexical options, but also on the cognitive load imposed through restrictive factors that inhibit the speaker's development of natural, fluent speech. This hypothesis is also in line with Goldman-Eisler's (1961: 20) observation that "greater conciseness in summarizing was associated with more hesitation". The finding that the production and frequency of UH/M are also influenced by lexical restriction is especially interesting because it can be used to infer that UH/M are not only determined by task-related aspects such as perceived difficulty and the associated cognitive load, but also by language-related aspects such as word choice or the extent to which a given task allows for the development of fluency. The strong link between linguistic factors and UH/M can be especially observed in the analysis of the production of UH/M in different linguistic settings and genres, which will be the topic of the following section.

#### **4.1.2. Genre, register, and linguistic conventions as determinants of UH/M**

As was shown in the previous section, it seems plausible to assume that UH/M are determined not only by a task's difficulty or complexity, but also by linguistic and contextual aspects of a given discourse situation; these include, for instance, genre, register, or linguistic conventions. This approach is relevant to the exploration of UH/M's planning function insofar as all of these three aspects may produce significant effects on the extent to which a speaker feels obliged to (carefully) plan and monitor his/her speech. This, in turn, contributes to the cognitive load imposed by the respective situation, the assumption being that speakers will find it more cognitively challenging to communicate in highly formal genres, which are characterized by strict linguistic conventions, than in informal conversations with family

members or friends. In other words, evidence confirming a relationship between genre, register, linguistic conventions and UH/M could be most probably regarded as evidence confirming the relationship between psycholinguistic processes and UH/M that is commonly referred to as their “planning function”. The following section aims at discussing evidence suggesting linguistic determinants of UH/M, which can then be used to draw conclusions about the question of whether or not UH/M are representative of the cognitive processes involved in language production.

Even though researchers now generally seem to agree that genre *does* affect the production rates of UH/M, the question of whether UH/M are more characteristic of formal or informal speech has still not been clarified. While some studies suggest that speakers are more prone to produce UH/M in informal registers such as small-talk, informal interviews, or conversations between family members and/or friends (cf. Swerts 1998: 485; Schachter et al. 1991: 365; Clark and Fox Tree 2002: 98), others, including Tottie (2014: 6) and Stenström and Svartvik (1994: 247-248), have found frequencies of UH/M to be higher in what Tottie (2014: 6) refers to as “non-private environments”, such as offices, classrooms and court proceedings. As a consequence, the question of contextual determinants of UH/M remains unanswered, for it cannot be reliably concluded whether the production of UH/M is more typical of formal or informal, of private or public speaking contexts.

One possible explanation for the discrepancy outlined in the above paragraph lies in the vagueness of the terms *formal* and *informal speech*. This is suggested, for example, by Tottie (2014: 26) who dismisses the terms *formal* and *informal* and rather proposes a distinction between *private* and *non-private speech*, which she justifies by pointing out that “[s]peech in a non-private space like a bank-meeting room, or in a public space like a university classroom does not have to be formal or qualify as public speaking. Public speaking is a totally different genre – it is done in broadcasting, at political meetings, at formal dinners [...]” (Tottie 2014: 26). Indeed, it appears that at the origin of the seemingly contradictory findings addressed above lies the misconception that the terms *private* and *non-private speech* can be used synonymously for *informal* and *formal speech* respectively; however, this assumption is ill-informed because, as Tottie (2014: 26) points out, “[...] speech can be non-private and informal at the same time”. For instance, a university professor holding a lecture is clearly constructing speech in a non-private environment, even though he/she might deliver it in what is classified as informal register. From this example, it can be seen that the distinction between the different degrees of privacy is indeed more meaningful than that of register. This

is especially true when analyzing the frequencies of UH/M across genres, because it allows for a more accurate terminological representation of what has been witnessed. After all, when Schachter et al. (1991: 365), for example, cite their observation that university lecturers tend to produce more instances of UH/M “while *formally* lecturing on his or her discipline [...] [than] while *informally* talking about his or her department’s graduate training requirements [my emphasis]”, one is ignorant of the extent to which the terms *formal* and *informal* refer to the respective speaking environments (where *formal* would represent a professional lecture with a large audience and *informal* a one-on-one interview in the lecturer’s office) or to the actual register of the language being produced. Moreover, the terms *formal* and *informal* are often not clearly defined and thus leave room for interpretation as to their specific characteristics; this significantly reduces their academic value in the context of attempting to comprehend the precise linguistic factors determining UH/M. As a consequence, it seems that the distinction between *non-private* and *private context* is preferable to that between *formal* and *informal language* because it allows a more accurate attribution of observations to specific characteristics of a given speaking environment.

Applying these theoretical considerations to the study of UH/M allows us to obtain a more detailed idea of the aspects determining the presence, frequency, or even absence of UH/M in spoken discourse. In her discussion of the functions fulfilled by UH/M in conversation, Tottie (2014: 26) points out that while register, i.e. the degree of formality, did not produce significant effects on the frequencies of UH/M, speakers were found to use UH/M more frequently in non-private environments (such as classrooms or offices) than in private settings (e.g. at home). This leads her to hypothesize that “UH[/]M use is governed by distinction between private and non-private rather than between formal and informal” (Tottie 2014: 26). This observation is in line with Stenström and Svartvik (1994: 247-248), who have found that court proceedings generally contained higher frequencies of UH/M than spontaneous conversation. One possible explanation for this phenomenon is that speakers tend to be more deliberate in their phrasing of non-private speech (e.g. when talking to colleagues, doctors, superiors, etc.) than they are in private discourse (Tottie 2014: 6); as a consequence, speakers spend more time planning their upcoming utterances and subsequently also face an elevated need of signaling their thinking and planning. The hypothesis that UH/M are determined by the degree of a speaker’s lexical deliberation in a given context seems further supported by Tottie’s observation that UH/M was also particularly frequent when discussing unpleasant

topics (Tottie 2014: 6); this, again, seems to suggest that the more time speakers spend on the careful phrasing and planning of speech, the more instances of UH/M they generally produce.

Summing up, it appears that there is indeed a relationship between (the need for) lexical planning, induced by the increased need for deliberation in specific discourse genres, and the production of UH/M, which can be seen as clear evidence in support of a link between linguistic planning and the uttering of UH/M.

#### **4.1.3. Syntactic evidence for UH/M's planning function**

In addition to the findings discussed in the previous section, the relationship between UH/M and linguistic planning has been further supported by a large body of syntactic evidence, which links the production of UH/M to a speaker's linguistic planning processes. One of the most important observations made in this context is that UH/M tend to occur at (or near) the boundaries of syntactic units rather than inside them. Based on the assumption that language is planned in units larger than the word (Boomer 1965: 155), and that the points of content- and form-related linguistic decision-making coincide with the boundaries of syntactic units (Christenfeld 1994: 192), this observation has led many scholars to argue in favor of a link between planning and the production of UH/M. However, it is often difficult to know how to interpret findings concerning the syntactic position of UH/M, because researchers do not seem to agree on their implications: while some believe that UH/M's tendency to occur at syntactic boundaries proves their planning function – thus assuming that the units of speech encoding are known, but that the functions of UH/M are not (cf. Kjellmer 2003: 171) – others argue that UH/M's frequent occurrence at syntactic boundaries should be interpreted as evidence in support of the respective unit's role in speech encoding; this latter approach is based on the supposition that UH/M certainly fulfill a planning function and may thus reveal the units of speech encoding (cf. Holmes 1988; Boomer 1965). In other words, while some have suggested that the syntactic position reveals the functions of UH/M, others have used the presumed function of UH/M to study the points of syntactic decision-making; this leads the discussion into a situation of circular argumentation that is difficult to resolve. For the purpose of the following discussion of syntactic evidence for UH/M's planning function, it will be assumed that syntactic boundaries mark points of conceptual and/or linguistic planning and decision-making, thus adopting the former of the two hypotheses.

The tendency of UH/M to appear at or near syntactic boundaries has most frequently been confirmed in relation to the clause. In their respective studies of the functions fulfilled by

UH/M, Kjellmer (2003: 180), Holmes (1988: 333) and Hawkins (1971: 286) have all shown that UH/M tend to occur at the boundaries of clauses rather than inside them. Given that the onset of clauses is generally associated with a broad range of options concerning the conceptual, syntactic, lexical, and organizational presentation of content (Hawkins 1971: 286-287), the finding that UH/M tend to occur at or near these points in speech seems to provide valuable evidence in support of the presumed planning function of UH/M. In addition, Kjellmer (2003: 180) and Holmes (1988: 336) have both pointed out that UH/M were more likely to occur before coordinate clauses than before subordinate clauses. This finding can be interpreted as further support of UH/M's planning function because, for, as Kjellmer (2003: 180) points out, "subordinated clauses are obviously more intimately integrated semantically with their matrix clauses" and thus require less immediate planning than the introduction of a new independent thought in the form of a coordinate clause.

Interestingly, Boomer (1965: 151) could not confirm the general observation that UH/M tend to occur at syntactic boundaries, for he observed that the most frequent position of UH/M was not at the direct onset, but after the first word of the clause. Based on the hypothesis of a "two-stage, hierarchical encoding process" where the larger, syntactic units of speech encoding are chosen before making lexical decisions (Boomer 1965: 156; cf. also Carroll 1953, cited in Boomer 1965: 156; Goldman-Eisler 1962, cited in Hawkins 1971: 286), Boomer (1965: 156) uses this observation to infer that UH/M are more characteristic of lexical planning than of syntactic choices. However, this interpretation of data was refuted by Hawkins (1971: 285), who points out that

[Boomer's] argument that position 1 pauses [i.e. at the clause boundary] select the grammatical frame while position 2 pauses [i.e. after the first word] select the 'key responses' only holds good if the first word really commits the speaker to a particular construction or set of constructions (as he claims). But the first word of a clause is highly likely to be a linking conjunction such as *and*, *but*, *so*, particularly in informal spontaneous speech; and these conjunctions do not commit the speaker in any way to the type of construction which must follow them. Even binding conjunctions (*if*, *when*, *because*, etc.) impose only a very loose framework on the subsequent clause, and so clause-initial adjuncts such as *then*, *one day*, etc.

Based on this argumentation, Hawkins (1971: 285) then moves on to suggest that UH/Ms in clause-initial position perform the same functions as those succeeding conjunctions, thus implying that they can neither be exclusively ascribed to syntactic nor to lexical planning. In light of these findings, it can be concluded that even though it does seem legitimate to assume a relationship between the uttering of UH/M and linguistic planning, one cannot be certain of the precise nature of these planning processes. What can still be retained from this paragraph

is that the existent syntactic evidence seems to further support the hypothesis that UH/M are indicative of a speaker's cognitive processes linked to speech production.

In the paragraph above, only evidence concerning the clause was considered; however, it should be noted that the tendency of UH/M to occur at unit boundaries has also been observed with regard to other entities relevant to the syntactic analysis of spoken discourse. For instance, Ford (1982: 805) has shown that speakers tend to utter UH/M at the beginning of sentences rather than at other boundaries within them; similarly, Clark and Fox Tree (2002: 94) found that the probability of UH/M to occur in a given *intonation unit*, i.e. “a stretch of speech under a single intonation contour” (Clark and Fox Tree 2002: 94), decreased significantly the further speaker proceeded into the units, thus suggesting that the frequency of UH/M is directly influenced by the number of options available to a speaker at a given point in speech. Finally, it has been proposed by some that UH/M also tend to occur at the boundaries of phrases rather than inside them (Hawkins 1971: 283; Kjellmer 2003: 175-179), even though the clause boundary seemed to be preferred over the phrase boundary inside a clause (Maclay and Osgood 1959: 33).

As has been stated above, the syntactic evidence presented in the course of this section has been interpreted to suggest a strong link between the uttering of UH/M and the cognitive processes they represent. In that sense, the tendency of UH/M to occur at syntactic boundaries has led many researchers to assume that they often indicate conceptual, i.e. content-related planning, information retrieval, and syntactic planning; moreover, the fact that UH/M might also appear within syntactic units has caused some to assume a link to lexical decision-making. The following sections aim at presenting evidence in support of the relationship between the production of UH/M and these different kinds of planning, focusing mainly on the conception of UH/M as markers of information retrieval and lexical access.

#### **4.1.4. UH/M as markers of information retrieval**

In the course of time, many researchers have supported the presumed causal relationship between UH/M and cognitive load by pointing out that speakers tend to produce them whenever they find the process of information retrieval particularly challenging (cf. Tottie 2016: 106-107; Christenfeld 1994: 192). This becomes especially obvious when considering the frequency of turn-initial UH/Ms in answers to questions, in particular when the speaker displays a low “feeling of knowing” (Smith and Clark 1993: 37; Swerts and Krahmer 2005: 91; Clark and Fox Tree 2002: 87). This relationship between self-consciousness with regard

to one's ability to correctly answer a question and the uttering of UH/M has been suggested, amongst others, by Smith and Clark (1993: 37) and Swerts and Krahmer (2005: 91), who found that speakers were most likely to utter a turn-initial UH/M when they were not sure whether they knew the answer to a question. Based on the assumption that speakers who are not certain of the correct answer will devote more cognitive effort into the process of information access, these findings seem to suggest that UH/M are indeed linked to the retrieval of information and, in more general terms, the speaker's cognitive processes. This conclusion is further supported by the observation that UH/M frequently represent points in speech at which a speaker is busy retrieving specific details such as proper names or locations (Tottie 2016: 107; Rühlemann, Bagoutdinov and O'Donnell 2011: 82). The fact that speakers utter UH/M whenever they are engaged in the process of information access seems to provide significant evidence for a relationship between cognitive processes and the uttering of UH/M, thus further consolidating the hypothesis of a planning function. However, as has been stated on numerous occasions throughout this paper, information access is not the only cognitive process linked to the production of UH/M, for there is substantial evidence that UH/M can also represent other cognitive processes such as lexical access; this will be discussed in the following section.

#### **4.1.5. UH/M as a marker of lexical retrieval**

If we assume that UH/M are representative of various planning processes linked to speech production, it would be surprising if lexical access did not constitute one of them. Even though it is now generally believed that speech is planned in syntactic units larger than the individual word (Levelt 1989: 202; Boomer 1965:155), it has been observed that speakers rather frequently encounter problems of lexical access within these units; this observation has led some to assume that the process of lexical retrieval is more or less independent from the planning of syntactic structures (cf. Shriberg 1994: 154). The present section seeks to briefly outline the most pertinent findings made with respect to the role of UH/M in the process of lexical retrieval.

The largest proportion of evidence suggesting a relationship between lexical retrieval and the uttering of UH/M concerns the frequently made observation that UH/M tend to occur before words that are assumed to be cognitively challenging to access. For instance, UH/M have been shown to precede lexical rather than function words (Quinting 1971: 39; MacLay and Osgood 1959: 32-33) and words of low frequencies rather than those that were characterized

as occurring frequently in spontaneous speech (Beattie and Butterworth 1979: 208). Given that the retrieval of both infrequent words and the open class of lexical words is thought to be more cognitively challenging than that of their more accessible counterparts (Beattie and Butterworth 1979: 208), these observations seem to suggest a link between lexical retrieval and the uttering of UH/M. Further evidence for this relationship stems from the frequent observation that UH/M tend to precede words of low transitional probability, i.e. words that are difficult to guess from their immediate surroundings (cf. Goldman-Eisler 1968: 41-42; Tannenbaum, Williams and Hillier 1965, cited in Levelt 1995 [1989]: 203; Beattie and Butterworth 1979: 202). Assuming that a word with a low transitional probability is not only difficult to guess for the listeners, but – due to the lack of semantic clues aiding the fast selection of a suitable element – is also more challenging to access for the speaker (Beattie and Butterworth 1979: 209), this tendency of UH/M to appear before unpredictable lexical items seems to further support the supposition of a link between lexical planning and the production of UH/M.

In addition to the above, the link between UH/M and lexical decision-making has been further supported by evidence suggesting that UH/M are particularly frequent in contexts that allow for lexical creativity. This has been proposed by Schachter et al. (1991: 565), who showed that UH/M are more frequently produced in lectures within the humanities than within the natural sciences. Based on the assumption that the humanities allow for a higher degree of lexical creativity than the natural sciences<sup>7</sup>, this finding was used to argue in favor of a relationship between the production of UH/M and lexical decision-making. Combining this argumentation with the other evidence presented in the course of this section, it seems safe to conclude from the aforementioned evidence that the production of UH/M is often a consequence of the quantity of lexical options available at a given point in speech as well as the extent to which a given task offers room for lexical variety and creativity; as a result, it appears that UH/M can indeed be linked to the process of lexical access and planning.

Summing up, it has been shown in the above paragraphs that there are indeed good reasons to believe that UH/M may act as markers of lexical retrieval. According to this view, speakers may – but do not have to – decide to utter UH/M whenever they encounter a problem of lexical retrieval and thus has to actively search his/her mental lexicon; one reason for which

---

<sup>7</sup> This assumption is based on the fact that natural sciences contain significantly higher proportions of technical terms and fixed expressions, while the humanities often allow for a broader range of synonyms and lexical creativity (cf. Schachter et al. 1991: 565).



they may decide to do so is to ensure that he/she does not expose his/her listener(s) to longer periods of unexpected silence. In that sense, the uttering of UH/M in periods of lexical retrieval could be interpreted to not only signal the speaker's cognitive processes (cf., e.g., Fischer 2006: 432) but also serve the purpose of turn-keeping or – if the speaker desires help from his/her interlocutor – even momentary turn-yielding (cf. Maclay and Osgood 1959: 41-42; Larcombe 1995, cited in Gilquin 2008: 121; Clark and Fox Tree 2002: 90; Ford 1993: 49; Kjellmer 2003: 185); this interpretation is also in line with the multifunctional view of UH/M proposed in section 3.4.

Having provided a brief discussion of evidence supporting the assumption that UH/M may indicate periods of lexical access, the following section will provide a brief interim conclusion which will summarize the most pertinent findings with regard to the planning function of UH/M.

#### **4.1.6. Interim conclusion: the planning function of UH/M**

In the course of the present section, it has been shown that there is a substantial amount of evidence supporting the idea that the uttering of UH/M is linked to the cognitive processes involved in speech production. Even though it is true that the precise nature of the respective planning process (i.e. conceptual, syntactic, or lexical) is often impossible to determine, it has been shown that the presence of UH/M in spoken discourse can most probably be linked to various types of planning, ranging from information retrieval over syntactic planning to the processes of phrasing and lexical retrieval. As a consequence, it seems safe to conclude that a main function of UH/M is that of “[...] helping the speaker to plan what to say next by getting a space for thinking [...]”, as has been proposed by Tottie (2016: 110). However, as has been briefly addressed earlier in this paper, the signaling of cognitive processes does not seem to be the only function of UH/M, for they have also been frequently linked to the turn-management system; this function shall be discussed in the upcoming section.

#### **4.2. UH/M and the turn-management system**

As was already briefly addressed in the discussion of the symptom-versus-signal debate in chapter 3, it is now a widely accepted hypothesis that UH/M do not only represent a speaker's cognitive processes, but also signal his/her intentions with regard to the turn-management system. This hypothesis is mainly based on the increased awareness that silent pauses and UH/M do not fulfill the same functions in conversation (Christenfeld 1994: 193) and the subsequent conclusion that there has to be a specific reason for a speaker to fill

his/her pauses with UH/M (cf., for instance, Gilquin 2008: 124), such as the desire to actively signal one's intention to take, keep, or yield one's turn. The present section examines the role of UH/M in the turn-management system by discussing the three main functions that UH/M have been linked to in this context. For this purpose, the present section is divided into four parts, where section 4.2.1 explores the turn-keeping function of UH/M, section 4.2.2 investigates the role of UH/M in turn-taking and turn-accepting, and section 4.3.3 examines the function of UH/M as turn-yielding devices. Finally, section 4.3.4 provides a brief interim conclusion on the role of UH/M in the turn-management system.

#### **4.2.1. UH/M and turn-keeping**

One of the most frequent propositions made with regard to the role of UH/M in the turn-management system is that they are used as turn-keeping devices that help a speaker to keep the floor during periods of thinking and planning. This hypothesis has been supported, amongst others, by Maclay and Osgood (1959: 41-42), who have suggested that a speaker's production of UH/M is directly determined by his/her fear to lose his/her turn in periods of excessive pausing. As they put it:

Let us assume that the speaker is motivated to keep control of the conversational "ball" until he has achieved some sense of completion. He has learned that unfilled intervals of sufficient length are the points at which he has usually lost this control – someone else has leapt into his gap. Therefore, if he pauses long enough to receive the cue of his own silence, he will produce some kind of signal ([m, əɹ], or perhaps a repetition of the immediately preceding unit) which says, in effect, 'I'm still in control – don't interrupt me!'" (Maclay and Osgood 1959: 41)

The hypothesis that speakers produce UH/M as a reaction to their own silence for the sake of turn-keeping has been confirmed by many, including Kjellmer (2003: 184-185), Gilquin (2008: 120-121) and Clayman (2013: 153), all of whom suggest that speakers use UH/M to signal that he/she "is preparing a new information unit, intends to go on speaking and is not willing to yield his turn" (Kjellmer 2003: 184-185). This assumption is also in line with Sacks, Schegloff and Jefferson (1974: 718-720), who have proposed that the uttering of semantically relatively blank elements such as conjunctions or UH/M at points that offer themselves to speaker transitions is a frequent means of preventing others from taking the turn. The idea of UH/M as turn-keeping devices is well summarized by Larcombe (1995, cited in Gilquin 2008: 121), who compares the uttering of UH/M in periods of planning to the "drap[ing] of coats on a seat at the cinema to prevent others from taking it". In that sense, the view of UH/M as turn-keeping devices might also explain why they tend to occur at or near syntactic boundaries (Kjellmer 2003: 180; Holmes 1988: 336; Kjellmer 2003: 185, cf. also

section 4.1.3), for it is generally assumed that these boundaries lend themselves to speaker transition and thus render the active signaling of turn-keeping intentions particularly relevant. In summary, it seems that there are indeed convincing reasons to assume that speakers tend to utter UH/M whenever they need time for planning and need to signal this to the listener so as not to lose their right to the floor. This conception of UH/M is interesting for two main reasons. On the one hand, it reinforces the idea that UH/M are important conversational strategies that help the listener and the speaker to attain their shared goal of efficient communication. On the other hand, this view of UH/M also provides further evidence in support of a multifunctional conception of UH/M, where speaker-based functions (such as winning time for planning) complement listener-based functions (e.g. the signaling of turn-keeping).

Despite the evidence presented above, some have pointed out that not every instance of UH/M can be associated with the signaling of a speaker's intention to hold a turn. In that sense, Cenoz (2000: 54) argues that the high frequency of UH/M even in non-dialogic genres such as university lectures, where a speaker does not need to actively signal his/her intention to continue his/her turn, seems to refute the idea of UH/M being linked to the turn-management system. Similarly, Tottie (2015: 397) suggests that most instances of UH/M in turn-central position appear in contexts where there are no observable attempts at turn-taking by other speakers; this leads her to the conclusion that "UH[/]M is not normally used to deliberately prevent speakers from taking over the turn" (Tottie 2016: 108). However, these arguments have been dismissed by Kjellmer (2003: 185), who points out that UH/M have been frequently argued to be multi-functional phenomena; in that sense, he refutes Cenoz' above argumentation by stating that "[...] filled pauses have several other functions that are manifested in most lectures" (Kjellmer 2003: 185). This multi-functionality of UH/M is also highlighted when considering the fact that UH/M have been linked not only to turn-keeping, but also to turn-taking and turn-yielding; these functions will be discussed in the following sections.

#### **4.2.2. UH/M, turn-taking and turn-accepting**

As has already been addressed several times in the course of this paper, UH/M have not only been linked to turn-keeping, but also to turn-taking and turn-yielding; the former of these functions – turn-taking – shall be discussed in this present section. In his corpus-based study of UH/M, Kjellmer (2003: 184) presents statistical data according to which "[e]very seventh or eighth *er(m)* [...] is used to introduce a speech turn". However, he points out that not all of

these turn-initial instances of UH/M serve the purpose of signaling a speaker's planning or thinking (Kjellmer 2003: 184); instead, he argues that an important proportion of these occurrences act as a means of directing attention towards oneself as well as announcing one's intention to contribute something to the conversation in the near future (cf. also Kjellmer 2003: 167; cf. also Shriberg 2001: 156). In that sense, a speaker wishing to interrupt his/her interlocutor might choose to utter an instance of UH/M so as to attract attention as well as prepare the current speaker for the loss of his/her turn; this way, he/she also avoids surprising the latter by simply commencing to utter the intended message. This use of UH/M is observable, for instance, in the following example from *Collins Cobuild corpus* cited by Kjellmer (2003: 184), where speaker 1 (S1) interrupts speaker 2 (S2) in order to enquire further information:

- (2) S2: And mining t ceased the mining lecturers er found other posts elsewhere.  
 S1: Mm.  
 S2: Rather interesting where =  
 S1: =Er did it became the department of minerals engineering.  
 (ukspok00833)

This example is a very good illustration of the turn-taking function of UH/M. In order to prepare the current speaker (S2) for the upcoming transition of which he/she is still unaware, S1 decides to utter a preparatory message with the core meaning of "I am planning to interrupt you now, please stop talking and listen to me", thus aiming at facilitating the envisaged transition of roles. Only when S1 receives a positive reaction to his turn-taking attempt (i.e. when S2 stops talking) does he continue to utter his intended contribution; this way, he seems to bypass the risk of having to repeat himself by surprising the current speaker with his sudden interruption and thus having to prolong the disturbance. This interpretation of the above data is particularly interesting insofar as it presents UH/M as an interpersonal signal that serves both the speaker *and* the listener, thus constituting a valuable contribution to the interlocutors' ultimate goal of achieving successful, smooth-flowing communication.

Interestingly, UH/M have not only been found to occur in the context of self-selection, i.e. when a speaker actively chooses to take the floor (cf. Sacks, Schegloff and Jefferson 1974: 716), but also when a speaker is (involuntarily) assigned the next turn; this behavior shall be referred to *turn-accepting*<sup>8</sup>. The turn-accepting function of UH/M becomes especially

---

<sup>8</sup> Note that the term *turn-accepting* stands opposed to the term *turn-taking* insofar as the latter designates self-selection while the former refers to situations in which a speaker is actively assigned the turn by another speaker.

obvious when analyzing turn-initial instances of UH/M in answers to questions, as can be found in the below examples taken from the *International Corpus of British English* (ICE-GB):

- (3) S1: How much <,> could you afford //  
S2: Uh <,> I don't know // No nothing really // [...]  
(S1A-015:42)
- (4) S1: Is it going to make serious inroads into your writing <,> //  
S2: Uh I should think so // [...]  
(S1B-047:017)

These two examples are very good illustrations of UH/M as means of turn-accepting. In both cases, S1 asks a question and thus allocates the turn to S2, who positively reacts by uttering UH/M and thereby signals his/her intention to accept the turn while still being able to plan his/her answer. The signaling of turn-accepting is important because, as Clark and Fox Tree (2002: 89) point out, speakers who have selected a speaker other than themselves to take the next turn always require some sign of reassurance that their allocation has been accepted and that communication will thus continue to flow smoothly, even if the newly allocated speaker still needs time to plan his turn. If he/she fails to provide such a signal, the initial speaker will probably continue talking or renew his attempt at initiating speaker-transition, e.g. by rephrasing the question (Clark and Fox Tree 2002: 89). Naturally, it could be argued that in these cases, UH/M also act as a means of signaling the speaker's consideration of the question and the planning of his/her answer; however, it is most likely that the truth lies somewhere in the middle. In that sense, UH/M could be argued to fulfill a kind of double function, where UH/M signal both the planning of the upcoming contribution and the acceptance of turn-allocation; this has also been suggested by Tottie (2016: 107), who ascribes UH/M a "[...] Janus-faced function [...], at the same time referring back to a question and the following answer".

Summing up the above, it can be concluded that turn-initial instances of UH/M do indeed often seem to be related to the turn-management system, signaling either a speaker's intention to actively take the turn or his/her acceptance of turn-allocation by another speaker. In both cases, UH/M can be interpreted to constitute useful resources for both the current and the new speaker. While in the case of active turn-taking, UH/M can be used to signal one's intention to take the floor by attracting attention or – in the case of interruption – preparing the current speaker for the unexpected loss of his/her turn, it also helps a speaker to gain time for planning one's turn without having to expose interlocutors to long periods of unexplained

silence in the case of turn-accepting. In summary, the consideration of the turn-taking and turn-accepting functions of UH/M further supports the hypothesis that UH/M are interpersonal resources that are helpful in communicating one's interactional intentions in a given speech situation without having to explicitly formulate them. This point will be further supported in the course of the next section, in which the last function of UH/M with regard to the turn-management system – turn-yielding – will be discussed.

#### **4.2.3. UH/M and turn-yielding**

As has been frequently mentioned throughout this paper, UH/M have been found to act not only as means of turn-holding and turn-taking, but also of turn-yielding. This function has been observed, for instance, by Clark and Fox Tree (2002: 90), who have claimed that speakers might utter UH/M to signal their desire to give up their turn. This might be the case, for instance, in situations where speakers assume that their interlocutors know a word they themselves cannot think of at the moment and thus wish to invite them to help complete their utterance (de Leeuw 2007: 86). Further support for the turn-yielding function of UH/M stems from Ford (1993: 49), who hypothesizes that “it could be that speakers [...] pause because they have nothing substantial to say” and subsequently require the help of their interlocutors to maintain the flow of conversation; this hypothesis also seems to be indirectly supported by Kjellmer (2003: 185), who asserts that UH/M occur in contexts of turnyielding whenever “the hesitation element is very prominent”. The underlying assumption of all of these propositions is that UH/M signal a speaker's difficulties in retrieving linguistic material to produce in a given context<sup>9</sup> and that listeners can interpret these instances of UH/M as an invitation to help, whether it be in the retrieval of a word or in upholding the conversation by asking further questions or taking the turn.

Summing up, there seems to be substantial evidence for the view that UH/M can function not only as turn-keeping and turn-taking devices, but also as signals of one's turn-yielding intentions; however, the question that remains unanswered is how listeners can distinguish these functions to draw conclusions on the behavior expected from them when the lexical cue – UH/M – is always the same. In their famous discussion of the functions of UH/M in spoken discourse, Clark and Fox Tree (2002: 90) claim that one factor determining a listener's understanding of UH/M as a turn-yielding device could be their pausal environment. In other

---

<sup>9</sup> Note that this assumption is in line with with evidence presented earlier in this paper, according to which the rates of UH/M can be linked to the cognitive load imposed on the respective speaker in a given situation (cf. section 4.1)

words, they argue that listeners are more likely to feel invited to take the turn from a speaker who has failed to produce more than instances of UH/M and silent pauses for a longer period of time. This seems to be in line with Ford's (1993: 46) and Kjellmer's (2003:185) aforementioned propositions that speakers produce UH/M to signal that they are either particularly hesitant or unsure about what to say. According to this conception, then, it would appear that listeners are able to interpret longer stretches of material that do not contribute to the primary message of an utterance – such as UH/M or silent pauses – as prompts to take over. Another approach is suggested by Kjellmer (2003: 185-186), who has proposed that a listener's interpretation of UH/M seems to rely to a large extent on the intonational realization of UH/M as well as other prosodic cues surrounding them. In that sense, he argues that if a speaker utters UH/M in a way that “[o]ne can almost hear the voice of the speaker trailing off at the end, hoping to be relieved” (Kjellmer 2003: 185), listeners will most likely interpret the uttering of UH/M as an invitation to take the turn. Indeed, both of these hypotheses seem plausible to anyone who has already experienced speakers desperate to give up their turn because they do not know what else to say; however, their empirical verification would require a thorough analysis of intonational patterns and pause duration, which would exceed the scope of the present discussion.

In summary, there seems to be substantial evidence supporting the view that UH/M act not only as means of turn-taking and turn-holding, but also of turn-yielding, and that listeners seem to have a remarkable intuition that helps them to identify the nature of specific instances of UH/M by interpreting intonational and prosodic cues. This conclusion does not only further support the assumption that UH/M are important signals to the listener, but also that the understanding of UH/M is highly situational, an observation that is particularly relevant to the multidimensional conception of UH/M which has been proposed in section 3.4. Having examined the individual functions of UH/M with regard to the turn-management system, the following section will provide a brief interim conclusion of this section.

#### **4.2.4. Interim conclusion: UH/M and the turn-management system**

It has been shown in the course of this section that apart from the planning function suggested in chapter 4.1, UH/M also seem to fulfill a variety of functions that are relevant to the turn-management system, namely turn-holding, turn-taking, turn-accepting, and, finally, turn-yielding. As previously addressed, it is by no means suggested that these four functions contradict the assumption that UH/M represent a speaker's cognitive processes; instead, it is argued that these two functions seem to complement each other in a way that allows both

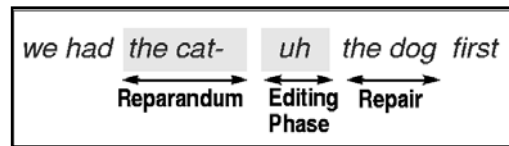
speakers and listeners to profit from the presence of UH/M in spoken discourse. While the former, by uttering UH/M, gains time to complete the cognitive processes he/she is currently engaged in – such as planning the upcoming stretch of speech in the case of turn-taking - holding and -accepting – without having to fear losing his/her right to the floor (cf., e.g. Maclay and Osgood 1959: 41-42), the latter obtains valuable information on the speaker's inner processes, which in turn can be used to draw conclusions on the expected behavior (e.g. not interrupting, helping the speaker find the right word, etc.). Moreover, it seems that speakers might also be able to use UH/M to signal that they are not busy planning, but that they do not know what else to say or require assistance in finding a word; this would suggest that speakers may use UH/M to ask for help. In summary, one could thus conclude that evidence supporting the turn-holding, -taking, or –yielding function of UH/M provides valuable proof of hypothesis that UH/M act as an important resource to the listener. This will further be supported in the following section, in which it will be shown that UH/M also constitute a valuable means of helping both speakers and listeners to overcome problematic stretches of speech such as false starts or error correction.

#### **4.3. UH/M as correction markers**

In addition to signaling planning on the one hand and one's intentions with regard to the turn-management system on the other hand, UH/M have also frequently been shown to serve as a correction marker in the context of self-repair, reformulations, false starts, and self-interruption in general (Kjellmer 2003: 188-189). Very simply put, this means that a speaker who has made a mistake – e.g. by mispronouncing a word, producing a wrong word or a grammatical misconstruction – may utter UH/M immediately before the correct, more suitable form, thus providing linguistic evidence of self-correction (Kjellmer 2003: 188). In other words, UH/M, in these cases, act as an *editing phrase*, a facultative word, phrase, or sound that a speaker might use to bridge the time span between the *interruption point*, i.e. the point where a speaker has interrupted his/her speech for the sake of self-correction, and the beginning of the *repair*, i.e. the corrected stretch of speech (Shriberg 2001: 160). For the sake of comprehensibility, Shriberg (2001: 160) provides the following visual representation of the anatomy of self-repair:



Figure 1: The anatomy of self-repairs (Shriberg 2001: 160)



As can be seen in Figure 1, there are two main lexical components within a self-repair: the *reparandum*, i.e. the element to be repaired, and the *repair*, i.e. the corrected version of the reparable. What comes in between, i.e. the *editing phrase*, is optional and may vary not only from speaker to speaker but also from self-repair to self-repair; possible editing phrases other than UH/M include expressions such as *I mean*, apologies (e.g. *I'm sorry*), or even non-verbal cues such as gestures or facial expressions (Shriberg 1994: 7; Shriberg 2001: 160). The purpose of these editing phrases is twofold: not only does their presence in the midst of a self-repair signal that “the immediately preceding sequence [i.e. the reparable] should be disregarded” (Kjellmer 2003: 189), but they also attract the listener’s attention towards the repair (Kjellmer 2003: 189). Therefore, editing phrases may play an important role in the comprehension of self-repair and self-interruption by helping the listener to focus on the more relevant stretch of speech.

Indeed, UH/M have been found to produce beneficial effects on the comprehension of self-repair and self-interruption. In their study of listener reactions towards different forms of self-repair, Brennan and Schober (2001: 277) found that in both between-word interruptions (e.g. *the cat uh the dog*) and mid-word interruptions (e.g. *the ca-uh-dog*) the repair word (i.e. *dog*) could be identified faster when the self-interruption was accompanied by an instance of UH/M; this suggests that their presence as editing phrases contributes significantly to a listener’s interpretation of self-repairs. Consequently, it seems that UH/M may indeed act as a valuable means of facilitating the comprehension of self-interruption and self-repair by indicating that listeners should ignore the preceding stretch of speech and rather consider the upcoming repair; in this way, they help listeners to understand these potentially problematic periods in speech production.

What seems surprising concerning the function of UH/M as an editing phrase is that listeners can interpret the same acoustic element they frequently experience as a signal of planning and/or turn-holding, -taking, or -yielding (see sections 4.1 and 4.2) as an invitation to “delete” the preceding stretch of speech from his/her mental record. In their discussion of the topic, Corley and Stewart (2008: 594) have suggested that this might partly be due to the presumed *highlighting function* of UH/M (see section 4.4 for a more detailed discussion of this

function); more specifically, they argue that UH/M act less as a signal to disregard the reparandum, but rather as a signal to focus attention on the upcoming repair. This view has been supported, amongst others, by Fox Tree (2001: 321), who has suggested that the presence of UH/M has highly beneficial effects on a listener's comprehension of immediately succeeding target words; similarly, Stenström (1990; cited in Kjellmer 2003: 187) claims that UH/M seem to have an "emphatic effect" which helps to "giv[e] extra prominence to the following word(s)" (Kjellmer 2003: 187). In this conception, then, UH/M are not signals to "delete" the previous material (which cannot be entirely accomplished anyhow, for the material produced will still be present in the listener's short-term memory directly after the occurrence), but rather heighten the semantic value of the succeeding stretch of speech; as a consequence, listeners perceive the repair to be more important than the reparandum and will thus successfully choose to consider the correct option. This hypothesis will be examined in more detail in section 4.4.2 of this paper.

Considering all of the above evidence, it seems that UH/M, in the context of self-repair, should be perceived as a kind of "Janus-faced" phenomenon (Tottie 2016: 107) which simultaneously refers to the preceding reparandum and the future repair; more simply put, UH/M could be argued to convey a message similar to "Oops, forget what I have just said, what I really mean is...". This meaning of UH/M has been argued to be largely dependent on the so-called "highlighting function" of UH/M; in the context of self-repair, then, this would suggest that the primary function fulfilled by UH/M is that of providing the repair with sufficient additional value to semantically outdo the reparandum in its direct vicinity, which in turn causes the listener to disregard the latter (Kjellmer 2003: 187). The hypothesis that UH/M can be used to highlight upcoming linguistic material is highly interesting because it contradicts the frequent assumption that UH/M "[...] go largely unnoticed in everyday comprehension" (Shriberg 1994: 28); for this reason, this hypothesis will be further explored in the following section, which will investigate the role of UH/M in speech comprehension.

#### **4.4. The role of UH/M in speech comprehension**

As was previously mentioned, there are good reasons to believe that UH/M produce positive effects on speech comprehension, for instance by helping listeners to understand a speaker's intention to take, keep or yield a turn or to correct a previously uttered item; thus, the present section seeks to further examine the role of UH/M for the listener. Generally speaking, this section will discuss two main questions, namely (a) whether or not listeners *notice* instances of UH/M (cf. section 4.4.1) and (b) whether or not (and how) they *use* them in the process of

speech comprehension (cf. section 4.4.2). The discussion of these questions will then be used to draw conclusions on whether or not listeners generally benefit from the presence of UH/M in spoken interaction.

#### **4.4.1. Do listeners notice UH/M?**

In the course of time, it has been suggested on numerous occasions that listeners do not notice every single instance of UH/M in the process of speech comprehension, but even seem to filter out the majority of their occurrences. A large proportion of evidence for this hypothesis stems from the finding that even listeners that are eager to heed all instances of UH/M – such as professional transcribers or researchers studying spoken discourse – find it hard to provide a complete record of every single occurrence (Deese 2010 [1980]: 70; Lindsay and O’Connell 1995: 101; Lickley and Bard 1996: 1876). According to Fox Tree (2001: 321), this frequently made observation seems to suggest the existence of a kind of “automatic filter” that identifies and removes instances of UH/M even before they become part of the usual procedure of identifying words and their grammatical roles that is characteristic of speech comprehension; consequently, UH/M are not consciously noticed by listeners. The existence of such a filter does not only prevent listeners from being distracted by the frequent uttering of UH/M, but also allows *speakers* to spontaneously construct speech without having to fear problems in communication; as Shriberg (1994: 29) argues, “[i]f we could not so easily process speech with DFs [i.e. *disfluencies* such as UH/M or false starts], we should have to devote great time and attention to the planning and delivery of each utterance”. In that sense, the filtering of UH/M does not only serve the listener, but also the speaker.

Even though it does indeed seem plausible to assume that listeners are able to filter UH/M in the course of speech comprehension, the *automatic filter-hypothesis* suggested by Fox Tree (2001: 321) does not seem to account for the fact that listeners *can* notice UH/M when they choose to do so. This becomes particularly obvious when considering the mere existence of research on and transcripts of UH/M, both of which could not exist if listeners could not notice UH/M even if they wanted to (Deese 2010 [1980]: 70; Lindsay and O’Connell 1995: 101; Lickley and Bard 1996: 1876). Moreover, several studies have shown that listeners may choose to use UH/M as a source of information, particularly when they are asked to focus on a speaker’s speaking style. This has been illustrated, for instance, by Brennan and Williams (1995: 396), who found that listeners seemed to show an increased awareness of UH/M when they were asked to judge how confident an unknown speaker was in answering questions to

general knowledge questions, a task which they assumed to cause listeners to increasingly focus on the stylistic features of a speaker's answer. Similarly, Martin and Strange (1968: 438) and Christenfeld (1995: 184 and 1995: 177) provided evidence that listeners tend to notice UH/M only when they are asked to pay attention to speaking style; in fact, they even argued that focusing on UH/M naturally inhibits the comprehension of content and vice versa. As Christenfeld (1995: 180) puts it, "When people pay attention to the *style* of the speaker, they are conscious of ums in the speech. However, when they attend to the *content* of the message, they show no awareness that an um is occurring roughly every seven seconds [italics added]". In summary, then, the above findings seem to suggest that listeners *can* choose to notice a speaker's uttering of UH/M; however, since focusing on UH/M may produce detrimental effects on the comprehension of content, listeners only do so when it is highly relevant to the respective listening task. This hypothesis suggests that the filtering of UH/M is not necessarily an automatic process, but under a listener's control (Fox Tree 200: 325); moreover, it is able to account for the fact that listeners do not generally seem to notice all instances of UH/M in everyday communication, where the content of a message is often more important than its stylistic features (cf. Christenfeld 1995: 185).

In addition to the above findings, the hypothesis that listeners are in control of the extent to which they pay attention to UH/M has been supported by evidence suggesting that they are consciously aware of the filtering process. In one of his studies of the effects of UH/M on speech comprehension, Christenfeld (1995: 180) has made the highly interesting observation that listeners who had been asked to guess the number of UH/Ms they had heard in a recording of a radio talk show estimated to have heard rather high frequencies of UH/M even when they had in fact been listening to a manipulated recording that contained not a single such instance. This seems to suggest that (a) listeners seem to be aware of UH/M being a highly frequent characteristic of everyday spoken discourse and thus expect their presence in spontaneous conversation and (b) that they seem to be aware of their filtering of UH/M during the listening process to such an extent that they assume to have heard some instances even though they had in fact not (Christenfeld 1995: 180).

Even though it seems that listeners are indeed able to actively manipulate the extent to which they notice UH/M, there is evidence suggesting that the filtering of UH/M in speech comprehension can also be determined by other factors that escape the listener's control. In his aforementioned experiment, Christenfeld (1995: 180) found that those listeners who had not actually heard any instance of UH/M in the audio recordings were rather accurate in

guessing the number of UH/Ms that would typically appear in such an interaction, while those who had heard the original recording with all occurrences of UH/M estimated to have heard significantly more. According to Christenfeld (1995: 180), this seems to suggest that listeners are aware of their filtering of UH/M and therefore assume to have filtered a certain proportion of UH/Ms on the one hand, and, on the other hand, that they seem to be able to notice unusually high densities of UH/M. This hypothesis is also in line with Deese (2010[1980]: 80), who proposed that listeners automatically start noticing UH/M when their frequency in a speaker's delivery is so high that they cannot be efficiently filtered anymore.

The finding that the filtering of UH/M may be determined by their density in a given speaking situation has been linked to the abovementioned observation that listeners only seem to notice UH/M when listening for style rather than content. This has been suggested, amongst others, by Christenfeld (1995: 184-185), who hypothesizes that “[w]hen an audience attends to style, it may well be a result of the content being unworthy of attention, or the speaker's style being distracting”. In other words, Christenfeld assumes that a certain frequency of UH/M in a given stretch of speech will most likely produce detrimental effects on the message's content, which in turn causes listeners to focus on stylistic features and subsequently also UH/M. A similar hypothesis has also been proposed by Lennon (1990: 391-392), who has suggested that listeners feel invited “to focus on the working of the production mechanisms” rather than the message itself when a speaker produces unusually high frequencies of distracting items (cf. also Gilquin 2008: 142-143). From these hypotheses, then, one could conclude that one possible reason for which a certain frequency of UH/M can trigger the deactivation of the listener's inner filter could be that high rates of UH/M might convey a message on their own that the listeners need to understand in order to efficiently manage communication, such as “I cannot think of this specific word, please help me!” or “I cannot think of anything else to say, please take the turn from me!”.

Having proposed that listeners are indeed able to notice UH/M in spoken interaction, the following section will investigate whether or not (and how) speakers use their awareness of UH/M to actively extract information that is relevant to their comprehension of speech.

#### **4.4.2. Do UH/M affect speech comprehension?**

In recent years, research has provided a growing body of evidence suggesting that UH/M positively influence speech comprehension even when listeners do not actively choose to notice them. This hypothesis has been supported, for instance, by Barr (1998), Fox Tree

(2001), and Arnold and colleagues (2004; 2007), all of whom have suggested that UH/M positively affect the speed at which listeners process a given message. In that sense, Fox Tree (2001: 322) has shown that listeners were faster in identifying given target words in speech recordings when they were preceded by UH/M, which has led her to hypothesize that “*ums* and *uhs* could benefit on-line speech comprehension by prompting listeners to pay more attention to upcoming speech” (Fox Tree 2001: 325). A similar phenomenon was also observed by Barr (1998: 599), Arnold et al. (2004: 581) and Arnold, Hudson-Kam and Tanenhaus (2007: 924), who have all shown that listeners were faster in identifying a picture or object when the provided oral descriptions included instances of UH/M. Interestingly, this was especially the case when the item that was being referred to was new to the discourse. According to the authors of the respective studies, this seems to be largely due to the fact that listeners appear more likely to expect an upcoming (lexical or conceptual) item to be new to the discourse when it is preceded by UH/M; this, in turn, facilitates the process of identifying the new element (Barr 1998: 599; Arnold et al. 2004: 581; Arnold, Hudson-Kam and Tanenhaus 2007: 924).

The finding that UH/M seem to facilitate the identification of new or important lexical items has led some scholars to assume that they act as a kind of “semantic booster” or “verbal italics” that are able to attribute particular importance to upcoming linguistic items (Kjellmer 2003: 187). This assumption has been supported, amongst others, by Stenström (1990, cited in Kjellmer 2003: 187), who claims that “[o]ne often gets the impression, when listening to [a] recording, that the speaker inserts a pause before or after a particular word or string of words to obtain a certain emphatic effect rather than using pauses as linguistic demarcators (and for breathing)”. In a similar line of argument, Kjellmer (2003: 187) claims that UH/M are often used to communicate that the following element is chosen with particular care, which in turn suggests that it is very important and thus deserves increased attention from the listener. The conception of UH/M as highlighting devices has been further supported by evidence suggesting that speakers utter UH/M whenever they introduce elements that are expected to cause difficulties for the listener. For instance, Goldman-Eisler (1958b: 67) observed that “[w]here a sequence ceased to be a matter of common conditioning or learning, where a speaker’s choice was highly individual and unexpected [...] speech was hesitant”; this finding led her to the conclusion that UH/M (along with silent pauses) should be perceived as important means of communicating one’s intention to utter a word, stretch of speech or piece of information of low transitional probability, thus helping to avoid confusion

and misunderstandings (cf. also Goldman-Eisler 1958a: 104). Taken together, the findings presented in the present paragraph seem to suggest that UH/M can benefit the process of on-line speech comprehension by highlighting important semantic elements; however, as will be shown in the following paragraph, there are also some limitations to this line of argument that need to be considered before drawing any final conclusions on the mechanisms underlying the effects of UH/M on speech processing.

Even though it is tempting to interpret the above findings to suggest a highlighting function of UH/M, it is important to point out that the existence of such a function is not the only possible explanation for the observed positive effects of UH/M on the speech decoding process (Corley and Hartsuiker 2003: 277). Instead, it has been argued by certain scholars that UH/M might benefit the process of linguistic decoding not by adding extra semantic value to certain stretches of speech, but rather by delaying the uttering of the following lexical item(s), thus “[...] improving intelligibility, in the same way that clear pronunciation does [...]” (Corley and Stewart 2008: 594; cf. also Brennan and Schober 2001: 293). The assumption that the beneficial effect of UH/M on speech comprehension can be attributed to the associated delays in speech rather than their specific linguistic nature has been supported by Bailey and Ferreira (2003: 197), who found that the observed positive effect of *uh* on a listener’s understanding of complex syntactic structures could be replicated when *uh* was replaced by environmental noises such as a ringing telephone. This observation has led the two authors to the conclusion that it might be the interruption itself rather than its content that contributes to the “disambiguat[ing of] syntactic structures” (Corley and Stewart 2008: 598; cf. also Bailey and Ferreira 2003: 197). One major disadvantage of this hypothesis is that it seems to equate UH/M with silent pauses, for both of these phenomena constitute a delay in speech; however, this is problematic insofar as this idea has been refuted on several occasions throughout research history (Mahl 2014 [1987]: 277; Christenfeld 1994: 193; MacLay and Osgood 1959: 41; Clark and Fox Tree 2002: 75; cf. also section 2.2.2). In summary, then, it seems that the beneficial effects of UH/M can neither be entirely attributed to a highlighting function (Kjellmer 2003: 187) nor to the mere fact that they constitute periods of semantic void.

Since the beneficial effects of UH/M on speech comprehension can neither be entirely explained by the existence of a highlighting function nor by their associated delays in speech delivery, certain scholars have attempted to offer other potential explanations for the observed effects. One very promising such explanation has been provided by Arnold and Tanenhaus (2011: 210-211), who have suggested that a listener’s successful interpretation of UH/M

largely depends on his/her intuitive understanding of the speech production process and the associated awareness of potential reasons for uttering UH/M. This hypothesis is based to a large extent on the observation that listeners were generally found to expect UH/M to occur before discourse-new elements as well as near syntactic boundaries, which could be interpreted as evidence for the assumption that listeners are aware of the planning function of UH/M (Arnold and Tanenhaus 2011: 200-201, 210; Bailey and Ferreira 2003: 197). Following this line of argument, it would seem likely that listeners might use this theoretical knowledge on the functions and determinants of UH/M to interpret the meaning of UH/M in a given situation. In other words, according to this approach, a listener hearing an instance of UH/M is able to infer from it that the speaker is currently engaged in a planning process; consequently, he/she will assume the upcoming stretch of speech to have required an additional amount of planning and will therefore expect the upcoming stretch of speech to be of particular semantic value. In summary, this approach suggests that listeners do not understand the message conveyed by UH/M because the sign itself carries a specific meaning, but because the listener is usually also a speaker who has acquired intuitive knowledge on psycholinguistic processes that he/she can transfer to the context of listening (Arnold and Tanenhaus 2010: 211). This view does not only account for the findings obtained in the abovementioned experiments, but also provides some interesting insights into the ways UH/M can act as a interpersonal sign that serve both the speaker and the listener (cf. chapter 5 for a more detailed discussion of this multidimensional conception of UH/M).

In summary, the present section has suggested that UH/M do produce beneficial effects on conceptual and linguistic comprehension even in content-based tasks that do not require the listeners to consider stylistic features of speech. Thus, it can be concluded that that UH/M seem to serve both the speaker and their listener in reaching their shared goal of smooth communication, even if the precise mechanisms underlying these positive effects remain unknown up to this point. Together with the aforementioned observation that listeners are able to interpret instances of UH/M to understand a speaker's intentions with regard to the turn-management system (cf. section 4.2), it thus seems safe to conclude listeners are able to use as a resource in UH/M in everyday conversation.

Having provided a brief discussion of the (mechanisms underlying the) positive effects of UH/M on speech comprehension, the following – and final – section of this chapter will be dedicated to a recapitulation of the functions of UH/M that have been identified and discussed in the present chapter.



#### **4.5. Interim conclusion: The functions of UH/M in conversation**

From what has been shown in the course of this chapter, it can be concluded that UH/M certainly do not seem to be one-dimensional phenomena that can be attributed to one single function. As has been discussed in the previous sections, there is a substantial body of evidence that supports not only the assumption that UH/M can be linked to the planning of upcoming speech on the conceptual, syntactic and lexical level, but also the idea that UH/M act as important devices of turn-management; moreover, the finding that UH/M frequently function as editing phrases in self-repair suggests that UH/M fulfill certain functions that are vital to the understanding of “problematic”, i.e. potentially confusing, stretches of speech. Finally, concerning the role of UH/M in speech comprehension, it has been shown that there is plenty of evidence suggesting that UH/M are partly noticed by listeners, especially when the respective listening task requires them to do so or when the frequency of UH/M in a given stretch of speech renders them impossible to ignore. Even more interestingly, UH/M appear to positively influence the speech comprehension process by directing a listener’s attention towards discourse-new elements and the beginning of syntactic units. The fact that these beneficial effects have been observed even in content-based listening tasks that did not explicitly require listeners to focus on UH/M seems to suggest that listeners appear to use UH/M for the purpose of speech comprehension even if they do not consciously notice them; as Bailey and Ferreira (2003: 197) put it:

[...] disfluencies may ultimately be ‘filtered’ in the sense that they do not become part of the sentence’s final representation; but because the presence of an interruption may affect the syntactic parser by causing it to predict a particular type of constituent, the disfluency may still have influenced language comprehension processes.

In other words, the fact that listeners are not consciously aware of the instances of UH/M they encounter in their interlocutors’ speech does by no means suggest that listeners do not subconsciously access their meaning; what remains unknown up to this point, however, are the precise mechanisms that underlie these effects of UH/M on comprehension. One of the most interesting propositions that has been made in this context stems from Arnold and Tanenhaus (2011: 211), who have asserted that “comprehension processes are influenced by representations of the mental processes of the speaker”, thus suggesting that listeners use their knowledge of the speech production process to deduce the meaning of certain passages. This hypothesis is highly attractive and certainly deserves academic attention; unfortunately, however, the exploration of such mechanisms significantly exceeds the scope of this paper. In an attempt to provide a more theoretical summary of the above, the following chapter will try

to embed the findings obtained in the course of the present chapter into different theoretical models of the linguistic sign, thus aiming at providing a theoretical representation of UH/M that might provide some interesting insights into the mechanisms underlying their production.

## 5. Towards a theory of UH/M

Up to this point, this paper has presented a large body of evidence suggesting that the many functions of UH/M are not mutually exclusive in a way that UH/M can act as *either* signals of turn-management *or* expressions of cognitive processes, but that they should rather be perceived to complement each other in such a way that both the speaker and the listener may profit from the presence of UH/M in spoken discourse. According to this approach, then, UH/M should be perceived as at least two-sided phenomena which, at least in dialogic contexts, often seem to act as both an expression of a speaker's cognitive processes *and* a signal to the listener (see chapter 4.4.3); moreover, it is argued that UH/M are multidimensional in the sense that they are influenced by various aspects of the respective speaking environment (e.g. genre, topic, participants, etc.; see section 4.1). This proposed view of UH/M is significantly more complex than what has been suggested in many – though not all – other publications, and it is thus deemed necessary to provide a theoretical framework that is able to account for it. To my knowledge, such a theoretical approach to the study of UH/M as a multi-dimensional linguistic sign has not been proposed before; therefore, the following chapter will predominantly be based on my own application of existing linguistic models that are deemed appropriate to represent UH/M.

An important first step in the enterprise of constructing a theoretical representation of UH/M is the consideration of existing models of communication and language that could be apt for the purpose. After a consideration of assumed characteristics of UH/M, it was decided that a potentially applicable model would need to comprise the following three characteristics: (1) it should present the linguistic sign as a multi-dimensional entity that equally serves the speaker and the listener; (2) it should present these different dimensions as complementing layers, and finally, (3) it should include contextual parameters that shape the sign in the same way the frequency of UH/M, for instance, is shaped by the privacy or audience of a given conversation (see chapter 4.1). A brief study of linguistic history based on these requirements revealed that there are three models which seem particularly worth considering for the present context: firstly, Karl Bühler's *Organon model*, secondly, Roman Jakobson's model of communication, and thirdly, M.A.K. Halliday's model of the metafunctions of language which he proposed in

the context of his Systemic Functional Linguistics. As will be shown below, The following sections will discuss all three of these models with regard to their applicability to UH/M, where section 5.1 focuses on Bühler's and Jakobson's framework, while section 5.2 will be dedicated to the examination of Halliday's model.

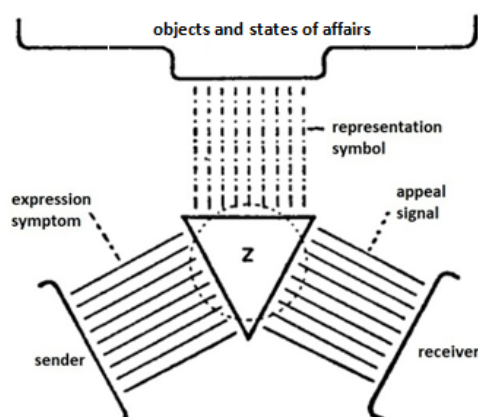
### **5.1. Karl Bühler, Roman Jakobson, and UH/M**

Of the many theoretical models that aim at representing the nature of the linguistic sign, Karl Bühler's *Organon model* (Bühler 1934: 28) was probably one of the first to propose that signs are multi-dimensional entities that are both symptomatic of a speaker's inner state and a signal to his/her listener (Mair 2015: 110). As will be shown below, Bühler (1934: 28) assumes a three-dimensional structure of the linguistic sign, where each dimension corresponds to one of the three basic functions of language: the *expression* of one's internal state, the *representation* of objects and states of affairs, and the *appeal* to one's listener. This three-dimensional model was further refined by Roman Jakobson in 1960 to include not three, but six functions of language, thus presenting the linguistic message as an even more complex entity that is able to fulfill multiple functions in a given context. It is precisely this multi-dimensionality of both of these frameworks that has led me to assume that both Bühler's *Organon model* and Jakobson's model of communication could provide an adequate framework for the theoretical representation of UH/M; however, as will be shown below, an application of the two models to UH/M poses several obstacles that render them significantly less useful for this purpose. The present section seeks to outline the advantages and disadvantages both of these models.

It has already been stated above that Bühler's *Organon model* (1934: 28) was perhaps one of the first to represent the linguistic sign as a multi-dimensional unit that may be both expressive of a speaker's thoughts and emotions and a signal to his/her listener (Mair 2015: 110). As is illustrated in Figure 2 below, Bühler (1934: 28) suggests that linguistic signs are communicative tools which fulfill three main functions: the *representation* of real-life objects and events in communicative messages, the *expression* of a speaker's thoughts and emotions, and the *appeal* to a listener so as to influence his/her behavior in a given context. Based on this assumption, Bühler (1934: 28) distinguishes three dimensions of the linguistic sign: the sign as an expressive *symptom* of a speaker's mental state, the sign as an appellative *signal* to a listener, and the sign as a *symbol* used to describe real-life objects and events. According to Bühler (1934: 30-31), each linguistic sign always consists of a combination of these three functions, even if he asserts that one usually dominates the other two. To provide an example

of this approach, the sentence *It is already late* could be dominated by either of the three defined functions. If dominated by the representative function, the sentence would act as a rough estimation of time (e.g. when a child asks his/her mother why he/she has to go to bed), whereas the same sentence could express a speaker's thoughts and feelings when dominated by the expressive function (e.g. when used to express "I am tired and want to go home now"). Finally, the sentence could also be dominated by its appellative function, in which case it could act as an indirect request to one's interlocutor to leave one's current environment in order to go home. Consequently, the sign – in this case, the sentence – could predominantly be *symbol* of a real-life circumstance (the time), *symptom* of a speaker's desire to go home, *signal* to the listener, or, even more probably, a combination of all of the three.

Figure 2: Bühler's *Organon model* of the linguistic sign (Bühler 1934: 28), translations copied from Nöth (1990: 186)



The most striking argument in favor of applying Bühler's framework to the study of UH/M is the fact that it contradicts neither the symptom nor the signal hypotheses (which are roughly represented by Bühler's expressive and appellative dimension respectively), but rather creates an environment in which both of them complement each other. As has been repeatedly pointed out in the preceding chapters, there is now significant evidence in favor of a multi-dimensional conception of UH/M, suggesting that they serve both the speaker (by creating time for planning without losing his/her right to the floor) and the listener (by informing him/her about the speaker's cognitive state and thus facilitating interpersonal communication); subsequently, Bühler's model could be argued to be highly useful to the study of UH/M because it presents the sign in precisely this light of multi-dimensionality.

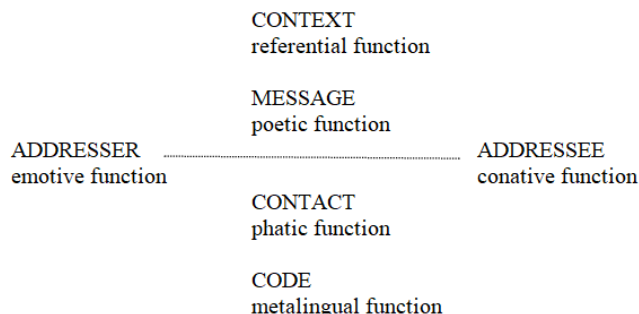
Despite the advantage outlined above, a closer investigation of Bühler's model reveals that it is in fact not entirely compatible with the common conception of UH/M, mainly because not

all of the proposed dimensions correspond with what is known about the functions of UH/M. While the expressive and the appellative functions proposed by Bühler (1934: 28) could indeed be argued to be applicable to UH/M for they roughly represent their ability to signal a speaker's cognitive state and their capacity to influence a listener's behavior with regard to the turn-management system respectively, the third dimension of the *Organon model* – the symbolization of objects and states of affairs – poses a serious obstacle for the application to the study of UH/M. This is because of the fact that even though the precise functions of UH/M are still subject to animated academic discussion, those scholars assuming them to serve certain purposes in communication generally agree that they do not designate occurrences, events, facts or emotions, but rather carry implicit information on the process of speech production and/or communication management that needs to be decoded by the listener through the consideration of contextual cues such as intonation or preceding discourse (see chapter 4.4.3). For this reason, UH/M do not appear to fulfill a representational function in the sense of Bühler's proposition. At most, the fact that UH/M are frequently associated with a speaker's thinking and planning could be interpreted to suggest that UH/M are able to represent that a speaker needs more time to plan the rest of the utterance; however, in my personal opinion, this argumentation constitutes a rather loose interpretation of the model that does not exactly correspond to Bühler's initial proposition as outlined above. In addition to the above, it could also be argued that Bühler's principle of the dominating function outlined above seems to contradict the perception of UH/M as *simultaneous* expressions of cognitive processes *and* signals of turn-keeping; however, this feature of the *Organon model* seems to be in line with the popular hypothesis that UH/M serve one basic function – the signaling of cognitive processes – which is frequently complemented by other, appellative functions such as the signaling of turn-keeping (cf. Fischer 2006: 432; Tottie 2016: 115), and is therefore not perceived as a major disadvantage of the model.

If the *Organon model*, in its basic structural representation of the sign, is thought to provide a good starting point for the enterprise of developing a theoretical framework of UH/M despite failing to account for several of their characteristics, it seems only logical to consider its successor, Roman Jakobson's model of communication (cf. Jakobson 1960: 357) as a potential alternative. Based on Bühler's *Organon model*, Jakobson developed a framework according to which each linguistic message may fulfill up to six different functions which can be roughly categorized to match Bühler's three dimensions of the sender, the receiver, and the message (Nöth 1990: 185; Innis 1998: 2201); in that sense, Jakobson's framework can be

thought of as an expanded version of Bühler's model. In order to render the upcoming discussion more comprehensive, a visual representation of the model is provided in Figure 3.

Figure 3: Roman Jakobson's model of communication (cf. Jakobson 1960: 353; Jakobson 1960: 357)



As can be seen from Figure 3, Jakobson (1960) expanded Bühler's *Organon model* to include not three, but six linguistic functions. This development is based on his assumption that there are six factors that are indispensable in order for the communicative act to succeed: (1) the presence of a *sender*, (2) the presence of a *receiver*, (3) a *context* the message can refer to, (4) psychological *contact* between the sender and the receiver that is manifested through a physical channel, (5) the *message* itself, and, lastly, (6) a common *code* that needs to be shared between the sender and the receiver (Jakobson 1960: 353). According to Jakobson (1960: 353-35), each of these factors is represented by a communicative function: (1) the *emotive* function, which establishes a link between a sender's thoughts and emotions and the linguistic message, (2) the *conative* function, which directly relates to the receiver, engaging him/her directly (through the use of imperatives, for instance) or indirectly (through implicit prompts or requests) (3) the *referential* function, which corresponds to the message's context and the representation of the real world within the linguistic message, (4) the *phatic* function, which centers around the establishment or prolongation of social contact, (5) the *poetic* function, where "the focus [lies] on the message for its own sake" (Jakobson 1960: 356), as is the case in poetry, and, finally, (6) the *metalingual* function, where language is used to comment on the message itself (as is the case, for instance, when providing a definition or asking for clarification, cf. Decreus et al. 1990: 19) (Nöth 1990: 187; Jakobson 1960: 353-357). What has been added to Bühler's *Organon model*, then, are the factor of the code and its corresponding metalingual function, the contact (phatic function), and the message (poetic function), while the referential, emotive and conative function seem to have been more or less directly adopted from Bühler's representational, expressive and appellative dimensions respectively. Thus, Jakobson's model offers a more integral representation of linguistic events

which Bühler's initial framework had not been able to account for, and could thus potentially be more relevant to the study of UH/M than its predecessor.

One of the main assets of Roman Jakobson's model is the newly introduced phatic function, which seems to represent the functions fulfilled by UH/M more accurately than Bühler's appellative function, which in turn is more or less reflected in Jakobson's conative function. While the definition of the conative or appellative function is based on the assumption that the linguistic sign acts as a communicative signal that aims at influencing the addressee's behavior (cf. Bühler 1934: 28; Jakobson 1960: 335), the phatic function focuses more on the interactional component of the sign. This is reflected in Jakobson's definition of the phatic function, according to which it "[...] primarily serv[es] to establish, to prolong, or to discontinue communication, to check whether the channel works ("Hello, do you hear me?"), to attract the attention of the interlocutor or to confirm his continued attention" (Jakobson 1960: 335). Following this definition, Jakobson's phatic function seems to be significantly more relevant to the theoretical representation of UH/M than his conative (or Bühler's appellative) function, because it more accurately represents the fact that UH/M are more often linked to the signaling of turn-keeping, turn-holding or turn-yielding intentions than to the transmission of explicit or implicit commands (cf., for instance, Künzel 1997; Maclay and Osgood 1959; Kjellmer 2003; Clark and Fox Tree 2002). In other words, Jakobson's phatic function seems to more accurately represent the role of UH/M in interaction management, whereas the inclusion of his conative function into a model of UH/M would require at least a slight redefinition of the function as well as a somewhat cumbersome argumentation that would need to justify the interpretation of turn-management behavior as an appellative command.

Apart from the phatic function, however, Jakobson's expansion of the contextual component of the *Organon model* does not seem to render his model any more useful to the study of UH/M than is its predecessor, for the large majority of functions included in the model are still irrelevant to this specific context. For instance, while Bühler has never specified the precise nature of his expressive function, Jakobson (1960: 354) defines his equivalent, the emotive function, as "[...] tend[ing] to produce an impression of a certain *emotion*, whether true or feigned [...] [my emphasis]", thus suggesting an emotion-based notion of a sign's expressivity that does not conform with what is known on the expressive functions fulfilled by UH/M (see chapter 4.1). Moreover, Jakobson's referential function, which relates the linguistic message to the real-life objects it refers to (Jakobson 1960: 355) does not seem any

more applicable to UH/M than Bühler's representational function. Lastly, neither the metalingual function, which is observable "[w]henver the addresser and/or the addressee need to check up whether they use the same code" (Jakobson 1960: 356) nor the poetic function, which centers around "the message for its own sake" (Jakobson 1960: 356) seem to correspond in any way to the functions of UH/M that have been identified in relevant research.

Summing up the above discussion of Bühler's *Organon model* and Jakobson's model of communication, it can be concluded that even though some of their central elements – such as the idea of the sign as a multi-dimensional entity that is situated between the speaker and the listener – seem promising to the enterprise of creating a theoretical framework of UH/M at first glance, their application to the intended context seems difficult. This is mainly due to the fact that not all of the two models' different dimensions are applicable to the study of UH/M, for at least the representational/ referential dimension would have to be entirely removed or replaced for the purpose of representing UH/M. Nevertheless, both Bühler's *Organon model* and Jakobson's model of communication comprise various aspects that seem relevant to the theoretical representation of UH/M, since they both share at least some characteristics with what is known about UH/M. In fact, even Bühler's principle of dominance, according to which a sign is always dominated by one function, with the other functions merely complementing the primary message (Hasan 2009: 19), could be argued to be in line with existing literature on UH/M, for it has sometimes been suggested that UH/M serve the primary function of signaling planning, which in turn is complemented by other functions such as signaling turn-keeping and/or turn-yielding, for example (cf. Fischer 2006: 432; Tottie 2016: 115). However, due to the inappropriateness of many of the two models' dimensions for the study of UH/M, the enterprise of applying these frameworks to UH/M has not been entirely successful; thus, the present discussion will now move on to an entirely new model that is hardly related to the above two frameworks, namely Halliday's model of the metafunctions of language, which he presented in the context of his Systemic Functional Linguistics.

## **5.2. UH/M and the three metafunctions of Systemic Functional Linguistics**

Since it has been shown above that neither Bühler's *Organon model* nor Jakobson's model of communication appear entirely apt for representing UH/M, the current section will be dedicated to a model that differs significantly from the other two in many ways and is



therefore hoped to provide a more suitable framework for the analysis of UH/M; this model is Halliday's model of the three metafunctions of language.

As has been briefly mentioned before, Halliday's model was proposed in the context of his *Systemic Functional Linguistics* (hereafter referred to as SFL), a theoretical approach to language that focuses on the construction of meaning in context (Caffarel 2010: 797). According to this approach, all linguistic acts are products of an interplay between three metafunctions<sup>10</sup> of language: the *ideational metafunction*, which concerns the ability of language to linguistically construct and express a speaker's experience of the world (Halliday 2004: 29), the *textual metafunction*, which enables speakers to do so in the form of "intelligible coherent discourse" (Hasan 2009: 19), and, perhaps most important to the study of UH/M, the *interpersonal metafunction*, which is related to the notion of language being a means of "enacting our personal and social relationships with the other people around us" (Halliday 2004: 29-30) by enabling speakers to build and/or maintain contact with their interlocutors. According to Halliday, all of these three metafunctions are equal in terms of importance and dominance, and they are inextricably intertwined in language in such a way that they are ubiquitous in every communicative act, with each metafunction being related to different aspects of the speech production process (Hasan 2009: 19). For instance, while the grammatical realization of the ideational metafunction concerns questions of apt lexicon, tense or transitivity (Hasan 2009: 19), the textual metafunction is linked to textual aspects of the construction of messages such as cohesion and coherence; finally, the interpersonal metafunction finds its realization in choices concerning the mood and modality (Hasan 2009: 19). As a final point, Halliday proposes that the choices related to the three metafunctions of language are often determined by the three dimensions of context, which he defines as the *Field*, i.e. the activity in which the speakers are involved (Caffarel 2010: 806), the *Tenor*, i.e. "the role and status relationships among participants" (Caffarel 2010: 806), and, lastly, the *Mode*, which concerns the "rhetorical function of language" as well as the type of language that is being used (Caffarel 2010: 807).

Even on a very general level, Halliday's model of the three metafunctions of language seems to account for several aspects of UH/M. For instance, Halliday's proposition that language is influenced by Field, Tenor, and Mode of a given context seems to be in line with findings

---

<sup>10</sup> Note that Halliday uses the term *metafunction* in an attempt to distinguish his approach from other theoretical models of language (such as Bühler's or Jakobson's aforementioned frameworks) that have studied the functions of individual utterances or signs rather than of language in general (Hasan 2009: 19).

suggesting that the presence and frequency of UH/M in a given discourse situation seems to be partially determined by contextual parameters such as genre, audience, and linguistic and/or social conventions.; moreover, Halliday's view of language as an interplay of various aspects and metafunctions seems to reflect to at least some extent the aforementioned multidimensionality of UH/M by accounting for a certain simultaneity of expressive, linguistic, and interpersonal functions. In addition to this general observations, Halliday's model also seems relevant to the study of UH/M when considering its more specific aspects; this will be discussed in the following paragraphs.

Applying Halliday's model to the study of UH/M, it seems plausible to argue that the uttering of UH/M is first and foremost a manifestation of the interpersonal metafunction of language. As has been proposed on numerous occasions throughout this paper, UH/M have been frequently shown to have a strong interpersonal component that seems to act as a source of information for the listener by expressing a speaker's inner state. In that sense, UH/M have often been found to facilitate conversation management by providing the listener with information on the cognitive processes the speaker is currently involved in on the one hand and, on the other hand, signaling a speaker's intention to take, keep, or yield his/her turn. Based on these findings, UH/M could be argued to contribute to the "enacting [of] our personal and social relationships with the other people around us" (Halliday 2004: 29-30) by helping speakers to build and/or maintain linguistic contact. Thus, it seems legitimate to conclude that the uttering of UH/M is a manifestation of the interpersonal metafunction of language, which serves the main purpose of facilitating and managing interpersonal communication.

Even though it appears plausible to assume that UH/M are embedded in the interpersonal metafunction of Halliday's model, it is interesting to observe that they also seem to be linked to the model's other dimensions. As has been suggested on numerous occasions throughout this paper, UH/M can often be associated with a speaker's ongoing cognitive processes, all of which can be ascribed to one of Halliday's three metafunctions of language. To provide only a few examples, it could be argued that information retrieval and the conceptual planning of messages could be ascribed to Halliday's ideational metafunction, whereas syntactic planning is embedded in the textual metafunction. In that sense, it seems that even though UH/M function as interpersonal signals, they can be directly associated with a speaker's decisions made with regard to the ideational and textual levels of speech production; according to this hypothesis, then, UH/M may be uttered whenever a speaker wishes to signal that he/she is

currently involved in making choices linked to either of Halliday's metafunctions. To provide a brief example of this approach, consider, for instance, the hypothetical situation in which a speaker is asked a question. As a consequence, the speaker starts to think about the question and formulate his/her answer; these processes can be ascribed to the ideational and textual dimensions of Halliday's model. Note that these processes do not necessarily cause the speaker to utter UH/M; what does, however, is his/her conscious or subconscious decision to signal that he/she accepts the turn and is currently formulating his/her answer. In that sense, UH/M could be argued to be an interpersonal signal that is able to indicate processes concerning the ideational and/or textual dimension. The decision of whether or not to utter UH/M seems to be mostly influenced by the respective conversational context; aspects potentially influencing this decision include the question of whether or not there is the need to actively signal turn-keeping (which is the case in spontaneous conversation among friends, but not in lectures or political speeches), whether or not the speaker can expect to receive assistance from his/her listeners (which is relevant in situations where UH/M marks difficulties in lexical retrieval or a speaker's desire to yield a turn), or whether or not the uttering of UH/M could help the listener in disambiguating potentially confusing stretches of speech such as self-corrections or complex linguistic structures.

Considering all of the above, it can be concluded that Halliday's model seems to provide a rather meaningful framework for studying UH/M. Following Halliday's assumption that language is shaped by an interplay of various decisions that are simultaneously made on the three metafunctional levels, it could be argued that UH/M are embedded in the interpersonal dimension of Halliday's model, for they predominantly serve the purpose of facilitating communication between two interlocutors; however, it could equally be suggested that the decision of whether or not to utter UH/M is based on the question of whether or not a speaker wants to actively signal that he/she is occupied with making choices concerning the ideational and textual dimension of speech production in a given context. To my mind, of the main asset of Halliday's model is that despite assuming a tripartite structure of language that is similar to Bühler's<sup>11</sup>, it is different from the latter insofar as it does not necessarily seem to suggest that

---

<sup>11</sup> Note that both Halliday's and Bühler's models do not only assume a tripartite structure of language, but also include more or less the same dimensions, which represent (1) the speaker, (2) the listener(s), and (3) the world surrounding them. However, what distinguishes these two models is the way they categorize these dimensions: while Bühler assumes that the sign consists of an *expressive*, an *appellative*, and a *representational* level, Halliday combines his rough equivalent of Bühler's expressive dimension with the representative dimension to form his *ideational* level (cf. Halliday 2004: 29). Moreover, as concerns Halliday's interpersonal level, it seems that it is more inspired by Jakobson's phatic function than by

UH/M have to actively fulfill all three of the defined metafunctions; instead, it allows for the hypothesis that UH/M are interpersonal signals that may – but do not always have to – indicate a speaker’s decision-making with regard to the ideational and textual dimensions.

Summing up the present chapter, it has been shown that both Bühler’s, Jakobson’s and Halliday’s models seem to be applicable to the study of UH/M, even though it is my personal belief that the view of UH/M as a manifestation of Halliday’s interpersonal metafunction seems to be the most promising theoretical approach to the representation of UH/M up to this point. This personal preference of Halliday’s model is based on the fact that this framework seems to most accurately represent the functions fulfilled by UH/M without having to adapt it to this specific purpose; however, this is by no means to suggest that Bühler’s and Jakobson’s models are not able to act as a theoretical representation of UH/M. Of course, the above discussion merely constitutes a rough preliminary to a theory of UH/M that will need to be further revised and refined in the future; still, I believe that this discussion has been able to provide a valuable insight into the nature of UH/M by applying recent findings to a theoretical framework. Given that the creation of such a theoretical representation of UH/M does not constitute the actual purpose of this paper, but has rather been attempted for the sake of providing a comprehensive summary of the complex analysis of the various functions of UH/M presented in chapter 4, this enterprise shall be discussed no further in the course of this paper; instead, the following chapter will mark the return to the actual topic of this paper by considering the question of whether or not there are functional differences between *uh* and *uhm*, which will also serve as a basis for the corpus-based study conducted for the purpose of this thesis.

## **6. Is there a difference between *uh* and *uhm*?**

After having provided an extensive global insight into the nature and functions of UH/M in general, the actual topic of this paper – the question of whether or not there are syntactic and/or functional differences between the vocalic *uh* and the vocalic-nasal *uhm* – can finally be addressed. To my knowledge, only few scholars have addressed this question up to this point, for *uh* and *uhm* have mostly been treated as two different realizations of the same

---

Bühler’s appellative function, for it focuses more on the maintenance of interaction than on the mere appellative influencing of a listener’s behavior (Halliday 2004: 29-30). Finally, it should be pointed out that the textual function, too, can only be found in Jakobson’s model of communication, but not in Bühler’s *Organon model*. In that sense, Halliday’s model seems to be a compromise between Bühler’s and Jakobson’s models.

element; those who have sought to study them separately, however, seem to generally agree that there are some systematic differences between *uh* and *uhm* (cf. Kjellmer 2003: 173; Fox Tree 2001: 320-321; Tottie 2016: 117; Clark and Fox Tree 2002: 86). One of the earliest hypotheses concerning potential differences between the two variants assumed that speakers utter *uhm* instead of *uh* for the same reason they produce the indefinite article *an* instead of *a*, namely to prevent the stutter caused by uttering two independent vowels right after each other; however, this hypothesis was refuted by Kjellmer (2003: 173). Instead, *uh* and *uhm* have been argued to distinguish themselves from one another by announcing different lengths of delay (Clark and Fox Tree 2002: 86) or representing different kinds of planning (Clark and Fox Tree 2002: 93-95; Shriberg 1994: 154). The present chapter summarizes the evidence provided in support of potential differences between the two, taking into account both syntactic and functional aspects of the discussion. This will then act as a basis for the formulation of the research questions to be presented later in this thesis (see section 7.1).

One of the most frequently cited propositions concerning the question of whether or not *uh* and *uhm* fulfill different functions stems from Clark and Fox Tree (2002: 86), who have famously suggested that “*uh* signals the initiation of what is expected to be a minor delay, and *um*, what is expected to be a major delay”. This hypothesis has been supported by the observation that *uhm* is not only more often followed by silent pauses than *uh*, but that the pauses succeeding *uhm* are also usually longer than those occurring after *uh* (Clark and Fox Tree 2002: 82-84). Further evidence for this view stems from de Leeuw (2007: 108), whose cross-linguistic study of UH/M and their respective counterparts in other languages suggested that *uhm* is significantly more often associated with long silent pauses (i.e. pauses that are longer than one second) than *uh*, which in turn seems to suggest that *uhm* indeed appears to be the variant that is more often linked to longer delays in speech. These findings also coincide with the frequently made observation that *uhm* shows a strong tendency to occur at or near major syntactic boundaries such as the onset of intonation units or sentences (Clark and Fox Tree 2002: 94-95), while *uh* seems to prefer medial positions in such units (Shriberg 1994: 154; Clark and Fox Tree 2002: 94-95). This, too, seems to suggest that *uhm* can be associated with more complex planning processes that require more cognitive effort and consequently also longer pausing (e.g. syntactic planning), while *uh* is a marker of local decision-making that can usually be accomplished within shorter periods of time (e.g. lexical retrieval) (Shriberg 1994: 154). Finally, Fox Tree (2001: 320) further supports the hypothesis that *uh* and *uhm* act as signals of minor and major delay respectively by pointing out that listeners

seem to benefit only from the presence of *uh* – but not from *uhm* – in the process of speech comprehension (Fox Tree 2001: 324). This argumentation is predominantly based on the assumption that listeners would interpret the indication of a minor delay as evidence for a speaker momentarily searching for a particularly important expression, word, concept, etc. that is about to follow and would thus heighten their attention in anticipation of this salient element, while the announcement of a long delay would cause the listener to expect a longer period of informational void throughout which he/she could not possibly maintain increased attention (Fox Tree 2001: 320).

Despite all of the above evidence, the view of *uh* and *uhm* as signals of minor and major delays respectively has been criticized by several researchers. For instance, O’Connell and Kowal (2005: 567) have presented data according to which only roughly a quarter of all instances of *uh* and *uhm* were followed by silent pauses, a proportion that “hardly suffices to render these fillers reliable signals of an upcoming silent pause”. Moreover, it should be pointed out that the majority of findings presented by Clark and Fox Tree (2002) is based on the pure impressionistic estimation of pause length, thus representing merely “how long the pauses would feel to listeners [...] [i.e.] the *perception* of pause length [...] [original emphasis]” rather than objectively measured criteria (Clark and Fox Tree 2002: 81; Corley and Stewart 2008: 592). In addition, the idea of a universally valid, systematic difference between the two variants seems to be refuted by findings suggesting that a speaker’s choice of *uh* or *uhm* seems to depend to at least some extent on personal preference and/or dialect (de Leeuw 2007: 109; Shriberg 1994: 155-156). Lastly, it has been briefly addressed by Clark and Fox Tree (2002: 80) themselves that instead of being deliberately planned signals of upcoming minor/major delays, there is a possibility that *uh* and *uhm* might be spontaneous reactions to one’s ongoing silence instead, with *uhm* acting as mere prolongation of *uh* (cf. also Maclay and Osgood 1959: 42). However, this approach seems easily refutable when considering not only evidence on their aforementioned different syntactic preferences, but also the fact that both *uh* and *uhm* may be prolonged to form *u:h* and *uh:m*, the former of which would be redundant if non-prolonged *uhm* was indeed only a prolongation of *uh*.

Considering all of the points mentioned above, it seems that, as is the case for many other aspects of the study of UH/M, existing literature does not yet provide a clear answer to the question of whether or not *uh* and *uhm* are indeed syntactically and/or functionally different from one another. It is true, of course, that their preference for different syntactic and pausal environments seems to suggest that they indeed represent different cognitive processes (such

as, for instance local lexical planning versus more complex syntactic planning); however, none of the findings presented in this discussion seem to act as concrete evidence for the “minor/major delay”-hypothesis suggested by Clark and Fox Tree (2002: 86). Nevertheless, there seems to be a lack of plausible alternative approaches, the exception being perhaps the assumption of dialectic and/or idiosyncratic differences that has been briefly suggested, yet not elaborated, by de Leeuw (2007: 109) and Shriberg (1994: 155-156). Consequently, it can be concluded that the examination of syntactic and/or functional differences between *uh* and *uhm* has not yet been able to provide a satisfactory answer and thus deserves more academic attention. In an attempt to contribute to the bridging of this research gap, I have chosen to dedicate the second part of this thesis to a corpus-based analysis of *uh* and *uhm* that contrasts the two variants with regard to their syntactic positions and subsequently also their functions. The precise research questions and methodology for this study will be presented in the course of the following chapters.

## **7. Studying functional differences between *uh* and *uhm*: an outline of research design**

In an attempt to contribute to the closing of the research gap outlined above, the second section of this thesis seeks to examine if there are observable functional differences between *uh* and *uhm*. For this purpose, a corpus-based study of the two variants of UH/M in British English was conducted; however, before presenting the obtained results in chapter 8, this chapter will provide relevant information concerning the methodological approach to the present study. To this end, the following chapter will be divided into four sections. The first section (section 7.1) will present the precise research questions which this study seeks to answer. After this introduction to the main research aims, the second section (section 7.2) will outline the corpus used for the investigation of *uh* and *uhm* in British English. The third section of this chapter (section 7.3) seeks to specify the levels of analysis on which *uh* and *uhm* will be investigated in the course of the present study as well as to provide information on the methodological steps taken in their exploration. Finally, the fourth and last section (section 7.4.) constitutes a brief summary of the method adopted in the analysis of syntactic and functional differences between *uh* and *uhm*.

### **7.1. Research questions**

As stated above, this thesis aims to investigate whether or not *uh* and *uhm* fulfill different functions in conversation. This question will be addressed on two different levels. Firstly, in an attempt to investigate potential contextual determinants, the production rates of *uh* and

*uhm* will be compared across different genres. This is deemed relevant to the present investigation not only because it may shed some more light on the previously addressed question of how context, setting and register affect the production of UH/M (see section 4.1.2), but especially because a comparison of the frequencies of *uh* and *uhm* across different speaking environments could provide some insights into their functional differences. For instance, if it was found that *uhm* appeared more often in telephone conversations than in face-to-face conversations, it would seem plausible to argue that *uhm* is strongly related to turn-keeping<sup>12</sup>, while an augmented rate of *uhm* in broadcast interviews could link it to lexical deliberation and/or information retrieval.

The second level of investigation concerns the question of whether or not *uh* and *uhm* prefer different syntactic positions within spoken discourse. The assumption underlying this syntactic approach is that *uh*'s and *uhm*'s location in discourse may offer a valuable insight into their functions; for instance, if *uh* and *uhm* are indeed markers of global and local planning respectively (as has been argued by Shriberg 1994: 154), the former should be expected to occur more often at syntactic boundaries, while the latter should be found to prefer unit-medial positions, e.g. directly in front of “problematic” lexical items. Based on these research aims and theoretical considerations, the following three research questions were defined for the purpose of the present study:

- (1) How does genre affect the frequency of *uh* and *uhm*?
- (2) Are there observable differences between *uh* and *uhm* with regard to their preferred syntactic positions, and if yes, what are they?
- (3) How do *uh* and *uhm* differ from each other with regard to the functions they fulfill in spoken discourse?

Having formulated these three research questions, the next step was to define the precise methodology to be applied in their examination. The decisions taken with regard to research design will be detailed in the following sections, where section 7.2 will explain the criteria according to which the corpus to be analyzed in the present study was defined, section 7.3 will outline the levels of syntactic analysis applied to the study of *uh*'s and *uhm*'s locations in spoken discourse, and section 7.4 will provide an overview on the precise methodological steps taken in the analysis of this data.

---

<sup>12</sup> This hypothesis is based on Künzel's (1997: 58; cited in de Leeuw 2007: 86) aforementioned finding that UH/M was more frequent in telephone calls than in face-to-face conversations, which he ascribed to the lack of visual means to signal turn-keeping.



## 7.2. Data acquisition

Since the present study seems to investigate differences between *uh* and *uhm* as observable in *everyday language*, it was obvious that a corpus-based approach would be needed. Having decided that the present study should be restricted to the exploration of British English, it was chosen to use the British component of the *International Corpus of British English* (hereafter referred to as ICE-GB). ICE-GB is a corpus of British English that contains 500 texts of around 2000 words each, where roughly 60% are transcripts of spoken language and the remaining 40% represent written genres such as student essays, newspaper articles, or fiction (Nelson, Aarts, Wallis 2002: 7). Given that UH/M are predominantly features of spoken language (O'Connell and Kowal 2005: 568), only spoken genres were considered in the present study.

The decision to use this specific corpus was based on two main reasons. Firstly, ICE-GB – unlike other corpora of British English – was primarily designed for the purpose of syntactic analysis (Nelson, Aarts, Wallis 2002: 4), meaning that it is fully tagged and parsed; this significantly facilitates the syntactic localization of *uh* and *uhm* in spoken discourse. Secondly, ICE-GB includes a wide variety of spoken genres and thus offers sufficient resources for the intended analysis of the frequencies of *uh* and *uhm* across different speaking environments. In addition to these two aspects, another advantage of ICE-GB is that audio recordings of all spoken texts are available. This feature is relevant to the present study insofar as listening to the chosen texts made it possible to add missing words on the one hand<sup>13</sup> and delete irrelevant material – such as the frequent transcription of the backchannel *mhm* as *uhm* – on the other hand; thus, the availability of audio recordings significantly contributed to the reliability of the obtained results.

As a first step, it had to be decided which genres would be analyzed in the present study. As was stated above, the spoken component of the ICE-GB contains both monologic and dialogic genres; however, since it has been argued in the theoretical part of this thesis that UH/M are to a large extent interpersonal phenomena that are directed towards the listener, it was deemed important to consider dialogic rather than monologic genres. Generally speaking, the dialogic section of ICE-GB is divided into *private* and *public* genres. The main distinction between these two lies in the fact that speakers in private settings (e.g. face-to-face conversations or

---

<sup>13</sup> Generally speaking, the transcriptions provided in ICE-GB are very reliable; however, sometimes it was found that prepositions or other short words that were audibly uttered by speakers were not transcribed, probably because they were confounded with background noise or simply overheard.

telephone calls) are only directed towards the other speaker(s) involved in the conversation, while speech in public genres such as classroom lessons, courtroom examinations or broadcast interviews addresses a broader audience (Nelson, Wallis and Aarts 2002: 6).

For the purpose of the present study, it was decided to compare fifteen texts from three different genres (i.e. five texts per genre): face-to-face conversations, telephone calls, and broadcast interviews. This choice of genres was predominantly motivated by the desire to examine two aforementioned hypotheses proposed in existing literature: firstly, Tottie's assumption that UH/M are more characteristic of non-private environments than of private genres (Tottie 2014: 6; cf. also Stenström and Svartvik 1994: 247-248), and, secondly, Künzel's finding that UH/M are more frequent in telephone calls than in face-to-face conversations (Künzel 1997: 58, cited in de Leeuw 2007: 86). The genre of broadcast interviews was chosen because it is deemed to be an interesting mixture of a private face-to-face conversation, where the speaker addresses only the present interlocutor, and non-private genres, where speech is directed towards a large anonymous audience.

Apart from belonging to one of the three relevant genres, the fifteen corpus texts to be analyzed in the present study had to fulfill three main criteria. Firstly, it was decided that the chosen texts should be around 2,000 words (+/- 10%) long. Secondly, for the sake of comparability, only texts with two speakers were included. The standardization of speaker numbers is based on the assumption that an increased number of speakers positively influences the frequency of UH/M within a conversation; not only because there are more speakers potentially producing them, but also because it is presumed that an increased number of speakers causes an augmented rate of turn-taking and turn-keeping attempts that are partly represented by UH/M. Consequently, the comparison of texts with different numbers of participants would be likely to distort the results obtained in the analysis of the frequencies of UH/M across different genres. The decision to include only texts with *two* speakers was based on the fact that the genre of telephone calls is usually restricted to two participants. Lastly, with regard to the private genres, it was decided to use only texts where the two speakers were friends so as to ensure that differences in the production rates of UH/M could not be ascribed to differences in speaker relationships.

Unfortunately, the control of speaker relationships and genre severely limited the range of eligible texts, and it was thus necessary to make certain compromises. In that sense, not all of the fifteen texts comply with the desired length of between 1,800 and 2,200 words, and the

corpus thus includes three texts that are between 2,200 and 2,300 words long. In order to resolve this problem, statistical figures reflecting UH/M frequency will be normalized to represent the average production per 1,000 words. Moreover, ICE-GB includes only four telephone calls between two friends, and it was thus chosen to include one conversation where the speakers were sisters (S1A-093). This choice was based on the assumption that sisterhood – of all interhuman relationships available in the corpus – resembles friendship the most; however, it is important to note that this supposition is purely intuitive and not based on scientific research. A table summarizing the main features of the chosen texts can be found in Table 1 in section 8.1.

In the end, a sample of fifteen texts containing 194 instances of *uh* and 335 instances of *uhm* was defined. According to de Leeuw (2007: 110), this clear dominance of *uhm* is little surprising because English speakers seem to show a general preference of the vocalic-nasal variant; thus, the chosen sample can be perceived as being more or less representative of common trends in the English language. Having chosen the data, the next step was to identify the precise levels of analysis for the investigation of syntactic differences between *uh* and *uhm*; this will be the subject of the following chapter.

### 7.3. Levels of analysis

It was already briefly stated above that the present thesis will adopt a syntactic approach to the comparative analysis of *uh* and *uhm*. As outlined in section 4.1.3 of this paper, this approach towards the study of UH/M is not a novel one and has already been used by numerous researchers who have sought to explore the functions of *uh* and *uhm* (cf., e.g., Clark and Fox Tree 2002; Kjellmer 2003; Hawkins 1971; Holmes 1988; Maclay and Osgood 1959). However, many of the existing publications on this topic focus on different levels of syntax, thus rendering their comparison difficult. For instance, while Maclay and Osgood (1959) have predominantly investigated the functions of UH/M via their immediate collocates, Clark and Fox Tree (2002) have chosen to primarily consider their location in the so-called *intonation units*, i.e. “stretch[es] of speech under a single intonation contour” (Clark and Fox Tree 2002: 93); yet another approach was proposed by Holmes (1988) who focused on clauses. Since this methodological diversity in existing literature renders it difficult to draw any reliable conclusions on the syntactic preferences of and functions fulfilled by UH/M, several scholars have stressed the importance of a comprehensive, multi-level analysis of UH/M that considers various syntactic ranks (cf. Hawkins 1971: 279; Kjellmer 2003: 174; Rochester 1973: 54). For this reason, the present study will attempt at adopting such a multi-layered analysis.

Having decided that the present study should include multiple syntactic levels, the question that remained was *which* ranks should be considered. In his seminal article “Categories of the Theory of Grammar”, Halliday (1961: 253) states that linguistic analyses should always consider five “ranks” of language: the sentence, the clause, the group/phrase, the word, and the morpheme; moreover, he proposes to consider these five ranks from the largest to the smallest unit (Halliday 1961: 251). The present study will loosely follow this framework of linguistic analysis. However, not all of Halliday’s five ranks are relevant to the study of UH/M; this is especially true for the morpheme and the sentence. While the former has been claimed to be irrelevant to the present context because UH/M generally occur at word boundaries rather than within them (cf. Hawkins 1971: 279), the latter has often been argued to be inapplicable to the analysis of spoken discourse (Selting 2005: 17, see below). Thus, Halliday’s framework needs to be slightly adapted to match the purpose of the present study; this process will be outlined in the following paragraphs.

One of the main problems of using Halliday’s framework for analyses of spoken discourse is that its highest rank – the *sentence* – is difficult to apply to spoken language (Selting 2005: 17). This is mainly because spoken discourse is typically characterized by more flexible – and often elliptical – structures that may be endlessly connected or interpolated to form lengthy and highly complex elaborations on one thought that would probably be separated into several shorter sentences in written texts (Crystal 1979: 155-159). Moreover, even those definitions of the sentence that include characteristics typical of spoken language, such as the possibility of elliptical constructions or the use of pauses, stress and pitch in lieu of punctuation<sup>14</sup>, are often difficult to apply to the study of spoken discourse because a segmentation according to these definitions would require a detailed analysis of information on prosody, stress and silent pauses, which is often not sufficiently provided in corpora. As a consequence of the issues related to the notion of *sentences*, conversation analysis has gradually turned towards another unit of segmentation: the *turn-construction unit*. As will be shown in section 7.3.2 of this chapter, the turn-construction unit is significantly more relevant to the investigation of spoken language than the sentence, and will thus be used in lieu of the latter for the purpose of the present study. In addition to this adaptation to Halliday’s model, it was also decided to

---

<sup>14</sup> Cf. Gardiner (1932: 208, cited in Fries 1952: 14), who defines the sentence as “an utterance which makes just as long a communication as the speaker has intended to make before giving himself a rest” or *Merriam Webster online dictionary*, where the sentence is defined as a “word, clause, or phrase or a group of clauses or phrases forming a syntactic unit which expresses an assertion, a question, a command, [...] that in writing usually begins with a capital letter and concludes with appropriate end punctuation, and that in speaking is distinguished by characteristic patterns of stress, pitch, and pauses [my emphasis]” (Merriam Webster: online)

consider the *turn*, i.e. the stretch of speech produced by a speaker between two points of speaker transition. This decision was mainly based on the assumption that the analysis of the positions of *uh* and *uhm* in turns could provide some insights into their respective roles in the turn-taking system.

Summing up, the present study will consider the following five levels of spoken discourse: the turn (section 7.3.1), the turn-construction unit (section 7.3.2), the clause (section 7.3.3), the phrase (7.3.4), and, finally, the word (7.3.5). Moreover, after the consideration of these units, *uh* and *uhm* will also be investigated with regard to their role in self-interruptions and repetitions (see section 7.3.6). The following six sections will provide an overview on the levels of analysis defined for the purpose of the present study, including definitions, justifications for their choice, and methodological information.

### 7.3.1. Turns

Given that existing literature has frequently linked UH/M to the turn-management system, it seems highly relevant to investigate *uh*'s and *uhm*'s positions in the conversational turn. Loosely following Tottie's (2015: 388) classification of turn-initial, turn-medial and turn-final UH/Ms, the following six possible positions were defined for the purpose of this study:

- (1) *Turn-initial*, where UH/M occupy the very first position inside a turn;
- (2) *Turn-medial*, where UH/M occur at any point within the turn that is not turn-initial or turn-final;
- (3) *Turn-final*, where UH/M occupy the very last position inside a turn. This position was further divided into the following three categories:
  - (a) *voluntary turn-yielding*, where a speaker interrupts him-/herself
  - (b) *other-interruption*<sup>15</sup>, where a speaker is interrupted by one of his/her interlocutors
  - (c) *other-continuation*, where an interlocutor interrupts the current speaker in order to continue the latter's turn, e.g. to assist him/her in the process of lexical retrieval.
- (4) *Standalone*, where UH/M do not belong to a turn but were uttered on their own

The above framework differs from the one proposed by Tottie (2015: 388) in two main respects. On the one hand, Tottie's framework was expanded by the *standalone* category. Even though it has been pointed out by O'Connell and Kowal (2005: 572) that UH/M do not typically construct turns on their own (which was also confirmed in the present study, cf.

---

<sup>15</sup> I use the terms "other-interruption" and "other-continuation" loosely following Sidnell (2012: 316).

section 8.2), the inclusion of this category is hoped to provide some interesting insights into the functions of UH/M in spoken discourse, e.g. by showing whether or not standalone instances of UH/M can act as (successful) turn-taking devices. On the other hand, Tottie's category of *turn-final* UH/Ms was further divided into three subcategories: *turn-yielding*, *other-interruption*, and *other-continuation*. This is predominantly based on Tottie's (2015: 402) assertion that turn-final UH/Ms might not always be deliberate signals of turn-yielding, but are often misinterpreted signals of turn-holding; the analysis of the different contexts of turn-yielding is thus hoped to shed some light on the relationship between UH/M and turn transition.

As concerns the second modification of Tottie's framework – the expansion of the *turn-final* position – it should be pointed out that the differentiation between voluntary turn-yielding, other-interruption and other-continuation is not always an easy task, for “it is often difficult to distinguish the function of turnyielding from that of turnholding, particularly when we have no information on the relevant prosodic elements” (Kjellmer 2003: 185). Even though the problem of lacking prosodic information is also given in the present study, it was attempted to distinguish between voluntary turn-yielding, other-interruption and other-continuation by listening to the audio recordings of the concerned instances as well as by considering the marked cases of overlap and the presence of (long) silent pauses in the immediate vicinity of UH/M. The three types of turn-final instances are exemplified below so as to provide the reader with a better insight into this classification.

#### (7) *Turn-yielding*

S1: Oh yeah yeah he thinks of himself as a writer // <1> He ju- </1> just doesn't write //  
 S2: <1> except that he doesn't write </1>  
 S2: Why doesn't he write //  
 S1: I don't know // I don't know // I mean I think of myself as a non-smoker // **uhm** <,,>  
 S2: [Well why doesn't why doesn't he] / [...]  
 (ICE-GB 2006: S1A-015: 125-133)

#### (8) *Other-interruption* (note that “=” marks interruption)

S2: Is he still writing //  
 S1: Well still isn't really the word // [It it stopped around **uh**]=  
 S2: =I know // it stopped years ago //  
 (ICE-GB 2005: S1A-015: 114-117)

#### (9) *Other-continuation*

S1: [And she is she's] / she spent the entire evening convincing her that Us Uts was desperately passionately in love with her <,> // [and it was **uhm**] /  
 S2: fatal <laugh>

Having defined the methodological approach towards the analysis of UH/M on the turn-level, the following section aims at outlining the methodology applied on the next rank, the *turn-construction unit*. To this end, the following section will not only present the methodological steps taken, but also the precise working definition of this unit.

### **7.3.2. The turn-construction unit**

As was already stated above, the *turn-construction unit* (henceforth referred to as TCU) was first proposed in the context of conversation analysis and designates the larger units of spoken discourse that combine to form *turns* in conversation. The concept of the TCU was first introduced by Sacks, Schegloff and Jefferson (1974: 701-702), who have defined the TCU as a “sentential, clausal, phrasal [or] lexical construction” that speakers use to construct a turn (Sacks, Schegloff and Jefferson 1974: 702). Moreover, according to the three scholars’ initial definition, each TCU ends in a so-called *transition relevance place* (henceforth referred to as TRP), i.e. a point where turn transition becomes “relevant but not necessary” (Selting 2005: 18) and at which interlocutors negotiate the right to the next turn (Houtkoop and Mazeland 1985: 596); consequently, TCUs may be defined as “possible complete turns” (Schegloff 1996: 55). In summary, then, the TCU could be defined as a unit of variable size and syntactic composition at the end of which turn-transfer between speakers becomes relevant and which thus acts as a potentially complete turn.

At first glance, the above definition of the TCU seems very useful to the segmentation of spoken discourse; however, attempts at application have shown that several aspects have to be further clarified in order to render the TCU an appropriate unit of linguistic analysis. One of the main problems associated with the definition of the TCU as proposed by Sacks, Schegloff and Jefferson (1974) is the fact that it is almost impossible to apply to larger turns in which a speaker tells a story, a joke, or the like (Houtkoop and Mazeland 1985: 597). Following the initial definition as proposed above, one would be obliged to treat these larger communicative projects as one single TCU, for turn-transition becomes relevant only *after* the story or joke has been completed (Selting 2000: 485); however, this conception is inappropriate to the syntactic analysis of spoken discourse because the units would be far too large and presumably too complex to allow for a meaningful analysis.

In an attempt to resolve the problem associated with the initial definition of the TCU outlined above, it has become popular to treat TCUs and TRPs as separate concepts, where the former

represents the “smallest linguistically possible unit in a given context” (Selting 2000: 487) and the latter designates “the endings of possibly complete turns” (Selting 2000: 485). More precisely, this approach suggests that TCUs are linguistic units of different composition that *may* – but do not have to – end in TRPs, and may – but do not have to – combine to form larger units which generally end in TRPs; these larger units are often referred to as *discourse units* (henceforth referred to as DUs, cf. Houtkoop and Mazeland 1985: 597). In that sense, not every TCU has to end in a TRP, but every discourse unit does. Generally speaking, this definition of the TCU has become the more accepted one (Selting 2000: 487), mainly because it allows for a more meaningful (syntactic) analysis of larger conversational activities. For this reason, this approach will also be adopted in the present paper.

Having provided a working definition of TCUs and TRPs, the next step was to identify the TCUs, TRPs and non-TRP *completion points* (henceforth referred to as CPs, cf. Selting 2000: 25) in the corpus used for the present study. Ideally, this step should be based on a consideration of prosodic information, because prosody constitutes an important factor in recognizing a speaker’s intention to continue his/her speech and thus helps to identify TCUs and TRPs (Selting 2005: 19). For this reason, the process of data segmentation into TCUs and TRPs was accompanied by a consideration of the audio recordings of all the source texts. It is important to note, however, that due to a lack of detailed prosodic information, the audio material was interpreted rather intuitively and might thus not be entirely trustworthy; still, together with the consideration of syntactic and contextual aspects, the segmentation of data is deemed to be sufficiently reliable for the purpose of the present study.

In order to provide an impression of the final result, examples (10) and (11) illustrate two examples which were segmented according to the above criteria<sup>16</sup>. Examples (10a) and (11a) show the respective passage’s original representation in *ICECUP*, while examples (10b) and (11b) illustrate the new segmentation into TCUs, TRPs, and CPs. Note that in these and all following transcripts, a capitalized *S* is used as shorthand for the word *speaker*, / marks a non-TRP CP, // marks a TRP, <, > and <., > represent short and long pauses respectively, and square brackets [ ] mark an aborted TCU, i.e. a TCU that is interrupted without being continued by the speaker.

---

<sup>16</sup> For the entire (segmented) corpus used in this study, cf. Appendices 1 and 2.



(10a) Representation of passage S1A-015: 039-044 in *ICECUP 3.1*.

S1A-015	039	A	You could commission prints of yourself
S1A-015	040	B	Yeah but I can't afford that kind of thing
S1A-015	041	A	How much <> could you afford
S1A-015	042	B	Uh <> I don't know
S1A-015	043	B	No nothing really
S1A-015	044	B	<del>I'm</del> → I've got too many debts <>

(10b) Segmentation of passage S1A-015: 039-044 into TCUs, TRPs and CPs

S1: You could commission prints of yourself //  
S2: Yeah // but I can't afford that kind of thing //  
S1: How much <, > could you afford //  
S2: Uh <, > I don't know // No nothing really // [I'm] / I've got too many debts //

This example illustrates the segmentation of spoken discourse into TCUs, TRPs and CPs. As can be seen from this example, turns may include multiple TRPs that do not involve speaker transition despite rendering it a relevant option. For instance, when S1 says, *You could commission prints of yourself*, S2 answers with *yeah*, which in itself could potentially constitute a sufficient reaction. However, after having reached the TRP, S2 decides to provide a reason for not doing as he/she was proposed by adding *but I can't afford that kind of thing*. After having reached the second TRP, S1 apparently considers S2's turn to be complete and thus reacts by asking another question, namely *How much <, > could you afford*, which S2 answers with three TCUs: *Uh <, > I don't know*, *No nothing really*, and *I've got too many debts*. After each of these three TCUs, S2's turn could potentially be regarded as complete, hence each of them is marked with a TRP. Note that after the second TCU in her last turn, S2 starts to construct the last TCU by uttering *I'm*, but then changes her mind and reformulates her message. This is marked as an *aborted TCU* because the speaker has abandoned the initial main verb *am* and substituted it by *have got*. The point of abortion is marked with a CP instead of a TRP; this is based on Clayman's (2013: 152) observation that "[w]hen a TCU is aborted before it is brought to completion, this is not usually treated as a locus for turn transition. Aborted TCUs tend to be initially followed, not by talk from a different speaker, but by further talk from the same speaker [...] or by silence".

A comparison between *ICECUP*'s segmentation illustrated in example 10a and the TCU segmentation shown in example 10b reveals that apart from the marking of TRPs and CPs, there are hardly any differences between the two approaches; consequently, the process of segmentation was rather unambiguous. However, as is shown in the below example, this is not the case for all passages:

(11a) Representation of passage S1A-093: 129-136 in *ICECUP 3.1*.

S1A-093	129	B	Oh what
S1A-093	130	A	I'm going to uhm <>
S1A-093	131	A	Fiona as you know does set design
S1A-093	132	A	she's just finishing this course and she's done the set design for a play up in Islington in a pub called the Red Lion
S1A-093	133	B	Oh +
S1A-093	134	B	I think I've been there actually
S1A-093	135	A	You might've done
S1A-093	136	A	it's quite a <> uhm well-known one

(11b) Segmentation of passage S1A-093: 039-044 into TCUs, TRPs and CPs

S2: Oh what //

S1: [I'm going to uhm <,>] / Fiona as you know does set design / she's just finishing this course / and she's done set design for a play up in Islington // in a in a pub called *The Red Lion* //

S2: Oh I think I've been there actually //

S1: You might've done // it's quite a <> uhm well-known one //

In this example, the TCU segmentation significantly differs from the segmentation proposed by *ICECUP*, especially with regard to S1's first turn. In this turn, S1 wants to tell his/her interlocutor where he/she is going in the evening (*I'm going to uhm <,>*) but then decides that his/her listener would probably need more information concerning her reasons for going there. Therefore, he/she aborts the present TCU and provides some relevant details, namely that he/she is going to see a play to which a friend of his/hers – *Fiona* – has contributed as set designer. As explained in the discussion of example (10), the abortion of the initial TCU makes speaker transition an irrelevant option, and S1 therefore keeps his/her right to the floor. After a short pause, he/she then resumes her turn by uttering *Fiona as you know does set design*. Given that this remark is not obviously linked to the initial TCU, it can be expected that S1 will continue talking; in other words, he/she has projected a multi-TCU project that can be expected to be complete as soon as a relationship between the unspecified event and her friend Fiona has been established. Based on this assumption, the TCUs *Fiona as you know does set design* and *she's just finishing this course* have been marked to end in CPs rather than TRPs because neither of them establishes this relationship. On the contrary, the fourth TCU in this turn – *and she's done set design for a play up in Islington* – contains the relevant information that allows S2 to deduce where S1 is going (to see the play for which Fiona has done the set design); therefore, the projected multi-TCU unit has reached its point of completion and speaker transition becomes a relevant option again. For this reason, the end of this fourth TCU has been marked with a TRP. The turn-final TCU *in a in a pub called 'The Red Lion'* can be perceived as supplementary information that is not indispensable for the primary message of this turn and has thus been marked as being surrounded by TRPs.

Example (11) is a very good illustration of various theoretical aspects associated with the concepts of TCUs, TRPs and CPs. For instance, it does not only exemplify the difference

between CPs and TRPs, but also represents the relevance of treating the TCU as the “smallest linguistically possible unit in a given context” (Selting 2000: 487). If this paper followed the traditional conception of the TCU as a potentially complete turn that has to end in a TRP, then the whole passage between *I’m going to uhm* <,> and [...] *up in Islington* would need to be treated as one single TCU. However, this approach would render the syntactic localization of UH/M within this larger stretch of speech extremely challenging, and it is thus more convenient to treat the TCU as the “smallest linguistically possible units” (Selting 2000: 487). Moreover, the above example also highlights the importance of considering the *implicit*, *explicit* and – whenever possible – *prosodic projection* of turn-length<sup>17</sup> in the process of segmentation.

Having segmented the corpus texts into TCUs, the next step was to identify the possible positions of *uh* and *uhm* in TCUs that would be considered in the analysis. Roughly based on an article by Hawkins (1971: 283-284), in which four broad different positions within a clause were identified, eleven different positions – which were grouped into five different categories – were defined; these are presented in Figure 4 below. Note that in the present study, the term *complete TCU* refers to any TCU that is neither interrupted by the current speaker (e.g. in self-correction) nor by his/her interlocutor, while the adjective *aborted* is used to refer to TCUs that have been abandoned by the current speaker or interrupted by his/her interlocutor. Finally, the term *Point of Abortion* refers to the point at which a current TCU is interrupted, i.e. *aborted*, by the speaker.

---

<sup>17</sup> I use the terms *implicit*, *explicit* and *prosodic projection* as follows:

- *implicit projection* refers to the implicit announcement of a larger project as observed in the above example, where the speaker is expected to add more TCUs because her statement stands in no relation to the primary message of the turn
- *explicit projection* refers to the explicit announcement of a larger project as achieved, for instance, by uttering, “Let me tell you a story”
- *prosodic projection* refers to the use of intonation to signal continuation, e.g. by using a rising intonation at the end of a TCU

Figure 4: Possible positions of *uh* and *uhm* in TCUs

Category 1: UH/M are situated at a TRP that does not involve speaker transition
<ul style="list-style-type: none"> <li>a) UH/M are situated at a TRP between two complete TCUs</li> <li>b) UH/M are situated between a complete and an aborted TCU</li> </ul>
Category 2: UH/M are situated at non-TRP CPs
<ul style="list-style-type: none"> <li>a) UH/M are situated at a CP between two complete TCUs</li> <li>b) UH/M are situated at a CP between a complete and an aborted TCU</li> </ul>
Category 3: UH/M are situated at a Point of Abortion (PoA)
<ul style="list-style-type: none"> <li>a) UH/M are situated between an aborted and a complete TCU (no speaker transition)</li> <li>b) UH/M are situated between two aborted TCUs (no speaker transition)</li> <li>c) UH/M are situated at a PoA that involves speaker transition</li> </ul>
Category 4: UH/M are situated at a CP/TRP involving turn transition
<ul style="list-style-type: none"> <li>a) UH/M are situated at the first position in a turn-initial TCU</li> <li>b) UH/M are situated after a complete turn-final TCU</li> </ul>
Category 5: UH/M are not situated at a TRP/CP
<ul style="list-style-type: none"> <li>a) UH/M occupy the second (or third) position after (a cluster of) connectives, conjunctions, interjections or discourse markers</li> <li>b) UH/M occur inside the TCU, i.e. after the speaker has lexically committed him-/ herself; this includes all positions that are not covered by the above list.</li> </ul>

The segmentation into these eleven potential positions is hoped to provide some interesting insights into potential functional differences between *uh* and *uhm*. For instance, if the two variants do indeed signal different levels of linguistic planning (cf. Shriberg 1994: 154), one should be expected to occur more frequently at TRPs and CPs, while the other should tend to occur inside the TCU. Moreover, an inclination of *uh* and/or *uhm* to occur at TRPs might support the aforementioned assumption that turn-management is a main function of UH/M.

Having outlined the key concepts related to the segmentation of spoken discourse into TCUs and TRPs, the next level of analysis to be considered in the course of this section is the clause level.

### 7.3.3. The clause level

Following Halliday's framework of the five ranks of linguistic analysis, the next level to be considered in the present study is the clause. The importance of the clause as a unit of analysis has been highlighted by Kjellmer (2003: 180), who asserts that "[i]f we think of F[illed] P[ause]s as signaling something new, a new thought unit, it would be natural to expect to find them at the beginning of clauses". Given the many types of clauses (cf. Holmes 1988: 330) as well as the relative syntactic flexibility within clauses (e.g. the possibility of adding pre-

nuclear adjuncts before the subject, cf. Kaltenböck 2009: 57-58), the localization of *uh* and *uhm* on the clause level can be a very complex enterprise; thus, a precise framework of possible positions inside a clause was defined, which allowed a more structured categorization of all instances of *uh* and *uhm*.

Figure 5 below shows the list of possible positions which were defined for the study of *uh*'s and *uhm*'s *positions* in clauses. This framework was to a large extent based on Kaltenböck's (2009: 57-58) list of possible syntactic positions of English comment clauses, with some modifications being made in order to match the purpose of the present study. For instance, while Kaltenböck (2009: 57-58) distinguishes between main clauses and subordinate clauses, such a distinction was deemed irrelevant for the present study because it is assumed that *uh* and *uhm* do not differ with regard to their preference of different clause types, but rather in different positions within clauses in general (e.g. clause boundary vs. inside the clause). As can be seen, the possible positions were classified into four main categories: *prenuclear positions*, *middle positions*, *postnuclear positions*, and, lastly, a *miscellaneous* category which comprises all instances of UH/M that cannot be classified according to the former three. Following Kaltenböck's (2009: 57-58) model, # refers to the point of insertion, i.e. the point at which UH/M occurs, and *MVB* refers to the main verb of the clause. The following list might seem very detailed and complex at first glance; however, this was necessary in order to provide a profound insight into the potential syntactic differences between *uh* and *uhm*<sup>18</sup>.

---

<sup>18</sup> Note that the very specific positions concerning "existential there" constructions as in the sentence *There are many flowers in my parents' garden*, interrogative clauses and/or zero relative/subordinate clauses were added in the process of the actual analysis because they were required for the classification of specific instances.

Figure 5: Possible syntactic positions of *uh* and *uhm* in clauses  
(based on Kaltenböck 2009: 57-58)

<b>Category A: PRENUCLEAR POSITIONS</b>	
(1) initial/ at clause boundary	
a) UH/M occupies a clause initial-position in a clause that is not introduced by a conjunction, a connective, an interjection, or a discourse marker (henceforth referred to by the abbreviation CCID)	
b) UH/M is situated before a clause-introductory CCID	
(2) CCID # subject	
a) CCID # subject	
b) CCID # adjunct + subject	
c) CCID # <i>there</i> (in the case of “existential-there” - clauses)	
d) CCID # CCID	
(3) Adjunct # subject	
a) Clausal adjunct	
b) Non-clausal adjunct	
(4) CCID # adjunct + subject	
a) Clausal adjunct	
b) Non-clausal adjunct	
(5) CCID # verb in clauses that lack a subject, interrogations, and imperatives	
<b>Category B: MIDDLE POSITIONS</b>	
(1) Subject # verb	
a) Subject # MVB	
b) Subject # copular verb	
c) Subject # auxiliary + MVB	
(2) auxiliary verb # MVB	
a) auxiliary verb # MVB	
b) modal verb # MVB	
c) auxiliary # auxiliary + MVB	
(3) MVB # non-clausal complementation	
a) MVB # object	
b) MVB # subject complement	
c) MVB # other complement	
d) MVB + Object # object complement	
(4) MVB # clausal complementation	
a) MVB # object	
b) MVB # subject complement	
c) MVB # other complement	
(5) <i>there</i> + MVB # subject	
(6) in interrogative clauses: auxiliary # MVB	
a) auxiliary + subject # MVB	
b) auxiliary # subject + MVB	
<b>Category C: POSTNUCLEAR POSITIONS</b>	
(1) MVB / Clause # adjunct (non-clausal)	
(2) MVB / Clause # adjunct (clausal)	
(3) MVB / Clause + adjunct (non-clausal) # adjunct (non-clausal)	
(4) MVB / Clause + adjunct (clausal) # adjunct (clausal)	
(5) MVB + object # object complement (non-clausal)	
<b>Category D: MISCELLANEOUS</b>	

### 7.3.4. The phrase level

The next linguistic level to be investigated is the phrase level. In his detailed study of the syntactic position of UH/M, Kjellmer (2003: 175-179) investigated only selected types of phrases, namely noun phrases (henceforth referred to as NP), adjective phrases (AJP), verb phrases (VP), prepositional phrases (PP) and *not*-phrases. Following this model, the present study will also focus on these phrase types. The only exception is the *not*-phrase, which was excluded because it was counted towards the verb phrases. In general, the analysis of UH/M on the phrase level was predominantly concerned with the question of whether *uh* and *uhm* tend to occur at phrase boundaries or rather inside them, and, in the case of the latter, to identify tendencies concerning the phrase-internal positions of the two variants. The identification of phrase positions is not always unambiguous, and some choices had to be made concerning the classification of specific occurrences; for the sake of transparency, some examples of such problematic cases are discussed below, with screenshots from the *International Corpus of English Corpus Utility Program (ICECUP) version 3.1* (2006) being provided to illustrate them.

In general, the “phrase boundary” position designates occurrences of UH/M that separate two phrases from each other. This is the case, for instance, in Figure 6, where *uh* precedes the noun phrase *she* in the TCU *and she hadn't seen him for nine months [...]* (ICE-GB 2006: S1A-093:84) as well as in Figure 7, where the second *uhm* is situated at the boundary of the verb phrase *buried* in the TCU *uhm oh God uhm buried under a couple of uhm playwrighting commissions* (ICE-GB 2006: S1A-096:19).

Figure 6: Example of *uh* at a noun phrase boundary (ICE-GB 2006: S1A-093:84; screenshot: *ICE-CUP 3.1*.)

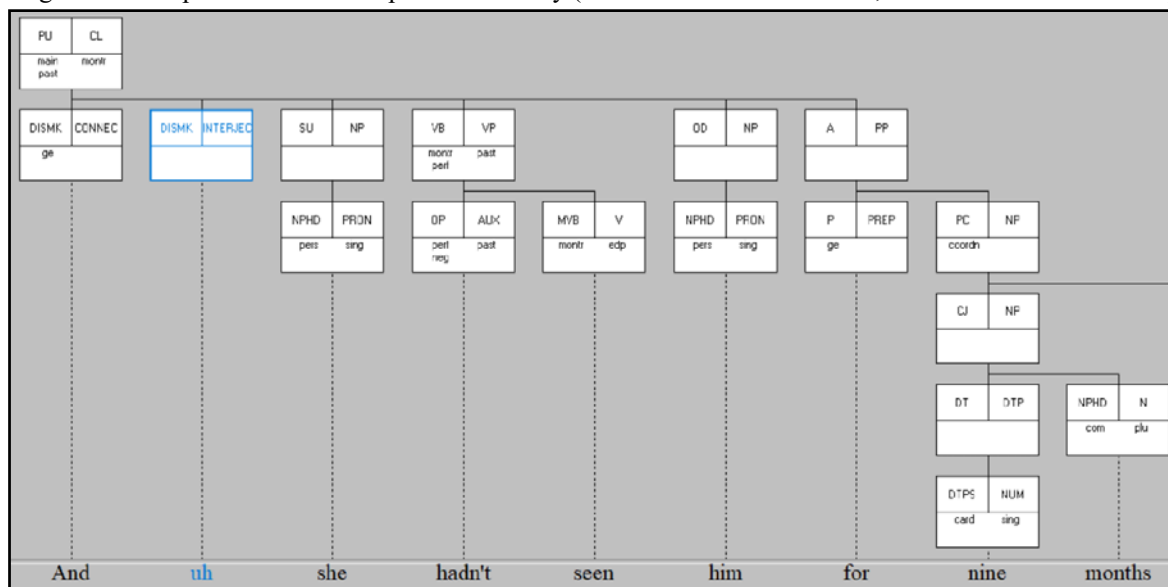
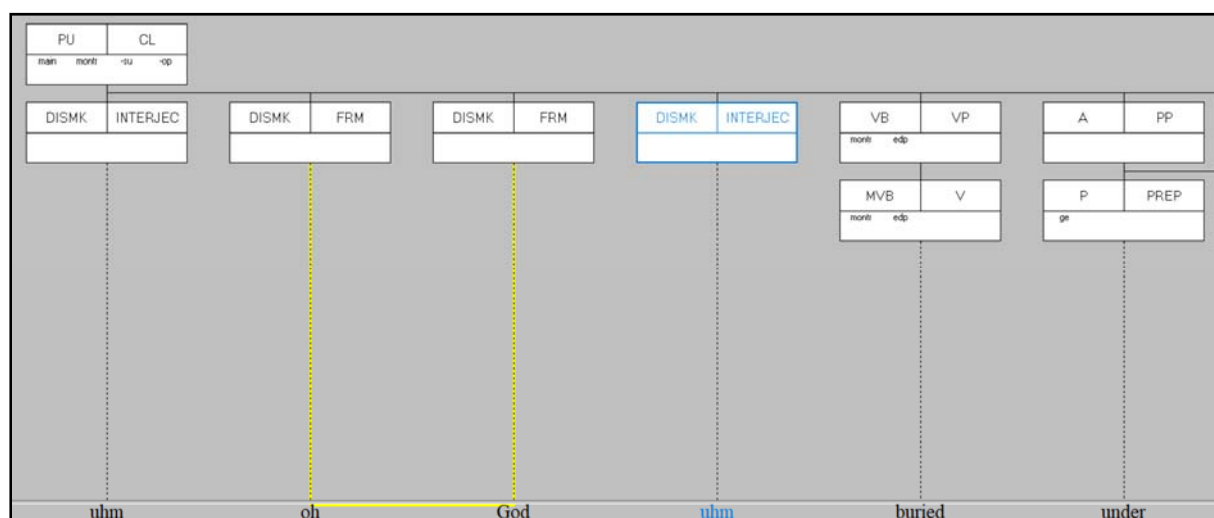
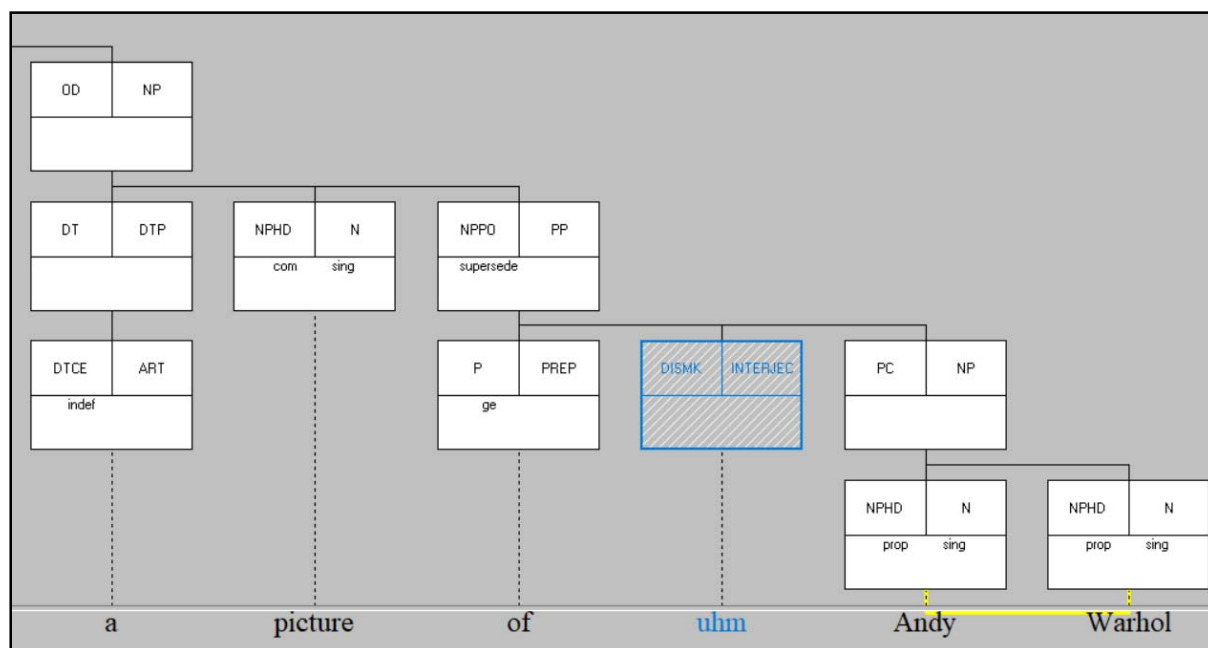


Figure 7: Example of *uhm* at a verb phrase boundary (ICE-GB 2006: S1A-096:19; screenshot: *ICE-CUP 3.1.*)



Both of these instances of UH/M were unproblematic to classify because they separate two phrases on the highest syntactic level of the TCU, the *parent level* (Nelson, Wallis and Aarts 2002: 74); however, obstacles were encountered when the respective instance of UH/M was situated in a so-called *child* of this first level. Figure 8 represents such a case. In this TCU (S1A-015:35), *uhm* is situated inside the PP *of Andy Warhol* but before the prepositional complement *Andy Warhol* – an NP. Consequently, the question arises of whether *uhm* should be classified as being situated inside a PP or at the boundary of an NP.

Figure 8: Example of *uhm* inside a preposition phrase (ICE-GB 2006: S1A-015:35; screenshot: *ICE-CUP 3.1.*)







these approaches have already been applied in existing literature and seem highly promising; therefore, both will be applied in the context of the present study.

As a first step, all instances of UH/M were classified according to their verbal and pausal environment; this is hoped to shed some light on Clark and Fox Tree's (2002: 82-84) hypothesis that *uh* marks a "minor delay" and *uhm* marks a "major delay". The relationship between UH/M and silent pauses has already been investigated in various publications; however, most of the studies of UH/M's pausal environment rely on (at least rough) time measurements of the surrounding silent pauses (cf. Clark and Fox Tree 2002: 83; O'Connell and Kowal 2005: 561-562) which are not provided in ICE-GB. In fact, the only distinction made in ICE-GB is between *short pauses*, i.e. "any perceptible break in phonation equal in length to one syllable" and *long pauses*, i.e. "any longer break in phonation" (Nelson, Aarts and Wallis 2002: 12). For this reason, it was chosen to follow de Leeuw's (2007: 94-95) approach towards studying the pausal environment of UH/M, which, too, does not consider exact time measurements. To this end, five different categories were defined which allowed the classification of UH/M with regard to the delays associated with their production. These categories were defined with the underlying assumption that a "major delay" would involve at least one long pause, while a "minor delay" would include only one short pause. The following list provides an overview of the defined five categories, where #, again, denotes UH/M, *w* stands for *word*, *s* stands for *short silent pause*, *ss* marks a *long silent pause*, and a hyphen signals the beginning/end of the respective turn (cf. de Leeuw 2007: 95 for similar practice):

- (1) No delay: *w#w*, *-#w*, *w# -*
- (2) Short delay: *- #s*, *s# -*, *w#s*, *s#w*
- (3) Medium delay: *s#s*
- (4) Long delay: *ss#ss*, *-#ss*, *ss# -*, *s#ss*, *ss#s*, *w#ss*, *ss#w*
- (5) Other: *- # -*

The second aspect to be studied on the word level concerns the most frequent collocates of *uh* and *uhm* in spoken discourse. The methodological approach to this level of analysis was largely based on Kjellmer (2003: 173-174), who compared pre- and post-UH/M collocates and used his findings to draw conclusions on the functions fulfilled by *uh* and *uhm*. The analysis of collocates constitutes the last step in the syntactic analysis of *uh* and *uhm*; the only aspect that remains to be studied, then, are the respective roles of *uh* and *uhm* in self-

correction, self-interruption and repetitions. The following chapter seeks to provide the reader with an insight into the investigation of this level of analysis.

### 7.3.6. *Uh* and *uhm* in self-interruption and repetitions

Apart from the levels of analysis outlined above, it was also deemed necessary to study *uh* and *uhm* in the context of self-interruption, self-correction and repetitions, because all of these phenomena have been frequently linked to the production of UH/M (see chapter 4.3 of this paper); thus, the consideration of these phenomena in the comparative study of *uh* and *uhm* is hoped to provide some interesting insights into potential functional differences between the two variants.

The investigation of the role of UH/M in self-interruptions and repetitions comprised two main steps. Firstly, all instances of *uh* and *uhm* occurring before, inside, and after repetitions were identified and analyzed with regard to the syntactic functions and/or word classes of the repeated passage. Secondly, for the purpose of studying those instances of UH/M that coincide with points of abortion (PoAs), four possible contexts of self-interruptions were defined; these are outlined below.

- (1) *self-interruption*, where a speaker interrupts his/her current TCU, clause or phrase in order to a) yield his/her turn or b) start a new TCU, clause or phrase that is not connected to the previous one in any obvious way; this is the case, for instance, in the passage *well <,> uh as far **uh** I mean <,> she won't see a psychiatrist [...]* (ICE-GB 2006: S1A-031:124)
- (2) *reformulation*, where a speaker aborts the current unit under construction in order to rephrase it (e.g.: *Well my impression [...] was that there wasn't **uhm** he wasn't filtered*; cf. ICE-GB 2006: S1A-015:091)
- (3) *correction*, where a speaker's interruption of a current unit for the sake of correcting a (linguistic or content-related) mistake (e.g. *[...] that she was no good in **uh** at that kind of relationship*; cf. ICE-GB 2006: S1A-031:123); and, finally,
- (4) *word-internal interruption*, which refers to a speaker's mid-word self-interruption, as can be observed in the example *in the end of Ju- **uh** July* (ICE-GB 2006: S1A-096:177)

Note that the above list is no more than a loose framework for classifying those instances of UH/M that appeared at PoAs, and that some categories are sometimes difficult to distinguish

objectively. For instance, the example provided for the *correction* category, which reads [...] *she was no good in uh at that kind of relationship* could equally be argued to be an example of a reformulation; similarly, mid-word interruption could, but does not necessarily have to, be linked to the correction of a mistake. In that sense, the below list should not be understood as an infallible or entirely reliable; however, it is still thought to provide a reasonable framework for studying UH/M in self-interruptions.

#### 7.4. Method

Having defined all levels of syntactic analysis, the present section provides a brief, yet detailed insight into the precise methodological steps taken in the analysis of UH/M. As previously stated, the first step in the analysis of the selected corpus texts was to listen to the audio recordings of the respective conversations so as to correct possible errors in the transcription. Afterwards, the data was segmented into TCUs; this was accomplished by considering the abovementioned criteria on the one hand (see section 7.3.2) and the audio recordings on the other hand. The latter were very helpful in disambiguating certain problematic passages, where segmentation was not obvious from syntactic or contextual criteria. The segmented data can be found in appendices 1 and 2.

After the segmentation of the transcripts, some instances of UH/M were excluded from specific levels of analysis. For instance, in one corpus text, there is a passage in which the two speakers play a game where they have to take turns in saying a word that starts with the letter *d*. As will be discussed in chapter 9.3, this passage is highly interesting to the present research because it establishes a link between the uttering of *uhm* and information retrieval; however, given that the words succeeding *uhm* were only chosen with regard to their initial letter and do thus not represent typical collocates of *uhm*, these instances were only considered on the turn and TCU levels, but not with regard to any other syntactic levels. Similarly, all instances of UH/M linked to self-interruption, self-correction or repetitions were extracted and analyzed separately.

Following these steps, each occurrence of *uh* and *uhm* received a number and was manually analyzed on each of the following levels of syntactic analysis: (1) turn position, (2) TCU position, (3) clause position, (4) phrase position and respective phrase type, (5) verbal and pausal environment, (6) preceding and succeeding collocates. The data acquired in these steps was then statistically processed, visualized, and finally interpreted with regard to their

implications for potential functional differences between *uh* and *uhm*. The results and insights obtained the analysis of *uh* and *uhm* will be presented in the following chapter.

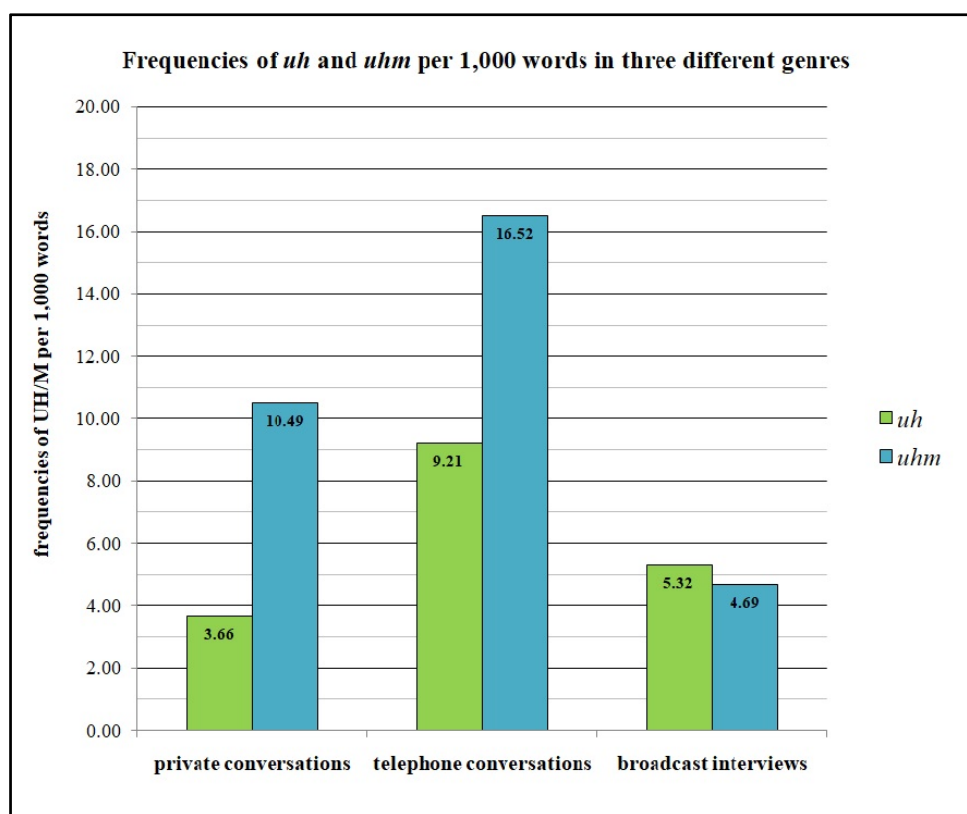
## **8. Results**

Having outlined the research design, the following chapter presents the results obtained for the different levels of analysis. For this purpose, the present chapter is divided into six sections. The first section – section 8.1 – will present the results obtained in the analysis of the effects of genre on the production of *uh* and *uhm*, while section 8.2 aims at identifying the preferred positions of *uh* and *uhm* in turns and TCUs. Following the investigation of these larger units of spoken discourse, sections 8.3 and 8.4 will be dedicated to the analysis of *uh*'s and *uhm*'s positions in clauses and phrases respectively. After this syntactic analysis, section 8.5 will present the results obtained in the investigation of common collocates of *uh* and *uhm*. Finally, section 8.6 provides a statistic representation of occurrences of *uh* and *uhm* in repetitions, self-interruptions and self-corrections. Note that the present chapter is merely a *descriptive* summary of the obtained results, whereas the interpretation of the data can be found in chapter 9.

### **8.1. *Uh, uhm* and genre**

The first section of the present chapter presents the results obtained in the analysis of the effects of genre on *uh* and *uhm*. As was detailed in chapter 7.2, this subject will be approached via a comparison of the frequencies of *uh* and *uhm* in three different speaking environments, namely private face-to-face conversations, telephone conversations, and broadcast interviews. For the purpose of this analysis, the instances of *uh* and *uhm* were counted for each text and were then normalized to obtain the average production rates per 1,000 words; furthermore, the average frequency of *uh* and *uhm* was also calculated for each genre as a whole, where each genre comprises five corpus texts. An overview on the results is presented in Figure 10 below.

Figure 10: Frequencies of *uh* and *uhm* per 1,000 words in three different genres

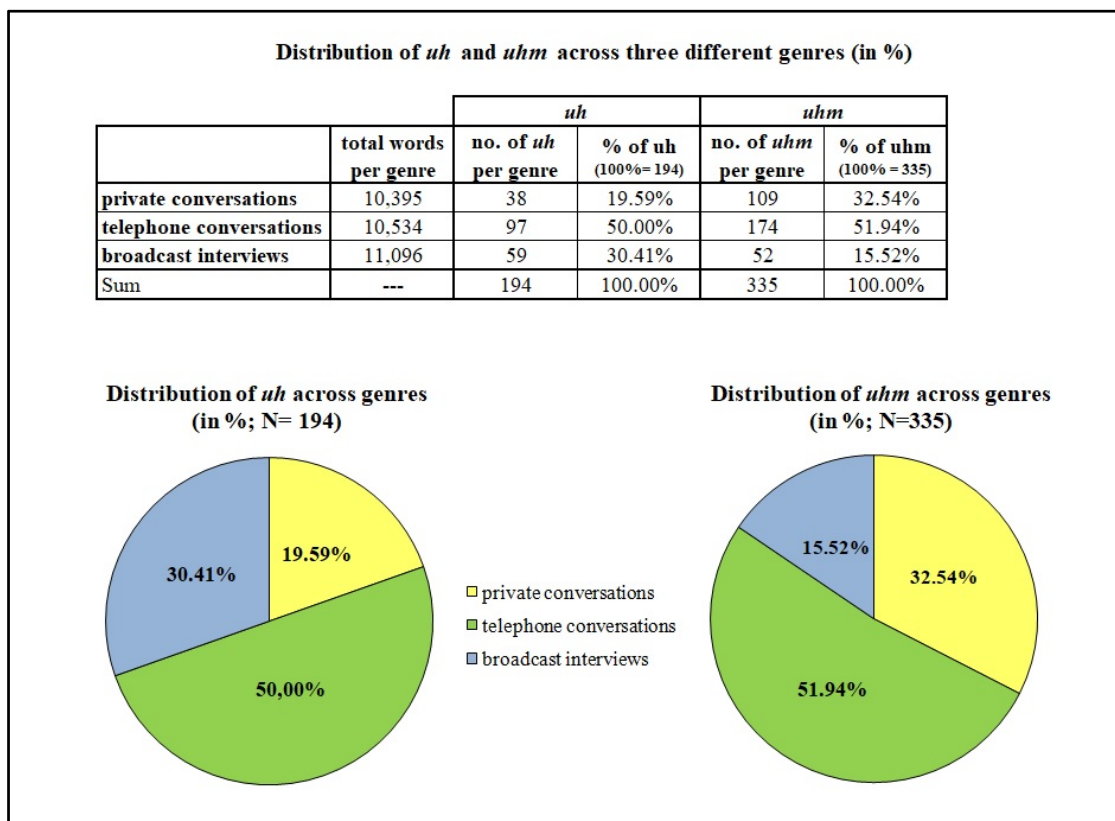


The above bar chart shows the frequencies of *uh* and *uhm* per 1,000 words in private face-to-face conversations, telephone conversations, and broadcast interviews. As can be seen, both *uh* and *uhm* were most frequent in telephone conversations, where they occurred at an average rate of roughly 9 and 17 instances per 1,000 words respectively. Moreover, the graph also shows that *uhm* significantly dominates *uh* in private settings (i.e. face-to-face and telephone conversations), with the former exceeding the latter by an average of about seven occurrences per 1,000 words. However, this dominance could not be observed in broadcast interviews, where the average production rates of *uh* and *uhm* seem to be more balanced, with *uh* even being slightly more frequent than *uhm*. Finally, it can be seen in the above graph that while *uhm* was found to appear considerably more often in private settings than in the public speaking setting, *uh* occurred more often in broadcast interviews than in private face-to-face conversations.

The different tendencies of *uh* and *uhm* with regard to genre outlined above are also visible in the pie chart in Figure 11 below, which represents the distribution all 194 instances of *uh* and 335 occurrences of *uhm* across the three different genres. Note again that *genre*, in the present study, does not refer to the complete collection of corpus texts included in ICE-GB, but

designates the small-scale corpus defined for the present study, where each genre comprises five corpus texts.

Figure 11: Distribution of *uh* and *uhm* across genres in percentages ( $N_{uh}=194$ ;  $N_{uhm}=335$ )



As can be seen in this chart, around half of all instances of both *uh* and *uhm* appeared in telephone conversations; this finding further highlights the tendency of both variants to occur in this genre. Moreover, it is shown that the second largest proportion of all observed instances of *uh* – roughly 30% – occurred in broadcast interviews and the remaining 20% were uttered in private conversations. The reverse is true for the vocalic nasal variant *uhm*, of which roughly every third instance in the corpus occurs in private conversations, with the remaining 16% occurring in broadcast interviews.

In summary, the results presented in Figure 10 and 11 do not only show that both *uh* and *uhm* were uttered more frequently in telephone conversations than in any other genre considered in the present study, but also that they are both more typical of private settings, with around 70% of all *uhs* and 85% of all *uhms* occurring during private face-to-face and telephone conversations; however, it was also shown that *uhm*'s tendency to appear in private contexts seems to be more pronounced than *uh*'s.

In addition to the findings presented above, an analysis of the individual texts revealed that there are very high fluctuations of UH/M frequencies within genres; this is illustrated in Table 1 below.

Table 1: Production rates of *uh* and *uhm* in individual corpus texts and genres

**Genre 1: Private conversations between friends**

Text code	Word count	<i>uh</i>	<i>uh</i> /1,000 words	<i>uhm</i>	<i>uhm</i> /1,000 words
S1A-015	2,013	21	10.43	41	20.37
S1A-031	2,045	10	4.89	31	15.16
S1A-067	2,049	5	2.44	21	10.25
S1A-080	2,198	2	0.91	4	1.82
S1A-085	2,090	0	0	12	5.74
<b>Total</b>	<b>10,395</b>	<b>38</b>	<b>3.66</b>	<b>109</b>	<b>10.49</b>

**Genre 2: Telephone conversations between friends**

Text code	Word count	<i>uh</i>	<i>uh</i> /1,000 words	<i>uhm</i>	<i>uhm</i> /1,000 words
S1A-091	2,129	4	1.88	28	13.15
S1A-092	1,959	26	13.27	15	7.66
S1A-093	2,020	9	4.46	25	12.38
S1A-096	2,281	41	17.97	65	28.50
S1A-097	2,145	17	7.93	41	19.11
<b>Total</b>	<b>10,534</b>	<b>97</b>	<b>9.21</b>	<b>174</b>	<b>16.52</b>

**Genre 3: Broadcast interviews**

Text code	Word count	<i>uh</i>	<i>uh</i> /1,000 words	<i>uhm</i>	<i>uhm</i> /1,000 words
S1B-045	2,258	28	12.40	16	7.09
S1B-046	2,101	5	2.38	12	5.71
S1B-047	2,270	15	6.61	5	2.20
S1B-048	2,267	5	2.21	11	4.85
S1B-049	2,200	6	2.73	8	3.64
<b>Total</b>	<b>11,096</b>	<b>59</b>	<b>5.32</b>	<b>52</b>	<b>4.69</b>

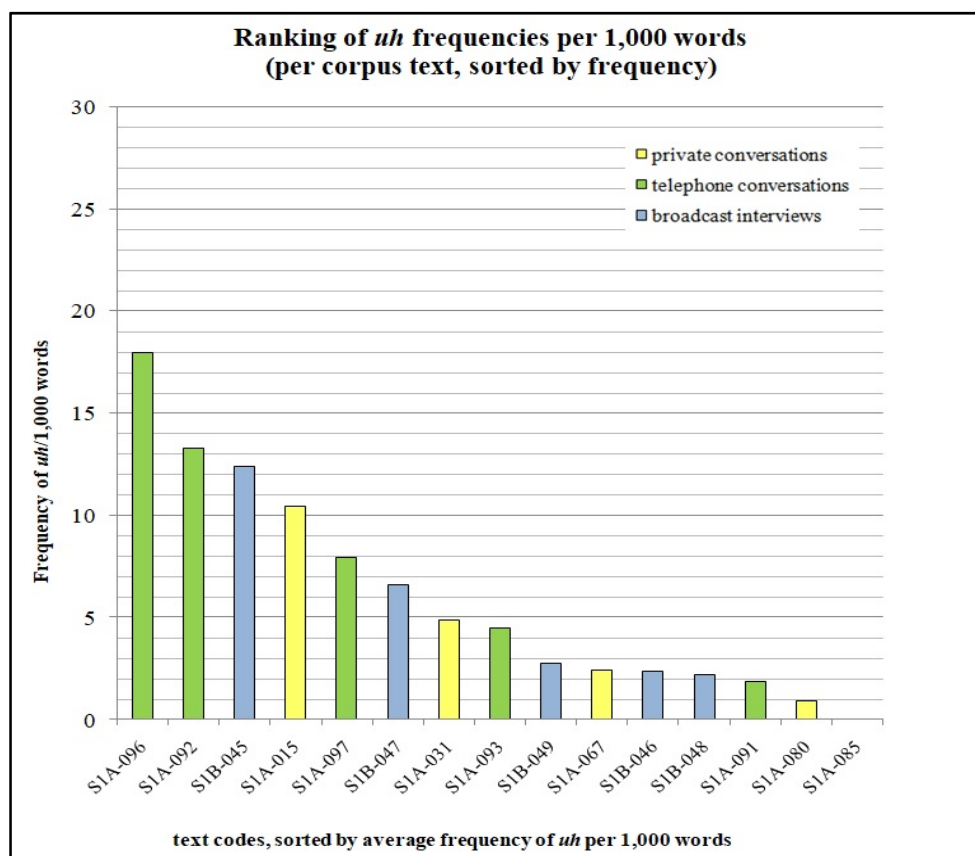
It can be seen in Table 1 above that the individual texts analyzed for each genre show substantial discrepancies with regard to their average rates of *uh* and *uhm*. This can be observed, for instance, in the case of private face-to-face conversations, where *uhm* is uttered roughly 20 times per 1,000 words in one text (S1A-015) but only less than two times in another one (S1A-080). Similarly, the average frequency of *uh* in telephone conversations ranges from 1.88 occurrences/1,000 words in S1A-091 to almost 18 instances in S1A-096. The existence of such wide discrepancies between minimum and maximum values within a genre in such a small-scale corpus constitute a challenge insofar as they render the identification of trends a rather difficult enterprise. This can be illustrated, for instance, by considering the average frequencies of *uhm* in private conversations between friends. It can be seen in Table 1 above that in three out of the five texts, *uhm* appeared between 10 and 20



times per 1,000 words, while the other two texts contained relatively low frequencies of *uhm* (roughly 2 and 6 instances per 1,000 words respectively); however, given the small sample size, it remains unknown which production rate is more representative of the general trends in the English language. Naturally, the high fluctuations of UH/M frequencies could be argued to illustrate that there are in fact no (or hardly any) general trends with regard to genre. In that sense, the results outlined in the present paragraph could be interpreted to suggest that the production of *uh* and *uhm* is at least to some extent a speaker-specific, idiosyncratic phenomenon; this hypothesis will be examined in more detail in section 9.4 of this paper.

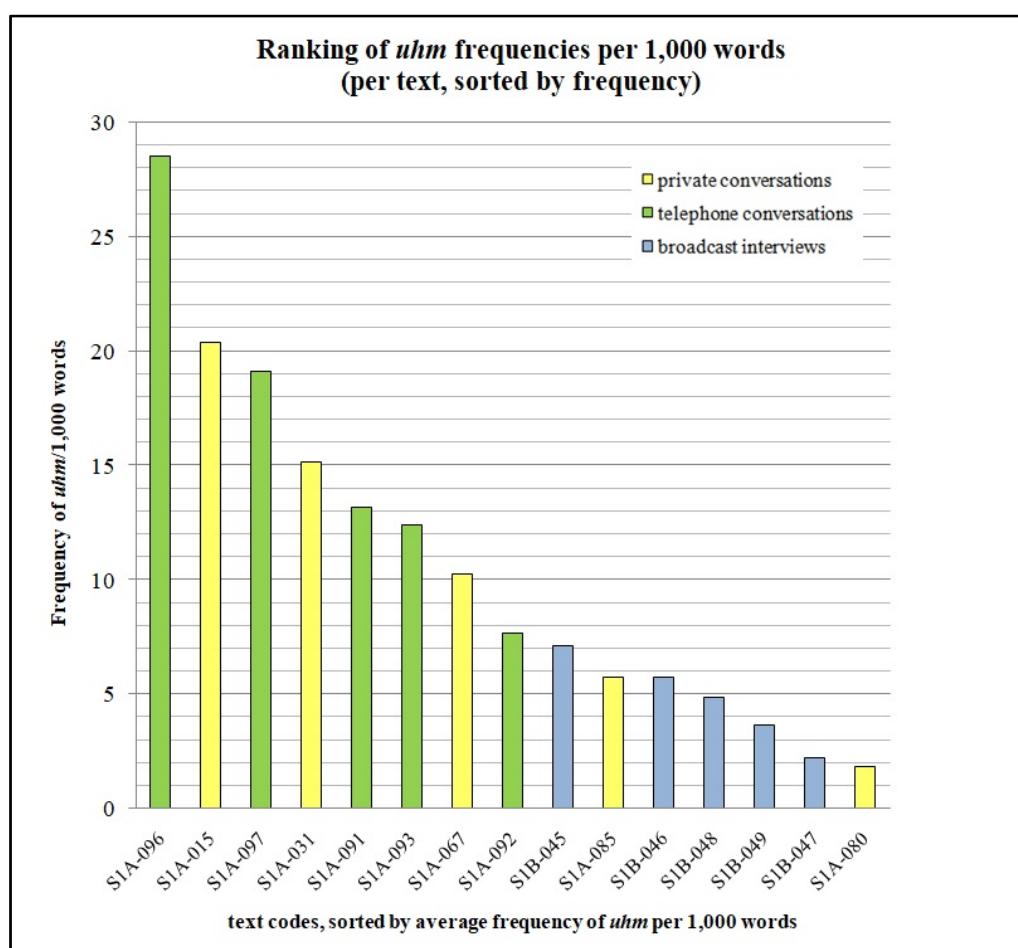
Due to these high fluctuations of UH/M frequencies within genres, it was deemed meaningful not to solely rely on the comparison of average values obtained for each genre, but to also consider the individual corpus texts. To this end, rankings of the texts ranging from highest to lowest average frequency per 1,000 words were created for both *uh* and *uhm*; these rankings were then investigated with regard to the distribution of the three chosen genres. Figure 12 and 13 show these rankings of corpus texts according to their average UH/M frequency per 1,000 words for *uh* and *uhm* respectively.

Figure 12: Ranking of *uh* frequencies per 1,000 words per corpus text (sorted by frequency)



The above bar chart shows that there does not seem to be a correlation between the production of *uh* and genre. Generally speaking, the different genre texts are more or less evenly distributed across the above ranking, even though there seems to be a tendency of *uh* to occur in telephone conversations. Moreover, it is shown that three out of the five texts chosen for the genre of private conversations include only a very small average frequency of *uh* per 1,000 words. Finally, the above chart relativizes the previously mentioned observation that *uh* tends to occur more frequently in broadcast interviews than in private conversations by showing that the former genre is in fact rather evenly distributed over the ranking of corpus texts. If the above figure does not provide much evidence for an influence of genre on the production of *uh*, Figure 13 below suggests that there is a much stronger correlation between setting and *uhm*.

Figure 13: Ranking of *uhm* frequencies per 1,000 words per corpus text (sorted by frequency)



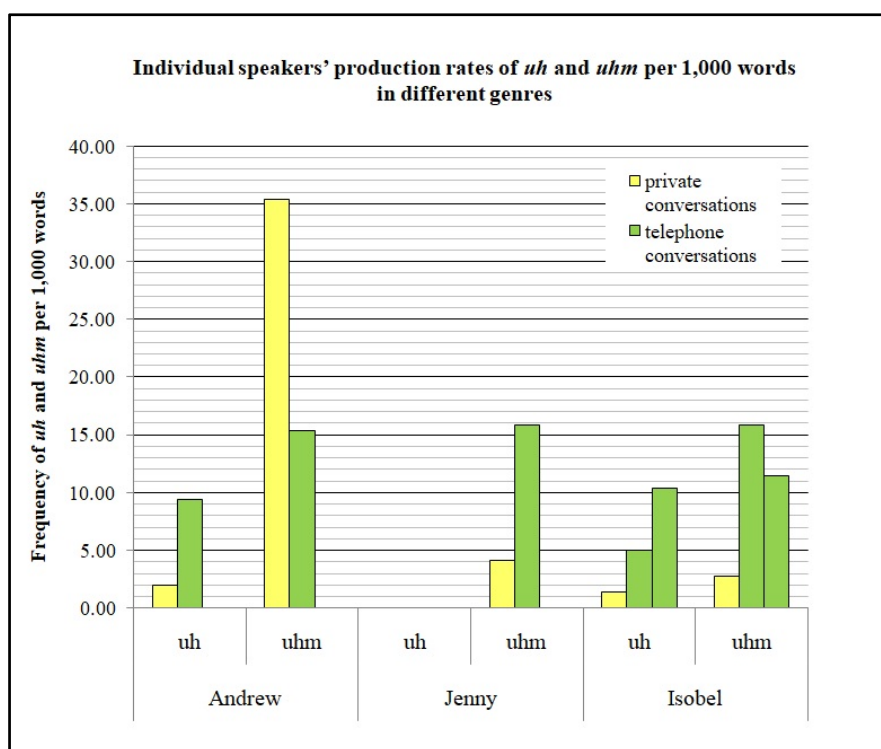
It can be seen in Figure 13 that the trends observed with regard to the production of *uhm* seems to be considerably more genre-dependent than that of *uh*. When looking at the above bar chart, the tendency of private settings – i.e. telephone conversations and private

conversations – to contain higher frequencies of *uhm* than the speech produced in the public speaking environment becomes immediately obvious. Moreover, it is shown that while there were clear tendencies of *uhm* to occur at higher rates in telephone conversations than in broadcast discussions, the genre of private face-to-face conversations is rather evenly distributed across the above ranking and therefore does not seem to produce an effect on the production of *uhm*.

Summing up the above results, it has been shown that both *uh* and *uhm* appear significantly more often in telephone conversations than in any other genre, while private face-to-face conversations have not been found to influence the production rates of *uh* or *uhm*. In addition, it has been found that genre seems to produce a more noticeable effect on *uhm* than on *uh*. Finally, it has also been shown that the individual corpus texts within genres tend to differ considerably from each other with regard to their average frequencies of *uh* and *uhm*. This finding might be interpreted to suggest that UH/M are not determined by genre, but rather by other factors such as a speaker's personal tendency to utter UH/M (cf. sections 9.1 and 9.4 for an examination of this hypothesis).

In order to further investigate the hypothesis that UH/M are idiosyncratic phenomena, a brief study of different speakers' production rates of *uh* and *uhm* across different genres was conducted. This was possible because – as was discovered during an analysis of the speaker profiles provided for the chosen corpus texts – three speakers were present in more than one conversation within the defined corpus; to facilitate reference, these speakers shall henceforth be referred to by their first names *Andrew*, *Isobel*, and *Jenny*. The average frequencies of *uh* and *uhm* per 1,000 words uttered by these three speakers were compared both across different genres and with each other; Figure 14 below illustrates the relevant data. Note that Andrew and Jenny both participated in only two corpus texts, while Isobel took part in three of the studied conversations; for this reason, Isobel is the only speaker for which three bars are shown in the following chart. Moreover, it should be pointed out that Jenny only participated in conversations where Isobel was also present; in other words, the two female speakers are interlocutors in both of Jenny's (and, subsequently, two of Isobel's) texts. The bars for each speaker are organized in such a way that the leftmost bars for *uh* and *uhm* refer to the same text; the same applies to Isobel's data set, where the leftmost bars for each variant refer to her private conversation with Jenny, the middle bars represent the data obtained in the analysis of her telephone conversation with Jenny and the rightmost bars show her average production rates of UH/M in another telephone conversation with a man called David.

Figure 14: Individual speakers' average production rates of *UH/M* in different genres



The data presented in the above chart shows that with the exception of Andrew, who produced fewer instances of *uhm* on the telephone than he did in private face-to-face conversations, all speakers uttered *uh* and *uhm* more frequently in telephone calls; moreover, all speakers generally produced more instances of *uhm* than *uh* within the same text. Apart from these findings, a comparison of Isobel's different production rates of *UH/M* in the three different conversations revealed that she not only produced significantly more instances of *uh* and *uhm* when talking on the telephone than in the face-to-face conversation, but also that her production rates in the two different telephone conversations varied significantly. In addition, it can be seen in Figure 14 that while Isobel produced only around 5 instances per 1,000 words in her telephone conversation with Jenny – which is represented by the second of Isobel's three bars in Figure 14 – she uttered twice as many *uhs* when talking to another interlocutor, David. Moreover, while *uhm* was considerably more frequent in her conversation with Jenny, Isobel produced more balanced rates of *uh* and *uhm* when talking to David. In addition, it was found that Jenny did not produce one single instance of *uh* in more than 2,000 words of speech despite uttering average frequencies of *uhm* in both conversations. As will be discussed in sections 9.1 and 9.4, these findings could suggest that the production of *UH/M* are also determined by other factors than genre, such as the precise context of a given speaking situation or a speaker's personal tendency to utter them.

Having presented the results obtained in the analysis of the effects of genre and setting on the frequencies of *uh* and *uhm* in spoken discourse, the following chapter will explore the preferred positions of *uh* and *uhm* in turns and TCUs.

## 8.2. *Uh* and *uhm* in turns and TCUs

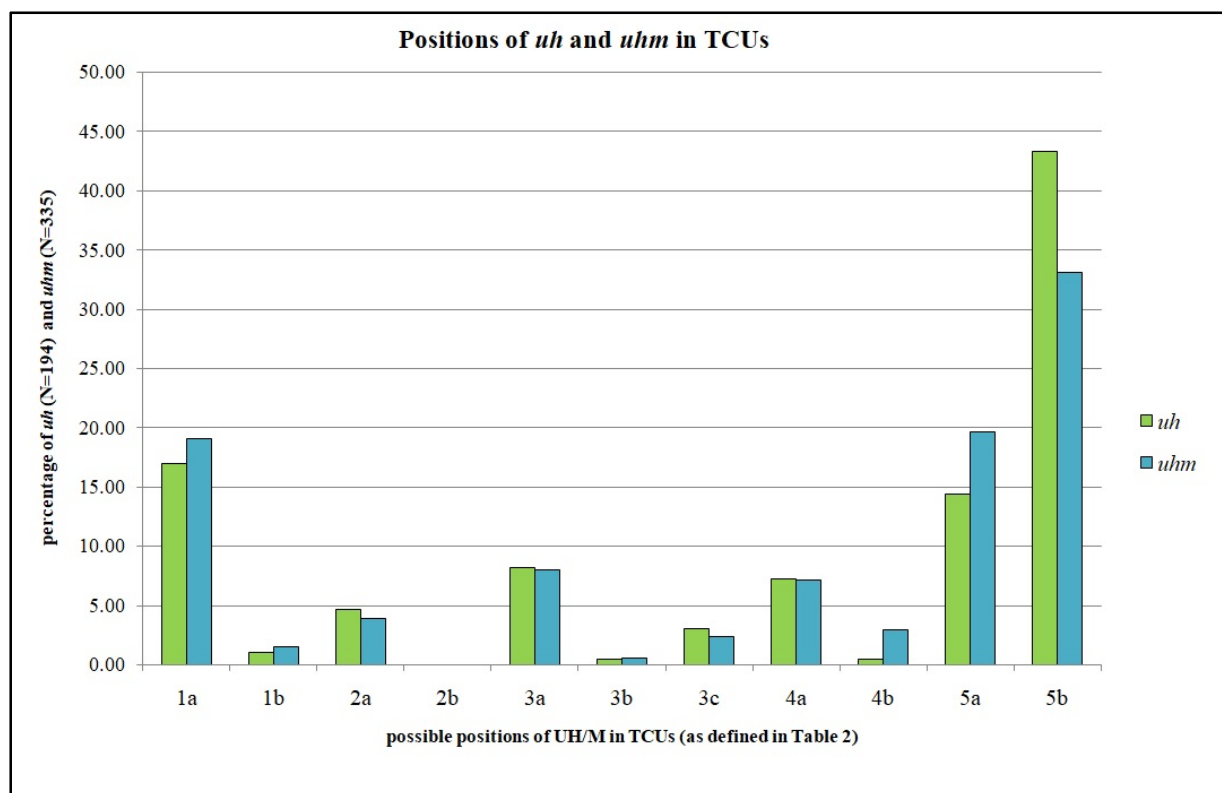
As stated above, the present chapter presents the results obtained in the analysis of *uh* and *uhm* in turns and TCUs. It was decided to discuss these two levels in one single chapter because they are inextricably linked to each other, TCUs being the units that construct the conversational turns of spoken interaction. As was detailed in chapter 7.3.2, the positions of *uh* and *uhm* in TCUs were investigated by localizing all 194 instances of *uh* and 335 *uhms* with respect to TCUs, TRPs, CPs and points of speaker transitions; the results obtained in this analysis are provided in Table 2 and Figure 15 below. Note that 5 instances of *uhm* were excluded from the below analysis because they occurred in standalone positions, and that the numbers provided for *uhm* do therefore not add up to 335 instances and a 100%, but to 330 occurrences and 98.51%<sup>19</sup>. The excluded instances will be investigated in the further course of this section.

Table 2: Positions of *uh* and *uhm* in turns and TCUs

	<i>uh</i> (N=194)		<i>uhm</i> (N=335)	
	N <sub>uh</sub>	% <sub>uh</sub>	N <sub>uhm</sub>	% <sub>uhm</sub>
<b>Category 1: UH/M is situated at a TRP within a speaker's turn</b>	<b>35</b>	<b>18.04</b>	<b>69</b>	<b>20.60</b>
1a UH/M are situated at a TRP between two complete TCUs	33	17.01	64	19.10
1b UH/M are situated between a complete and an aborted TCU	2	1.03	5	1.49
<b>Category 2: UH/M is situated at a non-TRP completion point (CP)</b>	<b>9</b>	<b>4.64</b>	<b>13</b>	<b>3.88</b>
2a UH/M are situated at a CP between two complete TCUs	9	4.64	13	3.88
2b UH/M are situated at a CP between a complete and an aborted TCU	0	0.00	0	0.00
<b>Category 3: UH/M is situated at a point of abortion (PoA)</b>	<b>23</b>	<b>11.86</b>	<b>37</b>	<b>11.04</b>
3a UH/M are situated between an aborted and a complete TCU	16	8.25	27	8.06
3b UH/M are situated between two aborted TCUs	1	0.52	2	0.60
3c UH/M are situated at a point of abortion that involves speaker transition	6	3.09	8	2.39
<b>Category 4: UH/M is situated at a CP/TRP involving turn transition</b>	<b>15</b>	<b>7.73</b>	<b>34</b>	<b>10.15</b>
4a UH/M are situated at the first position in a turn-initial TCU	14	7.22	24	7.16
4b UH/M are situated after a complete, turn-final TCU	1	0.52	10	2.99
<b>Category 5: UH/M is not situated at a TRP</b>	<b>112</b>	<b>57.73</b>	<b>177</b>	<b>52.84</b>
5a UH/M occupy the second (or third) position after (a cluster of) CCID	28	14.43	66	19.70
5b UH/M occur inside the TCU, after the speaker has committed him-/herself lexically	84	43.30	111	33.13
<b>Sum</b>	<b>194</b>	<b>100.00</b>	<b>330</b>	<b>98.51</b>

<sup>19</sup> Also note that from this point onwards, the capital letter *N* will be used to represent the overall number (i.e. 100%) of all instances relevant to a specific category. For instance, in Table 2, the remark  $N_{uh}=194$  refers to the fact that there are 194 instances of *uh* that were studied in this step of the analysis.

Figure 15: Positions of *uh* and *uhm* in turns and TCUs



It can be seen in the above bar chart that *uh* and *uhm* showed very similar tendencies with regard to their positions in TCUs. Both variants occurred most often inside TCUs (position 5b); however, while the second and third largest proportions of *uh* appeared at turn-internal TRPs (position 1a) and after TCU-introductory conjunctions, connectives, interjections or discourse markers (position 5a), the reverse was found to be true for *uhm*, which occurred more frequently after TCU-introductory elements than at TRPs between two complete TCUs. With regard to these three major positions, the above graph also illustrates that the proportion of TCU-internal *uhs* outweighed the respective occurrences of *uhm* by more than 10% (cf. position 5b); in turn, *uhm* was the variant that appeared slightly more often at turn-internal TRPs and after TCU-introductory elements (cf. positions 1a and 5a).

The above table and graph also present some interesting findings concerning *uhs* and *uhms* positions with regard to different types of TCUs and completion points. For instance, it is shown above that *uh* and *uhm* both tend to precede complete TCUs rather than aborted ones; this is illustrated by the low frequencies of both variants at TRPs or PoAs preceding aborted TCUs (cf. positions 1b, 2b, and 3b). Moreover, it can be seen that both variants appeared considerably more often at TRPs than at PoAs and CPs, with approximately 20% of each variant occurring at TRPs as opposed to the roughly 11% appearing at PoAs and the less than

5% that were situated at CPs. With regard to the instances of UH/M that occurred at PoAs, it was also found that they were only seldom associated with turn transition; in fact, it was generally found that only a relatively small proportion of both variants coincided with speaker change. These findings provide an interesting insight into the functions fulfilled by *uh* and *uhm*, particularly with regard to the question whether they are symptom, signals, or both; thus, they will be discussed in detail in chapter 9 of this paper.

The above findings are summarized in Table 3, in which the number of different positions were reduced and partly rearranged into new categories so as to facilitate the comparison of preferred positions of *uh* and *uhm* with regard to turns and TCUs. The decision to count position 5a – UH/Ms *occurring after TCU-introductory CCIDs* – towards the *boundary* category is based on Hawkins’ (1971: 285) aforementioned assertion that clause-introductory CCIDs “[...] do not commit the speaker in any way to the type of construction which must follow them” and should thus be counted towards the *boundary* position; this argument was also deemed valid in the context of TCUs and turns.

Table 3: Positions of *uh* and *uhm* in turns and TCUs (summary)

	<i>uh</i> (in %)	<i>uhm</i> (in %)
<b>at or near TCU boundaries</b>	<b>37.11</b>	<b>44.18</b>
at turn-internal TRPs	18.04	20.60
at turn-internal CPs	4.64	3.88
following TCU-introductory CCIDs	14.43	19.70
<b>at turn-internal PoAs</b>	<b>8.76</b>	<b>8.66</b>
<b>at points of turn transition</b>	<b>10.82</b>	<b>12.54</b>
first position in a turn-initial TCU	7.22	7.16
last position after a complete TCU	0.52	2.99
at a PoA involving turn transition	3.09	2.39
<b>inside TCUs</b>	<b>43.30</b>	<b>33.13</b>
excluded from analysis	0.00	1.49
<b>Sum</b>	<b>100.00%</b>	<b>100.00%</b>

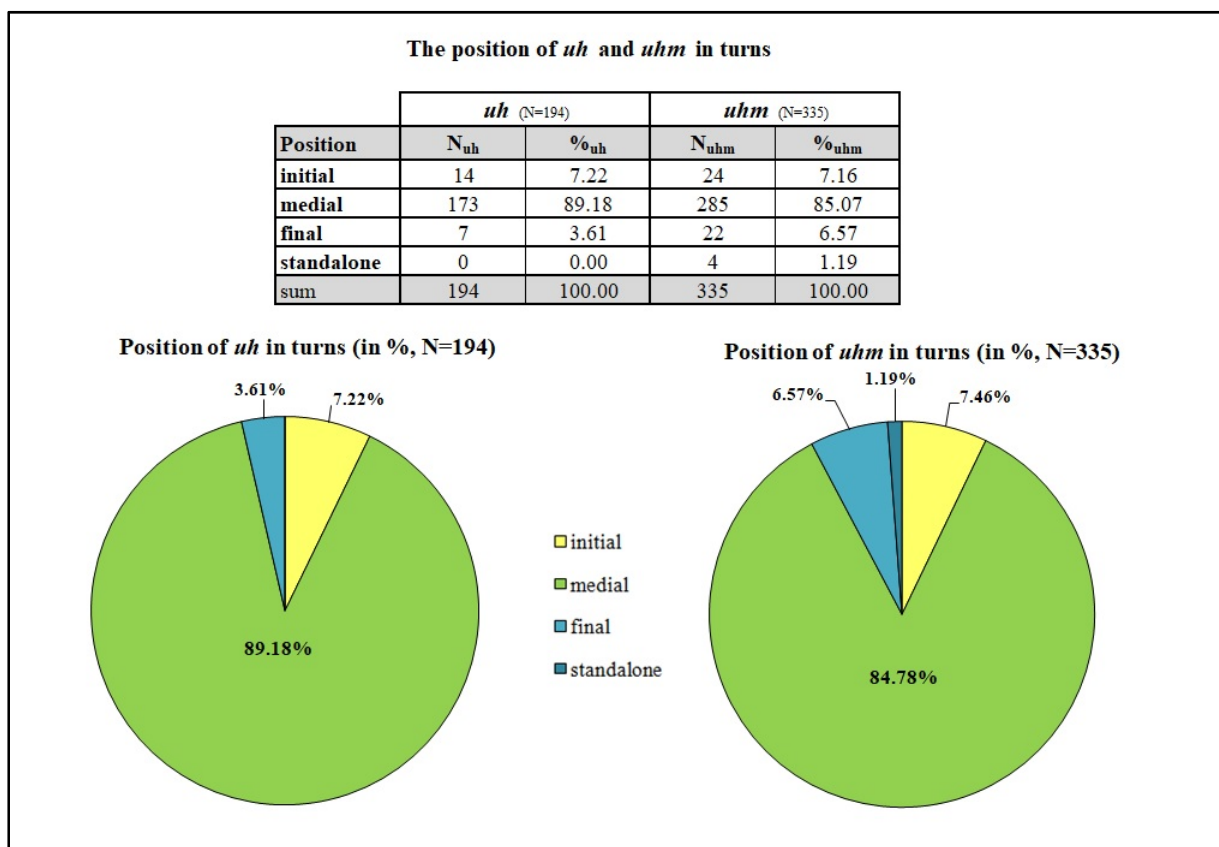
Table 3 does not only further highlight the trends outlined in the above presentation of the obtained results, but also provides some further insights into the differences between the two variants. For instance, if the bar chart in Figure 15 seemed to suggest a strong tendency of both variants to occur inside TCUs, Table 3 shows that *uh* slightly prefers TCU-internal positions, while *uhm* seems to be the variant that is more frequently associated with TCU boundaries. Moreover, even if only relatively small proportions of both variants— roughly 11% of *uhs* and 14% of *uhms* – were found to coincide with points of speaker change, it is worth noting that *uhm* was found to appear considerably more often in turn-final positions



than *uh*, while the proportions of *uh* and *uhm* in turn-initial positions were very similar. All of these findings provide some interesting insights into the differences between *uh* and *uhm*, and will therefore be returned to in the discussion offered in section 9 of this paper.

It was already addressed above that there seem to be some similarities and differences between *uh* and *uhm* with regard to their positions in turns; these are visually represented in Figure 16 below, which shows the percentages of both variants in turn-initial, turn-medial, turn-final, and standalone positions.

Figure 16: The position of *uh* and *uhm* in turns (in percentages)



The above pie chart shows that the vast majority of both variants occurred inside turns, while the other positions (turn-initial, turn-final, and standalone) only accounted for roughly 11% of all *uhs* and 15% of all *uhms*. This finding could suggest that *uh* and *uhm* are predominantly linked to the representation of a speaker's cognitive processes and his/her intention of keeping the turn rather than the signaling of turn-taking and turn-yielding; this hypothesis will be discussed in more detail in section 9.3.3. Despite these similarities, the above graph also shows that there were some differences between *uh* and *uhm* with regard to their tendency to occur in turn-final positions (i.e. after turn-final, complete TCUs and at PoAs that involved



turn transition), with *uhm* occurring nearly twice as often at the end of a turn than *uh*. Moreover, even though standalone positions were extremely rare in the corpus, it is worth noting that *uhm* was the only variant being uttered independently of a turn. Finally, as regards turn-initial instances of UH/M, it was found that very similar proportions of *uh* and *uhm* – roughly 7% of each – occurred in this position. In an attempt to gain a deeper understanding of why speakers sometimes utter *uh* and *uhm* in turn-initial, turn-final, and standalone positions, the respective instances were briefly studied with regard to their immediate conversational context; the results obtained in this step of the analysis will be outlined in the following paragraphs.

It has been shown in Figure 16 that very similar proportions of *uh* and *uhm* – roughly 7% of each – occurred at the beginning of turns. The categorization of these turn-initial *uhs* and *uhms* according to their precise contexts is very difficult, because most of them could be argued to fulfill more than one function at a time (cf. sections 3.4, and chapter 5). What could still be revealed, however, is that the largest proportion of turn-initial *uhs* and *uhms* did not occur in the context of active turn-taking but in answers to questions, accounting for 7/14 instances of *uh* (i.e. 50%) and 18/24 instances of *uhm* (i.e. 75%). Examples 12 and 13 illustrate two such answer-initial occurrences.

- (12) S1: Is it going to make some serious inroads into your writing <,> //  
 S2: **Uh** I should think so // but *uhm* <,> in the end I think a religious thinker is judged <,> by what he does in life // not just by the books that he writes //  
 (ICE-GB 2006: S1B-047: 016-017)
- (13) S1: Where is Weymouth //  
 S2: **Uhm** it's near do you know Bournemouth // it's in Dorset // [...]  
 (ICE-GB 2006: S1A-097: 044-047)

Apart from the finding that *uhm* occurred significantly more often in answers to questions than its purely vocalic counterpart *uh*, no other clear and statistically observable trends were found to exist for turn-initial UH/Ms.

Closely linked to the finding that turn-initial instances of *uhm* tend to occur in answers to questions is the fact that three out of the four standalone occurrences of *uhm* seemed to constitute failed attempts at turn-accepting after a proposition or question, as is shown in examples (14) and (15) below.

- (14) S2: So what did you do today <,> //  
 S1: **Uhm** =  
 S2: = swotting away //  
 S2: Yeah // I went to the library this morning // [...] (ICE-GB 2006: S1A-092:093-097)
- (15) S1: The other thing is *uhm* <,> do you confide in her <,> // does she feel excluded // because you don't <1> exactly confide in her </1> //  
 S2: <1> Well </1> <,> I don't // [But then <,> I haven't] / I mean I never have <,> // and I'm I'm rather scared that you know that would seem rather artificial to her // and as an attempt to win her over // and of course <,> you know <2> she's terribly alive to </2> things like that <,> //  
 S1: <2> But you could start slowly </2> //  
 S2: **uhm** <,>  
 S1: But I agree it's very difficult if you never have //  
 (ICE-GB 2006: S1A-031: 181-189)

In both of these passages, one speaker asks a question (*So what did you do today* in example (14)) or makes a proposition (*But you could start slowly* in example (15)), to which the other speaker wishes to react and thus begins her turn by uttering *uh* or *uhm*. However, before he/she begins the actual answer, the initial speaker adds another TCU (*swotting away* in example (14) and *But I agree it's very difficult if you never have* in example (15)). Thus, standalone occurrences appearing in contexts similar to the below examples are argued to constitute failed attempts at turn-accepting.

As regards turn-final instances, it was found that only very small proportions of both variants (7 out of 194 *uhs* and 22 out of 335 *uhms*) occurred in this position, even though *uhm* was found to appear almost twice as often at the end of turns than *uh*. For this reason, it was decided not to extensively study the precise contexts of these turn-final instances, for the numbers of *uhm* and especially of *uh* were far too small to allow any noteworthy conclusions; still, the most pertinent findings made in the brief study of turn-final instances shall be briefly summarized. Firstly, it was found that both turn-final *uhs* and *uhms* most frequently occurred in the context of other-interruption, i.e. at points where the respective instance of UH/M coincided with another speaker taking the turn without allowing his/her predecessor to finish his/her contribution; this context accounted for 6/7 instances of turn-final *uhs* and 11/22 *uhms*. Moreover, it was shown that the turn-final instances of *uhm* were more evenly distributed across the different contexts, while *uh* predominantly occurred in other-interruption (and never in the context of other-continuation); however, this finding is presumably linked to the very small number of turn-final *uhs* that were investigated in this step of analysis. In conclusion, the analysis of turn-final *uhs* and *uhms* has shown that *uhm* was more frequent in turn-final positions on the one hand and more often associated with voluntary turn-yielding

and other-continuation; however, since turn-final instances accounted for merely 3.6% of all *uhs* and 6.6% of all *uhms* in the corpus, one should not attach too much importance to these findings.

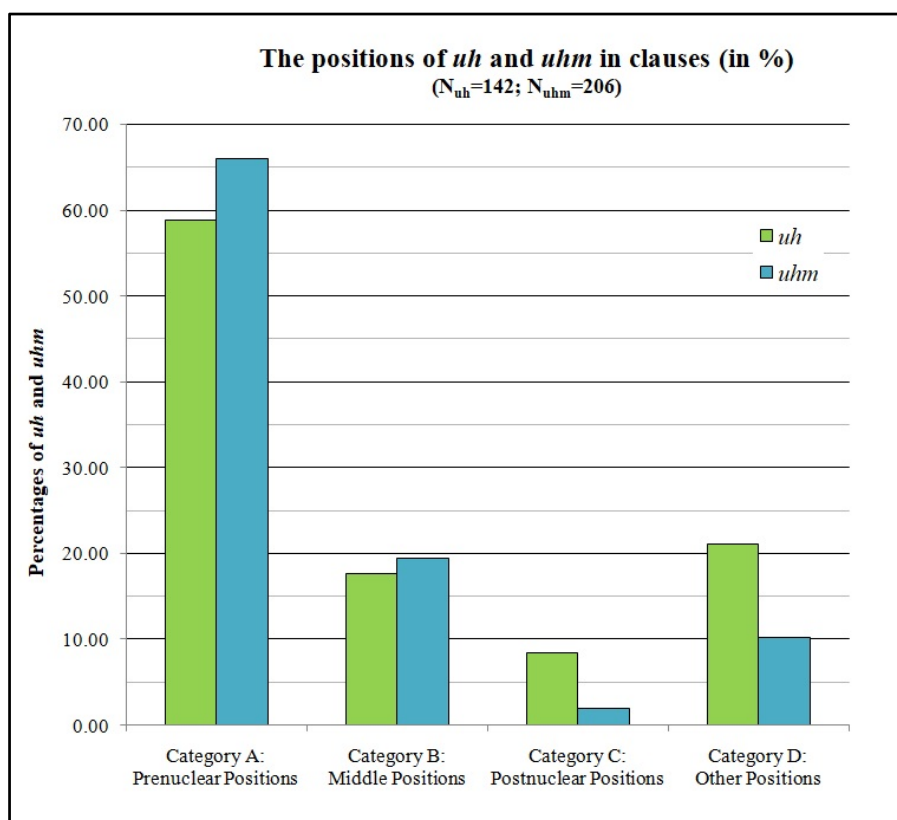
In summary, the present chapter has shown that even though *uh* and *uhm* show very similar tendencies with regard to their preferred positions in turns and TCUs, some differences between the two variants could be observed. For instance, *uh* was found to appear more often inside TCUs than *uhm*; in turn, *uhm* appeared more frequently at or near TCU boundaries. Moreover, as regards their positions in turns, it was found that even though both *uh* and *uhm* occurred most frequently in turn-medial positions, *uhm* was the variant that appeared more often in answers to questions and in turn-final positions. Having identified the major trends with regard to *uh*'s and *uhm*'s positions in turns and TCUs, this chapter will move on to present the results obtained on the clause-level.

### **8.3. The positions of *uh* and *uhm* in clauses**

The present section investigates the positions of *uh* and *uhm* in clauses. As was detailed in section 7.3.3, the results to be presented in this section were obtained by classifying each individual occurrence of *uh* and *uhm* according to their precise position within clauses. Table 4 and Figure 17 below offer an overview of the results obtained in this step of analysis, while a more detailed list including all possible positions defined in section 7.3.3 is provided in Appendix 3. As can be seen below, some instances of *uh* and *uhm* could not be described in terms of the predefined locations presented in section 7.3.3 and were thus assigned to the *other positions* category; an analysis of these instances showed that they all appeared inside constituents such as subjects or adjuncts. Moreover, it should be pointed out that several instances of *uh* and *uhm* were excluded from the clausal analysis because they occurred in the context of self-interruption, repetitions, or non-clausal turns; therefore, a 100% do not refer to the usual 194 *uhs* and 335 *uhms*, but to 142 *uhs* and 206 *uhms*.

Table 4: Positions of *uh* and *uhm* in clauses

	<i>uh</i> (in %, N= 142)	<i>uhm</i> (in %, N= 206)
<b>Category A: Prenuclear Positions</b>	<b>52.82</b>	<b>66.02</b>
1 Initial / at clause boundary	32.39	36.89
2 CCID # subject	19.01	23.79
3 Adjunct # subject	0.70	1.94
4 CCID # adjunct + subject	0.00	0.97
5 CCID # verb (in clauses that lack a subject / imperatives / interrogations)	0.70	2.43
<b>Category B: Middle Positions</b>	<b>17.61</b>	<b>19.42</b>
1 subject # verb	2.82	1.94
2 auxiliary verb # MVB	3.52	5.34
3 MVB # non-clausal complementation	7.75	6.80
4 MVB # clausal complementation	2.82	4.85
5 <i>there</i> + MVB # subject	0.70	0.00
6 auxiliary # MVB in interrogative clauses	0.00	0.49
<b>Category C: Postnuclear Positions</b>	<b>8.45</b>	<b>1.94</b>
1 MVB / Clause # adjunct (non-clausal)	7.75	1.46
2 MVB / Clause # adjunct (clausal)	0.70	0.49
3 MVB / Clause + adjunct (non-clausal) # adjunct (non-clausal)	0.00	0.00
4 MVB / Clause + adjunct (clausal) # adjunct (clausal)	0.00	0.00
5 MVB + object # object complement (non-clausal)	0.00	0.00
<b>Category D: other positions</b>	<b>21.13</b>	<b>12.62</b>
<b>Sum</b>	<b>100.00</b>	<b>100.00</b>

Figure 17: The distribution of *uhs* and *uhms* across different clausal positions ( $N_{uh}=142$ ,  $N_{uhm}=206$ )

The above table shows the frequencies of *uh* and *uhm* in different positions within clauses, while the bar chart in Figure 17 summarizes the presented data by illustrating the proportions

of *uh* and *uhm* occurring in prenuclear, middle, postnuclear and “other” positions. As can be seen, the frequencies of *uh* and *uhm* were generally found to decrease the further the clause proceeded, with the majority of both variants – around half of all *uhs* and two thirds of *uhms* – appearing in prenuclear positions, roughly 20% of each variant occurring in middle positions, and the minority of all *uhs* and *uhms* appearing in postnuclear positions. As will be discussed in section 9.3.2, this finding is interesting insofar as it supports the hypothesis that *uh* and *uhm* are associated with a speaker’s planning processes. Within the most frequent category – prenuclear positions – it was found that both *uh* and *uhm* tended to appear immediately at clause boundaries (accounting for approximately a third of each variant), while the second largest proportion occurred between clause-introductory CCIDs and the clause’s subject (cf. Table 4 above). Moreover, the above data also suggest that even though both *uh* and *uhm* occurred considerably more often at constituent boundaries than inside constituents (represented by category D), both variants appeared more frequently inside constituents than at postnuclear constituent boundaries. Finally, as regards the differences between the two variants, it can be seen above that *uhm* was considerably more frequent in prenuclear positions, while *uh* noticeably exceeded *uhm* both in postnuclear positions and inside constituents, occurring more than four times as often in postnuclear positions and more than twice as often inside constituents. Having identified these main trends, the following paragraphs will examine the precise positions favored by *uh* and *uhm* in prenuclear, middle, and postnuclear positions; this analysis is to a large extent based on the extensive table of results provided in Appendix 3.

As has already been shown above, the largest proportion of *uh* and *uhm* was found to occur in prenuclear positions, with the majority of prenuclear *uhs* and *uhms* occurring directly at clause boundaries. Apart from this finding, the extensive analysis presented in Appendix 3 revealed that both *uh* and *uhm* showed a slight tendency to appear at clause boundaries that were not introduced by a CCID rather than preceding such a clause-introductory element. Concerning syntactic differences between *uh* and *uhm*, it was found that *uh* occurred slightly more often after clause-introductory CCIDs than before them, while the reverse was true for *uhm*. In summary, the analysis of *uh* and *uhm* in prenuclear positions yielded three main results: firstly, that both *uh* and *uhm* strongly favored clause-initial positions; secondly, that the probability of uttering UH/M in prenuclear positions seemed to gradually decrease the further a speaker committed him-/herself to the clause; and, thirdly, that *uh* tended to occur slightly more often after CCIDs, while *uhm* more often occurred before them.

If it was stated above that the frequency of UH/M in prenuclear positions decreases the further the speaker committed him-/herself to the clause, the reverse was found to be true for middle positions, where the frequencies of *uh* and *uhm* seemed to increase the further the clause proceeded. In that sense, Table 4 above shows that the smallest proportions of *uh* and *uhm* in middle positions occurred between the subject and the verb, while the largest proportion occurred between the main verb and its (non-clausal or clausal) complements. As concerns postnuclear positions, it was not only found that *uh* occurred considerably more often in these locations than *uhm*, but also that almost all postnuclear instances of *uh* and *uhm* appeared before non-clausal adjuncts (cf. Appendix 3).

Summing up the above findings, it has been shown in the present chapter that *uh* and *uhm* generally showed similar tendencies with regard to their positions within clauses, even though some differences could be observed between the two variants. For instance, while it was shown above that both *uh* and *uhm* occurred most frequently in prenuclear positions, *uhm* was more often associated with clause boundaries, while *uh* was the variant that occurred more frequently inside constituents and in postnuclear positions; moreover, the present data was able to illustrate some trends concerning UH/M's general tendencies with regard to clause positions. Following the investigation of clauses, *uh* and *uhm* will now be analyzed with regard to their positions in phrases.

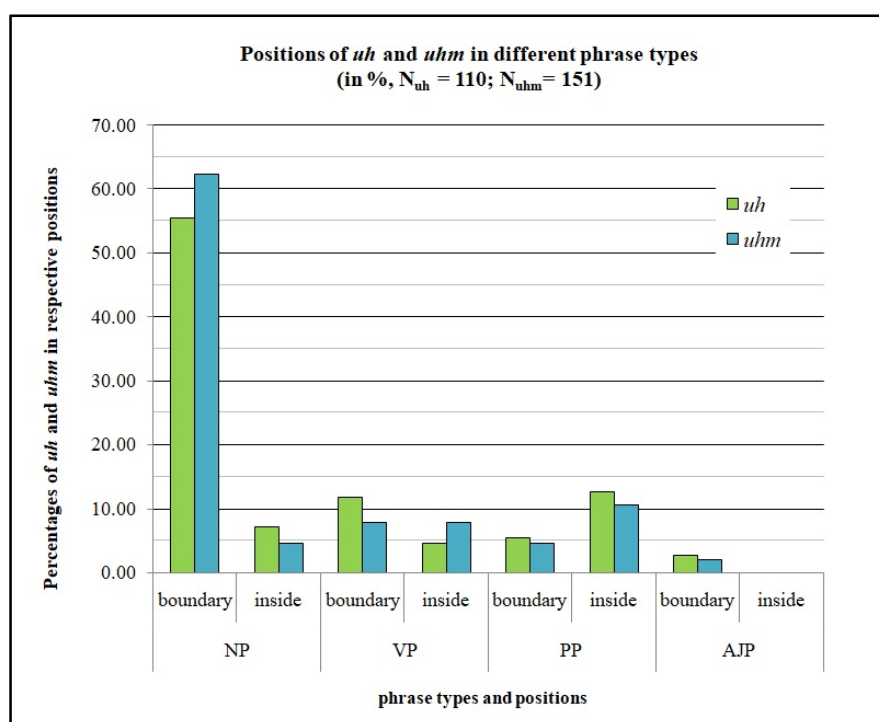
#### **8.4. The positions of *uh* and *uhm* in phrases**

The next linguistic level to be examined is the phrase level. As was detailed in chapter 7.3.4, four phrase types were analyzed: noun phrases (NPs), adjective phrases (AJs), verb phrases (VPs), and prepositional phrases (PPs). After excluding all those instances of *uh* and *uhm* that occurred in contexts that were not relevant to this analysis – e.g. before conjunctions or discourse markers, before/inside adverb phrases, in repetitions or in standalone positions – a sample of 110 *uhs* and 151 *uhms* remained; these were then categorized according to their positions in the respective phrase types. Table 5 below presents the data obtained in this step of analysis; moreover, Figure 18 provides a visual representation of these results.

Table 5: Positions of *uh* and *uhm* in phrases ( $N_{uh}=110$ ;  $N_{uhm}=151$ )

phrase type	position	$N_{uh}$	$\%_{uh}$	$N_{uhm}$	$\%_{uhm}$
NP	boundary	61	55.45	94	62.25
	inside	8	7.27	7	4.64
VP	boundary	13	11.82	12	7.95
	inside	5	4.55	12	7.95
PP	boundary	6	5.45	7	4.64
	inside	14	12.73	16	10.60
AJP	boundary	3	2.73	3	1.99
	inside	0	0.00	0	0.00
Sum		110	100.00	151	100.00

Figure 18: Positions of *uh* and *uhm* in phrases (in percentages)



The above table and bar chart show that *uh* and *uhm* display very similar tendencies as regards the syntactic level of phrases. As can be seen above, the majority of the 110 *uhs* and 151 *uhms* analyzed – approximately 60% of each variant – occurred at the boundary of noun phrases, while the other positions – inside noun phrases, before or inside verb phrases, before or inside preposition phrases, and before or inside adjective phrases – were significantly less popular. In addition, it was found that both variants generally tended to appear at phrase boundaries rather than inside the respective phrases; however, there were two exceptions to this trend. On the one hand, it was found that the instances of *uhm* associated with verb phrases were evenly split between the two positions, in such a way that half of the relevant occurrences were located at verb phrase boundaries and the other half appeared inside verb phrases. On the other hand, it can be seen in the above graph that both *uh* and *uhm* occurred

more frequently inside preposition phrases than at their boundaries. As regards syntactic differences between *uh* and *uhm*, it was found that *uhm* considerably exceeded *uh* at noun phrase boundaries, while the latter appeared more frequently at verb phrase boundaries and inside noun phrases; the other differences between *uh* and *uhm* were only minor.

In an attempt to discover a pattern in *uh*'s and *uhm*'s phrasal positions, all instances of UH/M occurring inside phrases were investigated with regard to their precise position; the results of this enterprise are presented in Appendix 4. It was found in this step of the analysis that the NP-internal occurrences of *uh* and *uhm* were frequently situated between a determiner and the noun phrase head (*DT* # *NPHD*); this is the case, for instance, in the NPs *a uh misidentification* (ICE-GB 2006: S1B-047: 044) and *the uhm sort of board* (ICE-GB 2006: S1A-067:093). As concerns verb phrases, it was found that the vast majority of VP-internal *uhs* and *uhms* appeared between an auxiliary and the main verb (as is the case in the TCUs *I'm just about uh seeing the light [...]* in ICE-GB 2006: S1A-096:031 and *he's uhm working on his language* in ICE-GB 2006:S1A-015:153); other positions – such as between a modal and the main verb or between two main verbs – merely constituted individual exceptions to this trend. Finally, with regard to preposition phrases, it was found that all PP-internal instances of *uh* and *uhm* occurred between the preposition and the preposition complement (as is the case in the TCU *he went to uh Colchester* in ICE-GB 2006: S1A-015:110); therefore, most PP-internal *uhs* and *uhms* equally occurred inside a PP and at a NP boundary.

### 8.5. The lexical/ pausal environment of *uh* and *uhm*

Following the syntactic analysis, the next aspect to be considered in the course of this study is the lexical and pausal environment of *uh* and *uhm*. As was outlined in chapter 7.3.5, this step comprises an analysis of the delays associated with the production of *uh* and *uhm* on the one hand and the investigation of common collocates of the two variants on the other hand. The present section will first present the results obtained in the analysis of *uh*'s and *uhm*'s lexical environments, while the second part of this section is dedicated to *uh*'s and *uhm*'s most frequent collocates.

Figure 19 below illustrates the findings obtained with regard to *uh*'s and *uhm*'s pausal environments, while the table provided above the bar chart summarizes once more the precise positions comprised in the different categories. An extensive list of the frequencies of UH/M in each individual position is provided in Appendix 5.



Figure 19: The lexical/pausal environment of *uh* and *uhm*

category	positions comprised within category
no delay	w#w, -#w, w# -
short delay	- #s, s# -, w#s, s#w
medium delay	s#s
long delay	ss#ss, -#ss, ss# -, s#ss, ss#s, w#ss, ss#w
other	- # - , laugh#w, w#laugh

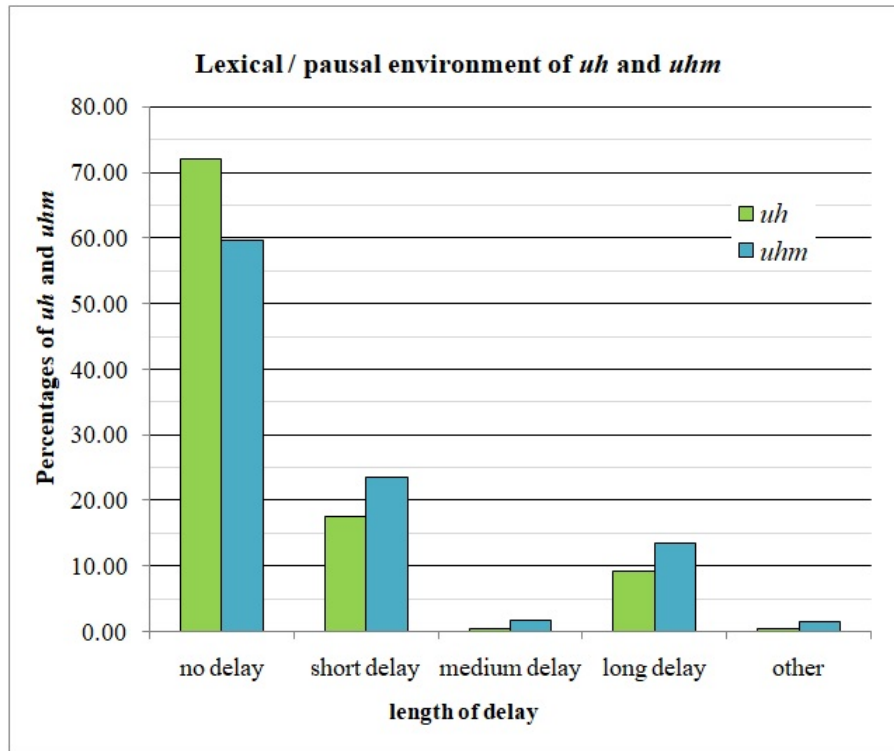


Figure 19 above shows the positions of all instances of *uhs* and *uhms* with regard to their lexical and pausal environment. As can be seen, the majority of both variants – around 72% of *uhs* and 60% of *uhms* – did not co-occur with pauses and could therefore not be associated with any delays in speech delivery. The second largest proportions of both *uh* and *uhm* appeared with short delays, and both variants were only seldom associated with long delays. The categories *medium delay* and *other types of delay* occurred only rarely in the studied corpus. Even though both variants showed the same trends with regard to their lexical and pausal environments, it was observed that *uh* dominated *uhm* in the *no delay*-category, while *uhm* outweighed its vocalic counterpart in all other categories; therefore, the present data seems to suggest that even though the majority of variants could be assigned to the *no delay*-category, *uhm* is the variant that is slightly more frequently associated with delays in spoken interaction.

Following the analysis of pausal environment, *uh* and *uhm* were examined with regard to their most frequent collocates. To this end, the words preceding and succeeding each variant were investigated; afterwards, the lists of pre- and post UH/M collocates were compared with each other. Table 6 comprises four sections, representing the most frequent collocates preceding and succeeding each variant. Note that in these tables, only those collocates occurring a minimum of 5 times were included. Moreover, it should be pointed out that the percentages were calculated only after irrelevant occurrences (e.g. turn-final instances in the case of the analysis of post-UH/M collocates) had been removed from the respective data sets; therefore, the percentages presented in the following tables were calculated on the basis of different overall numbers, which are indicated above each table (marked by  $N_{uh}$  or  $N_{uhm}$ ).

Table 6: Most frequent collocates of *uh* and *uhm* ( $N \geq 5$ )

Pre-uh collocates

(N<sub>uh</sub>=138)

Rank	collocates	N	%
1	<i>and</i>	25	18.12
2	<i>is</i> (including 's)	11	7.97
3	<i>uh</i>	6	4.35
4	<i>but</i>	6	4.35

Post-uh collocates

(N<sub>uh</sub>=174)

Rank	collocates	N	%
1	<i>I</i>	16	9.20
2	<i>and</i>	10	5.75
3	<i>the</i>	9	5.17
4	<i>it</i>	7	4.02
5	<i>uh</i>	5	2.87

Pre-uhm collocates

(N<sub>uhm</sub>=220)

Rank	collocates	N	%
1	<i>and</i>	35	15.91
2	<i>is</i> (including 's)	20	9.09
3	<i>but</i>	18	8.18
4	<i>so</i>	10	4.55
5	<i>said</i>	8	3.64
6	<i>the</i>	7	3.18
7	<i>that</i>	6	2.73

Post-uhm collocates

(N<sub>uhm</sub>=278)

Rank	collocates	N	%
1	<i>I</i>	29	10.43
2	<i>and</i>	21	7.55
3	<i>it</i>	13	4.68
4	<i>well</i>	10	3.60
5	<i>you know</i>	8	2.88
	<i>so</i>		
	<i>the</i>		
	<i>it</i>		
6	<i>yeah</i>	6	2.16
7	<i>that</i>	5	1.80
	<i>uh</i>		
	<i>I mean</i>		
	<i>no</i>		
	<i>he</i>		
	<i>a</i>		

As can be seen in the above tables, no strong tendencies could be observed with regard to the collocates of *uh* and *uhm*, mainly because the two variants were found to only seldom co-occur with the same collocate more than five times; as a consequence, even the most frequent

collocates accounted for rather small proportions of *uh* and *uhm*. Interestingly, the range of frequent collocates was considerably more restricted for *uh* than for *uhm*. This can be illustrated by drawing attention to the tables representing post-UH/M collocates. In these tables, it can be seen that while there were only five words that preceded *uh* more than five times, three times as many preceding collocates were identified for *uhm*; consequently, the range of *uhm*'s frequent collocates (and especially of post-*uhm* collocates) is also considerably more varied.

In addition to the observation to the above, the comparison of frequent collocates was able to yield some interesting insights into the syntactic position of UH/M on the one hand and the similarities and differences between the two variants on the other hand. For instance, the comparison of common collocates of *uh* and *uhm* revealed that both variants most frequently followed the conjunction *and* and often preceded the personal pronoun *I*; this coincides with the aforementioned finding that both *uh* and *uhm* frequently appeared between clause-introductory CCIDs and the respective clause's subject. In addition, it was also found that *uh* and *uhm* occurred more often after than before the conjunction *and*. The tendency of UH/M to occur after conjunctions can also be observed when considering that both variants were rather frequently situated after the subordinating conjunction *but* and – in the case of *uhm* – *so*, the latter of which was almost exclusively used as a connective in the defined corpus. Moreover, it is also interesting to observe that *uhm* rather frequently co-occurred with the reaction signals *yeah* and *no* as well as with the discourse markers *well*, *you know* and *I mean*; neither of these collocates was found to frequently appear in the vicinity of *uh*. In turn, *uh* seemed to be the variant that more often co-occurred with other instances of *uh*. Finally, as regards pre-UH/M collocates, it was found that *uhm* rather frequently followed the verb *said*, which seems to suggest that *uhm* could be linked to the introduction of reported speech.

In summary, the present data have shown that *uh* and *uhm* showed very similar tendencies with respect to their most frequent collocates, even though some differences – such as *uh*'s tendency to co-occur with other instances of *uh*, or *uhm*'s inclination to precede discourse markers and reaction signals – could be observed. Having considered the last of Halliday's levels of syntactic analysis – the word level – this paper shall now turn to the final section of this chapter, which examines *uh*'s and *uhm*'s role in repetitions and self-interruptions.

## 8.6. *Uh* and *uhm* in repetitions and self-interruptions

As was stated in section 4.3, the existing literature has frequently linked the production of UH/M to self-interruption, self-correction, and repetitions; this final section of the present chapter examines this relationship on the basis of the studied corpus. For this purpose, all instances of *uh* and *uhm* situated inside, directly before or immediately after repetitions as well as those occurrences of UH/M located at points of self-interruption were counted; the results obtained in this step of the analysis are summarized in Table 8 below. However, before turning to the statistical analysis of these findings, Table 7 provides examples for each of the defined categories, so as to clarify what exactly they refer to.

Table 7: Examples of *uhs* and *uhms* occurring in the context of repetition and self-interruption

	Definition	Examples
<i>inside repetitions</i>	UH/M are located between two identical repeated items, where <i>items</i> may refer to individual words or larger repeated clusters	<p><i>And she uh &lt;,&gt; she 's so sophisticated</i> (ICE-GB 2006: S1A-091:345)</p> <p><i>it's uhm it's your opinion of yourself that one</i> (ICE-GB 2006: S1A-067: 369)</p>
<i>before repetitions</i>	UH/M occur directly before the repetition of an item	<p><i>Uh I I never saw myself as a costume designer</i> (ICE-GB 2006: S1B-045:003)</p> <p><i>We were talking about uhm &lt;,,&gt; that that that man at college</i> (ICE-GB 2006: S1A-091: 002)</p>
<i>after repetitions</i>	UH/M immediately follow the repetition of an item	<p><i>Bi-Bim and I always planned to have some done based on on uh photos of Auden and Isherwood</i> (ICE-GB 2006: S1A-015:52)</p> <p><i>And he said he said uhm well you're obviously slow</i> (ICE-GB 2006: S1A-091:30)</p>
<i>combined positions</i>	UH/M occur both before/after and inside a repetitions	<i>uh he he he's uhm he's uhm working on his language</i> (ICE-GB 2006: S1A-015:153)
<i>at points of self-interruption</i>	UH/M occur at points of abortion, where the speaker interrupts him-/herself to start a new unit that is not (syntactically and/or content-wise) related to the previous unit in any obvious way	<p><i>Now the middle August &lt;,&gt; is uh I mean I might still come up to London [...]</i> (ICE-GB 2006: S1A-097:225)</p> <p><i>I think uhm I don't know what my mother would have done [...]</i> (ICE-GB 2006: S1B-046: 44)</p>

Table 8: Proportions of *uh* and *uhm* appearing in the context of repetitions and self-interruptions

	<i>uh</i> (N=194)		<i>uhm</i> (N=335)	
	N <sub>uh</sub>	% <sub>uh</sub>	N <sub>uhm</sub>	% <sub>uhm</sub>
<b>in, before or after repetitions</b>	21	10.82	43	12.84
<b>inside repetitions</b>	11	5.67	29	8.66
<b>before repetitions</b>	5	2.58	8	2.39
<b>after repetitions</b>	5	2.58	5	1.49
<b>combined positions</b>	0	0.00	1	0.30
<b>at points of self-interruption</b>	22	11.34	37	11.04
UH/M that did not co-occur with repetitions or self-interruptions	151.00	77.84	255.00	76.12
<b>sum</b>	194	100.00	335	100.00

As can be seen in the above table, very similar proportions of *uh* and *uhm* – between 10% and 13% – occurred in the context of repetitions and self-interruptions. As concerns those instances of *uh* and *uhm* that co-occurred with repetitions, it was found that the largest proportion appeared inside repetitions, while they appeared significantly less frequently before or after such passages. The only difference that could be identified with regard to repetitions was that *uhm* was found to appear slightly more frequently inside repetitions than *uh*, while the latter appeared slightly occur slightly more often before and after repetitions; however, these differences were only very subtly pronounced and do thus not suffice to draw any conclusions on functional differences between the two variants.

Having identified the main trends with regard to *uh*'s and *uhm*'s positions in the context of repetitions, the next step was to consider the syntactic locations of those instances of UH/M that appeared inside, before, or after repetitions. The results obtained in this step of analysis show that the largest proportion of repetition-internal *uhs* (5 out of 11 instances) occurred between two instances of the same subject-verb cluster (e.g. *I'm uh I'm* in ICE-GB 2006:S1A-096:031), while *uhm* showed a tendency to occur between repeated conjunctions/connectives and especially *ands* (accounting for 12/29 and 8/29 instances respectively). Moreover, it was found that some instances of *uhm* (3/29) could be associated with the repetition of longer stretches of speech (e.g. *he rang me up when uhm yeah he rang me up when uh Julie was away*, cf. ICE-GB 2006: S1A-080:278), while no such example could be observed for *uh*. No specific trends could be identified with regard to those instances of *uh* and *uhm* occurring at repetition boundaries.

Following the study of *uh* and *uhm* in repetitions, the final step of analysis was to investigate those instances occurring at points of self-interruption. As was detailed in chapter 7.3.6, this was done by categorizing the relevant instances of *uh* and *uhm* according to their precise contexts, where *self-interruption* refers to a speaker's interruption of his/her current TCU, clause, or phrase with the purpose of either yielding his/her turn or starting a new TCU, clause or phrase that is not connected to the previous one in any obvious way, *reformulation* refers to a speaker's abortion of the current unit under construction for the sake of rephrasing his/her delivery of content, *correction* concerns a speaker's interruption of a current unit for the sake of correcting a (linguistic or content-related) mistake, and *word-internal interruption* refers to a speaker's mid-word self-interruption. It is shown in the extensive table of results provided in Appendix 6 that the majority of instances of *uh* and *uhm* occurring at points of interruption could be linked to the category of self-interruption, accounting for 14/22 *uhs* (i.e. 63.64%) and 29/37 *uhms* (i.e. 78.38%). Examples of the other categories – *correction*, *reformulation* and *word-internal interruption* – were also observed, even though they were found to be very rare for both variants. Finally, it was found that *uh* was the only variant that occurred in the context of word-internal interruptions (e.g. *Ju-uh July*).

In summary, the present chapter has presented the results obtained in the analysis of *uh* and *uhm* with regard to genre, turns, TCUs, clauses, phrases, pauses and collocates as well as repetitions and self-interruptions. The key findings outlined above will serve as a basis for the discussion of the research questions offered in the following chapter.

## 9. Discussion

Following the above description of the results, this chapter discusses the key findings obtained in the analysis of syntactic and functional differences between *uh* and *uhm*. The chapter will be divided into five main sections, where the first three sections correspond to the three research questions presented in section 7.1 of this thesis. Therefore, section 9.1 explores the extent to which the production of *uh* and *uhm* is determined by genre and speaking environment, section 9.2 investigates the syntactic differences between the two variants, and section 9.3 examines whether or not *uh* and *uhm* fulfill different functions in spoken discourse. As will be discussed below, the present study could not support the assumption that there are substantial functional differences between the two variants; thus, section 9.4 briefly investigates an alternative approach to the study of *uh* and *uhm*, according to which at least one of the two variants is an idiosyncratic phenomenon. Finally, section 9.5 briefly

summarizes the most pertinent findings obtained in the present study, not only with regard to the syntactic and functional differences between the two variants, but also concerning more general features of *uh* and *uhm* such as their determinants and their multi-functionality in spoken discourse.

### **9.1. Does genre affect the production of *uh* and *uhm*?**

The first section of this chapter examines the effects of genre on the production of *uh* and *uhm*. To this end, the results described in section 8.1 will be interpreted; moreover, potential implications for the study of *uh*'s and *uhm*'s functions will be discussed whenever this is deemed relevant.

One of the most noticeable findings presented in section 8.1 is that *uh* and *uhm* tend to occur considerably more often in telephone conversations than in private face-to-face conversations and broadcast interviews. This trend is not only illustrated by the observation that telephone conversations included the highest frequencies of *uh* and *uhm* per 1000 words, but also by the fact that around half of all studied instances of *uh* and *uhm* occurred in this genre (cf. section 8.1). Furthermore, this tendency was also generally observed in the comparison of the three individual speakers Andrew, Isobel, and Jenny, even though one exception to this trend could be observed with regard to Andrew's production of *uhm* (cf. the following paragraph for a discussion of possible explanations for and implications of this exception). The finding that *uh* and *uhm* generally occurred more frequently in telephone conversations than in any other genre is in line with Künzel (1997: 58, cited in de Leeuw 2007: 86), who, too, observed that speakers usually produce higher rates of UH/M when talking on the telephone. As has been stated in section 3.2, Künzel ascribes this trend to the increased need to *verbally* signal one's cognitive processes and turn-keeping intentions. Thus, following Künzel's argumentation, the findings presented in section 8.1 could be interpreted to suggest that both *uh* and *uhm* can be associated with the verbal signaling of cognitive processes for the sake of turn-keeping.

Even though the tendency of *uh* and *uhm* to occur more frequently in telephone conversations than in any other genre seemed to be rather strongly pronounced in the studied corpus, it is important to point out that it could not be observed for all speakers. It was shown in section 8.1 that one of the individually studied speakers – Andrew – produced fewer instances of *uhm* on the telephone than in face-to face-conversations; to my mind, this finding can be interpreted in two different ways. On the one hand, it is possible that Andrew faced an untypically high need to signal planning, turn-keeping, or any other process of turn-

management behavior associated with *uhm*, which caused him to produce these high rates of *uhm*. On the other hand, the unusually high frequency of *uhm* in Andrew's private face-to-face conversation could be linked to contextual determinants, such as his social relationship to his listener, the precise setting of the conversation or maybe even the consumption of alcohol (cf. sections 2.1 and 4.1.2). What both of these hypotheses have in common is that they suggest that *uh* and *uhm* are not solely determined by genre, but also by other factors such as topic or the intimacy of interpersonal relationships; this interpretation could also explain why another speaker – Isobel – produced considerably more *uhms* than *uhs* in her telephone conversation with Jenny, while the two variants occurred at more balanced rates in her telephone conversation with another interlocutor (cf. section 8.1). In summary, then, it can be concluded that even though there is a clearly observable trend of *uh* and *uhm* to occur more often in telephone calls than in face-to-face conversations, genre does not seem to be the only factor that influences the frequencies of *uh* and *uhm* in spoken discourse; this hypothesis also reinforces the conception of UH/M as highly situational phenomena, which has been proposed in section 3.4 and chapter 5.

With regard to differences between the two variants, the results presented in section 8.1 suggest that genre produces a stronger effect on *uhm* than on *uh*. This can be observed by the fact that the production rates of *uhm* per 1000 words within the same genre – and especially in telephone conversations and broadcast interviews – seemed to be more consistent than those of *uh* (cf. tables 13 and 14 in section 8.2). This finding is highly interesting for two main reasons. On the one hand, it suggests that genre seems to have an effect on *uhm*, but not – or at least to a much lesser extent – on *uh*. This, in turn, could mean that *uh* and *uhm* are either determined by different contextual factors, or fulfill different functions in spoken discourse, or both. On the other hand, the finding that genre seems to produce more noticeable effects on the production of *uhm* than on the frequency of *uh* could also explain why some scholars (e.g. Schachter et al. 1991: 365; Clark and Fox Tree 2002: 98) have found UH/M to be more characteristic of private contexts, while others (e.g. Tottie 2014: 6; Stenström and Svartvik 1994: 247-248) have argued them to occur more frequently in non-private settings (cf. section 4.1.2). Assuming that genre influences *uh* and *uhm* to different extents, and given that most of the respective publications do not differentiate between *uh* and *uhm*, it might well be that these contradicting results have been considerably influenced by the precise proportions of *uhs* and *uhms* present in the respective corpora.



Apart from the finding that *uhm* was considerably more influenced by genre than *uh*, it was also found that *uhm* noticeably dominated *uh* in private settings. It was already mentioned in section 7.2 of this thesis that the studied corpus contains much more instances of *uhm* than of *uh*, which is in line with de Leeuw's (2007: 110) finding that English speakers generally seem to prefer the vocalic-nasal variant over its purely vocalic counterpart. However, in the course of the present study, it was found that this dominance of *uhm* was only observable in private settings, whereas the non-private genre contained roughly equal rates of *uh* and *uhm* (cf. section 8.1). This is not to say, of course, that *uh* did not frequently appear in private genres, for it is in fact the case that the majority of both variants occurred in private settings. Still, it has been clearly shown in the present study that *uhm* is the dominant variant in private settings; this is also closely linked to the finding that *uhm* was much more influenced by genre than *uh* (cf. Figures 12 and 13 in section 8.1). Thus, it can be concluded that even though both variants occurred most frequently in private settings and especially in telephone conversations, the tendency to appear more often in private genres seems to be more deeply pronounced for *uhm* than for *uh*. The reasons for and implications of this finding will be discussed in the following paragraphs.

There are various possible explanations for the finding that the production of *uhm* seems to be more affected by the privacy of a given speaking setting than that of *uh*. In the following, I wish to propose three hypotheses which could potentially account for the different effects of genre on the production of *uh* and *uhm*. As can be seen below, all of these hypotheses are based on the assumption that genre and setting are not the only determinants of *uh* and *uhm*, and that their production is also influenced by other aspects, such as setting, context, or a speaker's personal tendency to utter *uh* and *uhm*. For reasons outlined below, none of these hypotheses can be entirely examined in the context of the present paper; however, they are considered to be relevant to the present discussion because they illustrate the vast range of possible explanations for the obtained results.

*Hypothesis 1: uh and uhm fulfill different functions in discourse, and the functions fulfilled by uhm are more typical of private settings than those fulfilled by uh.*

For instance, it is possible that *uhm* is more closely linked to turn-taking and turn-yielding, both of which are presumed to be considerably less relevant in an interview setting than in private conversations. In another line of argument, *uh* and *uhm* might indeed signal the initiation of short and long delays respectively (cf. Clark and Fox Tree 2002: 79), and speakers might try to avoid long delays in non-private speaking context. This hypothesis will

be partly examined in section 9.3, which investigates functional differences between the two variants; however, the detailed examination of this idea would also require a thorough study of the linguistic conventions associated with the respective genres, which cannot be sufficiently provided in the present thesis.

*Hypothesis 2: uhm is more publicly stigmatized than uh. This causes speakers to restrict their production of uhm – but not of uh – in non-private contexts.*

This hypothesis could be argued to account for two findings. Firstly, it could explain why *uhm* occurred considerably more often in private conversations, where speakers did not have to excessively monitor their speech, and least frequently in broadcast interviews, where speakers are presumably more careful about what they say. Secondly, it could also account for the observation that individual speakers produced different rates of *uh* and *uhm* when talking to different interlocutors, for it might be argued that a speaker's desire to monitor his/her speech could also be determined by the intimacy of an interpersonal relationship; however, this is a purely intuitive interpretation that would require further examination. One possible reason for which *uhm* could be more stigmatized than *uh* is the popular assumption that *uhm* indicate longer delays in speaking (cf. Clark and Fox Tree 2002: 86). In order to examine this hypothesis, it has to be clarified whether or not *uhm* may indeed be linked to excessive periods of pausing and hesitation; this will be examined in section 9.3.1.

*Hypothesis 3: uhm is determined by context, while uh is an idiosyncratic phenomenon.*

This hypothesis can be supported by the finding that there was no clear correlation between genre and the production of *uh*, while genre did seem to affect the frequency of *uhm*. Moreover, the view of *uh* as an idiosyncratic phenomenon could potentially explain why one of the individually analyzed speakers – Jenny – was found to produce not a single instance of *uh* in more than 2000 words of speech despite uttering average frequencies of *uhm* in both conversations. The possibility of *uh* (and *uhm*) being an idiosyncratic phenomenon will be further examined in section 9.4.

Having summarized and discussed the most pertinent trends observed in the genre-based analysis of UH/M, we are now in the position to answer the question of how genre affects the production of *uh* and *uhm* in spoken discourse. The present section has presented three key findings that are relevant to this enterprise: firstly, it has been shown that both *uh* and *uhm* appeared significantly more often in telephone conversations than in private face-to-face conversations or broadcast interviews; secondly, it was found that genre and setting seem to

produce a considerably stronger effect on *uhm* than on *uh*; and thirdly – and closely linked to the second finding – it was shown that *uhm* significantly dominates *uh* in private settings, but not in broadcast interviews. With regard to the defined research question, these findings allow two main conclusions. On the one hand, it appears that the production of *uh* and *uhm* seems to be influenced by the fact that a speaker is not able to see his/her interlocutor when he/she is talking on the telephone; this, in turn, suggests an effect of genre and setting on the production of the two variants and reinforces the conception of UH/M as verbal signals to one's listener(s) (cf. Künzel 1997: 58; cited in de Leeuw 2007: 86). On the other hand, however, it seems that apart from the observed effect of telephone conversations on the production rates of *uh* and *uhm*, *uhm* is the only variant that is influenced by genre. This finding is interesting insofar as it might suggest that *uh* and *uhm* are either differently determined or associated with different functions in discourse; a more detailed investigation of this hypothesis will be provided in section 9.3. In conclusion, then, the abovementioned research question can be answered by stating that setting seems to influence the production of *uh* and *uhm* considering that both variants occurred more frequently in telephone conversations than in the other studied genres, but that apart from this trend, *uhm* appears to be the only variant that is influenced by genre.

Following the analysis of the effects of genre on the production of *uh* and *uhm*, this chapter will now move on to examine the second research question, which concerns the syntactic differences between the two variants.

## 9.2. Are there syntactic differences between *uh* and *uhm*?

As stated above, the second section of this chapter examines the question of whether or not there are any observable syntactic differences between *uh* and *uhm* by discussing the most pertinent results presented in chapter 8. Note that this section is exclusively dedicated to the analysis of *syntactic* differences and that findings concerning *uh*'s and *uhm*'s positions in turns and their pausal environment will be addressed in the next section (cf. section 9.3). However, the TCU level *will* be considered, for – as has been argued earlier in this paper – it is considered to be a rough equivalent to the sentence which often comprises multiple syntactic units (cf. section 7.3.2). The syntactic findings presented in this section will serve as a basis for the discussion of *functional* differences provided in section 9.3.

With regard to the largest linguistic rank to be considered in this section – the TCU – it was shown that even though *uh* and *uhm* generally displayed very similar tendencies, there was at

least one relevant, yet rather subtle, difference between the two variants. Concerning the similarities, it was shown, for instance, that *uh* and *uhm* both preceded complete TCUs significantly more often than aborted ones; moreover, it was also found that both variants were more frequently associated with TRPs and PoAs than with non-TRP CPs. Most interestingly, however, the results presented in section 8.2 also show that both *uh* and *uhm* occurred significantly more often inside TCUs – i.e. after a speaker had lexically committed him-/herself to the respective syntactic construction – than at TRPs (cf. section 8.2); nevertheless, it is important to point out that the significance of this finding depends very much on the definition of the concepts *medial position* and *boundary position*. If we interpret the *boundary position* to comprise only those instances of UH/M that occurred directly at a TRP, CP, or PoA, then the findings presented in section 8.2 indeed indicate that both *uh* and *uhm* occurred most frequently in TCU-medial position. However, loosely following Hawkins' (1971: 285) aforementioned assertion that clause-introductory conjunctions, connectives, interjections and/or discourse markers (CCIDs) "[...] do not commit the speaker in any way to the type of construction which must follow them" and should thus not be counted as a commitment to the respective unit, I wish to propose that those instances of UH/M occurring after TCU-introductory CCIDs should also be counted towards the *boundary position*<sup>20</sup>. The application of this system of classification revealed a difference between *uh* and *uhm* that had not been observed before: while *uh* was found to appear slightly more frequently inside TCUs than at or near their boundaries, the opposite was found to be true for its vocalic-nasal counterpart, which displayed a considerable tendency to appear in boundary position. However, it must also be noted that this difference was not very strongly pronounced, for each variant was also found to occur rather frequently in the respective other position. Summing up, the results obtained in the analysis of *uh* and *uhm* on the TCU level seem to suggest that even though both variants showed very similar trends with regard to their preferred positions, there appears to be a slight tendency of *uh* to appear in TCU-medial positions, whereas *uhm* occurred more often at or near TCU boundaries.

The abovementioned observation that *uh* and *uhm* were generally found to occur in very similar positions was also made with regard to their location in clauses and phrases. Concerning the clause level, it was shown in section 8.3 that both variants most frequently occurred in prenuclear positions and least frequently in postnuclear positions. Moreover, with

---

<sup>20</sup>Note that the adoption of this approach on the TCU level was also important to the achievement of consistency in this study, for on the clause level, too, the *CCID # clause* position was counted as a *boundary position* (cf. section 8.3).

regard to the prenuclear instances of *uh* and *uhm*, it was found that both variants showed a strong tendency to appear immediately at clause boundaries on the one hand and, on the other hand, that they both tended to succeed rather than precede clause-introductory CCIDs. These findings are also in line with the results obtained in the analysis of *uh*'s and *uhm*'s most frequent collocates (cf. section 8.5), where it was shown that both variants most frequently followed the conjunction *and* and most often preceded the personal pronoun *I*; this observation further supports the finding that both *uh* and *uhm* frequently appear between clause-introductory CCIDs and the respective clause's subject (cf. Kjellmer2003: 180 for a similar interpretation). With regard to the phrase level, it was also shown that both *uh* and *uhm* appeared at phrase boundaries rather than inside them; the only exception to this trend was the PP, for it was found that both variants appeared more frequently inside them than at their boundaries. In addition, the finding that all PP-internal *uhs* and *uhms* occurred between the preposition and the NP serving as preposition complement reinforces the observation that the majority of all *uhs* and *uhms* occurred before NPs.

The above paragraph has already shown that there are many similarities between *uh*'s and *uhm*'s preferred clausal and phrasal positions; and indeed, there were only very few differences that could be observed to exist between the two variants on these two syntactic levels. In addition, especially on the phrasal level, the observed differences were only minor and are thus not deemed sufficiently significant to draw conclusions on systematic syntactic differences between the two variants. The only exception to this observation concerns the position of *uh* and *uhm* with regard to VPs, where it was found that *uh* significantly preferred the boundary position, whereas *uhm* was evenly split between the boundary and the VP-medial positions. With regard to the clause level, the only difference that is deemed relevant to the present study is that the tendency to occur in prenuclear position was slightly more pronounced for *uhm*, while *uh* considerably dominated *uhm* in postnuclear positions and inside constituents. The tendency of *uh* to occur slightly more frequently inside the clause than *uhm* is also illustrated by the fact that those instances of *uh* occurring inside repetitions were generally found to be associated with repetitions of subject-verb clusters (e.g. *I'm uh I'm*), while mid-repetition *uhms* were typically located between repeated conjunctions or connectives (e.g. *and uhm and*). These findings also appear to be in line with the abovementioned observation that *uh* seemed to slightly prefer TCU-medial positions over TCU-boundaries, while the reverse was found to be true for *uhm*.

In conclusion of this section, it can be said that only minor syntactic differences between *uh* and *uhm* could be observed. The above summary of the results obtained in the present study has shown that the two variants generally seem to occur in the same syntactic positions, not only considering the clause and phrase level, but also the TCU level. In fact, the only syntactic differences that are deemed relevant to the present study could be observed on the TCU level and the clause level, on both of which it was found that *uh* seemed to display a slight tendency to occur in medial positions, while *uhm* was the variant that dominated in boundary positions. The implications of these findings will be further discussed in the following section, which is dedicated to the question of whether or not there are *functional* differences between the two variants.

### **9.3. Are there functional differences between *uh* and *uhm*?**

The third section of this chapter examines whether or not there are any observable functional differences between *uh* and *uhm*. For this purpose, the results presented in chapter 8 will be discussed with regard to their implications for the study of *uh*'s and *uhm*'s functions; moreover, the obtained findings will be used to examine existing hypotheses concerning functional differences between the two variants (cf. chapter 6 for a summary of these hypotheses). In order to allow for a structured analysis of the abovementioned research question, the present section will be divided into five sections. The first section investigates Clark and Fox Tree's (2002: 86) famous proposition that *uh* and *uhm* serve as markers of minor and major delays respectively, while section 9.3.2 will analyze the relationship between *uh*, *uhm* and the signaling of cognitive processes. Following these two sections, section 9.3.3 discusses the roles of *uh* and *uhm* in the turn-management system, before section 9.3.4 will briefly examine their function in self-interruption. Finally, section 9.3.5 comprises a brief interim conclusion on the question of whether or not there are observable functional differences between *uh* and *uhm*.

#### **9.3.1. *Uh* and *uhm* as markers of delay**

The first question to be addressed in this chapter is whether or not *uh* and *uhm* do indeed indicate different lengths of delays in speech delivery, as has been proposed by Clark and Fox Tree (2002: 86). Generally speaking, the results obtained in the present study could not confirm Clark and Fox Tree's (2002: 86) hypothesis that *uh* and *uhm* can be associated with different lengths of delay. It was shown in section 8.5 that the majority of both variants – almost three quarters of all *uhs* and more than half of all *uhms* – did not co-occur with pauses,

while only significantly smaller proportions of both could be assigned to the *short delay* and *long delay* categories. Therefore, the data obtained in the present study does not seem to suggest a strong link between delays in speech delivery and the production of *uh* and *uhm*. This finding is in line with those obtained in O’Connell and Kowal’s (2005: 567) study of *uh* and *uhm*, in which they also showed that only a few instances of both variants were followed by silent pauses (cf. chapter 6).

Even though *uh* and *uhm* could not generally be linked to the “initiation of delays” (Clark and Fox Tree 2002: 84), it is interesting to note that *uhm* was found to co-occur slightly more often with delays than *uh*. As was shown in section 8.5, *uhm* dominated its vocalic counterpart in all categories except the *no delay*-group; this could be interpreted to suggest that *uhm* is the variant that is more closely linked to the signaling of delays in spoken discourse. However, as shown in the graphs presented in section 8.5, these differences between the two variants were only subtle and do thus not seem to confirm Clark and Fox Tree’s hypothesis that *uh* and *uhm* are typically associated with minor and major delays respectively (Clark and Fox Tree 2002: 86).

In summary, the present study provides only weak support of Clark and Fox Tree’s hypothesis. Even though it is true that *uhm* was slightly more often associated with silent pauses than *uh*, it was also shown that the vast majority of both variants did not co-occur with silent pauses; this observation further reinforces O’Connell and Kowal’s (2005: 570) finding that *uh* and *uhm* do not seem to primarily serve the function of signaling delay. At most, it could be argued that the uttering of *uh* and *uhm* constitutes a slight delay in speech delivery in itself; according to this view, then, *uh* and *uhm* are not ways to indicate one’s intention to delay speech production, but a means of buying time for completing the creative and cognitive processes involved in speech production. This hypothesis will be discussed in the following section, which will analyze the relationship between *uh*, *uhm*, and the signaling of cognitive processes.

### **9.3.2. *Uh* and *uhm* as signals of thinking and planning**

As stated above, the present section examines the view of *uh* and *uhm* as signals of thinking and planning. More precisely, it will be investigated whether it is indeed the case that *uhm* is the variant that is more closely linked to complex planning processes, such as conceptual or syntactic planning, while *uh* serves as a marker of “local lexical decision-making” (cf. Shriberg 1994: 154). To this end, the syntactic and non-syntactic evidence presented in

chapter 8 will be discussed with regard to its implications for the planning function of *uh* and *uhm*.

As has been stated earlier in this paper (cf. section 4.1.3 and chapter 6), it is frequently assumed that the planning function of *uh* and *uhm* can be investigated by analyzing their syntactic positions in spoken discourse; this approach is based on two main assumptions. The first assumption is that speech is not planned word-by-word, but in larger syntactic units (Boomer 1965: 155), and that syntactic boundaries – and especially clause boundaries – thus usually coincide with points of linguistic decision-making (cf. Hawkins 1971: 286-287). The second assumption underlying the syntactic approach towards studying *uh*'s and *uhm*'s functions is that speech is planned in a “two-stage, hierarchical encoding process” (Boomer 1965: 156), in which decisions concerning content, syntax and discourse organization are made *before* the precise lexical items to be used in a given context are chosen (Boomer 1965: 156). Following these two assumptions, it is often believed that those instances of *uh* or *uhm* that occur at syntactic boundaries represent conceptual, organizational, or syntactic planning, while unit-medial occurrences can usually be associated with lexical planning. This hypothesis has most frequently been applied on the clause level, but has also been argued to concern the phrase level (cf. section 4.1.3). Following many notable researchers in the field (e.g. Christenfeld 1994; Holmes 1988; Hawkins 1971; Kjellmer 2003) the present paper, too, will base the discussion of functional differences concerning *uh*'s and *uhm*'s planning functions on the idea that speakers generally use syntactic boundaries to plan larger upcoming stretches of speech. Having outlined the theoretical background of the upcoming discussion, this section will now move on to investigate whether or not there are any functional differences between *uh* and *uhm* with regard to their planning functions.

Generally speaking, the present study was able to provide a considerable body of evidence supporting the view that *uh* and *uhm* are indeed symptoms of a speaker's cognitive processes; however, the obtained results could only provide weak support for the view that the two variants typically signal different planning processes. This is largely due to the fact that both variants were found to display very similar tendencies with regard to their positions in turns, TCUs and clauses. For instance, it was not only found that very similar proportions of *uh* and *uhm* occurred in turn-initial positions, but also that both variants were found to appear most frequently at or near clause boundaries. Following the line of argument outlined above, these findings seem to suggest that both variants can be linked to the signaling of a speaker's cognitive processes associated with the onset of larger syntactic and conceptual units. This



hypothesis can be further supported by pointing out that both variants preceded complete TCUs considerably more often than aborted ones, the assumption being that speakers are less likely to abort a TCU when they carefully planned it beforehand. Finally, the fact that both variants were found to co-occur with repetitions rather frequently could be interpreted to further reinforce the idea that they are linked to a speaker's cognitive processes; this interpretation is predominantly based on Kjellmer's (2008: 46-48) finding that repetitions are very often associated with a speaker's desire to gain time for planning his/her upcoming stretch of speech. In addition, it was also found that both variants could be linked to the process of lexical retrieval. One very interesting observation that was made in this respect is that even though *uh* and *uhm* generally displayed a strong tendency to occur at phrase boundaries, both variants occurred considerably more often inside PPs than at their boundaries. Considering that all PP-internal instances of UH/M occurred before the NP acting as a preposition complement, this observation does not only reinforce the observed tendency of both variants to occur most frequently before NPs, but also confirms Maclay and Osgood's (1959: 32-33) aforementioned finding that UH/M tend to appear before content rather than function words. This, in turn, could be argued to suggest that both variants can be linked to lexical decision-making.

The finding that there are only few differences between *uh* and *uhm* with regard to their planning functions can be further supported by pointing out that *uhm* was frequently found to serve functions that are typically assumed to be fulfilled by *uh*, and vice versa. One example for such a case is illustrated in example (16) below. In this passage, two speakers play a game in which they have to take turns in saying words that start with the letter *D*; thus, the game involves a great extent of lexical retrieval, which is often thought to be typically associated with the production of *uh* by those who assume functional differences between the two variants (cf., e.g. Shriberg 1994: 154). However, as illustrated in example (16) below, speakers were found to frequently utter *uhm*, but not a single instance of *uh*, in the course of this passage:

- (16) S2: Right let's see how many words we can think of beginning with D <,,> // I'll start / destitute <,,> //  
 S1: drug <,,> //  
 S2: demon //  
 S1: drink <,,>  
 S2: **uhm**<,,> dyslexic  
 [...]  
 S2: donkey // it's your turn <,,> //  
 S1: dick <,,> //  
 S2: devil <,,> //

S1: **uhm**<,> dick //  
 [...]  
 S2: If you stick a de in fr- in front of something // like with dehumanize <,> //  
 S1: OK / I'll try that <,> / OK <,> / **uhm**<,> dehumanize <,> //  
 S2: [your imagination is] / I love it when you work your brain // **uhm**<,> drunk <,> //  
 (ICE-GB 2006:S1A-085: 260-295)

The example above contains several instances of *uhm* which seem to be linked to lexical retrieval; this is especially observable when considering the two instances produced by S1. In both cases, S1 repeats a word that has already been previously said, namely *dick* (which S1 said in his/her last turn) and *dehumanize* (which was proposed by S2 in the immediately preceding turn). The fact that S1 produces these previously uttered items could be interpreted to suggest that he/she cannot think of any other word that starts with *D*; this assumption can be further supported by the fact that the respective words are both preceded by long silent pauses in which S1 is seemingly thinking about potential next words. Supposing that S1 is still trying to successfully participate in this game (which seems rather likely because he/she has not previously produced any signs of not wanting to play), it can be assumed that he/she is engaged in periods of (seemingly unsuccessful) lexical retrieval directly before uttering *dick* or *dehumanize*; thus, the above passage seems to suggest a link between lexical access and the uttering of *uhm*. Apart from illustrating the relationship between lexical retrieval and the vocalic-nasal variant *uhm*, it is also interesting to note that both speakers S1 and S2 tend to utter *uhm* in the vicinity of (short and long) silent pauses; this seems to support Maclay and Osgood's (1959: 42) hypothesis that speakers use UH/M to react to their own ongoing silence. In that sense, the above passage could also be argued to support the hypothesis that *uhm* (or UH/M in general) serves as both an indication of a speaker's cognitive processes and a signal to the listener(s), thus reinforcing the multidimensional conception of UH/M proposed earlier in this paper.

It was already addressed above that it was not only found that *uhm* is able to indicate processes that are often assumed to be linked to the production of *uh*, but, similarly, also that *uh* often performs functions that are typically associated with *uhm*. For instance, if we follow the assumption that *uhm* is indeed the variant that is responsible for the signaling of conceptual and syntactic thinking and planning, we should expect *uhm* to be the only variant that occurs in turn-initial positions; however, it has been shown in section 8.2 of this paper that very similar proportions of *uh* and *uhm* were found to occur in this position. Similarly, if we assume that speakers only utter *uhm* when engaged in content-related thinking, it is surprising to find that *uh* was also found to rather frequently precede answers to questions.

Finally, the aforementioned finding that a large proportion of the investigated *uhs* was also observed to occur at or near the boundaries of TCUs, clauses and phrases seems to contradict the hypothesis that *uh* is typically associated with “local lexical decision-making” (Shriberg 1994: 154).

Concerning the ability of *uh* to signal cognitive processes that are usually associated with *uhm*, there was one particularly interesting passage in the analyzed corpus. In this passage, S1 (a woman called Isobel) asks S2 (a man called David) where he thinks someone called *Yibin* could come from. Apparently not knowing the correct answer, S2 replies by making a joke:

- (17) S1: Yibin //  
[...]  
S2: Very unusual name <,> //  
S1: Uh guess where he's from with a name like that //  
S2: **U:h**Yorkshire  
S1: <laughs> oh a laugh a minute here <,,>  
(ICE-GB 2006: S1A-092: 245-247)

What makes the above passage so interesting is that the speaker seems to deliberately utter *uh* to signal thinking, even though he does not in fact seem to seriously consider answering the question. This does not become obvious in the above transcript, but it is clearly audible in the audio recording, where S2 produces a seemingly deliberately exaggerated instance of *uh* before giving his answer, *Yorkshire*. The assumption that S2's answer was meant as a joke is not only based on my own interpretation of the passage, but can also be supported by drawing attention to the fact that S1, too, understands it as such; this becomes obvious by considering her reaction, which consists of laughter and the phrase *oh a laugh a minute here*. Due to the fact that S2 immediately commits himself to making this joke by beginning to utter the exaggerated instance of *uh* right after being handed the turn, it can be assumed that the joke is intended as a means of communicating that S2 has no idea where the name *Yibin* could come from. Following this line of argument, it seems that the exaggerated instance of *uh* highlighted in the above passage is not really a signal of S2's cognitive processes, but rather a way of pretending to consider a question when in fact he is not. This is interesting insofar as it suggests that speakers are not only aware of their uttering of *uh*, but also that they seem to associate this variant with the signaling of information retrieval and conceptual thinking, both of which are generally associated with *uhm* rather than *uh* by those assuming functional differences between the two variants (cf. Shriberg 1994: 154).

Despite the findings outlined above, it would be wrong to conclude that no functional differences between *uh* and *uhm* could be observed *at all*; instead, the case is rather that the differences that were found to exist were significantly less pronounced than is suggested in existing literature. For instance, as was already addressed in section 9.2, it was found that *uh* and *uhm* seemed to display slight preferences of medial and boundary positions respectively; however, the statistic differences observed in this respect are deemed to be too subtly pronounced to follow Shriberg (1994: 154) in concluding that *uhm* typically appears at or near syntactic boundaries and is thus more closely linked to conceptual and syntactic planning whereas *uh* usually occurs more often in medial positions and is thus typically linked to lexical retrieval. Similarly, even though it was shown that the largest proportions of turn-initial instances of *uh* and *uhm* occurred in answers to questions, it was found that this trend was more pronounced for *uhm* than for *uh*, thus supporting Tottie's (2016: 113) finding that *uhm* is the variant which appears in answers to questions more often. Finally, the abovementioned finding that *uhm* could be slightly more frequently associated with silent pauses seems to suggest that speakers tend to utter *uhm* whenever they need more time than usual to plan a specific stretch of speech, while *uh* is more typically associated with minor decisions that are easier to make. This hypothesis could also explain why the speakers in example (16) above uttered *uhm* instead of *uh*, for especially S1 seemed to find it very difficult at times to search his/her mental lexicon for new words starting with the letter *D* (see above). In this sense, it is possible that *uhm* is associated with more "difficult" planning processes than *uh*; however, as has already been mentioned earlier in this paper, there are hardly any means of objectively measuring the difficulty and/or cognitive effort associated with a given situation.

In summary, then, it can be concluded that the present study seems to provide only weak evidence in support of Shriberg's (1994: 154) hypothesis that *uh* and *uhm* can be linked to different planning processes. As outlined above, it has not only been found that *uh* and *uhm* appear in very similar syntactic environments, but also that both variants are able to signal the "larger" projects of conceptual or syntactic planning on the one hand and lexical retrieval on the other hand. Consequently, it seems that there are hardly any systematic differences between *uh* and *uhm* with regard to their signaling of different planning processes. However, it is still interesting to observe that *uh* was slightly more frequently associated with unit-medial positions while *uhm* was found to appear more often at unit boundaries, in answers to questions, and in the vicinity of silent pauses, for these tendencies seem to be in line with

existing hypotheses on the presumed functional differences between *uh* and *uhm*. In conclusion, I therefore wish to argue that *uh* and *uhm* do not seem to systematically fulfill different functions with regard to the signaling of thinking and planning, but that there are still some slight differences between the two variants, even if these differences are considerably less pronounced than is suggested in existing publications on the topic.

### 9.3.3. *Uh* and *uhm* as turn-management devices

As was stated in the introductory paragraph, the third section of this chapter examines the question of whether or not *uh* and *uhm* fulfill different functions within the turn-management system; this includes not only turn-keeping, but also turn-taking, turn-accepting, and turn-yielding. To my knowledge, the question of whether or not *uh* and *uhm* fulfill different functions in the turn-taking system has never been explicitly studied before, and the present section is thus hoped to provide some interesting insights into the potential functional differences between *uh* and *uhm*.

With regard to the turn-keeping function of *uh* and *uhm*, the results obtained in the present study seem to suggest that both variants act as turn-keeping devices rather frequently, and that only minor differences exist between the two variants in this respect. This conclusion is based on three main findings. Firstly, both variants occurred considerably more often in telephone calls than in face-to-face conversations (cf. sections 8.1 and 9.1). Following Künzels (1997: 58, cited in de Leeuw 2007:86) line of argument, this tendency can be ascribed to the fact that speakers have no other choice than to *verbally* signal their turn-keeping intentions on the telephone which can thus be regarded as evidence in support of a turn-keeping function of *uh* and *uhm*. Secondly, it was shown in chapter 8 that *uh* and *uhm* display a strong tendency to occur at or near the boundaries of TCUs and clauses (cf. also section 9.2). Assuming that syntactic boundaries and especially those coinciding with TRPs particularly lend themselves to turn transition because they mark the completion of a previous unit, the frequent occurrence of UH/M at or near such boundaries could suggest that speakers produce them in order to signal their intention to continue their turn. This interpretation can further be supported by the finding that both variants occurred considerably more often at TRPs, which by definition render turn-transition a relevant option, than at non-TRP CPs and PoAs, where speakers usually keep their right to the floor; however, the significance of this finding could be relativized by pointing out that CPs were generally less frequent in the corpus than TRPs. Finally, it was found that speakers sometimes utter *uh* and especially *uhm* in periods of pausing in order to signal their intention to say something in the near future (cf. section 8.5

and example (16) in section 9.3.2); this finding seems to support Maclay and Osgood's (1959: 41-42) hypothesis that speakers may use UH/M to fill his/her own ongoing silence for the sake of turn-keeping. Concerning functional differences between *uh* and *uhm*, it could be argued that *uhm*'s tendency to occur slightly more frequently at or near clause boundaries and TRPs than its purely vocalic counterpart could be interpreted to suggest that *uhm* is the variant that is more closely linked to turn-keeping; however, these differences are too subtle to conclude a systematic functional difference between *uh* and *uhm*.

Even though the above paragraph has shown that turn-keeping seems to be one function that is often fulfilled by *uh* and *uhm*, the present study could not support the view that it is their *main function*; instead, I wish to argue that it is a function that frequently complements the main function of signaling a speaker's cognitive processes (cf. also Clark and Fox Tree 2002: 77). This hypothesis is largely based on the finding that both variants frequently occurred inside TCUs or clauses, where the active signaling of turn-keeping does not seem necessary. This interpretation is also in line with Tottie's (2015: 397) assertion that most instances of turn-medial *uhs* and *uhms* appear in contexts where there are no observable attempts at turn-taking by other speakers, which leads her to the conclusion that "UH[/]M is not normally used to deliberately prevent speakers from taking over the turn". Closely linked to this argument, the fact that *uh* and *uhm* often occur at or near syntactic or TCU boundaries does not necessarily suggest that a speaker utters them for the purpose of turn-keeping; instead, it is also possible that they are mere symptoms of the cognitive processes in which the speaker is currently engaged. Finally, the finding that neither *uh* nor *uhm* are typically associated with delays in speech delivery does not confirm the hypothesis that *uh* and *uhm* are first and foremost used to signal one's turn-keeping intentions in periods of excessive pausing (cf. Maclay and Osgood 1959: 41-42). For these reasons – and given the vast body of evidence supporting the view that *uh* and *uhm* can be linked to a speaker's thinking and planning (cf. section 9.3.2) – I wish to propose that *uh* and *uhm* predominantly act as representations of a speaker's cognitive processes, but that listeners are able to interpret the indicated planning processes as an announcement of turn-continuation; in that sense, the turn-keeping function of *uh* and *uhm* could be seen as a side-effect of their planning function. This view of UH/M has also been proposed by Tottie (2014: 25), who has suggested that "uh and um [sic] originated in situations of cognitive load, where speakers needed time to pause to think and plan, but that they [...] have now also acquired pragmatic meanings".

Summing up the discussion on the turn-keeping function of *uh* and *uhm*, it can be concluded that both variants can be linked to this function, even though it is possible that the turn-keeping function is only a side-effect of UH/M's main function of signaling cognitive processes. The two paragraphs above showed that there are indeed good reasons to assume that UH/M act as signals of turn-keeping; however, it has also been shown that due to *uh*'s and *uhm*'s frequent occurrence in environments that do not render turn-transition a relevant option, it seems likely that turn-keeping is not *the one main* function of UH/M, as has been proposed, for instance, by Maclay and Osgood (1959: 41-42). In that sense, I wish to reinforce the multidimensional conception of UH/M that has already been proposed earlier in section 3.4 and chapter 5 of this paper, and according to which *uh* and *uhm* act as symptoms *and* signals at the same time. Concerning functional differences between *uh* and *uhm*, it has been argued that *uhm* could be perceived as the variant that is slightly more closely linked to the signaling of turn-keeping because it shows a stronger tendency to occur at clause boundaries and TRPs (cf. section 8.2); however, it has equally been shown that these differences are too subtly pronounced to conclude a systematic functional difference between *uh* and *uhm*. Having discussed the turn-keeping function of *uh* and *uhm*, the present section will move on to examine the functions fulfilled by turn-initial *uhs* and *uhms*.

As stated in section 8.2, the majority of turn-initial instances of *uh* and *uhm* could not be linked to a speaker's active turn-taking, but to turn-accepting in answers to questions<sup>21</sup>; thus, the present study contradicts the popular hypothesis that turn-taking is one of the major functions of UH/M (cf. Kjellmer 2003: 167; Shriberg 2001: 156). Regarding the question of functional differences between *uh* and *uhm*, it was found that *uhm* appeared slightly more often in the context of turn-accepting than *uh* (cf. section 8.2); this reinforces the finding that *uhm* is the variant that is slightly more closely linked to the planning of larger conceptual and/or syntactic units (cf. section 9.3.2). In order to examine whether or not answer-initial *uhs* and *uhms* could be linked to different levels of cognitive load, it was investigated whether one of the two variants could be more closely linked to yes/no questions or open-answer questions; however, no trends were discovered to exist in this respect. Therefore, it can be concluded that the only difference found to exist between *uh* and *uhm* with regard to their turn-accepting function is that *uhm* seems to be slightly more likely to appear in this context than *uh*; however, it should also be pointed out that the proportion of answer-initial *uhms* still

---

<sup>21</sup> Two examples for instances of *uh* that were categorized as turn-accepting devices can be found in section 4.2.2 of this paper.

only accounted for a very minor proportion of all instances of *uhm* in the corpus. In summary, this paragraph not only suggested that neither turn-taking nor turn-accepting seem to be major functions of *uh* and *uhm*, but also that there are no considerable functional differences between turn-initial *uhs* and *uhms*.

Having discussed the functions fulfilled by turn-medial and turn-initial *uhs* and *uhms*, the final part of this section aims at examining the question of whether or not there are any observable differences between the turn-final instances of both variants. As was shown in section 8.2, only very small proportions of both variants were found to occur in turn-final position; this suggests that turn-yielding, too, does not constitute a major function of *uh* and *uhm*. Nevertheless, it is interesting to note that *uhm* appeared almost twice as often in turn-final positions than *uh*. As can be seen in the results presented in section 8.2, this difference seemed to be mostly due to the fact that *uhm* appeared more often after a completed, turn-final TCU; this could either suggest that *uhm* more frequently acts as a device of voluntary turn-yielding or that listeners regard *uhms* at TRPs as slightly more convenient points for speaker change than *uhs* in the same position. However, one should be careful not to attach too much importance to this latter finding, for only very small proportions of both variants were found to appear in the turn-final positions, and the analysis of their distribution across different contexts is thus not thought to yield any significant results. Moreover, the fact that both *uh* and *uhm* only seldom coincided with PoAs seems to suggest that listeners do not generally perceive *uh* and *uhm* as convenient points for interruption; this, in turn, reinforces the view that both variants are first and foremost signals of planning and turn-keeping. Summing up, the analysis of turn-final instances of *uh* and *uhm* allows two main conclusions. On the one hand, the small proportions of *uh* and *uhm* occurring in this position suggests that turn-yielding does not constitute a major function of *uh* and *uhm*. On the other hand, it was shown that *uhm* seems to be the variant that is more likely to occur in the context of speaker transition and especially of turn-yielding; however, the studied corpus includes too few examples of turn-finals *uhs* and *uhms* to draw any valid conclusions based on this phenomenon.

Summing up the present section, it can be concluded that there are hardly any functional differences between *uh* and *uhm* with regard to the turn-management system. As stated above, the vast majority of both variants appeared in turn-medial positions, while turn-initial and turn-final instances only seldom occurred in the studied corpus. Along with the findings presented in the above paragraphs, this leads me to hypothesize that *uh* and *uhm*



predominantly act as means of turn-keeping rather than turn-taking, turn-accepting, or turn-yielding; however, as has been argued in the first part of this section, even the function of turn-keeping does not seem to constitute the main function of UH/M. Concerning functional differences between *uh* and *uhm*, it was found that *uhm* is the variant that slightly more closely associated with the signaling of turn-keeping, turn-accepting, and turn-yielding, which could be linked to the finding that *uhm* also seems to be the variant that indicates the planning of larger conceptual and syntactic units slightly more often than *uh*. Nevertheless, as has been pointed out on numerous occasions, the differences observed in this respect are too subtle to support the hypothesis that *uh* and *uhm* systematically fulfill different functions. Having discussed the functions of *uh* and *uhm* with regard to the turn-management system, the last section of this chapter will examine whether or not *uh* and *uhm* play different roles in the context of self-interruption.

#### 9.3.4. *Uh* and *uhm* as editing phrases

The final section of this chapter is dedicated to *uh*'s and *uhm*'s role as editing phrases in self-interruptions. As was shown in section 8.6, very similar proportions of *uh* and *uhm* – roughly 11% of each – could be linked to mid-turn points of interruptions; moreover, the majority of those instances where *uh* and *uhm* were found to act as editing phrases occurred in the context of self-interruption rather than self-correction and/or reformulation (cf. section 8.6 for examples of the different categories). Concerning differences between the two variants, it was not only found that *uhm* appeared slightly more often inside self-interruptions, but also that *uh* was the only variant that occurred in the context of mid-word interruptions. Assuming that self-interruption, as defined in section 7.3.6 of this paper, usually involves the abortion of a current TCU or clause for the sake of starting a new one that is not linked to the previous in any obvious way, this finding could be argued to offer some, albeit weak, support for Shriberg's hypothesis that *uhm* is more typically associated with the planning of larger units, while *uh* usually indicates lexical decisions (cf. Shriberg 1994:154). Apart from this finding, no other significant differences or trends could be identified with regard to their functions in self-interruptions, especially because the differences that *were* identified (cf. section 8.6) were too subtly pronounced to be interpreted as substantial evidence for functional differences between the two variants.

In conclusion, neither *uh* nor *uhm* was found to frequently act as editing phrases in self-interruption, even though it is interesting to note that they were more frequently associated with this function than with turn-taking or turn-yielding. Moreover, hardly any differences

were found to exist between the two variants with regard to self-interruption. Having discussed the most pertinent functions fulfilled by *uh* and *uhm* as observed in the studied corpus, the following section will provide a brief interim conclusion on whether or not *uh* and *uhm* can be said to serve different functions in spoken discourse.

### **9.3.5. Interim conclusion: Are there any functional differences between *uh* and *uhm*?**

As has been stated on numerous occasions throughout this section, the present study could provide only weak evidence for the hypothesis that *uh* and *uhm* fulfill different functions in spoken discourse. Even when the obtained results could be interpreted to suggest some functional differences, these were usually too subtly pronounced to consider them as conclusive evidence for systematic differences between *uh* and *uhm*. Interestingly, however, many of the differences that *were* found to exist are in line with the hypotheses on functional differences presented in existing literature; for instance, the finding that *uhm* was slightly more likely to appear at syntactic boundaries than *uh* seems to support to at least some extent Shriberg's (1994: 154) hypothesis that *uhm* typically marks the planning of larger syntactic and/or conceptual units, while the fact that *uh* dominated *uhm* in unit-medial positions could suggest that it is indeed slightly more closely linked to local lexical retrieval. Moreover, the finding that *uhm* showed a slight tendency to occur at or near syntactic boundaries could be argued to support Shriberg's (1994: 154) hypothesis that *uhm* is more closely linked to conceptual and syntactic planning, while *uh* can be more closely associated with local decisions. However, this is not to say that the present study could be found to really support either of these hypotheses, for – as has been mentioned before – the observed differences between *uh* and *uhm* were far too subtly pronounced to be interpreted as conclusive evidence in support of the view that *uh* and *uhm* signal different lengths of delays, planning processes, or turn-management behavior. Thus, in summary, I wish to argue that *uh* and *uhm* do not systematically seem to fulfill different functions with regard to their capacity to signal planning, turn-management behavior and/or self-interruption, but that there are still some slight differences between the two variants; however, these seem to be significantly less pronounced than is suggested in existing literature on the topic.

The finding that there are no – or hardly any – systematic functional differences between *uh* and *uhm* raises the question of how speakers decide which variant to use at specific points in their everyday conversations. As suggested in section 9.1, one possible explanation could be that one or both of the two variants are idiosyncratic phenomena. According to this view, the production of *uh* and/or *uhm* is not determined by contextual factors, such as topic, genre, or

the need of signaling turn-keeping, but are rather speaker-specific phenomena. Even though the present study cannot provide extensive research on this hypothesis, this approach shall still be briefly discussed in the upcoming section.

#### **9.4. Are *uh* and/or *uhm* idiosyncratic phenomena?**

Given that hardly any functional differences could be found to exist between *uh* and *uhm*, the question arises whether a speaker's production and choice of the two variants is in fact an idiosyncratic phenomenon. For the purpose of briefly examining this approach, the present section comprises not only a summary of existing publications that assume the production of UH/M to be idiosyncratic, but also a brief analysis of the corpus studied in the present thesis with regard to this hypothesis.

Only few scholars have explicitly proposed that *uh* and *uhm* are idiosyncratic phenomena up to this point. One of those is Elizabeth Shriberg, who, in her article "To err is human: ecology and acoustics of speech disfluencies" (2001) observed high fluctuations across different speakers' production rates of the two variants; this led her to argue that the production of *uh* and *uhm* is to a large extent dependent on the individual speaker's personal tendencies and preferences (Shriberg 2001: 157). Similarly, Tottie (2014: 23) proposes that "[s]peaker personality and idiosyncratic preferences as well as extra-linguistic setting all play a part both when quantity and types of pragmatic markers [such as UH/M] are considered". Finally, Friginal (2009: 324) offers an interesting perspective on the discussion by asserting that "some filled pauses may have already become part of many speakers' idiosyncratic speech mannerisms", thus suggesting that *uh* and *uhm* are idiosyncratic phenomena that many speakers now use in very similar ways, such as in the context of turn-accepting or as editing phrases in self-interruption.

As previously addressed, some findings obtained in the present study could be argued to suggest that the production of *uh* and *uhm* is at least partly determined by a speaker's individual tendencies. Up to this point, this hypothesis has been restricted to *uh* for two main reasons. Firstly, as was discussed in section 9.1, genre was found to affect the production of *uhm*, but not of *uh*, which could suggest that *uh* is more influenced by a speaker's idiosyncratic mannerisms than *uhm*. Secondly, the analysis of individual speakers showed that one speaker – Jenny – did not produce a single instance of *uh* in more than 2000 words of speech despite uttering average frequencies of *uhm* in both conversations; this, again, could suggest that the production of *uh* depends to a large extent on the individual speaker him- or

herself. However, a closer consideration of the obtained data reveals that *uhm*, too, might be at least partly idiosyncratic. This assumption is largely based on the finding that the production rates of *both* variants – not only *uh* – varied significantly across the different corpus texts within the genre of private face-to-face conversations. Assuming that speakers do not usually strictly monitor their speech when engaged in private face-to-face conversations with friends, this finding could be interpreted to suggest that the production of *uh* and *uhm* is a highly speaker-specific phenomenon. However, given that both variants were found to increase in telephone conversations, and considering the finding that *uhm* was shown to display clear tendencies with regard to its frequency in telephone conversations and broadcast interviews, it appears that idiosyncratic mannerisms do not seem to be the only factor influencing their production rates. Instead, it is possible that the linguistic conventions associated with a different genre, such as the increased need to verbally signal one's cognitive processes for the sake of turn-keeping in telephone conversations, are also able to influence the frequencies of *uhm*. This influence of other, genre-specific factors seems to be significantly less pronounced for *uh*, which did not display any noticeable correlations between genre and its frequencies in spoken discourse apart from its augmented production rates in telephone conversations.

So as to further investigate the view of *uh* and *uhm* as idiosyncratic mannerisms, a brief analysis of the rates of *uh* and *uhm* produced by the twenty-six different speakers<sup>22</sup> in the studied corpus was conducted. For this purpose, the numbers of *uh* and *uhm* produced by each speaker were counted and normalized so as to represent the respective speaker's production rates of *uh* and *uhm* per 1000 words of speech. In order not to distort the obtained results, each genre was analyzed separately. The results obtained in this step of the analysis are visually presented in Appendix 8. Generally speaking, it was found that even within genres, the production rates of *uh* and *uhm* per 1000 words varied considerably across speakers. For instance, within the genre of private face-to-face conversations, it was found that some speakers produced no instances of *uh* at all, while others uttered it 19 times per 1000 words; similarly, some speakers produced only roughly 3 *uhms* per 1000 words, while others produced as many as 35 instances per 1000 words. Similar findings could also be obtained in the other two genres, even if the observed differences were not as deeply pronounced as in the private face-to-face conversations.

---

<sup>22</sup> Note that the studied corpus contains fifteen texts with two speakers each; however, as was stated in section 8.1, 2 speakers were found to participate in two corpus texts, and one speaker even took part in three conversations. Therefore, the studied corpus does not include thirty, but twenty-six different speakers.

In addition to these general observations, it was also found that speakers partly displayed very different tendencies with regard to the choice of the two variants. For instance, while some were found to produce rather balanced rates of *uh* and *uhm*, others showed very high discrepancies between their production rates of the two variants; in addition, some speakers – such as Jenny (cf. the discussion in the above paragraph) – were found to produce only one of the two variants. Moreover, it was also shown that some speakers produced considerably higher rates of *uhm* than of *uh* and vice versa. Finally, it was found that the three speakers that participated in more than one conversation largely displayed the same preferences with regard to their production of *uh* and *uhm* across the different corpus texts; for instance, it was shown that Jenny did not produce a single instance of *uh* in neither of her two conversations, and that both Andrew and Isobel produced more instances of *uhm* than of *uh* in their respective texts. All of these findings seem to suggest that *uh* and *uhm* are at least to some extent influenced by a speaker's individual preferences; this seems to concern the general production of UH/M on the one hand and the choice of variants on the other hand.

The analysis of individual speaker's production rates also yielded some interesting results with regard to differences between the two variants. For instance, it was found that the largest proportion of speakers (11 out of 26) produced more *uhms* than *uhs*, while only 6 of the 26 speakers uttered *uh* more often than *uhm*. Moreover, the vast majority of speakers produced at least some instances of *uhm*, while only one speaker was found to *exclusively* utter *uh*. These findings are in line with de Leeuw's (2007: 110) claim that native speakers of English generally seem to prefer *uhm* over *uh*. Plausible explanations for this are manifold: for instance, it is possible that speakers of English prefer *uhm* over *uh* because they find it more convenient to produce the phonetically longer variant so as to buy more time for completing their planning processes (cf. also O'Connell and Kowal 2005: 563). However, this is merely one of many possible explanations for the above findings, and more research is needed to determine the extent to which *uh* and *uhm* can be argued to be idiosyncratic phenomena.

One challenge to the perception of UH/M as idiosyncratic phenomena could be that *uh* and *uhm* were found to display certain syntactic tendencies, which could suggest that their production underlies some kind of system that is not influenced by a speaker's individual preferences. However, there are in fact several possible explanations for this finding. Firstly, it has already been suggested in section 9.1 that it is possible that *uh* and *uhm* are only partly idiosyncratic, and that their production is still also determined by other factors such as genre, setting, or the specific speaking situation; moreover, it is also conceivable that the production

of *uh* is more strongly influenced by idiosyncratic mannerisms than that of *uhm*. Yet another approach is based on Friginal's (2009: 324) assertion that "some filled pauses may have already become part of many speakers' idiosyncratic speech mannerisms". According to this view, it is possible that *uh* and *uhm* are idiosyncratic phenomena, but that speakers, over time, acquire communicative experience which leads them to develop a certain system of using *uh* and *uhm* for their purposes. For instance, it might be that a speaker's tendency to utter UH/M at or near clause boundaries or TRPs is caused by his/her experience of needing to signal the continuation of his/her turn. Assuming that each speaker develops his/her own system for producing UH/M based on his/her experience in spoken communication, it seems very likely that many speakers would develop similar systems of using *uh* and *uhm*; this, in turn, would explain why *uh* and *uhm* show strongly pronounced tendencies with regard to their syntactic location in spoken discourse. The importance of individual experience for the use of *uh* and *uhm* has also been highlighted by Maclay and Osgood (1959: 41-42), who claim that a speaker decides to utter UH/M in periods of excessive pausing because "*he has learned that unfilled intervals of sufficient length are the points at which he has usually lost this control [of the turn] – someone else has leapt into his gap [my emphasis]*". In summary of this approach, it seems possible that *uh* and *uhm* are indeed at least partly idiosyncratic, and that speakers have developed similar systems of using them in spoken discourse in the course of time.

Another explanation for the considerable differences in UH/M production rates across speakers is that they are determined by sociolinguistic factors such as a speaker's sex, age, and/or education. This hypothesis has been supported, for instance, by Tottie (2011), who has not only found that men seem to use more instances of UH/M than women (Tottie 2011: 180), but who has also shown that older speakers generally preferred to utter *uhm*, while younger speakers preferred *uh* (Tottie 2011: 186-187). Moreover, Tottie (2011: 192) also suggests that "speakers from higher socio-economic and more well-educated strata use more fillers than speakers from lower strata". Indeed, with regard to speaker sex, the analysis of individual speakers' production rates of *uh* and *uhm* was able to support some of Tottie's findings. For instance, it was found that men produced considerably more instances of *uh* and *uhm* per 1000 words than women<sup>23</sup>. In addition, it was found that of the 11 speakers who produced more *uhms* than *uhs*, 10 were female, while 5 out of the 6 speakers who produced higher rates of *uh* were male; this observation seems to support Tottie's (2011: 183) finding that "[...] women

---

<sup>23</sup>A calculation of *uhs* and *uhms* per 1000 words per gender showed that men produced 10.50 *uhs* and 12.97 *uhms* per 1000 words, while women only uttered 3.95 *uhs* and 9.21 *uhms* per 1000 words.

use higher proportions of nasalized fillers than men”. Unfortunately, the other sociolinguistic factors investigated by Tottie (2011) – age and education – could not be examined in the present study because ICE-GB does not include sufficiently precise information on these two variables. Note that the conception of UH/M as sociolinguistic markers does not necessarily contradict their perception as idiosyncratic phenomena, for – if we assume that a speaker’s use of UH/M is indeed shaped by his/her communicative and interactional experiences (see above) it seems likely that this experience is also shaped by sociolinguistic factors; for instance, it seems likely that “speakers from higher socio-economic and more well-educated strata” (Tottie 2011: 192) often interact with other speakers of similar socio-economic classes, and would thus gain similar linguistic experiences.

In summary, the present section has provided some support for the hypothesis that the production of *uh* and *uhm* are at least partly determined by a speaker’s individual tendency to utter *uh* and *uhm*, even if it remains unknown whether this individual tendency can be ascribed to idiosyncratic mannerisms or sociolinguistic factors. Either way, the finding that the uttering of *uhm* and especially *uh* seems to be a very speaker-specific phenomenon carries some interesting implications for the general conception of *uh* and *uhm*. While up to this point the assumption in this thesis was that *uh* and *uhm* are symptoms and signals that are determined by the given linguistic and extra-linguistic context, I now wish to propose a fourth dimension to this theoretical model of UH/M: the individual speaker. More precisely, I wish to argue that the production of *uh* and *uhm* is determined by four different factors: (1) a speaker’s personal need to buy time for an ongoing cognitive process (the *symptom function* of UH/M); (2) a speaker’s need or desire to verbally signal this ongoing process *to the listener* (the *signal function* of UH/M); (3) the linguistic and extra-linguistic context of the given discourse situation, including, for instance, the level of privacy or the linguistic conventions and/or necessities associated with a given genre; and, finally, (4) the speaker as an individual who demonstrates personal tendencies with regard to his/her willingness to signal his/her cognitive processes by uttering UH/M on the one hand and his/her choice of variants on the other hand. Interestingly, the results obtained in the present study seem to suggest that the influence of these four dimensions is differently pronounced for the two variants; this can be seen, for instance, by the fact that *uhm* is more influenced by genre than *uh* (cf. section 9.1) or that *uhm* appears slightly more often at points of potential turn-transition such as TRPs (cf. section 9.2), while the production of *uh*, in turn, seems to be more dependent on a speaker’s idiosyncratic mannerisms.

Having examined the possibility that *uh* and *uhm* are idiosyncratic phenomena, this section marks the end of the discussion of the results obtained in the present study of syntactic and functional differences between *uh* and *uhm*. The following and final chapter will conclude this paper by summarizing the key findings of this thesis on the one hand and suggesting possible directions for future research on the other hand.

## 10. Conclusion

As was detailed in chapter 7 of this thesis, the primary aim of this paper was to examine the question of whether or not there are any observable syntactic and/or functional differences between *uh* and *uhm* in British English. In pursuit of this objective, it was sought to answer the following three questions:

- (1) How does genre affect the frequency of *uh* and *uhm*?
- (2) Are there any observable differences between *uh* and *uhm* with regard to their preferred syntactic positions, and if yes, what are they?
- (3) How do *uh* and *uhm* differ from each other with regard to the functions they fulfill in spoken discourse?

In order to answer these questions, this thesis was divided into two parts: a comprehensive summary of the existing literature on various aspects of the study of *uh* and *uhm* on the one hand, and an empirical examination of a small-scale corpus study of *uh* and *uhm* in ICE-GB on the other hand. The first part of the thesis showed that a large proportion of research conducted on the study of *uh* and *uhm* can be embedded in the context of the so-called symptom-versus-signal debate, which assumes that *uh* and *uhm* function *either* as symptomatic expressions of a speaker's cognitive processes *or* as a signal to one's listener. The present thesis constitutes a break with this popular one-dimensional conception of *uh* and *uhm*, however, by following the idea that they are in fact multi-functional phenomena that serve both the speaker *and* the listener in attaining their shared goal of efficient – i.e. fluent and effortless – communication. Based on this assumption, this paper also attempted to construct a theoretical representation of *uh* and *uhm* by applying three existing linguistic models – Karl Bühler's *Organon model*, Roman Jakobson's *Model of Communication*, and Halliday's model of the three metafunctions of language – to the description of *uh* and *uhm*. In the course of this enterprise, it was repeatedly stressed that *uh* and *uhm* seem to be highly complex phenomena that fulfill various functions at the same time: while they may help the speaker to buy time for completing his/her ongoing processes linked to information retrieval,



the linguistic realization of pre-verbal messages, or lexical access, they may simultaneously act as a signal of turn-keeping, turn-taking, or turn-yielding for the listener. This two-sided functional structure of *uh* and *uhm* becomes especially obvious when considering the specific cases of turn-accepting and self-interruption, in both of which *uh* and *uhm* act as signals of a speaker's mental processes *and* as a communicative signal to his/her listener(s). Based on these findings, *uh* and *uhm* should consequently no longer be perceived as disruptive disfluencies; instead, it seems plausible to consider them to be so-called *collateral signals* (Clark and Fox Tree 2002: 105-107), i.e. lexical phenomena such as pragmatic markers which allow a speaker to communicate messages concerning his/her cognitive processes without adding to the primary message under construction.

Following the extensive discussion of existing literature, the second part of this thesis was dedicated to the empirical examination of the abovementioned research questions. For this purpose, a small-scale corpus study of fifteen texts taken from ICE-GB was conducted; this study sought to investigate the syntactic and functional differences between *uh* and *uhm* on the one hand as well as the extent to which genre affects the production of the two variants on the other hand. So as to allow for the most comprehensive analysis possible, the location of the two variants was investigated with regard to five different linguistic ranks: the turn, the TCU, the clause, the phrase, and the immediate lexical/pausal environment. The results obtained in the present study carried some interesting implications not only with regard to the defined research questions, but also concerning the linguistic nature and functions of *uh* and *uhm* in general.

Overall, the present study could not support the hypothesis that *uh* systematically fulfills different functions in spoken discourse than *uhm* (and vice versa); this conclusion is largely based on the finding that the two variants were observed to occur in very similar syntactic environments and conversational contexts. For instance, it was shown that the vast majority of studied instances occurred in turn-medial positions, which has led me to contradict the assumption that turn-keeping and turn-yielding constitute major functions of *uh* and *uhm*. Moreover, it was found that a very large proportion of both variants appeared in contexts that were assumed to involve the completion of one or multiple cognitive processes linked to the production of speech, such as at or near the boundaries of TCUs and clauses, in contexts of lexical retrieval, in answers to questions, or in self-repairs. Along with the finding that large proportions of both variants occurred at or near points of possible turn-completion, this latter finding has been interpreted as evidence for the assumption that *uh* and *uhm* are

predominantly interpersonal signals of a speaker's ongoing mental processes. Further evidence in favor of this assumption stems from the observation that both variants appeared more often in telephone conversations than in any other of the studied genres, which seems to reinforce the conception of *uh* and *uhm* as verbal signals to the listener. In conjunction with the observation that *uh* was rather frequently associated with functions that have been argued to be typically fulfilled by *uhm* and vice versa, the findings outlined in the present paragraph have led me to conclude that *uh* and *uhm* do not generally serve different purposes in spoken discourse. In a different line of argument, the results obtained in the present study also contradict Clark and Fox Tree's (2002: 9.3.1) famous hypothesis that *uh* and *uhm* announce minor and major delays respectively, for neither of the two variants was found to frequently co-occur with silent pauses.

Even though the present study seems to suggest that *uh* and *uhm* largely fulfill the same functions in spoken discourse, some interesting, yet only weakly pronounced, differences could be observed to exist between the two variants. One of these differences concerns the preferred syntactic position of *uh* and *uhm*. As was discussed in chapter 9.2 of this thesis, it was found that *uhm* generally dominated its vocalic counterpart at or near TCUs and clause boundaries, while the number of *uhs* typically exceeded that of *uhms* in unit-medial positions. Assuming that syntactic boundaries and especially clause boundaries often coincide with points of conceptual and syntactic decision-making (cf. Hawkins 1971: 286-287), this observation has been regarded as evidence in support of Shriberg's (1994: 154) hypothesis that *uhm* is more closely linked to conceptual and/or syntactic planning, while *uh* typically represents local lexical retrieval. Nevertheless, it is important to point out that these differences are only minor, for each variant was also observed to frequently occur in the respective other position; this becomes especially obvious when considering that even though *uh* dominated *uhm* in the middle and postnuclear positions in clauses, the largest proportion of this variant was still found to occur in prenuclear positions. Similarly, as concerns Clark and Fox Tree's (2002: 86) aforementioned hypothesis that *uh* and *uhm* signal different lengths of delays, it was found that *uhm* was the variant that was slightly more often associated with delays in speech delivery; however, these differences were far too subtly pronounced to confirm their hypothesis. Summing up the above findings, I wish to argue that there seem to be some minor syntactic and functional differences between *uh* and *uhm*; however, these differences appear to be significantly less pronounced than is suggested in existing literature

on the topic and do thus not suffice to interpret them as conclusive evidence for the assumption that *uh* and *uhm* fulfill different functions in spoken discourse.

Since hardly any syntactic and functional differences could be found to exist between *uh* and *uhm*, it has been proposed that a speaker's production and choice of the two variants is in fact determined by idiosyncratic phenomena. This idea is supported by various findings obtained in the present study, such as the observation that the production rates of both variants varied significantly across the different corpus texts within the genre of private face-to-face conversations, in which speakers are not assumed to excessively monitor their own speech. In addition, it was also discussed in section 9.4 that speakers appeared to differ considerably with regard to their production rates and preferences of *uh* and/or *uhm*; this seems to further support the assumption that *uh* and *uhm* are idiosyncratic mannerisms. As was pointed out in section 9.4, one main challenge to the view of *uh* and *uhm* as idiosyncratic phenomena is that they seem to display certain syntactic tendencies that are similar for a large proportion of speakers; however, it has been argued that this can likely be explained by the assumption that a speaker's idiosyncratic use of *uh* and *uhm* is shaped by his/her personal experiences with regard to the efficient management of spoken interaction. Moreover, it was also pointed out that idiosyncratic mannerisms do not seem to be the only determinants of *uh* and *uhm*, for the vocalic-nasal variant *uhm* has also been shown to be considerably influenced by the contextual factors of genre, setting, and especially the privacy of a given speaking situation.

Apart from relativizing existing theories on potential differences between *uh* and *uhm* on the one hand and suggesting that they are at least partly idiosyncratic on the other hand, the originality of this thesis lies in the construction of a theoretical model of *uh* and *uhm* that seeks to outline the mechanisms and functions underlying their production as inferred from the obtained results. According to this theory of UH/M, the production of *uh* and *uhm* is determined by four different factors: (1) a speaker's need to buy time for an ongoing cognitive process, (2) a speaker's need or desire to verbally signal this process to the listener, (3) the linguistic and extra-linguistic context of a given discourse situation, and (4) the speaker as an individual who demonstrates distinct idiosyncratic mannerisms and preferences with regard to the production of UH/M, which have been shaped by his/her respective communicative and interactional experiences. Based on the findings obtained in the present study, it seems that the influence of these four dimensions is differently pronounced for the two variants. This is illustrated by the highly interesting observation that *uhm* was found to be considerably more influenced by genre than *uh*, while the production of *uh*, in turn, was argued to be more

dependent on a speaker's idiosyncratic mannerisms. In summary then, the findings obtained in the present study of functional differences between *uh* and *uhm* have led me to suggest that the two variants do not generally differ from each other with regard to the functions they fulfill, but that they are determined by different contextual factors; the further exploration of this hypothesis constitutes an interesting direction for future research.

Naturally, the findings obtained in the present study have to be regarded in the light of various limitations, which in turn can be used to formulate suggestions for future research. The biggest limitation of this study is certainly the size of the corpus. As was detailed in section 7.2 of this paper, the corpus studied in the present analysis comprised only roughly 30,000 words, of which 194 were *uhs* and 335 were *uhms*. In future research, it would be interesting to test the hypotheses formulated in the course of this thesis with regard to a larger sample, for it is assumed that the study of more instances of *uh* and *uhm* might provide an even clearer picture of (the differences between) *uh*'s and *uhm*'s functions in spoken discourse. The second important limitation is that the present thesis could provide only a relatively superficial examination of the idea that *uh* and *uhm* are determined by idiosyncratic mannerisms; this is mainly due to the fact that this hypothesis has only been formulated in the course of the discussion section of this paper. However, given that even this very brief analysis has suggested that the production and choice of *uh* and *uhm* seems to be at least partially determined by a speaker's individual tendencies, preferences, and – as has been argued in section 9.4 – experiences, it seems highly relevant to further examine the conception of *uh* and *uhm* as idiosyncratic phenomena. Finally, the finding that the production of *uhm* seems to be considerably more influenced by genre than that of *uh*, while the uttering of the latter seems to be more dependent on idiosyncratic mannerisms also deserves further consideration in the future.

Despite the fact that further research is certainly necessary to verify the findings obtained in the present study, this paper constitutes a valuable contribution to the study of *uh* and *uhm*. Not only did it contribute to the closing of the outlined research gap by showing that *uh* and *uhm* do not generally seem to fulfill different functions in spoken discourse, but it has also proposed a new theoretical model of *uh* and *uhm* that could act as a valuable framework for future research on the topic. Moreover, this thesis has also provided suggestions for future research by formulating the hypothesis that *uh* and *uhm* might not differ from each other with regard to their functions, but with regard to their contextual and speaker-specific determinants. Finally, this thesis also marks a valuable contribution to the general study of *uh*

and *uhm* by reinforcing the view that *uh* and *uhm* are not in fact disruptive disfluencies, but should rather be perceived as important communicative signals that greatly facilitate interpersonal communication. Given these conclusive findings, it is to be hoped that *uh* and *uhm* will soon gain recognition as important communicative resources, not just in linguistic research, but also in the broader public.

## 11. References

- Alami, Manizheh. 2015. "Pragmatic Functions of Discourse Markers: A Review of Related Literature". *International Journal on Studies in English Language and Literature* 3(3), pp. 1-10.
- Allwood, Jens; Nivre, Joakim; Ahlsén, Elisabeth. 1990. „Speech management – On the non-written life of speech". *Nordic Journal of Linguistics* 13(1), pp. 1-45.
- Archer, Dawn; Aijmer, Karin; Wichmann, Anne. 2012. *Pragmatics. An Advanced Resource Book for Students*. Abingdon: Routledge.
- Arnold, Jennifer E.; Tanenhaus, Michael K.; Altmann, Rebecca J.; Fagnano, Maria. 2004. „The old and thee, *uh*, new: disfluency and reference resolution". *Psychological Science* 15, pp. 578-582.
- Arnold, Jennifer E.; Hudson-Kam, Carla L.; Tanenhaus, Michael K. 2007. „If you say thee *uh* you are describing something hard: the on-line attribution of disfluency during reference comprehension". *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(5), pp. 914-930.
- Arnold, Jennifer E.; Tanenhaus, Michael K. 2011. „Disfluency Effects in Comprehension: How New Information Can Become Accessible". In: Pearlmutter, Neal J.; Gibson, Edward (eds.). *The Processing and Acquisition of Reference*, pp. 197-217.
- Bailey, Karl G.D.; Ferreira, Fernanda. 2003. "Disfluencies affect the parsing of garden-path sentences". *Journal of Memory and Language* 49, pp. 183-200.
- Barr, Dale J. 1998. "Trouble in mind: paralinguistic indices of effort and uncertainty in communication". In: Santi, S.; Guaitella, I.; Cave, G.; Konopczynski, G. *Oralité et gestualité. Communication multimodal, interaction*. Paris : L'Harmattan, pp.597-600.
- Beattie, Geoffrey W.; Butterworth, B.L. 1979. "Contextual Probability and Word Frequency as Determinants of Pauses and Errors in Spontaneous Speech". *Language and Speech* 22(3), pp. 201-211.
- Berkley, Susan. 2002. "How to Cure the Verbal Virus: A Five Step Treatment Plan". <http://www.careerknowhow.com/improvement/virus.htm> (8 March 2018).
- Boomer, Donald S. 1965. "Hesitation and grammatical encoding". *Language & Speech* 8(3), pp. 148-58.
- Brennan, Susan E.; Schober, Michael F. 2001. "How Listeners Compensate for Disfluencies in Spontaneous Speech". *Journal of Memory and Language* 44(2), pp. 274-296.

- Brennan, Susan E.; Williams, Maurice. 1995. "The Feeling of Another's Knowing: Prosody and Filled Pauses as Cues to Listeners about the Metacognitive States of Speakers". *Journal of Memory and Language* 34, pp. 383-398.
- Broen, Patricia A.; Siegel, Gerald M. 1972. "Variations in Normal Speech Disfluencies". *Language and Speech* 15(3), pp. 219-231.
- Bühler, Karl. 1934. *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena: Fischer.
- Caffarel, Alice. 2010. „Systemic Functional Grammar and the study of meaning”. In: Heine, Bernd; Narrog, Heiko (eds.). *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press, pp. 797-825.
- Carroll, J.B. 1953. *The Study of Language*. Cambridge, Massachusetts: Harvard University Press.
- Cenoz, Jasone. 2000. "Pauses and hesitation phenomena in second language production". *ITL: Review of Applied Linguistics* 127(8), pp. 53-69.
- Chafe, Wallace L. 1980. "Some reasons for hesitating". In: Dechert, H.W.; Raupach, M. (eds.). *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler*. The Hague: Mouton, pp. 169-180.
- Christenfeld, Nicholas. 1994. "Options and Ums". *Journal of Language and Social Psychology* 3(2), pp. 192-199.
- Christenfeld, Nicholas. 1995. "Does it hurt to say um?". *Journal of Nonverbal Behavior* 19(3), pp. 171-186.
- Christenfeld, Nicholas. 1996. "Effects of a Metronome on the Filled Pauses of Fluent Speakers". *Journal of Speech and Hearing Research* 39(6), pp. 1232-38. Accessed online: Retrieved March 23, 2018, from <http://web.b.ebscohost.com/uaccess.univie.ac.at/ehost/detail/detail?vid=0&sid=3d935d49-00a6-4bcb-9102cf6422a84e54%40sessionmgr102&bdata=JnNpdGU9ZWwhvc3QtbGl2ZQ%3d%3d#AN=9701275294&db=ufh>.
- Christenfeld, Nicholas; Creager, Beth. 1996. "Anxiety, Alcohol, Aphasia, and Ums". *Journal of Personality and Social Psychology* 70(3), pp. 451-460.
- Clark, Herbert H.; Fox Tree, Jean E. 2002. „Using *uh* and *um* in spontaneous speaking". *Cognition* 84, pp. 73-111.
- Clayman, Steven E. 2013. "Turn-Constructional Units and the Transition-Relevance Place". In: Sidnell, Jack; Stivers, Tanya (eds.). *The Handbook of Conversation Analysis*. West Sussex: Blackwell Publishing, pp. 150-166.

- Corley, Martin; Hartsuiker, Robert J. 2003. "Hesitation in speech can...um...help a listener understand". *Proceedings of the Annual Meeting of the Cognitive Science Society* 25(25), pp. 276-281.
- Corley, Martin; Stewart, Oliver W. 2008. "Hesitation Disfluencies in Spontaneous Speech: The Meaning of *um*". *Language and Linguistic Compass* 2(4), pp. 589-602.
- Croucher, Stephen M. 2004. "Like, You Know, What I'm Saying: A Study of Discourse Marker Frequency in Extemporaneous and Impromptu Speaking". *National Forensic Journal* 2, pp. 38-47.
- Crystal, David. 1979. "Neglected grammatical factors in conversational English". In: Greenbaum, S.; Leech, G.; Svartvik, I. (eds.). *Studies in English Linguistics*. London: Longman, pp. 153-166.
- Decreus, Freddy; Hallyn, Fernand; Musschoot, Anne-Marie; Tritsmans, Bruno. 1990. „Poétique“. In: Delcroix, Maurice; Hallyn, Fernand. 1990. *Introduction aux études littéraires. Méthodes du texte*. Paris : Duculot, pp. 11-28.
- Deese, James. 2010 [1980]. "Pauses, prosody, and the demands of production in language". In: Dechert, Hans W.; Raupach, Manfred (eds.). 2011. *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*. Berlin: De Gruyter Mouton, pp. 69-84.
- De Leeuw, Esther. 2007. "Hesitation Markers in English, German, and Dutch". *Journal of Germanic Linguistics* 19(2), pp. 85-114.
- Duez, Danielle. 1982. "Silent and Non-Silent Pauses in Three Speech Styles". *Language and Speech* 25(1), pp. 11-28.
- Erard, Michael. 2008. *Um...: slips, stumbles, and verbal blunders, and what they mean*. New York, NY: Anchor Books.
- Filler [Def. 1]. In *Merriam Webster Online*. Retrieved March 16, 2018, from <https://www.merriam-webster.com/dictionary/filler>.
- Fischer, Kerstin. 2006. "Frames, constructions, and invariant meanings: the functional polysemy of discourse particles". In: Fischer, Kerstin (ed.). *Approaches to discourse particles*. Amsterdam: Elsevier, pp. 427-447.
- Ford, Marilyn. 1982. "Sentence Planning Units: Implications for the Speaker's Representation of Meaningful Relations Underlying Sentences". In: Bresnan, Joan. *The mental representation of grammatical relations*. Cambridge, Mass.: MIT Press, pp. 797-828.
- Ford, Marilyn. 1993. "Characteristics of Hesitation Phenomena". In: Blanken, Gerhard et al. (eds.). *Linguistic Disorders and Pathologies: An International Handbook*. Berlin: de Gruyter, pp. 46-52.



- Fox Tree, Jean E. 1995. "The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech". *Journal of Memory and Language* 34, pp. 709-738.
- Fox Tree, Jean E. 2001. "Listeners' uses of *um* and *uh* in speech comprehension". *Memory & Cognition* 29(2), pp.320-326.
- Fraser, Bruce. 1996. "Pragmatic Markers". *Pragmatics* 6(2), pp. 167-190.
- Friginal, Eric. 2009. *The Language of Outsourced Call Centers. A corpus-based study of cross-cultural interaction*. Amsterdam/Philadelphia: John Benjamins.
- Gilquin, Gaëtanelle. 2008. "Hesitation markers among EFL learners: Pragmatic deficiency or difference?". In: Romero-Trillo, Jesus (ed.). *Pragmatics and Corpus Linguistics: A Mutualistic Entente*. Berlin/New York: Mouton de Gruyter, pp. 119-149.
- Gilquin, Gaëtanelle; De Cock, Sylvie. 2013. "Errors and disfluencies in spoken corpora: Setting the scene". In: Gilquin, Gaëtanelle; De Cock, Sylvie (eds.). *Errors and disfluencies in spoken corpora*. Amsterdam/Philadelphia: John Benjamins, pp. 1-32.
- Goldman-Eisler, Frieda. 1958a. "Speech production and the predictability of words in context". *Quarterly Journal of Experimental Psychology* 10(2), pp. 96-106.
- Goldman-Eisler, Frieda. 1958b. "Speech Analysis and Mental Processes". *Language and Speech* 1(1), pp.59-75.
- Goldman-Eisler, Frieda. 1961. "A Comparative Study of two Hesitation Phenomena". *Language and Speech* 4(1), pp. 18-26.
- Goldman-Eisler, Frieda. 1962. "Speech and Thought". *Discovery* 23(4), pp. 36-41.
- Goldman-Eisler, Frieda. 1968. *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.
- Greene, John O.; Lindsey, A.E.; Hawn, John J. 1990. "Social Goals and Speech Production: Effects of Multiple Goals on Pausal Phenomena". *Journal of Language and Social Psychology* 9(1-2), pp. 119-134.
- Halliday, M.A.K. 1961. "Categories of the Theory of Grammar". *Word* 17(2), pp. 242-292.
- Halliday, M.A.K. 2004. *An introduction to functional grammar* (3<sup>rd</sup> edition). London: Hodder Arnold.
- Hasan, Ruqaiya. 2009. "Wanted: A theory for integrated sociolinguistics". In: Webster, J.J. (ed.). *Semantic Variation: Meaning in Society and in Sociolinguistics*. Vol. 2 of *The Collected Works of Ruqaiya Hasan*. London & Oakville: Equinox, pp. 5-40.

- Hasselgren, Angela. 2002. "Learner corpora and language testing. Smallwords as markers of learner fluency". In: Granger, Sylviane; Hung, Joseph; Petch-Tyson, Stephanie (eds.). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Testing*. Amsterdam/Philadelphia: John Benjamins, pp. 143-173.
- Hawkins, P.R. 1971. "The syntactic location of hesitation pauses". *Language and Speech* 15(3), pp. 277-288.
- Heine, Bernd; Kaltenböck, Gunther; Kuteva, Tania; Long, Haiping. 2013. "An Outline of Discourse Grammar". In: Bischoff, Shannon; Jany, Carmen (eds.). *Functional Approaches to Language*. Berlin: de Gruyter, pp. 155-206.
- Hesitate [Def. 1]. In *Merriam Webster Online*. Retrieved March 16, 2018, from <https://www.merriam-webster.com/dictionary/hesitating>.
- Holmes, Virginia M. 1988. "Hesitations and Sentence Planning". *Language and Cognitive Processes* 3(4), pp. 323-361.
- Houtkoop, Hanneke; Mazeland, Harrie. 1985. "Turns and Discourse Units in Everyday Conversation". *Journal of Pragmatics* 9, pp. 595-619.
- ICECUP. 2006. *The International Corpus of English Corpus Utility Program, version 3.1*. Wallis, Sean (ed.). [CD-ROM]. London: Survey of English Usage, University College London.
- ICE-GB. 2006. *The British component of the International Corpus of English, Release 2*. Aarts, Bas; Nelson, Gerald; Wallis, Sean (eds.). [CD-ROM]. London: Survey of English Usage, University College London.
- Innis, Robert E. 1998. "Bühler and his followers". In: Posner, Roland; Robering, Klaus; Sebeok, Thomas A. (eds.). *Semiotik. Ein Handbuch zu den zeichentheoretischen Grundlagen von Natur und Kultur*. Volume 2. Berlin/New York: de Gruyter, pp. 2198-2203.
- Jakobson, Roman. 1960. "Closing Statement: Linguistics and Poetics". In: Sebeok, Thomas A. (ed.). *Style in Language*. Cambridge, Massachusetts: MIT Press, pp. 350-377.
- James, D. 1972. "Some aspects of the syntax and semantics of interjections". *Papers from the eighth regional meeting of the Chicago Linguistic Society*. Chicago, IL: Chicago Linguistic Institute, pp.162-172.
- Kaltenböck, Gunther. 2009. "English comment clauses: position, prosody, and scope". *Arbeiten aus Anglistik und Amerikanistik* 34(1), pp. 49-75.

- Kaltenböck, Gunther; Keizer, Evelien; Lohmann, Arne. 2016. „Extra-clausal constituents: An overview”. In: Kaltenböck, Gunther; Keizer, Evelien; Lohmann, Arne (eds.). *Outside the Clause. Form and function of extra-clausal constituents*. Studies in Language Companion Series. Vol. 178. Amsterdam: John Benjamins, pp. 1-28.
- Kjellmer, Göran. 2003. “Hesitation. In Defence of ER and ERM”. *English Studies* 84(2), pp. 170-198.
- Kjellmer, Göran. 2008. „Self-Repetition in Spoken English Discourse”. In: Nevalainen, Terttu; Taavitsainen, Irma; Pahta, Päivi; Korhonen, Minna (eds.). *The Dynamics of Linguistic Variation: Corpus Evidence on English Past and Present*. Amsterdam: John Benjamins, pp. 37-60.
- Künzel, Hermann J. 1997. “Some general phonetic and forensic aspects of speaking tempo”. *Forensic Linguistics* 4, pp. 48-83.
- Larcombe, Michael. 1995. “Why do people say ‘um’ and ‘er’ when hesitating in their speech?” *New Scientist*. Online source. Cited in: Gilquin, Sylvie. 2008. Retrieved from [www.newscientist.com/backpage.ns?id=lw125](http://www.newscientist.com/backpage.ns?id=lw125) (last accessed on 27 February 2007 by Gilquin 2008).
- Lay, C. H.; Paivio, A. 1969. “The effects of task difficulty and anxiety on hesitations in speech”. *Canadian Journal of Behavioral Science* 1, pp. 25-37.
- Lennon, Paul. 1990. “Investigating fluency in EFL: a quantitative approach”. *Language Learning* 40(3), pp. 387-417.
- Levelt, Willem J.M. 1989. *Speaking: from intention to articulation*. Cambridge, Mass.: MIT Press.
- Levelt, Willem J.M. 1995 [1989]. *Speaking: from intention to articulation* (4<sup>th</sup> edition). Cambridge, Mass.: MIT Press.
- Lindsay, Jean; O’Connell, Daniel C. 1995. „How do transcribers deal with audio recordings of spoken discourse?”. *Journal of Psycholinguistic Research* 24(2), pp. 101-115.
- Lickley, Robin J.; Bard, Ellen G. 1996. „On not recognizing disfluencies in dialogue”. *Proceedings of the Fourth International Conference on Spoken Language Processing* 3, pp. 1876-1879.
- Lounsbury, Floyd G. 1954. “Sequential Psycholinguistics”. In: Osgood, Charles E. Sebeok, Thomas A. (eds.). *Psycholinguistics: A survey of theory and research problems. With “A survey of psycholinguistic research, 1954-1964”*. Bloomington/London: Indiana University Press, pp. 93-125.
- Maclay, Howard; Osgood, Charles E. 1959. “Hesitation Phenomena in Spontaneous English Speech”. *Word* 15(1), pp. 19-44.

- Mahl, George F. 2014 [1987]. *Explorations in Nonverbal and Vocal Behavior*. New York, NY: Psychology Press.
- Mair, Christian. 2015. *English Linguistics. An Introduction*. 3<sup>rd</sup> edition. Tübingen: Narr Francke Attempto.
- Martin, James G.; Strange, Winifred. 1968. "The perception of hesitation in spontaneous speech". *Perception and Psychophysics* 3(6), pp. 427-438.
- Nelson, Gerald; Aarts, Bas; Wallis, Sean. 2002. *Exploring Natural Language: Working with the British Component of the International Corpus of English*. Amsterdam: John Benjamins.
- Nöth, Winfried. 1990. *Handbook of Semiotics*. Bloomington, Indianapolis: Indiana University Press.
- O'Connell, Daniel C.; Kowal, Sabine. 2004. "The history of research on the filled pause as evidence of *The Written Language Bias in Linguistics* (Linell, 1982)". *Journal of Psycholinguistic Research* 33(6), pp. 459-474.
- O'Connell, Daniel C.; Kowal, Sabine. 2005. "*Uh* and *Um* Revisited: Are They Interjections for Signaling Delay?". *Journal of Psycholinguistic Research* 34(6), pp. 555-576.
- O'Connell, Daniel C.; Kowal, Sabine; Ageneau, Carie. 2005. "Interjections in Interviews". *Journal of Psycholinguistic Research* 34(2), pp. 153-171.
- Quinting, Gerd. 1971. *Hesitation Phenomena in Adult Aphasic and Normal Speech*. The Hague: Mouton.
- Reynolds, A.; Paivio, A. 1968. "Cognitive and emotional determinants of speech". *Canadian Journal of Psychology* 22(3), pp.164-175.
- Richard, Mark. 2005. "Non-Truth-Conditional Meaning". *Encyclopedia.com: Encyclopedia of Philosophy*. <https://www.encyclopedia.com/humanities/encyclopedias-almanacs-transcripts-and-maps/non-truth-conditional-meaning> (21 March 2018).
- Rochester, S.R. 1973. "The Significance of Pauses in Spontaneous Speech". *Journal of Psycholinguistic Research* 2(1), pp. 51-81.
- Rühlemann, Christoph; Bagoutdinov, Andrej; O'Donnell, Matthew Brook. 2011. "Windows on the Mind: Pauses in Conversational Narrative". *International Journal of Corpus Linguistics* 16(2), pp. 198-230.
- Sacks, Harvey; Schegloff, Emanuel A.; Jefferson, Gail. 1974. "A simplest systematic for the organization of turn-taking for conversation". *Language* 50(4), pp. 696-735.

- Schachter, Stanley; Christenfeld, Nicholas; Ravina, Bernard; Bilous, Frances. 1991. "Speech Disfluency and the Structure of Knowledge". *Journal of Personality and Social Psychology* 60(3), p.362-367.
- Schegloff, Emanuel A. 1996. "Turn organisation: one intersection of grammar and interaction". In: Ochs, Elinor; Schegloff, Emanuel A.; Thompson, Sandra A. (eds.). *Interaction and grammar*. Cambridge: Cambridge University Press, pp. 52-133.
- Selting, Margret. 2000. "The construction of units in conversational talk". *Language in Society* 29, pp. 477-517.
- Selting, Margret. 2005. "Syntax and prosody as methods for the construction and identification of turn-constructive units in conversation". In: Hakulinen, Auli; Selting, Margret (eds.): *Syntax and Lexis in Conversation*. Amsterdam: Benjamins, pp. 17-44.
- Sentence. In *Merriam Webster Online*. Retrieved April 16, 2018, from <https://www.merriam-webster.com/dictionary/sentence>.
- Shriberg, Elizabeth. 1994. "Preliminaries to a Theory of Speech Disfluencies". Doctoral dissertation, University of California at Berkeley.
- Shriberg, Elizabeth. 2001. "To 'errrr' is human: ecology and acoustics of speech disfluencies". *Journal of the International Phonetic Association* 31(1), pp. 153-169.
- Siegmán, Aron Wolfe; Pope, Benjamin. 1966. „Ambiguity and verbal fluency in the TAT". *Journal of Consulting Psychology* 30, pp. 239-245.
- Smith, Vicki L.; Clark, Herbert H. 1993. "On the Course of Answering Questions". *Journal of Memory and Language* 32, pp. 25-38.
- Stenström, Anna-Brita. 1990. "Pauses in monologue and dialogue". In : Svartvik, Jan (ed.). *The London-Lund Corpus of Spoken English*. Lund Studies in English 82. Lund University Press, pp. 211-252.
- Stenström, Anna-Brita; Svartvik, Jan. 1994. "Imparsable speech: Repeats and other nonfluencies in spoken English". In: Oostdijk, N; de Haan, P. (eds.). *Corpus-based Research into Language. In Honour of Jan Aarts*. Amsterdam/Atlanta: Rodopi, pp. 241-254.
- Swerts, Marc; Wichmann, Anne; Beun, Robbert-Jan. 1996. "Filled pauses as markers of discourse structure". *Proceeding of Fourth International Conference on Spoken Language Processing* (Vol. 2), pp. 1033-1036.
- Swerts, Marc. 1998. "Filled pauses as markers of discourse structure". *Journal of Pragmatics* 30, pp. 485-496.

- Swerts, Marc; Krahmer, Emiel. 2005. "Audiovisual prosody and feeling of knowing". *Journal of Memory and Language* 53, pp. 81-94.
- Symptom. In *Merriam Webster Online*. Retrieved March 22, 2018, from <https://www.merriam-webster.com/dictionary/symptom>
- Tannenbaum, Percy H.; Williams, Frederick; Hillier, Carolyn S. 1965. "Word predictability in the environments of hesitations". *Journal of Verbal Learning and Verbal Behavior* 4(2), pp. 134-140.
- Tottie, Gunnel. 2011. "Uh and Um as sociolinguistic markers in British English". *International Journal of Corpus Linguistics* 16(2), pp. 173-197.
- Tottie, Gunnel. 2014. "On the use of uh and um in American English". *Functions of Language* 21(1), pp. 6-29.
- Tottie, Gunnel. 2015. "Turn management and the fillers uh and um". In: Aijmer, Karin; Rühlemann, Christoph (eds.). *Corpus Pragmatics. A Handbook*. Cambridge: Cambridge University Press, pp. 381-407.
- Tottie, Gunnel. 2016. "Planning what to say : Uh and um among the pragmatic markers". In: Kaltenböck, Gunther; Keizer, Evelien; Lohmann, Arne (eds.). *Outside the Clause. Form and function of extra-clausal constituents*. Studies in Language Companion Series. Vol. 178. Amsterdam: John Benjamins, pp. 97-122.
- Uh. In *Merriam Webster Online*. Retrieved March 22, 2018, from <https://www.merriam-webster.com/dictionary/uh>.

## 12. Appendices

### Appendix 1: ICE-GB corpus data studied in the analysis of *uh* (segmented into TCUs)

no.	ICE-GB code	segmented data (number in brackets indicate the number of the instance)
1	S1A-015:23	S2: [...] But the thing is you always have to write them in a slight code // so people don't know exactly what you are talking about // [and uhm <,> and it's (1) <b>uh</b> ] / you know so so that if a financial director comes in and happens to read it then you know you're in the clear // [...]
2	S1A-015:42	S1: How much <,> could you afford // S2: (2) <b>uh</b> <,> I don't know // No nothing really // [I'm] / I've got too many debts <,> //
3	S1A-015:52	S1: Bi- Bim and I always planned to have to have some done based on on (3) <b>uh</b> photos of Auden and Isherwood <,> // S2: Really // <laughs>
4	S1A-015:54.1	
5	S1A-015:54.2	
6	S1A-015:54.3	S1: Yes // but I I think (4) <b>uh</b> <,> uhm <,> we don't look <,> we don't look as uhm <,> as spruce and (5) <b>uh</b> <,> elegant and (6) <b>uh</b> <,> as pure // especially good-looking like him //
7	S1A-015:82.1	S2: Do you think uhm he's really good <,> //
8	S1A-015:82.2	S1: (7) <b>uh</b> <,> well (8) <b>uh</b> I think he is really good // [I mean as a as a sort of] / Not his pictures yet <,> //
9	S1A-015:91	S1: [...] Well my impression of the pictures I saw at the exhibition was that there wasn't uhm he wasn't filtered (9) <b>uh</b> by an actual filter // but he was just painting what he wanted to paint <1> <,> </1> // and the quality was variable // and there was no one standing around say- saying to him you should junk this // [and and well you know] / uhm and and (10) <b>uh</b> so some of it wasn't was not very good quality S2: <1> Yes </1> S2: Mm S1: [(11) <b>uh</b> I c-] / I mean uhm I can sort of read it (you) see / because I I knew his life // the personal (12) <b>uh</b> signification significance of the pictures //
10	S1A-015:97	
11	S1A-015:100.1	
12	S1A-015:100.2	
13	S1A-015:110	S2: Does he live out of London //
14	S1A-015:111.1	S1: [Yeah he went to live in] / he went to (13) <b>uh</b> Colchester // He's training to be (14) <b>uh</b> don't know what you call it sort of the step above (15) <b>uh</b> an EFL teacher // S2: Really //
15	S1A-015:111.2	
16	S1A-015:113	
17	S1A-015:116	S1: (16) <b>uh</b> a sort of cour- course director : I think it is // S2: Is he still writing // S1: Well still isn't really the word // [it it stopped around (17) <b>uh</b> ] = S2: = I know // it stopped years ago //
18	S1A-015:152	S2: But does he does he not just like write for himself // S1: Internally driven I meant // No he is not // no <,> //
19	S1A-015:153	S2: Yeah // yeah <,> // S1: (18) <b>uh</b> <,> I don't know though // (19) <b>uh</b> he he he's uhm he's uhm working on his language <,> //
20	S1A-015:175	S1: Yeah well I've written down bits of the grammar and the phonology <,> // and the semantics // but it's not all merged into a lexicon <,> // S2: Yeah // S1: [(20) <b>uh</b> but his uhm his] = S2: = I thought you meant language like a language // [like you know you have different] / [...]
21	S1A-015:204	S1: And uhm so I've I've written a few short texts in it // like the Lord's Prayer : I've

		<p>&lt;1&gt; translated &lt;/1&gt; //</p> <p>S2: &lt;1&gt; Really &lt;/1&gt;</p> <p>S1: Yes // uhm (21) <u>uh</u> &lt;,,&gt; but we don't actually converse in it //</p>
22	S1A-031:53	<p>S2: [...] // So I think that's something I'll just have to you know sit back and weather : more or less // and uhm as somebody put it &lt;,&gt; just be (22) <u>uh</u> try and be a mattress // [and &lt;laugh&gt; sort of] /</p>
23	S1A-031:102	<p>S2: [...] // I mean &lt;,&gt; you know at the bottom it was sort of &lt;,&gt; (23) <u>uh</u> sexual problems &lt;unclear-words&gt; //</p>
24	S1A-031:123.1	<p>S1: [...] // and when she failed that obviously meant that she was no good &lt;,,&gt; //(24) <u>uh</u> in (25) <u>uh</u> at that kind of relationship // so that if there's if there are a couple around (26) <u>uh</u> an older couple you know &lt;,&gt; then uhm &lt;,&gt; you know they they need a sort of cementing bond of a child</p>
25	S1A-031:123.2	
26	S1A-031:124	
27	S1A-031:138.1	<p>S1: [...] What's the answer for her //</p>
28	S1A-031:138.2	<p>S2: &lt;1&gt; I could ask somebody &lt;/1&gt; // what would you make of that //</p> <p>S2: &lt;2&gt; Mm &lt;/2&gt;</p> <p>S2: [Well &lt;,&gt; (27) <u>uh</u> as far (28) <u>uh</u>] // I mean &lt;,&gt; she won't see a psychiatrist or anything remotely approaching that // uhm</p>
29	S1A-031:167	<p>S1: [...] // Well &lt;,&gt; I do I do things to show that I want her need her miss her &lt;,&gt; // uhm &lt;,&gt; completely &lt;,&gt; (29) <u>uh</u> &lt;,,&gt; gratuitously &lt;,&gt; // uhm &lt;,&gt; which at the moment is difficult // [...]</p>
30	S1A-031:192	<p>S1: It is // but that surely would improve with time &lt;,,&gt; //</p>
31	S1A-031:193	<p>S2: [Well I think (30) <u>uh</u>] / yes // I mean I think it will // But you know that's (31) <u>uh</u> that's not something I can &lt;,&gt; start in on straight away //</p>
32	S1A-067:89	<p>S1: Two of wands //</p> <p>S2: Ah wands &lt;,,&gt; //</p> <p>S1: Hmm // Hmm // But it's called sticks in mahjong // (32) <u>uh</u> bamboo sticks //</p>
33	S1A-067:154	<p>S1: Do you think that applies to you / at &lt;1&gt; the moment &lt;/1&gt; //</p> <p>S2: &lt;1&gt; Well it &lt;/1&gt; does in a way // cos I am changing // thinking about change a lot // and having done this massage course // which I've got my (33) <u>uh</u> assessment on Sunday //and then I do my anatomy and physiology and all the rest of it &lt;,&gt; // [...]</p>
34	S1A-067:273	<p>S2: Oh yes // And handing it out // Alms // Maund- (34) <u>uh</u> maundy money // Is he feeding beggars or something // [...]</p>
35	S1A-067:320.2	<p>S1: Ah that's the Great Mother // or one of the forms of the Great Mother &lt;,,&gt; //</p> <p>S2: There's a &lt;,,&gt; an atmosphere of uhm consolidation and (35) <u>uh</u> &lt;,,&gt;=</p> <p>S1: Yes // Agreed //</p> <p>S2: = confirmation in this hand : isn't there //</p>
36	S1A-067:327	<p>S1: Mm // So far //</p> <p>S2: I remember last time there were a couple of sort of slightly anxious ones &lt;,&gt; / (36) <u>uh</u> the the woman the weeping woman in a boat // [being] / and ghastly things like that // which I remember //</p>
37	S1A-080:179	<p>S2: Anyway just just OK just for the sake of the exercise // should you not have him / and should you want to design yourself a new one / what (37) <u>uh</u> what particular model would you like //</p> <p>S1: Someone specific //</p> <p>S2: No no // Just design yourself a new man //</p>
38	S1A-080:278	<p>S1: Nick's body Billy's money Murray's mind &lt;,&gt; //</p> <p>S2: Yes // Oh Murray's so sweet / [He rang me up when uhm &lt;,&gt;] / Yeah he rang me up when (38) <u>uh</u> Julie was away / and we had a chat / Oh he was so funny / [...]</p>
39	S1A-091:180	<p>S1: I've had I've had (39) <u>uh</u> two exams so far // but they were the easy ones // I I did quite well in those I think &lt;,&gt; // [...]</p>



40	S1A-091:324	S2: Is he still cooking // S1: Yeah <,> // S2: <laughs> // Yeah <,> // Good // S1: [And (40) <u>uh</u> ] = S2: = Have you have you still got the other two then // [...]
41	S1A-091:340	S2: Oh good // And how's Emily //
42	S1A-091:345	S1: She's wonderful // She's great // She's just had she's just had her school photo // you know her her <1> school photo </1> for this year // so sweet // she looks so big <,> // (41) <u>uh</u> I mean these children of seven / when I compare them with the children of four it's just absolutely unbelievable <,> // S2: <1> Oh that's sweet </1> <,> <laughs> // S2: [Oh well she is ] / Yeah // S1: And she (42) <u>uh</u> <,> she's so sophisticated and grown-up // And she plays the piano // It's just amazing // [...]
43	S1A-092:19	S2: Same thing really // S1: <laughs> S2: <laughs> // (43) <u>uh</u> and oh I got a job today // S1: Oh good // Oh not the four lines one <,> // [...]
44	S1A-092:24	S2: And (44) <u>uh</u> to add insult to injury it's two lines <,> // [...]
45	S1A-092:40	S2: [...] Well // And it was delightful // There were Linda Butcher and (45) <u>uh</u> Mike Vardy //
46	S1A-092:45	S1: Yeah // I think I <1> met them </1> //
47	S1A-092:52.1	S2: <1> Most </1> delightful interview // [but (46) <u>uh</u> <,>] / And they obviously didn't ask me to read / And I thought well that's <,> you know what a joke [...] / But uhm they were saying <,> (47) <u>uh</u> <,> they want to enhance these uhm <,> sce- the scene of crimes officer department <,> / (48) <u>uh</u> into the plot / and therefore may well be asking me back later in the year <,> // to do some more work //
48	S1A-092:52.2	
49	S1A-092:63	S1: Oh well <,> it's better than a bang on the nose I suppose // poke in the eye with a sharp stick // S2: Yeah // yeah <,> // I suppose so // [It is (49) <u>uh</u> ] / Well I don't want to sound blasé or <1> <unclear-words> </1> //
50	S1A-092:68	S1: <1> Oh no </1> // I mean no don't worry about it // I <2> would feel exactly </2> the same as <3> you </3> //
51	S1A-092:70	S2: <2> Yeah // [I really] / </2> S2: <3> The </3> money's great // And <,> and (50) <u>uh</u> <,> the rehearse's next Thursday next Friday <,> // Shooth thirtieth of June // [But (51) <u>uh</u> ] / and I just smiled about something // and you know <,> // well //
52	S1A-092:76	S2: Did something on From Little Acorns // You can't get much littler than this : can you //
53	S1A-092:79	S1: (52) <u>uh</u> you know <unclear-words> a few years ago <,> TV series // No one'll ever see it // No one'll know (any different) <,> // Should've been the bank robber in that <,> // <1> [in the (53) <u>uh</u> ] </1> = S2: =<1> Oh I don't know </1> // [...]
54	S1A-092:102	S1: [...] // and accidentally saw two videos vaguely related to the subject I'm studying <laugh> // Sat and watched telly for the afternoon <,> // watching these programmes <,> //
55	S1A-092:103.1	
56	S1A-092:103.2	S2: Ah // Brilliant // S1: [About (54) <u>uh</u> ] / One's about the human brain and language <,> / and the other's about (55) <u>uh</u> this guy called Chomsky // <,> who's (56) <u>uh</u> <,> well one of the world's most important human beings // if you happen to be interested in linguistics <,> //

57	S1A-092:124	S1: So you're still painting your mum this weekend //
58	S1A-092:126	S2: Yeah <,> // yeah // [I think <,>] / (57) <u>uh</u> (going) down tomorrow // I'm meeting her at Windsor at twelve <,> // for lunch // And then we're going to <,> buy paints and (58) <u>uh</u> work on it // And then Saturday night stay down at <,> you know you know <,> Dan and Jane's <,> // Work again on Sunday // And (59) <u>uh</u> <,> may stay Sunday night // or just come back <,> //
59	S1A-092:129	
60	S1A-092:148	S1: Bo-fucking-ring // uhm <,> yes // Interesting one // it's putting a word and putting an expletive in the middle // [that's (60) <u>uh</u> <,>] / yeah //
61	S1A-092:245	S2: Very unusual name <,> //
61	S1A-092:246	S1: (61) <u>uh</u> guess where he's from with a name like that // S2: (62) <u>uh</u> Yorkshire // S1: <laughs> Oh a laugh a minute here <,> //
62	S1A-092:275	S2: oh <,> right <,> // say no more <,> // (63) <u>uh</u> what was I going to say //
63	S1A-092:300	S1: <1> Yes I've </1> had my eye on him for for for some time <,> // Had my eye on him since the beginning of the year really // But uhm <,> I decided that I really ought to be wise for months and months and months and <,> not (64) <u>uh</u> <,> ask lecturers out <,> //
64	S1A-092:313.1	S2: <1> Right </1> <chuckles> // so much for that <laughs> //
65	S1A-092:313.2	S1: So uhm <,> eventually a couple of weeks ago I plucked up my courage and <,> rang him up / [...] / And he said that sounds like a good idea <,> // <laughs> S2: <unclear-syllables> <laughs> S1: [So <2> uhm </2> <,>] / [so my (65) <u>uh</u> ] / [so I was (66) <u>uh</u> <,>] / Bravery won the day // S2: <2> <laughs> </2> S2: So this is for Yibin then // S1: Sorry // S2: This (67) <u>uh</u> linguistics stuff <,> //
66	S1A-092:318	
67	S1A-092:353	
68	S1A-093:62	S1: Yeah // S2: And (68) <u>uh</u> how are you // and why is the man opposite me picking his nose // [...]
69	S1A-093:83	S2: Oh how's Lee // S1: Well he's OK // uhm <,> <1> (69) <u>uh</u> </1> S2: <1> Did he </1> get that job // about teaching English //
70	S1A-093:84	S1: [Well I mean it's] / The thing is he was supposed to be coming over for her birthday / coming back for her birthday <,> / (70) <u>uh</u> because he's over there now <,> / And (71) <u>uh</u> she hadn't seen him for nine months or something / Her birthday's in May <,> / And he phoned and yes he was going to get a f- a standby flight <,> / and he never turned up // So she managed to get hold of his Dad / who's living over here / you know in Montague Terrace <,> / And (72) <u>uh</u> <,> he said oh yes you know Uts is uhm he w- he wasn't able to get a standby flight / and uhm <,> it's really difficult and he'll have to wait till he goes via South Korea or something // And she swallowed every word of it // and I mean the thing it's so obvious that he'd (73) <u>uh</u> that he was putting her off <,> : you know //
71	S1A-093:88	
72	S1A-093:91	
73	S1A-093:198	
74	S1A-093:221	S2: [And they a- they] / All of them didn't feel like doing their exams //cos he died like the night before <1> a lot of <1> (74) <u>uh</u> the exams <,> // and uhm they said they they just wanted to go and sort of talk about that // [...]
75	S1A-093:258	S1: Have you heard from anyone from college <,> // S2: No // I've got to phone Liz / because she said she was going to phone me <,> on (75) <u>uh</u> Monday night / but she hasn't [...] //
76	S1A-093:288	S2: <laughs> Maybe he forgot // and that's an excuse // S1: Maybe // Maybe // <1> [but (76) <u>uh</u> ] </1> = S2: = <1> Silly boy </1> //

77	S1A-096:26	S2: So that is good news // S1: Yes // yes // [Ginny] / oh (77) <u>uh</u> I told you yeah // Don't worry // No // I was going to tell you Ginny's got engaged // but you knew that anyway //
78	S1A-096:29.1	S1: So what have you been up to // S2: uhm oh God uhm buried under a couple uhm playwriting commissions uhm <1> <,> </1> // which are <,> are driving currently driving me a bit potty <laughs> // <2> [Well one of them is] </2>
79	S1A-096:29.2	S1: <2> What are they about </2> // S2: Sorry // S1: What are they about : I mean //
80	S1A-096:31.1	S2: [uhm well uhm one of them's a play which is] / the one I'm sort of trying to finish at the moment is a a play for a company called Quicksilver // It's actually the second play I've written for them // uhm <,> which is kind of based on the uhm the Browning poem the Pied Piper // It's sort of about dreams and possibilities inside a mountain // But (78) <u>uh</u> the problem is I've kind of I started with fantasy / and uhm and (79) <u>uh</u> I don't think I'll ever try that again // I I think unless you start with uhm with something real <,> it's uhm you you you end up having too many possibilities open to you // and it just uhm it becomes a bit of a kind of a <,> uhm a fantastically complicated mind maze really <,> // but uhm anyway I'm (80) <u>uh</u> I'm just about (81) <u>uh</u> seeing the light at the other side at the moment // but it's taken me a hell of a long time <,> //
81	S1A-096:31.2	
82	S1A-096:44	S2: [...] // Well I'm I'm writing a a play for them // so uhm <,> so I've kind of <,> been doing that // I did I did a bit of storytelling for uhm a shadow puppet theatre company from September to December // which was really nice <1> <,> </1> // and (82) <u>uh</u> <,> that's all I've been doing really <2> <,> </2> // been pretty busy <,> // how about you // S1: <1> Mm </1>
83	S1A-096:51	S2: [...] // been pretty busy <,> // how about you //
84	S1A-096:52	S1: Well I've <,> I've changed what I've been doing uhm <,> / I kind of kind of see myself as a writer these days / (83) <u>uh</u> I don't know exactly what I'm doing exactly / (84) <u>uh</u> it's very difficult to explain / I I joined a group about uhm <,> over a year ago now <,> / (85) <u>uh</u> called A School for Prophets <,> / [which is uhm ] <,> / uhm Prophets not Profits <,> / (86) <u>uh</u> which is based in Warmesley / which is quite near <,> us in Chatham / it's run by a catholic priest called Bert White <,> / And the group's all <,> ecumenical / uhm mainly catholic because it was started by Bert / but it's got all kinds of other things connected to it / and we've been examining well (87) <u>uh</u> things like Matthew Fox / who's a theologian who's been incredibly liberating <,> / and uhm there's a whole sort of movement happening in this country called creation-centred spirituality / It's not a cult or anything / It's just where people that have particular views on spirituality are actually sort of coming together // not necessarily Christians <,> // (88) <u>uh</u> and it's kind of the sane end of the New Age that are actually quite sensible // uhm and quite positive // but there's also a lot of nutty stuff as well <,> // uhm <,> and that's basically it // But we've we've had uhm I think a very interesting year // we meet every Tuesday // uhm and last year we had people like Sidney Carter come down and do some stuff with us uhm // We had uhm <,> stuff on liberation theology and stuff on feminist theology <,> // and (89) <u>uh</u> <,> we had Julie Felix come down and do some stuff for us a few weeks ago //
85	S1A-096:53.2	
86	S1A-096:53.2	
87	S1A-096:57	
88	S1A-096:61	
89	S1A-096:66	
90	S1A-096:88	S1: And we're now going into a second phase which is starting with a magazine <,> //
91	S1A-096:91	We want a magazine that (90) <u>uh</u> with quite a wide brief on it // But the idea is that
92	S1A-096:92.1	uhm <,> it's a magazine that's going to be a kind of forum for a lot of the things that

93	S1A-096:92.2	are happening // [uhm we've got uhm] / (91) <u>uh</u> it's called Paradigm Shift // because we will be examining the paradigm shift that's actually occurring <> // (92) <u>uh</u> so we've got a whole lot of (93) <u>uh</u> clergy scientists poets <> artists all kinds of people involved in it <> // and uhm <> I've been put in charge of editing this magazine // and I'm preparing the first issue at the moment // (94) <u>uh</u> we've got so far we've got an article by John Medcalf // who's uhm <> written a lot of books about Nicaragua and Peru // He was in N- both of those countries as a Catholic priest // (95) <u>uh</u> we've got uhm a woman called Caitlin Matthews /who is a pagan priestess // a K- of a Celtic <> persuasion // uhm she's written seventeen books so far // on Celtic mysteries (96) <u>uh</u> Arthurian legends uhm <> all kinds of stuff feminism the feminine aspect of God // Her last one was called Sophia the Feminine Aspect of Godness // [It's] / She's (97) <u>uh</u> published by Harper Rowe // They're quite amazing sort of stuff // Anyway she's (98) <u>uh</u> she's given us an article for free // which is good //
94	S1A-096:96	
95	S1A-096:98	
96	S1A-096:99	
97	S1A-096:101	
98	S1A-096:103	
99	S1A-096:110.1	S1: He was with the Cheyenne and the Navaho <> Mohawk and the <unclear-word> people // and he does a lot of liberation theology stuff with them <> // So it's it's sounding quite interesting // uhm <> (99) <u>uh</u> I want to get some more <> (100) <u>uh</u> kind of right brain stuff into it // [it was you know] <> / literature and poetry <,,> // [uhm <,,> wou-] / do you think maybe in the future you could envisage doing a story for us //
100	S1A-096:110.2	
101	S1A-096:141	S1: [...] // which is what I've always liked about the best writing from Gabriel Garcia Marquez // because his best work is really when he sets things in the real world <> // (101) <u>uh</u> and yet there's always this element of mystery and the fantastic and amazing events happening <,,> //
102	S1A-096:177.1	S1: We've been doing it sort of quarterly // But the thing is uhm <> Ann and myself hopefully are going to be in in America in <> the end of Ju- (102) <u>uh</u> July and all of August <1> <> </1> // because uhm I'm trying to sort of go and meet some people out there // and (103) <u>uh</u> <,,> and network with them <2> <,,> </2> // So <,,> (104) <u>uh</u> if you wanted to do something I'm looking for between six hundred and eight hundred words // but I will stretch to a thousand// S2: <1> Right </1> S2: <2> Yeah </2> S2: Right // So quite short // S1: quite short // yeah // S2: [I'm I'm sh-] // [I'm] // you know that that would be quite some nice little challenge really // S1: I want something that's really mysterious and (105) <u>uh</u> and magical <> but real //
103	S1A-096:177.2	
104	S1A-096:178	
105	S1A-096:186	
106	S1A-096:194	S1: [...] There are magazines that are like out there for intellectuals and unfortunately academics <> / There are also (106) <u>uh</u> magazines that are in the New Age / and there are also Christian magazines / but there's nothing that's actually cutting between the three <> // and we're hoping like to do stuff on holistic medicine uhm the new physics // In fact I think there's a number of scientists that we think might be interested in contributing as well // [because <> we went to see] / when Matthew Fox came over we went up to Cardiff and heard him speak on <> uhm <,,> (107) <u>uh</u> what was it now (108) <u>uh</u> Celtic spirituality / and when you think in Cardiff you know there were seven hundred people in this thing <> / and they were over that (109) <u>uh</u> when he came to Saint James // (110) <u>uh</u> <> when he was with Rupert Sheldrake // who's (111) <u>uh</u> a biologist //
107	S1A-096:196.1	
108	S1A-096:196.2	
109	S1A-096:197.1	
110	S1A-096:197.2	
111	S1A-096:197.3	
112	S1A-096:201.1	S1: And you had a scientist up there talking about pilgrimages and (112) <u>uh</u> <laughs> and (113) <u>uh</u> (114) <u>uh</u> the possibility of angels // And then you had a theologian talking about the Big Bang <laughs> // which I thought was brilliant // (115) <u>uh</u> it's been really amazing : the things // [...]
113	S1A-096:201.2	
114	S1A-096:201.3	
115	S1A-096:203	

116	S1A-096:221.2	S2: Yeah // well you sound a lot brighter than the last time I heard you anyway // [so (116) <u>uh</u> ] =
117	S1A-096:224	S1: =[Oh yeah I've had] / (117) <u>uh</u> maybe there may be some way of of changing what I do and getting back into things // [<1> but </1>] S2: <1> Mm </1>
118	S1A-096:228	S2: Mm // Well I mean you know we can have a chat on the twentieth // when we see each other anyway <,,> // (118) <u>uh</u> it is the twentieth : isn't it // she says off the top of her head //
119	S1A-097:6	S1: Hello S2: Hi Leo // <1> sorry </1> // sorry // have I got you out of the bath // or or suchlike // S1: <1> Oh hi there </1> S1: Hi // (119) <u>uh</u> <,,> no // I was just cleaning my teeth // S2: Oh right //
120	S1A-097:24	S2: Well how how's it going // S1: (120) <u>uh</u> well not too bad // I I've finished my exams <1> <,,> </1> // [so the] /
121	S1A-097:27.1	S2: <1> Oh all of them </1> // S1: Well there were only two // so uhm <,,> (121) <u>uh</u> there was one on Wednesday and one on (122) <u>uh</u> Thursday last week <,,> // So the whole course and and <,,> and everything's finished now // <2> [And (123) <u>uh</u> ] </2> =
122	S1A-097:27.2	S2: =What // everything //
123	S1A-097:29	S1: Sorry // yeah // yes // everything // [...]
124	S1A-097:58	S1: Well Bournemouth's quite a major town : isn't it // [Weymouth's think about a f-] S2: M- m- major in what sense // S1: In a sense of size // [and sort of (124) <u>uh</u> ] / and so on <,,> // S2: Y- y- well yeah maybe //
125	S1A-097:73	S2: Is that is that the thing that you'd qualified yourself to be // S1: (125) <u>uh</u> aiming to be // yeah // [...]
126	S1A-097:88	S2: But you kn- you know when we were teaching together <1> a couple of years ago </1> and there was a w- I think a woman called Margaret // <2> is she a course director </2> // S1: <1> Yeah yeah </1> S1: <2> That's right // yeah // yeah // </2> S1: (126) <u>uh</u> yeah // yeah // yeah // S2: Well that was a hellish job <,,> //
127	S1A-097:117	S2: [...]Oh right // S1: So uhm so that's not too bad // [it's] S2: So what's two twenty <,,> uhm <1> <,,> </1> //d- d- deduct about seventy quid off that // S1: <1> Three quarters </1> // [For me the basic rate is (127) <u>uh</u> ] / I'll I'll end up with eleven hundred and sixty-eight <,,> // S2: So that's about a hundred and fifty a week // a bit less //
128	S1A-097:135	S2: So are you going to be in London at all // S1: Well it's very very unlikely // uhm because the other thing I'm doing is try is trying to pass a driving test <,,> // [I've got it] / It's really hell / (128) <u>uh</u> cos I've got to have these exams and like loads of assignment things to do / plus doing driving lessons //
129	S1A-097:175	S2: [uhm well that sounds] / I I mean I don't associate <,,> you with (129) <u>uh</u> you know one of these sort of skills // like like driving // S1: Well it's kind of an essential though / isn't it now
130	S1A-097:198.1	S1: [...] // But <,,> uhm <,,> the thing is I was hoping to sort of <,,> sort of <,,> visit some places around here // and also some some people // [who uhm]0 S2: What // Colchester //

		S1: Yeah // Well (130) <u>uh</u> more than that // Essex Suffolk // S2: But but you're not coming up much from Weymouth //
131	S1A-097:225.1	S1: It's a long one as well actually <,> // Yeah // [I mean I I]
132	S1A-097:225.2	S2: uhm // S- ah yes // No // Well yeah // (131) <u>uh</u> and so I mean where is when is (132) <u>uh</u> when is nine weeks' time // Where does that take us to //
133	S1A-097:229	S1: <unclear-syllable> // nine weeks' time is in the middle of August // [Now that] / [now the middle of August <,> is (133) <u>uh</u> ] / I mean I might still come up to London between now and next week // I don't know //
134	S1A-097:253	S2: And is Japan the one that doesn't like beards or long hair // S1: So they say // Beards // Yes so they say // Yeah // They said (134) <u>uh</u> apparently they don't mind long hair so much // It's beards // [...]
135	S1A-097:295	S2: Yes // But it's something to twiddle in your case // am I right // S1: Yes it is // yeah // yeah // yeah // [uhm except that I I don't really] / I haven't really twiddled the moustache so much since I've had the beard as well // because somehow <,> like aft- attention is taken away from it // sort of thing // I mean when I just had the moustache it was very visible <,> // and very much (135) <u>uh</u> yes something to be twi-/ I'm twiddling it now actually// [...]
136	S1B-045:001	S1: What is the exact mixture do you think that brought forth Derek Jarman the the artist into the world // because you are a are a product of the the famous British public school system for one thing / (136) <u>uh</u> and then you went to art school and then you went into costume and set design // and all this really happened before you started picking up a camera and becoming a movie maker S2: [Well it's the] / you know I I became a movie maker after several years' work in in other areas // (137) <u>uh</u> I I never saw myself as a costume designer // it was just that at the Slade it was a a place where you could work on a communal project // whereas as a painter you worked for yourself by yourself in a corner // In the theatre design department there which I did as my ancilliary course (138) <u>uh</u> you were given say Prokofiev's Prodigal Son some you know a ballet or something like this and we all worked on it together //
137	S1B-045:003	
138	S1B-045:006	
139	S1B-045:014	S2: [...]/ I I mean other people liked it // but (139) <u>uh</u> I felt it was it it failed somewhere // and I was actually rath- rather pleased to have this other one as a bolt hull hole really <,> //
140	S1B-045:015	S1: What what was the exact sort of sequence (140) <u>uh</u> in terms of the creation of your style // [...]
141	S1B-045:026	S2: [...]/ And I I suppose later on in this particular film in The Garden the new film uhm it is quite like (141) <u>uh</u> a dressing up box // and I quite like that // it is rather like kids in the attic // It has (142) <u>uh</u> <,> a slightly naïve quality in that in that area // [I think it's something I] / [on the other hand I look at design films and they they seem to have you know] / I always think of them as the sort of Christmas cake th- that's got the wrapping round it / and all that people have done is put a sort of pretty wrapping <,> // (143) <u>uh</u> uhm the design comes from within my films // in a way that I I don't see in other British cinema actually <,> //
142	S1B-045:029	
143	S1B-045:032	
144	S1B-045:036	S1: [...] // I mean in what way is it related to that // I mean not just in terms of the sort of cliché thing of (144) <u>uh</u> you know gays like to put on disguises <1> of various kinds </1> / but also in the fact that the themes of your movies are very much self-portraits kind of obliquely disguised // uhm I mean you've got the theme of homosexuality <2> <,> </2> (145) <u>uh</u> in the guise of Romans versus Christians in Sebastiane or in the guise of the New Testament in The Garden <,> // S2: <1> Yes </1>
145	S1B-045:037	
146	S1B-045:042.1	
147	S1B-045:042.2	

148	S1B-045:042.3	S2: <2> Yeah </2>
149	S1B-045:042.4	S2: [Well in] / I would have thought that culture is homosexual // I mean it's as simple as that // If you read (146) <u>uh</u> the first speech in Edward the Second / as I have been / cos' we're going to make that into a film / Gaveston's <, > description of the King of England as being interested in plays and masks and poetry (147) <u>uh</u> (148) <u>uh</u> immediately makes him as far as I can (149) <u>uh</u> see in the eyes of most people who read that play effeminate <, > // It is quite interesting that // uhm <, > does it really // of course it doesn't //(150) <u>uh</u> but there is a way of reading it like that //
150	S1B-045:045	
151	S1B-045:050	S2: Well I think (151) <u>uh</u> yes I think that that's been done in The Garden / It's very complicated / I think one of the things about uhm being openly gay in the cinema in this country meant that you had accepted certain limitations // [financially and <, >] // I I mean I I knew after I'd made Sebastiane that there wasn't much chance of going to Hollywood // even if I'd wanted to <, > //
152	S1B-045:056	
153	S1B-045:058	S1: Did did you want to // S2: No I didn't // because <, > I was too interested in my own life // another great conflict in the whole all of this was to you know during those decades the sixties and the seventies to discover (152) <u>uh</u> just how I I should lead my life // If you think it was a whole generation doing that // a whole generation of y- of young gay men and women // And (153) <u>uh</u> that was much more important than any work that I could do // that sort of that coming to terms with my sexuality // [...]
154	S1B-045:085	S2: [...]/ I mean Pasolini was one of the few people who did // and actually turned his camera round onto the young men // (154) <u>uh</u> it's very hierarchical the cinema // [...]
155	S1B-045:098	S1: Wh- wh- what what is collective activity on the Jarman set like // I mean clearly the same sort of rigid hierarch- hierarchies don't obtain // But uhm are you very much the (155) <u>uh</u> the the person in charge // or as I understand from some reports about The Garden uhm you've con- con- confessed that a lot of the film was shot by other people and indeed edited by other people //
156	S1B-045:100	
157	S1B-045:114.1	
158	S1B-045:114.2	S2: I I I think in this particular film I was <, > not in control of it in the way that you would be in the traditional feature film // I think I I intervene in certain areas where I I suppose areas that I I think or under the illusion that I know more // (156) <u>uh</u> s- you know one of the reasons why I was able to make all these low-budget films was the realisation after The Devils that if you put a film in one place <, > you save money and you could do it quickly // [...] // I think it's really s- for other people to answer that // but certainly I mean because I was ill this (157) <u>uh</u> earlier on this year I wasn't able to (158) <u>uh</u> be at the edit really // although (159) <u>uh</u> Peter Cartwright who edited The Garden brought it in to the hospital a couple of times // (160) <u>uh</u> and certainly Si- Simon Turner who did the music and sound track which in that particular type of film is very very important did it after a few discussions // cos I wasn't there at the sound dub //
159	S1B-045:114.3	
160	S1B-045:115	
161	S1B-045:119	S1: (161) <u>uh</u> it's famous that some film makers make the film <, > before they've made the film // like Hitchcock // cos <1> he </1> it's in his head // it's it's on the storyboard // and then he's rather bored when the cameras actually start rolling // other people perhaps make the film in the editing room // someone like Peckinpah probably makes it in the editing room // [when does the film]/ when is the film actually born for you
162	S1B-045:125	
163	S1B-045:128	S2: <1> Yes </1> S2: I have to say with that sort of film you're you're working on it all the time and you never know where it's going to go (162) <u>uh</u> until it's finished // Now everyone says oh well then it's an editor's film / well of course the editing is extremely important / but that isn't quite <, > (163) <u>uh</u> the story // because without a script the various scenarios in The Garden are an invention themselves // [...]

164	S1B-046:022	S1: And what about the music // (164) <u>uh</u> what does it mean to you // music // anything <,> //
165	S1B-046:033.1	S2: [...]which is how both of us used to sing // because nobody could hear us // because of the noise //
166	S1B-046:033.2	S1: (165) <u>uh</u> (166) <u>uh</u> tell me Antonia Byatt about this family who recited Wordsworth over the Sunday lunch // I mean was that a typical scene //
167	S1B-046:119.1	S1: Why didn't you speak to anybody outside class //
		S2: uhm I had no social graces // I had no <,> capacity to make friends I think //
168	S1B-046:119.2	S1: But didn't the teachers take any pride in you // (167) <u>uh</u> (168) <u>uh</u> this this shining scholar //
169	S1B-047:014.1	S1: Now you've been a scholar all your life and now suddenly you're going to be catapulted into public life as as Chief Rabbi <,> // That's not a prospect which all scholars would relish <,> //
170	S1B-047:014.2	S2: It isn't a prospect that I relish// but (169) <u>uh</u> there's a Jewish tradition <,> of <u>uh</u> (170) scholarship in which after a certain stage you just have to get stuck in to public life <,> // and uhm that's the tradition which I reluctantly have to follow //
171	S1B-047:017	S1: Is it going to make serious inroads into your writing <,> //
		S2: (171) <u>uh</u> I should think so // but uhm <,> in the end I think a religious thinker is judged <,> by what he does in life // not just by the books that he writes //
172	S1B-047:019	S1: What does your wife think about the prospect of your becoming Chief Rabbi //
		S2: I think we're worried about the loss of privacy // because we're both very private people // and (172) <u>uh</u> I think that's going to take a great deal of mental adjustment // I think we'll manage it //
173	S1B-047:026	S1: When your appointment was as Chief Rabbi was announced one of your fellow rabbis who as they say preferred to remain anonymous was reported as saying he isn't Jewish enough <,> // Are you Jewish enough <,> //
		S2: Well I hope so // I mean (173) <u>uh</u> I've (174) <u>uh</u> for the last ten or so years (175) <u>uh</u> been involved in training rabbis // Of course I've been a congregational rabbi // (176) <u>uh</u> what I think is happening to the Jewish community in this country and maybe <,> in other diaspora communities as well is that there's a certain inward turn // A century ago one of the great <,> challenges to Jews was to combine <,> Jewish values and a secular education // and somehow to make a creative synthesis out of the two // nowadays we tend to see those things as perhaps incompatible // and therefore you're <,> Jewish only to the extent that you're exclusively Jewish // Well I still feel that as Jews we're challenged to participate in the wider society // to share our values with our society // and in turn to be inspired or provoked by the values that we find elsewhere //
		S1: There was another comment made (177) <u>uh</u> (178) <u>uh</u> about the time of your appointment by Mr. Stanley Kalms / who I think until recently was chairman of Jew's College <1> <,> </1> / and he said that you'd have to walk a tightrope // What did he mean <,> //
		S2: <1> Right </1>
174	S1B-047:027.1	S1: I notice that wrapped up in in all the press coverage about your appointment was a certain amount of comment on the <,> the position taken by your predecessor on a number of political and social issues // One commentator getting his his metaphors slightly mixed wrote about the the virtual canonization of Lord Jakobovits by the Prime Minister <,> // Now would would you anticipate that uhm in the fullness of time you would be similarly canonized // Does orthodoxy in fact incline to the right politically <,> //
175	S1B-047:027.2	
176	S1B-047:029	
177	S1B-047:033.1	
178	S1B-047:033.2	
179	S1B-047:044	S2: No I think that's a (179) <u>uh</u> misidentification // perhaps a half century ago the identification would have been made in completely the opposite direction <,> // [...]
180	S1B-047:054	S1: You didn't always mean to be a rabbi // did you //



181	S1B-047:055	S2: No <,> / I (180) <u>uh</u> started life wanting to be <,> an accountant really // I think that was in those days every Jewish boy's dream // (181) <u>uh</u> f- for reasons which elude me <,> // but then of course I went into secular philosophy // and (182) <u>uh</u> <,> I don't know // I suppose I experienced a very minor kind of conversion at the time of the Six Day War in nineteen-sixty seven // [...]
182	S1B-047:057	
183	S1B-047:124.1	S1: [...] I was much struck by a phrase in in one of your books where you say to lose the argument <,> is as enlightening as to win it // Now I uhm stand lost in admiration before a formulation like that // but it seems to me that (183) <u>uh</u> to be able to <,> indulge in debate on those terms <,> calls either for mm a saintliness of character which few possess <,> or for quite remarkable reserves of self-confidence // or possibly for both <,> //
184	S1B-048:021	S1: You always read poetry presumably // S2: Yes // always // I lectured in English at Durham University / and (184) <u>uh</u> <,> well I studied English all my life / so of course <,> I love poetry //
185	S1B-048:104	S1: Was that first one a great success <,> // S2: Yes // it it was a bestseller // which was (185) <u>uh</u> rather marvelous // S1: And it was a novel of suspense //
186	S1B-048:151	S2: [...] // Yes I think Josephine Tey was a wonderful mixture of <,> knowledge l- learning and also <,> r- romanticism <,> // I remember <,> long long ago telling my publishers that that's who I would like to be like // (186) <u>uh</u> Josephine Tey // who was not only romantic in the proper use of the word <,> but also intelligent // and wrote well <,> //
187	S1B-048:186	S1: [...] // they're always people who find themselves <2> in difficult </2> situations // S2: <1> Yes </1> S2: <2> Yes yes </2>
188	S1B-048:188	S2: [Because I think of the way] / (187) <u>uh</u> as we were saying earlier because I start with the place <,> / and if you choose your place like Delphi or Crete or wherever else I've set my books you you send someone into them / an amateur as you say / (188) <u>uh</u> the situation arises from the place that they're in [...]
189	S1B-049:010	S2: [...] // And uhm they said well I'd just have to proceed with the pregnancy / and (189) <u>uh</u> get it over and done with / and go and start again <,> on the next one // like a baby machine //
190	S1B-049:140	S1: So you set up house together // S2: Yes <,> // After (190) <u>uh</u> it causing a great deal of trouble / because <,> in those days it just was not done <,> for a white woman to go with a coloured man <,> // it wasn't heard of // and especially in the area where I lived // I mean well <,> you were nothing but a tart if you did that <,> // which is something unfortunately he (191) <u>uh</u> seemed to arrange easily for me not so long after I'd met him //
191	S1B-049:144	
192	S1B-049:157	S2: No // no // it wasn't quite that easy actually // it well it was sort of when we had got no food <,> / and (192) <u>uh</u> there was just enough sort of for Neil <,> / he was marvelous with Neil / can't take that away from him / and it was a case of well one of us has got to go out and get some money <,> / and seeing as how you're the better looking out of the two of us I think it should be you <,> / but I hadn't got a clue / [...]
193	S1B-049:188	S1: Sheila I mean you're tough enough to have said no to him // if you'd really wanted to // S2: Yes // I could've said no <,> / I realize now I could've said no <,> / [...] / but saying that he was marvelous with Neil <,> // and (193) <u>uh</u> as you say I could've said no / but I must tell you I was very very frightened of him <,> / he controlled me <,> / sexually / and he also controlled me with fear <,> //
194	S1B-049:198	S1: What // You mean he was violent // S2: I knew he could be <,> // I'd stood up to him on two occasions / and I realized you

		<p>just didn't stand up to him &lt;,&gt; // [he was] / I mean I'm not even five foot and he was six foot four &lt;,&gt; // not that that's any excuse &lt;,,&gt; // (194) <u>uh</u> and I think we're forgetting the biggest factor / I did love him &lt;,&gt; // And I was a very very stupid twenty-two-year-old girl &lt;,&gt; //</p>
--	--	--

## Appendix 2: ICE-GB corpus data studied in the analysis of *uhm* (segmented into TCUs)

no.	ICE-GB code	segmented data (number in brackets indicate the number of the instance)
1	S1A-015:1.1	S1: Well (1) <u>uhm</u> <,> certain maniacs transmit by electronic mail (2) <u>uhm</u> binary encoded pornographic pictures S2: Really S1: But they are encoded // [So it's] / [when it's] / when I receive them it just looks like (3) <u>uhm</u> <,> (4) <u>uhm</u> <,> garbage // (5) <u>uhm</u> lots of dollar signs and things // But (6) <u>uhm</u> you know if you're a software freak you can no doubt (7) <u>uhm</u> <,> find out what they're of //
2	S1A-015:1.2	
3	S1A-015:4.1	
4	S1A-015:4.2	
5	S1A-015:4.3	
6	S1A-015:5.1	
7	S1A-015:5.2	
8	S1A-015:12	S2: [...] Do you send electronic mail to people that you know though // S1: Yes and that I don't know // It's quite marvelous / because because it's (8) <u>uhm</u> it's glo- it's global //
9	S1A-015:22	S2: [...] // but the thing is you always have to write them in a slight code // so people don't know exactly what you are talking about //[and (9) <u>uhm</u> and it's uh] / you know so so that if a financial director comes in and happens to read it then you know you're in the clear // (10) <u>uhm</u> <,> yeah they're really good // really nice // [...]
10	S1A-015:24	
11	S1A-015:35	S1: You send pictures of you // S2: No // I don't send pictures of me // Other people send pictures of me // I <unclear-word> // I send pictures I like // Like like this friend of mine / Andy / I always send a picture of (11) <u>uhm</u> Andy Warhol // I get an Andy Warhol card from the Tate / like a self-portrait thing / and then I fax it through to him // so it just says to and then he's got this picture of Andy Warhol //
12	S1A-015:54.1	S1: Yes // but I I think uh <,> (12) <u>uhm</u> <,> we don't look <,> we don't look as (13) <u>uhm</u> <,> as spruce and uh <,> elegant and uh <,> as pure // especially good-looking like him // S2: Who would be who // S1: (14) <u>Uhm</u> oh I don't know <,> // But (15) <u>uhm</u> <,> I suppose we're not really like either of them //
13	S1A-015:54.2	
14	S1A-015:56	
15	S1A-015:57	
16	S1A-015:77	S2: Is Bim at the Slade now : or not // did he ever get <1> there </1> // S1: <1> At Camberwell </1> S2: Oh Camberwell // S1: Mhm // (16) <u>uhm</u> had an exhibition // which I forgot to invite you to // S2: Really // S1: Yeah// Just didn't sell anything you know <,> // S2: Do you think (17) <u>uhm</u> he's really good //
17	S1A-015:81	
18	S1A-015:87	
19	S1A-015:90	
20	S1A-015:91	S2: Do you think he has it in him // S1: Yeah <,> / [I don't think there's anybody (18) <u>uhm</u> ] / [There's just] / [There's no one] / [There's no (19) <u>uhm</u> ] / Well my impression of the pictures I saw at the exhibition was that there wasn't (20) <u>uhm</u> he wasn't filtered // uh by an actual filter // but he was just painting what he wanted to paint <1> <,> </1> // and the quality was variable // and there was no one standing around say- saying to him you should junk this // [and and well you know] / (21) <u>uhm</u> and and uh so some of it wasn't was not very good quality <,> // S2: <1> Yes </1> S2: Mm
21	S1A-015:96	
22	S1A-015:100	
23	S1A-015:101	

24	S1A-015:132	S2: Why doesn't he write // S1: I don't know // I don't know // I mean I think of myself as a non-smoker // (24) <u>uhm</u> <,,> S2: Well why doesn't why doesn't he // [...]
25	S1A-015:145	S2: Why do you think he doesn't write then // Does he not have the time // or do you just think he's not <unclear-syllables> // S1: I don't know // I mean I think really (25) <u>uhm</u> <,,> it's very difficult to to to produce any form of art unless you are driven <,,> //
26	S1A-015:153.1	S1: Uh <,,> I don't know though // uh he he he's (26) <u>uhm</u> he's (27) <u>uhm</u> working on his language <,,> // S2: How do you mean //
27	S1A-015:153.2	S1: Well he he and I invent lang- lang- languages // Each of us have our our own languages //
28	S1A-015:158	S2: Can you understand each other's languages <,,> // S1: [ (28) <u>Uhm</u> <,,> well] = S2: = [Well is it a re-] / is it recognizable // I mean is it complete nonsense or is it just a language <,,> //
29	S1A-015:171.1	S2: It's another language // S1: Yeah // it's another language //(29) <u>uhm</u> he can't understand my language // because in my language at the moment there aren't (30) <u>uhm</u> there's no the words aren't matched up with the sounds // so he can't actually speak or write (it) <,,> //
30	S1A-015:171.2	S2: So it's just like <unclear-syllable> S1: Yeah well I've written down bits of the grammar and the phonology <,,> // and the semantics // but it's not all merged into a lexicon <,,> //
31	S1A-015:175	S2: Yeah // S1: [Uh but his (31) <u>uhm</u> <1> his </1> ]= S2: = <1> I thought </1> you meant I thought you meant language like a language //
32	S1A-015:190.1	S1: No // no // no // no // It's a it's a completely different language // <1> I mean a real language </1>
33	S1A-015:190.2	S2: <1> Yeah // yeah // yeah // </1>
34	S1A-015:191	S1: As different from English as Turkish is // (32) <u>uhm</u> that was what I meant // when it's like Tur- Tur- Turkish // [and (33) <u>uhm</u> <,,>] / [well <,,> I (34) <u>uhm</u> ] / there's a <unclear-word> // his his <unclear-syllable> his language is called Yathoyua // [...]
35	S1A-015:197	S2: How do you know how to write it // S1: (35) <u>Uhm</u> <,,> well I've learnt it //
36	S1A-015:201	S2: From him // S1: Yeah // S2: Right //
37	S1A-015:204	S1: And (36) <u>uhm</u> so I've I've written a few short texts in it // like the Lord's Prayer : I've <1> translated </1> // S2: <1> Really </1> S1: Yes // (37) <u>Uhm</u> uh <,,> but we don't actually converse in it //
38	S1A-015:216	S2: Well could you think in your language // S1: Oh I can think in it // <1> yeah </1> S2: <1> Can he </1> // S1: No // [because of it's] / it's so unlike normal natural language that (38) <u>uhm</u> I find it extremely difficult to explain to anybody how it works // S2: Mm //
39	S1A-015:227.1	S1: Mm <,,> // [but that's only] / only part of that has changed / I mean basically it's
40	S1A-015:227.2	the same / the phonology and the syntax would be the same <,,> // [(39) <u>uhm</u> <,,>

41	S1A-015:228	and all the (40) <u>uhm</u> / so (41) <u>uhm</u> <,> I mean really the probably the most famous example of what we do is Tolkien's invented languages // Do you know do you know Tolkien //
42	S1A-031:7	S1: He's at home // [But I mean he's] / he (42) <u>uhm</u> <,> he plays the piano all day // and he <,> plays at a ballet class // [...]
43	S1A-031:18.1	S1: And Sue how is it in Muswell Hill // S2: Well Muswell Hill itself I en- can't say I've really (43) <u>uhm</u> explored got to know (44) <u>uhm</u> worked out at all //
44	S1A-031:18.2	S1: Do you have one of those houses with a view // S2: [It] / no it doesn't // S1: Oh //
45	S1A-031:22	S2: (45) <u>Uhm</u> except that it looks out over a bowling green at the back // which is rather nice <,> // and it it faces south // and it has big rooms // and it's a nice house // [...]
46	S1A-031:32.1	S1: Right // So it's not g- gloomy dark and depressing //
47	S1A-031:32.2	S2: No // It's a lovely house //
48	S1A-031:37	S1: Good //
49	S1A-031:39.1	S2: So that's very nice // And we've been (46) <u>uhm</u> concentrating on (47) <u>uhm</u> trying to get i- it organized //
50	S1A-031:39.2	S1: And how're the children // Andy //
51	S1A-031:40.1	S2: The children / oh well I don't think happy // (48) <u>uhm</u> at least Dick seems to be more or less OK // he's being very fair // [...] // and of course it means you know having to work out that you've got your games kit in the right place on Sunday <,> // (49) <u>uhm</u> and Nell's completely anti // and (50) <u>uhm</u> has decided that she's going to be with Bernard // and so <,> that's that's pretty difficult // [because it's not (51) <u>uhm</u> ] / I mean it doesn't actually suit Bernard terribly well // and you know there he is // (52) <u>uhm</u> <,> holding down his job (53) <u>uhm</u> and running a house and running children // [...]
52	S1A-031:40.2	
53	S1A-031:40.3	
54	S1A-031:45	S2: [...] and he also I think wants time and space to himself // to sort himself out // S1: Absolutely // and (54) <u>uhm</u> <,> I mean <,> he'll have to think about who will look after Nell when she gets home from school //
55	S1A-031:53	S2: So I think that's something I'll just have to you know sit back and weather : more or less // and (55) <u>uhm</u> as somebody put it <,> just be uh try and be a mattress // [and] // <laughs // sort of //
56	S1A-031:66	S1: [But it must] // I mean Chris and I have always been incredibly careful <,> // (56) <u>uhm</u> although we kiss sometimes in public // so to speak //
57	S1A-031:122.1	S1: [She must] // She may just feel hurt by you <1> <,> </1> // [and just because you're her mother doesn't] = S2: <1> Mm </1> S2: = Oh I think there's a lot of that // In fact (57) <u>uhm</u> one of the most revealing things has been that (58) <u>uhm</u> she said twice at quite long-spaced intervals but if Dad finds somebody where will my home be then <,> // and <,> to me that seems as though she feels somehow that she's got to earn herself a place in her home // and that she was doing that because she was <,> you know keeping Dad and me together // and when she failed that obviously meant that she was no good <,> // uh in uh at that kind of relationship // so that if there's if there are a couple around uh an older couple you know <,> then (59) <u>uhm</u> <,> you know they they need a sort of cementing bond of a child //
58	S1A-031:122.2	
59	S1A-031:124	
60	S1A-031:139	S2: [Well <,> uh as far uh] // I mean she won't see a psychiatrist // or anything remotely approaching that // <1> (60) <u>uhm</u> </1> S1: <1> No // no </1> //but you could ask someone //
61	S1A-031:144	S2: I could ask someone <,> // Mm // What I what I've said to myself so far <,> doing

		my own bit on it is that I think for some reason or other she's (61) <u>uhm</u> <,> got a feeling and I can you know see reasons for it way back that I didn't <,> ever really love her // [and that when Dick came along] / I mean this this was all bound up with you know Dick coming along //
62	S1A-031:154	S2: [...] I think we should have you know done a bit of digging at that point // sort of you know make her feel particularly special // and find out if she was feeling at all bothered about it // or whether she thought she had to be nice to this child because again she was earning her place // already she was earning her place // and (62) <u>uhm</u> you know while Dickon now can sort of you know say I want this I want that I don't like this I don't like that and he may not get what he wants out of it but at least he feels he can I think she feels too threatened to // I think she feels probably that (63) <u>uhm</u> you know unless I need her for <,> (64) <u>uhm</u> you know making my life happy then I don't really want her // And clearly I've got what I need to make my life happy // so she can't see a place for herself //
63	S1A-031:155.1	
64	S1A-031:155.2	
65	S1A-031:167	S2: So again what do you do about it // Well <,> I do I do things to show that I do want her need her miss her <,> // (65) <u>uhm</u> <,> completely <,> uh <,> gratuitously <,> // (66) <u>uhm</u> <,> which at the moment is difficult // because it's it means sort of saying things like (67) <u>uhm</u> let's go and buy you a new coat // or let's go and have a shopping expedition // just see what we can find // or let's go to a film // or something like that // and kissing her //
66	S1A-031:168.1	
67	S1A-031:168.2	
68	S1A-031:171	
69	S1A-031:174	
70	S1A-031:179	
71	S1A-031:181	
72	S1A-031:187	S1: And what about (68) <u>uhm</u> holidays // just you and her // would she do that // S2: (69) <u>uhm</u> well I <1> su- suggested </1> that but she she didn't want to // and and somebody else said well maybe that was a bit of a mouthful // and you know try just an evening <2> on your own </2> together // S1: <1> like skiing or something </1> S1: <2> or a day trip </2> S2: Yeah // or something like that // (70) <u>uhm</u> <,> so you know so far <laugh> none of those suggestions <unclear-words> S1: The other thing is (71) <u>uhm</u> <,> do you confide in her <,> // does she feel excluded // because you don't <3> exactly confide in her </3> // S2: <3> Well </3> <,> I don't // But then <,> I haven't // I mean I never have <,> // [...] // S1: But you could start slowly // S2: (72) <u>uhm</u> <,> S1: But I agree it's very difficult if you never have // S2: It's well it's also very difficult if if <laugh> communication is <,> virtually nil //
73	S1A-067:67	S2: [My last] / Shall I shuffle the cards <,> // OK <,> // That's it //
74	S1A-067:72	
75	S1A-067:76	S1: Right // (73) <u>uhm</u> S2: I've had two bananas // Is that all right // S1: Yes // Since you're not having anything else you can have two of everything <,> // (74) <u>uhm</u> <,> right there's the significator of you <,> // now then you turn up the first card <,> from there // and you cover your card // (75) <u>uhm</u> <,> that's what covers you // that's the general influence in your environment at the moment <,> //
76	S1A-067:93	S1: Oh yes // What the four different corners of the (76) <u>uhm</u> <1> sort of board </1> // or the wall that you see // S2: <1> (77) <u>Uhm</u> </1> S2: No // They have a suite // yes // [and one of the] / they have (78) <u>uhm</u> Chinese characters circles and bamboo sticks // are the three hands // as it were like // S1: Ah there are three of them // [...]
77	S1A-067:94	
78	S1A-067:97	
79	S1A-067:111	S2: [...] and he's up on a parapet // looking out // S1: Mm // yeah // yes // S2: Well I'm up in my eyrie // (79) <u>uhm</u> <,> contemplating the world at the moment : in a way //

80	S1A-067:116	S2: I wonder why he's holding a globe // S1: [So you're in a quiet (80) <u>uhm</u> <,>] / it says here it looks like the sadness of Alexander amidst the grandeur of the world's wealth // S2: Yes // Mmm // Ah well // I'm broke <,> //
81	S1A-067:141	S2: Mm // Queen of Cups <,> // S1: Ah // well she's the Queen of water // and she's a rather dreamy <,> individual // <,> (81) <u>uhm</u> dreamy and imaginative <,> // see what it says here <,> // all you see is visions // S2: Mm
82	S1A-067:156	S1: Do you think that applies to you at <1> the moment </1> //
83	S1A-067:164	S2: <1> Well it </1> does in a way // cos I am changing // thinking about change a lot // and having done this massage course // which <2> I've got my uh </2> assessment on Sunday // and then I do my anatomy and physiology and all the rest of it <,> // (82) <u>uhm</u> and I am doing a year at the year psychotherapy course at Spectrum next year // <3> [So I am <,>] </3>= S1: <2> Mm </2> S1: <3> Mm </3> S1: =Well that all seems to go with him actually : I think : somehow // (83) <u>uhm</u> <,> S2: and imaginative <,> // I mean I do <,> dream a lot I suppose : in a way //
84	S1A-067:185	S1: Oh // Who's that // S2: J- Jo // Jo Lumley // <1> who's you know doing the Revenger's Tragedy </1> // and (84) <u>uhm</u> I'm being taken to see that <unclear-words> _ to go and see her afterwards <,> // (85) <u>uhm</u> <,>
85	S1A-067:186	S1: <1> Oh yes // oh yes // Mm </1> S1: Mm // well S2: So I don't know why I thought of her // [...]
86	S1A-067:234.1	S2: [Are those] / those a little look like the double snake of the medicine sign // S1: yes // yes // it is <1> <,> </1> // I think // it's also the sign of Hermes : isn't it // the messenger // (86) <u>uhm</u> who (87) <u>uhm</u> is a guide to the occult world <,> //
87	S1A-067:234.2	S2: <1> Ah </1> S2: yes // it could be <,> // the idea of change and new directions <,> //
88	S1A-067:286	S2: present prosperity // S1: Mm S2: well present prosperity is (88) <u>uhm</u> <,> not brilliant // [I'm not in the bl-] / I'm not in the red // [but I'm (89) <u>uhm</u> ] /
89	S1A-067:287	S1: well perhaps you're giving too much to other people // giving too much away //
90	S1A-067:320.1	S1: Ah that's the Great Mother // or one of the forms of the Great Mother <,> // S2: There's a <,> an atmosphere of (90) <u>uhm</u> consolidation and uh <,>] S1: Yes // agreed //
91	S1A-067:356	S1: yes // all right // one one goes down here now // S2: d- does this top half of the deck (91) <u>uhm</u> disappear altogether // S1: yes <,> // we're only interested in that //
92	S1A-067:366	S1: oh // ooh // that's a that's an interesting <1> one <,> </1> // [it's (92) <u>uhm</u> ] = S2: <1> is it </1> //
93	S1A-067:369	S2: = it looks more energetic than the rest of them // S1: it's (93) <u>uhm</u> it's your opinion of yourself that one : isn't it //
94	S1A-080:39	S1: Well I don't need a man at the moment <,> // S2: Why // S1: Cos I've got one // S2: Oh of course <laugh> <,> // [well I think I don't know that we (94) <u>uhm</u> should go and do it at the wh-] / where did you suggest //

95	S1A-080:98	S2: Shut up // oh well I'm sure the thing the thing about nose-picking is that nobody ever admits to doing it // S1: I have done it // but just not not when there's been anybody else around // S2: Well no // but that's the joke about traffic jams // when I saw my friend picking her nose in a traffic jam <,> I (95) <u>uhm</u> I thought no no that's not the place to pick your nose <,> //
96	S1A-080:264	S1: <laughs> Yes // not Billy's bald head // S2: God I'm frozen // I must go and get a cardigan // S1: [And <1> (96) <u>uhm</u> </1> ] = S2: = <1> Are you </1> cold // S1: Well I'm a bit chilly // whose mind //
97	S1A-080:278	S1: Nick's body Billy's money Murray's mind // S2: Yes // Oh Murray's so sweet / [He rang me up when (97) <u>uhm</u> <,>] / yeah he rang me up when uh Julie was away / and we had a chat / oh he was so funny / [...]
98	S1A-085:18	S2: You don't remember it // you were talking about this guy at Embankment tube station / who stood there in in skintight jeans / [and <,> he was] / apparently he was really extremely gay / [and you were going] / apparently he would (98) <u>uhm</u> <,> say choo choo choo or something <,> //
99	S1A-085:99	S2: Right look there's nothing wrong with a bit of dust // it's (99) <u>uhm</u> no worse than what we breathe in every day <,> // you know Mum is buying a new car with (100) <u>uhm</u> for diesel <,> // which has no lead in it at all <,> // so she's finally turned green // you'll be very pleased to hear // after talking about lead
100	S1A-085:100	
101	S1A-085:125	S2: Oh with a bicycle bit // oh <1> right </1> // S1: <1> no // no // </1> no// he pulls them // he runs and holds <,> the (101) <u>uhm</u> <,> holds two poles // and you sit in the carriage // [...]
102	S1A-085:203	S2: [...] She keeps that sort of goldfish jar of cigarette butts in water <,> // she gave up smoking and then when she started again she <1> kept them </1> // S1: <1> apparently that's a pretty good </1> idea // S2: Yeah but she kept it in her bedroom even after she started <unclear-words> started (102) <u>uhm</u> <,> smoking again <,> // it was a little bit pointless <,> //
103	S1A-085:266	S2: Right let's see how many words we can think of beginning with D <,> // I'll start // destitute <,> // S1: drug <,> // S2: demon // S1: drink <,> // S2: (103) <u>uhm</u> <,> dyslexic <,> // S1: dirigible <,> //
104	S1A-085:279	S2: donkey // it's your turn <,> // S1: dick <,> // S2: devil <,> // S1: (104) <u>uhm</u> <,> dick // S2: you did that one // you can't do the same one twice //
105	S1A-085:292	S2: If you stick a de in fr- in front of something // like with dehumanize <,> // S1: OK // I'll try that <,> // OK <,> //(105) <u>uhm</u> <,> dehumanize <,> // S2: [your imagination is] / I love it when you work your brain //(106) <u>uhm</u> <,> drunk <,> //
106	S1A-085:295	
107	S1A-085:304	S2: detoxicate // detoxicate <,> // S1: detoxicate <,> // I don't know if that exists // I don't know if I can let you have that <,> // S2: All right // (107) <u>uhm</u> <,> it does exist // S1: No // it doesn't //
108	S1A-085:328	S2: [...] Shut up // you haven't been able to come up with much // dick twice // I mean



		Christ almighty <,> // S1: (108) <u>uhm</u> damnable <,> //
109	S1A-085:341	S2: I love it when you act like a four-year-old // Go on // S1: depreciate // that's not a four-year-old word <,> // that's a good old thirty-two-year-old stockbroker word // S2: <unclear-words> two-year-old <unclear-words> four-year-old //(109) <u>uhm</u> <,> diminished <,> // but is it your go // or is it my go // it's my go <,> //
110	S1A-091:2	S1: No // we were talking about (110) <u>uhm</u> <,> that that that man at college // [who told you] / who who was saying all these <1> outrageous things to you </1> // S2: <1> Oh god // oh my god </1> // oh my god // that got worse // I don't believe it // it was so embarrassing //
111	S1A-091:21	S1: Why not //
112	S1A-091:23	S2: Well there's <laugh> there's nowhere to go <,> // well no // [he said (111) <u>uhm</u> / <,> what did he say / [he said <,> (112) <u>uhm</u> <,>] /
113	S1A-091:27	S1: No but it was the one that the one that made me laugh // [I mean I remember when he] / I mean I remember all the rude ones // but there was the one in the canteen : that he said to you //
114	S1A-091:30	
115	S1A-091:34	
116	S1A-091:35.1	S2: <unclear-words> // [what about the (113) <u>uhm</u> <laugh>] / yes / he said he was a palm-reader / and he said he said (114) <u>uhm</u> well you're obviously <laugh> slow <,> / and <laughter> he said you're obviously terribly slow / very imperceptive / and you've got an appalling memory <laughter> <,> / my hackles are really rising by this time / and (115) <u>uhm</u> and he said and if you want to pass your exams you'd really better start now /and I and I said oh yes and (116) <u>uhm</u> and <,> where did you (117) <u>uhm</u> learn to do this then / palm-read / he said from a book / I said well you're obviously not very good at it : are you <,> / yeah / [and then he said (118) <u>uhm</u> / and then I said I was going to aerobics / [he] / and he said <laugh> why / [...]
117	S1A-091:35.2	
118	S1A-091:39	
119	S1A-091:53.1	S2: [...] Oh my god // it was so embarrassing / [there was this (119) <u>uhm</u> <,>] / I was talking to this this guy at college / and (120) <u>uhm</u> he's really really really boring / and he always always says the same thing <,> / [and (121) <u>uhm</u> <,> the other] / and I must've told Albie this guy I was telling you about that he was boring / [although I I don't] / because he knows him / I don't know why I told him that / [but (122) <u>uhm</u> <,> he said to him the other day] / we had exams the other day <,> / and (123) <u>uhm</u> <,> he said Jenny I think you're really nice / I said oh I really that's really sweet of you / thank you / and he said (124) <u>uhm</u> <,> do you think I'm nice <,> / [and I said (125) <u>uhm</u> ] <1> <,> </1> / yes / this is the one that is really boring / [...]
120	S1A-091:53.2	
121	S1A-091:54	
122	S1A-091:56	
123	S1A-091:57	
124	S1A-091:60	
125	S1A-091:61	S1: <1> This is the one that you think is boring </1> //
126	S1A-091:76	S2: And I said oh really <,> / and (126) <u>uhm</u> and he said <,> well what do you think Jenny / and I thought oh god /
127	S1A-091:111	S2: <laugh> oh it's a shame the other thing didn't come out <,> // S1: [(127) <u>uhm</u> I've just been pa-] / I've just been handed something to give you here // S2: oh what // my tape //
128	S1A-091:148	S1: stick to your guns // stick to your gun girl <,> // [anyway (128) <u>uhm</u> <,>] = S2: = did you want me for something // S1: no // no // no // I was just phoning up for one of our silly conversations really //
129	S1A-091:159	S1: and there's people in the room // so you can't talk anyway // S2: (129) <u>uhm</u> yes // well there isn't any more actually <laugh> // S1: have they gone now //
130	S1A-091:162	S2: [there's (130) <u>uhm</u> ] / no // well yes // no // there's people here // [the] / there's no news // I meant to say <laugh> //

		S1: Oh I see <laugh>
131	S1A-091:175	S2: Mm // [So nothing very] / in fact nothing has happened to me at all over the last week <,> // [(131) <u>uhm</u> I had my exams last <,>] / [I had the] / well I had three tests before I came up here //
132	S1A-091:186	S1: [...] I've got the French written paper on Thursday // which I'm not looking forward to at all // S2: Oh // S1: [But (132) <u>uhm</u> ] = S2: = But I'm sure I had something to tell you <,> // I can't think what it was though //
133	S1A-091:264	S2: I'm just going to put it in a Walkman and stick it over his ears <laugh> // S1: Oh I see // S2: Yeah // next lesson // no (133) <u>uhm</u> // well I don't know // he might not want to : might he // [...]
134	S1A-091:284	S2: Oh <laugh> // why // S1: Oh nothing // S2: Oh OK then //
135	S1A-091:289	S1: (134) <u>uhm</u> so that's it then : is it // that's all the gossip for this week // S2: I'm afraid I haven't got any // nothing juicy anyway // [ (135) <u>uhm</u> or d-] / well I could try and think of some // [...]
136	S1A-091:316	S2: Oh that's terrible // give me a warning next time <,> //
137	S1A-091:318	S1: No // that was fine // (136) <u>uhm</u> no // no // there's nothing // I mean you know I'm still (137) <u>uhm</u> <,> living with this man // and it's all going well //
138	S1A-092:48	S1: Most delightful interview // [but uh <,> ] / and they obviously didn't ask me to read / and I thought well that's <,> you know what a joke / (138) <u>uhm</u> <,> but then they said listen we need to you know <,> decide very promptly / [so] / and by the time I got home they'd already phoned my agent / So my agent phoned me and said listen they they talked at length <,> / and I think <,> they were just sort of <,> making up for <,> for the embarrassment of having to <,> ask me <,> such an experienced artiste to come in for those two lines / but (139) <u>uhm</u> they were saying <,> uh <,> they want to enhance these (140) <u>uhm</u> <,> sce- the scene of crimes officer department <,> / uh into the plot / and therefore may well be asking me back later in the year <,> // to do some more work //
139	S1A-092:52.1	
140	S1A-092:52.2	
141	S1A-092:94	S2: So what did you do today <,> // S1: (141) <u>Uhm</u> = S2: = Swotting away //
142	S1A-092:98	S1: Yeah // I went into the library this morning // and swotted <,> // and (142) <u>uhm</u> <,> and then I strolled round the library this afternoon // [...]
143	S1A-092:152	S1: [...]Interesting one // it's putting a word and putting an expletive in the middle // [that's uh <1> <,> </1>] / yeah // S2: <1> yeah </1> // S2: [Ah but (143) <u>uhm</u> ] <,> / Tuesday I'm going down to the theatre in Farnham <,> // see a show <2> Jamai- </2> Jamaica Inn <,> //
144	S1A-092:163	S1: <2> Oh </2> S1: <laugh> Jamaica // S2: Jamaica // I'm sorry // I can never say that word without thinking of Jamaica no she came by herself <,> // S1: <laugh> S2: (144) <u>Uhm</u> what is that // [that's] / [is that <,>] / I know that title very well but I can't place it <,> //

145	S1A-092:253	S1: Oh hold on a minute <,,> // hello <,> // S2: hello //
146	S1A-092:254	S1: yes <,> // (145) <u>uhm</u> <,,> // S2: [well listen <1> (146) <u>uhm</u> </1>] // S1: <1> I feel </1> silly now // S2: No // that's OK //
147	S1A-092:259	S1: <laughs> S2: (147) <u>Uhm</u> I'd better leave you to your <,,> conversations //
148	S1A-092:300	S2: You didn't mention a name // but <1> you mentioned him </1> // S1: <1> Yes I've </1> had my eye on him for for for some time <,> // had my eye on him since the beginning of the year really // but (148) <u>uhm</u> <2> <,,> </2> decided that I really ought to be wise for months and months and months and <,> not uh <,> ask lecturers out <,> //
149	S1A-092:304	S2: <2> Right </2> // <chuckles> // so much for that <laughs> // S1: So (149) <u>uhm</u> <,,> eventually a couple of weeks ago I plucked up my courage and <,> rang him up / and said (150) <u>uhm</u> <,,> well first of all I said listen you've got to pretend you don't know who it is on the telephone / and he said but it's Isobel / I said no no no can you pretend you don't know who it is <,> / I said I've just phoned you to ring you up and discuss the ethics of whether students should invite their lecturers out to see films <,> / and he said that sounds like a good idea <,> // <laughs>
150	S1A-092:305	S2: <unclear-syllables> <laughs>
151	S1A-092:313	S1: [So (151) <u>uhm</u> <3> <,,> </3> so my uh] / [so I was uh <,,>] / bravery won the day //
152	S1A-092:321	S2: <3> <laughs> </3> S2: So this is for Yibin then // S1: sorry // S2: This uh linguistics stuff <,,> // S1: Yes // yes // [That's what he <,>] / That's how he earns his keep <,,> // doing linguistics <,,> // [so (152) <u>uhm</u> ] = S2: = So he's going to punch the details into a <unclear-word> screen <,> //
153	S1A-093:28	S1: Mm // are you actually going to bother getting a job <,> // S2: Well not for the next two weeks probably [...] // S1: Mm <,> S2: Probably // [but I've got to do like] / (153) <u>uhm</u> I might go back to Cambridge early or something // because I've got to write an extended essay <,> //
154	S1A-093:52	S1: Yes I mean the trouble is I really thought oh you know brilliant when I've finished my deadline and everything <,> (154) <u>uhm</u> I'll go out loads <,> // but all I want to do is completely collapse // <1> and just flop in front </1> of the telly // which is really boring <,> // I don't know // I think sort of intellectually completely boring you know <2> <,> </2> // I mean I went out with Joey and Lee on Monday evening //
155	S1A-093:62	S2: <1> You're really tired </1> S2: <2> Yeah </2> S2: Oh how's Lee // S1: Well he's OK // (155) <u>uhm</u> <,> <3> uh </3> S2: <3> Did he </3> get that job about teaching English //
156	S1A-093:71	S1: Yeah I mean he's OK // he's just always so depressed and and and <,> (156) <u>uhm</u> <,> you know complaining and <,> miserable about everything : yeah <,> // and he really likes Joey but he thinks Joey's a complete waster // so that evening was really strange you know //
157	S1A-093:80	S1: [And she is she's] / she spent the entire evening convincing her that Us Uts was

158	S1A-093:88	desperately passionately in love with her <,> // [and it was (157) <u>uhm</u> ] /
159	S1A-093:89	S2: fatal <laugh> S1: [Well I mean it's] / the thing is he was supposed to be coming over for her birthday / coming back for her birthday <,> / [...]// So she managed to get hold of his Dad / who's living over here / you know in Montague Terrace <,> / And uh <,> he said oh yes you know Uts is (158) <u>uhm</u> he w- he wasn't able to get a standby flight / and (159) <u>uhm</u> <,> it's really difficult and he'll have to wait till he goes via South Korea or something // And she swallowed every word of it // and I mean the thing it's so obvious that he'd uh that he was putting her off <,> : you know //
160	S1A-093:120	S1: yeah // [well I'm] / I mean it's all quite depressing really / because I phoned Harriet up to go out / and I don't know if Mum told you but she told me her Mum has cancer <,> // so I mean we're still going out // [but I mean that's not exactly <,> (160) <u>uhm</u> you know] //
161	S1A-093:130	S1: <1> You </1> don't want to come to a play tomorrow : do you // S2: <1> <unclear-words> </1> S2: Oh what //
162	S1A-093:136	S1: [I'm going to (161) <u>uhm</u> <,>] / Fiona as you know does set design / she's just finishing this course / and she's done the set design for a play up in Islington // in a in a pub called the Red Lion // S2: Oh I think I've been there actually // S1: You might've done // it's quite a <,> (162) <u>uhm</u> well-known one //
163	S1A-093:175	S2: Oh that's brilliant about Karen though //
164	S1A-093:181	S1: <1> It's lovely isn't it </1> //
165	S1A-093:187	S2: <1> Say congratulations </1> from me // S1: Yes I will do // yes I will do // <2> (163) <u>uhm</u> </2>
166	S1A-093:188	S2: <2> Only </2> Mum's sending a card from all of us <4> anyway </4> S1: <4> Yes </4> // well that's fair enough // yeah // I'm seeing Karen // well Karen and Andre are coming round for dinner after Simon's exams // so it's sort of <,> something like two weeks on Thursday or something // <5> (164) <u>uhm</u> </5> S2: <5> yeah </5> // yeah // oh that's the other thing // I suppose you've got to wait for his exams // but yeah there you go // S1: yeah <,> // S2: [But it was weird actually (165) <u>uhm</u> at King's / because in the third year there's there's] / you know that guy who died / the rower <,> / and he was (166) <u>uhm</u> very pally with Jim the third year classicist <,> and another girl / a classicist called Camilla / a third year / and they said that they couldn't celebrate at all // [...]
167	S1A-093:199	S1: I think you would feel like that : wouldn't you // S2: Yeah // [and they a they] / all of them didn't feel like doing their exams // cos he died like the night before a lot of uh the exams <,> // and (167) <u>uhm</u> they said they they just wanted to go and sort of talk about that // [...]
168	S1A-093:208	S1: Were they just first years <,> //
169	S1A-093:209	S2: (168) <u>Uhm</u> no // no // [The] / [these] / all the people who were friends are in the third year // because he was a (169) <u>uhm</u> a graduate //
170	S1A-093:233	S1: She's bound to have lost it // S2: yeah // so I'll phone her // and (170) <u>uhm</u> <,> yeah // because we've got to sort that out //
171	S1A-093:238	<i>S1: &lt;talks off the telephone, with her boyfriend/husband&gt;</i> S1: He's no good // honestly // S2: No // and apart from that I mean my results are supposed to have come out today <,> // (171) <u>uhm</u> <i>S1: &lt;talks off the telephone, with her boyfriend/husband&gt;</i>

		S1: <commenting on what she has discussed with boyfriend/husband> crap for dinner // [...]
172	S1A-093:249	S1: Oh it's so funny / Simon's on this revision course / and I asked him to get some veg on the way home from Shepherd's bush market / because he's right there <1> <,> </1> / and it's much cheaper than the local shop / and we'd run out / as you can imagine / And (172) <u>uhm</u> <,> so what does he do <,> / he comes back without any / I said oh typical you didn't go / And he said oh no I did go <,> / but they were they were wrapping them up in (173) <u>uhm</u> newspaper / and he didn't have a plastic bag <,> / so he didn't get any <laughs> // #
173	S1A-093:252	
175	S1A-093:265	S1: Oh Delia's funny as well / [she (174) <u>uhm</u> <,>] // you know the painting she did for me for my birthday // the green thing //
175	S1A-093:270	
176	S1A-093:272	S2: Oh yeah // that's really nice //
		S1: Yeah / she phoned me up (175) <u>uhm</u> yesterday when I was at work / and she said oh you know panic panic panic / I've got this (176) <u>uhm</u> exhibition coming up and I haven't got enough time to do enough paintings / can I borrow yours <laugh> //
177	S1A-093:280	S2: excellent // yes because it's quite big : isn't it // if I <1> remember </1> //
		S1: <1> No </1> // it's not // it's (177) <u>uhm</u> maybe a foot square at the most // [...]
178	S1A-096:19.1	S1: So what have you been up to //
179	S1A-096:19.2	
180	S1A-096:19.3	
181	S1A-096:19.4	
182	S1A-096:26.1	
183	S1A-096:26.2	
184	S1A-096:27.1	
185	S1A-096:27.2	
186	S1A-096:29.	
187	S1A-096:30.1	
188	S1A-096:30.2	
189	S1A-096:30.3	
190	S1A-096:30.4	
191	S1A-096:31	
192	S1A-096:32	
193	S1A-096:42	S2: (178) <u>uhm</u> oh God (179) <u>uhm</u> buried under a couple of (180) <u>uhm</u> playwriting commissions // (181) <u>uhm</u> <,> which are <,> are driving currently driving me a bit potty <laugh> // <1> [well one of them is] </1>
194	S1A-096:43	
195	S1A-096:49	
196	S1A-096:53.1	
197	S1A-096:53.2	
198	S1A-096:53.3	
199	S1A-096:56	
200	S1A-096:58	
201	S1A-096:61	
202	S1A-096:62	
203	S1A-096:63	
		S1: <1> oh right </1> // what are they about //
		S2: sorry //
		S1: what are they about : I mean //
		S2: (182) <u>uhm</u> well (183) <u>uhm</u> one of them's a a a play // [which is] / the one I'm sort of trying to finish at the moment is a a play for a company called Quicksilver // it's actually the second play I've written for them (184) <u>uhm</u> <,> // which is kind of based on the (185) <u>uhm</u> the Browning poem the Pied Piper // It's sort of about dreams and possibilities inside a mountain // but uh the problem is I've kind of I started with fantasy // and (186) <u>uhm</u> and uh I don't think I'll ever try that again // I I think unless you start with (187) <u>uhm</u> with something real <,> it's (188) <u>uhm</u> you you you end up having too many possibilities open to you // [and it just (189) <u>uhm</u> ] / it becomes a a bit of a kind of a <,> (190) <u>uhm</u> a fantastically complicated mind maze really <,> // but (191) <u>uhm</u> anyway I'm uh just about uh seeing the light at the other side at the moment // but it's taken me a hell of a long time <,> // [and the other is (192) <u>uhm</u> ] / I don't know if you've heard of a company called Monstrous Regiment //
		S1: No // I don't know much about the theatre //
		S2: The theatre // right // Well I'm I'm writing a a play for them // so (193) <u>uhm</u> <,> so I've kind of <,> been doing that // I did I did a bit of storytelling for (194) <u>uhm</u> a shadow puppet theatre company from September to December // which was really nice <1> <,> </1> // and uh <,> that's all I've been doing really <2> <,> </2> // been pretty busy <,> // how about you //
		S1: <1> Mm </1>
		S1: <2> Mm </2>
		S1: S2: [...] // been pretty busy <,> // how about you //
		S1: Well I've <,> I've changed what I've been doing (195) <u>uhm</u> <,> / I kind of kind of see myself as a writer these days / uh I don't know exactly what I'm doing exactly / uh it's very difficult to explain / I I joined a group about (196) <u>uhm</u> <,> over a year ago now <,> / uh called A School for Prophets <,> / [which is (197) <u>uhm</u> ]

204	S1A-096:64.1	<,>/(198) <u>uhm</u> Prophets not Profits <,> / uh which is based in Warmsley / which is quite near <,> us in Chatham / it's run by a catholic priest called Bert White <,> / And the group's all <,> ecumenical / (199) <u>uhm</u> mainly catholic because it was started by Bert / but it's got all kinds of other things connected to it / and we've been examining well uh things like Matthew Fox / who's a theologian who's been incredibly liberating <,> / and (200) <u>uhm</u> there's a whole sort of movement happening in this country called creation-centred spirituality / It's not a cult or anything / [...] // because there are lots of things I think that are in in the New Age that are actually quite sensible // (201) <u>uhm</u> and quite positive // but there's also a lot of nutty stuff as well <,> // (202) <u>uhm</u> <,> and that's basically it // But we've we've had (203) <u>uhm</u> I think a very interesting year // we meet every Tuesday // (204) <u>uhm</u> and last year we had people like Sidney Carter come down and do some stuff with us // (205) <u>uhm</u> We had (206) <u>uhm</u> <,> stuff on liberation theology and stuff on feminist theology <,> // and uh <,> we had Julie Felix come down and do some stuff for us a few weeks ago
205	S1A-096:64.2	
206	S1A-096:65	
207	S1A-096:83	S2: D'you know why the name's really familiar // why <1> do I know it </1> // S1: <1> She was </1> a folk singer // in the sixties // she did it on television // she had long sort of dark hair <,> // and she was half Mexican <unclear-words> // [and she came down and (207) <u>uhm</u> <,>] / oh it's been a quite an interesting in- quite an interesting year //
208	S1A-096:89	S1: And we're now going into a second phase which is starting with a magazine <,> // we want a magazine that uh with quite a wide brief on it // but the idea is that (208) <u>uhm</u> <,> it's a magazine that's going to be a kind of forum for a lot of things that are happening // [(209) <u>uhm</u> we've got (210) <u>uhm</u> ] / uh it's called Paradigm Shift // because we will be examining the paradigm shift that's actually occurring <,> // uh so we've got a whole lot of uh clergy scientists poets <,> artists all kinds of people involved in it <,> // [and (211) <u>uhm</u> it's going to] / one thing it's going to do is to foster a whole sort of new generation of grass roots creative intellectuals // which would be partly one of the things I see happening <,> // but also to examine paradox // and (212) <u>uhm</u> <,> I've been put in charge of editing this magazine // and I'm preparing the first issue at the moment // uh we've got so far we've got an article by John Medcalf // who's (213) <u>uhm</u> <,> written a lot of books about Nicaragua and Peru // He was in N- both of those countries as a Catholic priest // uh we've got (214) <u>uhm</u> a woman called Caitlin Matthews / who is a pagan priestess a K- of a Celtic <,> persuasion // (215) <u>uhm</u> she's written seventeen books so far // on Celtic mysteries uh Arthurian legends (216) <u>uhm</u> <,> all kinds of stuff feminism the feminine aspect of God // Her last one was called Sophia the Feminine Aspect of Godness // [It's] / She's uh published by Harper Rowe // They're quite amazing sort of stuff // Anyway she's uh she's given us an article for free // which is good // So Sidney Carter's promised something Julie Felix is going to write something for us // (217) <u>uhm</u> and also Richard Solly from CAFNNA // [who we] / who who is loosely connected with the group but may become closer to it // [he] / that's Catholic Action for Native North Americans <,> // [he does a lot of work for] / [...]
209	S1A-096:90.1	
210	S1A-096:90.2	
211	S1A-096:93	
212	S1A-096:94	
213	S1A-096:96	
214	S1A-096:98	
215	S1A-096:99.1	
216	S1A-096:99.2	
217	S1A-096:105	
218	S1A-096:110	S1: So it's it's sounding quite interesting // (218) <u>uhm</u> <,> uh I want to get some more <,> uh kind of right brain stuff into it // [it was you know] <,> / literature and poetry <,> // [(219) <u>uhm</u> <,> wou-] / do you think maybe in the future you could envisage doing a story for us // S2: (220) <u>uhm</u> yeah // possibly // S1: Yeah but <,> (221) <u>uhm</u> now at the moment we're not paying // but when Bert gets back I'm going to ask him if I can have like some smallish fees // so that we
219	S1A-096:112	
220	S1A-096:134	
221	S1A-096:135	
222	S1A-096:139.1	

223	S1A-096:139.2	can at least pay writers // because they are struggling // [and what have you <,>] / I was thinking about thirty or forty quid // or something like that // yeah <,> // the kind the type of story that I'm actually looking for / (222) <u>uhm</u> <,> I like things that are real <,> // (223) <u>uhm</u> and I like things that are quite strong <,> // but <,> I like things also with a fantastic element to them <,> // so they stretch the imagination a bit <,> // [...]
224	S1A-096:154	S2: I could have a go //(224) <u>uhm</u> <,> yeah // I mean you know I'd I'd I'd love to I'd dearly love to (225) <u>uhm</u> you know to be spending time writing poetry and fiction // and kind of this last year's been been (226) <u>uhm</u> <,> kind of commissioned work //[which which in one sense is] / I guess on paper looks great // the fact that I'm actually getting getting work like that // but (227) <u>uhm</u> in another sense it's very sort of stultifying // because (228) <u>uhm</u> you tend to sort of write for other people and and and do very much what they want // which(229) <u>uhm</u> is is very difficult // so I mean you know (230) <u>uhm</u> I'm I'm really hoping to clear some space in my life very soon // to be able to do all that // [so (231) <u>uhm</u> <,> (232) <u>uhm</u> at the moment I can't I can't really sort of (233) <u>uhm</u> ] / yeah / I've kind of got to to just (234) <u>uhm</u> reach reach reach the top really // of all of this //
225	S1A-096:156	
226	S1A-096:157	
227	S1A-096:160.1	
228	S1A-096:160.2	
229	S1A-096:160.3	
230	S1A-096:161	
231	S1A-096:162.1	
232	S1A-096:162.2	
233	S1A-096:162.3	
234	S1A-096:164	
235	S1A-096:175	S1: [...] I'm looking really towards <1> the autumn </1> // S2: <1> to the autumn </1> // yeah // well that would be about the right time for me // [so (235) <u>uhm</u> ] =
236	S1A-096:177.1	S1; = We've been doing it sort of quarterly // but the thing is (236) <u>uhm</u> <,> Ann and myself hopefully are going to be in America in <,> the end of Ju- uh July and all of August <2> <,> </2> // because (237) <u>uhm</u> I'm trying to sort of go and meet some people out there // and uh <,> and network with them <3> <,> </3> // so <,> uh if you wanted to do something I'm looking for between six hundred and eight hundred words // but I will stretch to a thousand //
237	S1A-096:177.2	
238	S1A-096:185	
239	S1A-096:195	S1: [...] and we're hoping like to do stuff on holistic medicine (239) <u>uhm</u> the new physics // In fact I think there's a number of scientists that we think might be interested in contributing as well // [because <,> we went to see] / when Matthew Fox came over we went up to Cardiff and heard him speak on <,> (240) <u>uhm</u> <,> uh what was it now uh Celtic spirituality / and when you think in Cardiff you know there were seven hundred people in this thing <,> / and they were over that uh when he came to Saint James // uh <,> when he was with Rupert Sheldrake // who's uh a biologist //
240	S1A-096:196	
241	S1A-096:206	S1: [...] My own career of course has gone right down the tube // but there you go // [(241) <u>uhm</u> <,> but then I'm] / I've I've learnt to let that go a bit // because there's nothing you can do about it // [so]
242	S1A-096:231	S2: Well I mean you know we can have a chat on the twentieth // when we see each other anyway <,> // uh it is the twentieth : isn't it // she says off the top of her head // S1: [Let's] / wait / (242) <u>uhm</u> no it's the fifteenth : isn't it //
243	S1A-097:26	S2: Hi // Well how how's it going // S1: Uh well not too bad // I I've finished my exams <,> // [so the] / well there were only two // [so (243) <u>uhm</u> ] uh there was one on Wednesday and one on uh Thursday last week <,> // [...]

244	S1A-097:45	S1: [...] Where is Weymouth // S2: (244) <u>uhm</u> it's near do you know Bournemouth // it's in Dorset <,> // <1> sort of </1> south-west <,> ish // S1: <1> Yes </1> // right // it's pretty dreary //
245	S1A-097:51	S2: (245) <u>uhm</u> S1: Well my parents live in Bournemouth for a couple of years // S2: Oh really // S1: Mm //
246	S1A-097:62	S1: But I mean I don't know about sort of culture or whatever // I've never been there // [but (246) <u>uhm</u> yeah I mean I suppose] / I dunno don't know // [...]
247	S1A-097:70.1	S1: And how long are you there for //
248	S1A-097:70.2	S2: Well seven weeks // [(247) <u>uhm</u> it's] / [but it's good though because it's
249	S1A-097:70.3	(248) <u>uhm</u> (249) <u>uhm</u> ] / I've got a job as a course director <,> // which is a sort
250	S1A-097:71	of step up // you know from teacher <,> // and (250) <u>uhm</u> so that's quite good you know //
251	S1A-097:94	S1: Yeah // it's very daunting // yeah <,> // but it's a real first step on the ladder sort of thing //
252	S1A-097:95	S2: But (251) <u>uhm</u> d-d- do you get paid decently //
253	S1A-097:97.1	S1: (252) <u>Uhm</u> the pay's OK // [I mean it's in terms of <,>] / well the the basic pay is
254	S1A-097:97.2	(253) <u>uhm</u> about fifteen hundred and fifty for seven weeks <,> // (254) <u>uhm</u> which works out at about just over two twenty a week : right // so it isn't //
255	S1A-097:111	S1: So (255) <u>uhm</u> so that's not too bad // [it's] S2: And so what's two twenty <,> (256) <u>uhm</u> <1> <,> </1> / d-d- deduct about seventy quid off that //
256	S1A-097:114	S1: <1> Three quarters </1> // S1: [For me the basic rate is uh] / I'll I'll end up with eleven hundred and sixty-eight <,> //
257	S1A-097:131	S2: And you you're going there when // in a couple of weeks //
258	S1A-097:134	S1: [In a couple] / well yeah (257) <u>uhm</u> yeah a couple of weeks // just over a week and a half in fact //
259	S1A-097:137	S2: So are you going to be in London at all //
260	S1A-097:139	S1: Well it's very very unlikely // (258) <u>uhm</u> because the other thing I'm doing is try is trying to pass a driving test <,> // [I've got it] / It's really hell // uh cos I've got to have these exams and like loads of assignment things to do // plus doing driving lessons // And I've hardly had a minute to sort of look at highway codes and driving manuals and stuff // but it's getting better // and I've got a test next week // so I kind of got to stay (259) // <u>uhm</u> hang around for that // S2: So you you can drive : can you // S1: (260) <u>Uhm</u> I'm pretty getting <,> getting up to the point where I feel sort of fairly confident about the test // [...]
261	S1A-097:155.1	S1: Well why // most people learn to drive by the time they're seventeen you know // S2: well not me // [I mean it's completely] S1: Not me either // [but]
262	S1A-097:155.2	S2: I must say I have driven once <1> <,> </1> // (261) <u>uhm</u> in Australia // [(262) <u>uhm</u> <,> and (263) <u>uhm</u> ]
263	S1A-097:155.3	S1: <1> Yeah </1> S1: with someone else <,> //
264	S1A-097:164	S2: [...] But I smashed off the gear box on a rock //
265	S1A-097:165.1	S1: What <,> // how did you do that //
266	S1A-097:165.2	S2: [(264) <u>uhm</u> well I wasn't driving] / well I was driving partly on the road but also i-
267	S1A-097:165.3	on through open country <,> / and (265) <u>uhm</u> there was this rock in the path / and (266) <u>uhm</u> and I sort of assumed I could go over it / and (267) <u>uhm</u> and it in fact it just smashed a hole in the gear box <,> // I don't even know what a gear box is



		actually // [...]
268	S1A-097:179	S2: Right // but don't don't you even manage some eccentricity in in in executing it // S1: oh (268) <u>uhm</u> <,> well well there's a measure of incompetence <,> // there's a measure of my sort of natural left-handedness you know //
269	S1A-097:193.1	S2: And so you're not you're not coming to London at all //
270	S1A-097:193.2	S1: Well I think it's unlikely // the only time I could come is at the weekends <,> // but <,> (269) <u>uhm</u> <,> the thing is I was hoping to sort of <,> sort of <,> visit some places around here // [and also some some people who (270) <u>uhm</u> ] = S2: = What // Colchester //
271	S1A-097:210	S2: But don't you get time off // S1: Well I do // I mean I will do // but not enough time to come up to London // I shouldn't think // unless I come up on an ex- // Oh no there aren't any excursions to London // (271) <u>uhm</u> it's very unlikely // do you fancy coming down to Weymouth <,> //
272	S1A-097:232	S1: <unclear-syllable> // nine weeks' time is in the middle of August // [Now that] / [now the middle of August <,> is uh] / I mean I might still come up to London between now and next week // I don't know // I I've still got to sort of decide what I want to do // but (272) <u>uhm</u> but in August yeah// (273) <u>uhm</u> did did I tell you that I was hoping to go to Japan // to teach // S2: Well no // [it was a vague <1> sort of</1>] S1: <1> Oh </1> right // well it's now definite <,> // (274) <u>uhm</u> and I I definitely want to go to Japan // (275) <u>uhm</u> until I uh <unclear-words> //
273	S1A-097:233	
274	S1A-097:240.1	
275	S1A-097:240.2	
276	S1A-097:246	S2: [...] The desire is definite // or you've actually got a job // S1: Oh the just the desire is definite // But it shouldn't be too difficult / because there are always loads of jobs going // and between now and August <,> // (276) <u>uhm</u> <,> well I won't be able to come up to London for interviews // because <unclear-words> //
277	S1A-097:255	S1: They said uh apparently they don't mind long hair so much // it's beards // but (277) <u>uhm</u> <,> I just heard a couple of people saying that // and I've heard other people say it doesn't really matter // so I really don't know //
278	S1A-097:266	S1: Oh right // yeah // [did I have a] / of course I did // yeah // right // yeah // (278) <u>uhm</u> yeah I'm kind of hovering on the verge of shaving it off anyway though <,> // S2: Have you still got the handlebar moustache // S1: Y- oh yes // yeah // S2: But it's kind of obscured by the beard really // [it's like (279) <u>uhm</u> ] / I mean it's a sort of minor version of (280) <u>uhm</u> you know Paul Vining's moustache cum beard // So the moustache is very much a sort of bushy <,> sort of outgrowth //
279	S1A-097:273.1	
280	S1A-097:273.2	
281	S1A-097:294	S2: But it's something to twiddle in your case // am I right // S1: Yes it is // yeah // yeah // (281) <u>uhm</u> except that I I don't really I haven't really twiddled the moustache so much since I've had the beard as well // because somehow <,> like attention is taken from it from it // sort of thing // I mean when I had the moustache it was very visible <1> <,> </1> and very much uh yes something to be twi- / I'm twiddling it now actually while as as you mention it // [(282) <u>uhm</u> <2> but with the beard </2>] S2: <1> Mm </1> S2: <2> [but] // yes // [I'm I'm] /</2> I tend to sort of tug on the (283) <u>uhm</u> on on the chin //
282	S1A-097:296	
283	S1A-097:300	

284	S1B-045:12.1	S2: [...] And o- oddly enough it's what I liked about the cinema eventually is you know you know there was a group of people making it rather than one individual // (284) <u>uhm</u> I think I was always (285) <u>uhm</u> nervous of my own painting // [I di-] / I I never felt that that it quite made the grade : whatever that might be // I I mean other people liked it // but uh I felt it was it it failed somewhere // and I was actually rath- rather pleased to have this other one as a bolt hull hole really <,> //
285	S1B-045:12.2	S1: What what was the exact sort of sequence uh in terms of the creation of your style // in terms of moving on from costume and set design // [...] // cos there's a strong element almost (286) <u>uhm</u> I mean not in an infantile sense but of dressing up i- i- in your movies : isn't there // [...]
286	S1B-045:17	
287	S1B-045:25	S2: Oddly enough I saw it as dressing down really <,> // because I'm <unclear-words> // and you saw people who were dressed up // and I I ha- you know I had t-t- this grounding in theatre design / and so i- you know even on a film like The Tempest it was actually to explode that whole idea of design concept : (287) <u>uhm</u> it might be called / and to put people into the costumes of any period really but the costumes you felt that they might wear // and I I suppose later on in this particular film in The Garden the new film (288) <u>uhm</u> it is quite like uh a dressing up box // and I quite like that // it is rather like kids in the attic // It has uh <,> a slightly naïve quality in that in that area // [I think it's something I] / [on the other hand I look at design films and they they seem to have you know] / I always think of them as the sort of Christmas cake th- that's got the wrapping round it / and all that people have done is put a sort of pretty wrapping <,> // uh (289) <u>uhm</u> the design comes from within my films // in a way that I I don't see in other British cinema actually <,> //
288	S1B-045:26	
289	S1B-045:32	
290	S1B-045:34	S1: [Wh- wh-] / [In what way was is it related to] / I I mean you haven't exactly (290) <u>uhm</u> concealed the fact that you're gay from film-goers and film critics / I mean in what way is it related to that // I mean not just in terms of the sort of cliché thing of uh you know gays like to put on disguises <1> of various kinds </1> / but also in the fact that the themes of your movies are very much self-portraits kind of obliquely disguised // (291) <u>uhm</u> I mean you've got the theme of homosexuality <2> <,> </2> uh in the guise of Romans versus Christians in Sebastiane or in the guise of the New Testament in The Garden <,> //
291	S1B-045:36	
292	S1B-045:43	S2: <1> yes // yes </1> S2: <2> yeah </2> S2: [Well in] / I would have thought that culture is homosexual // [...] // It is quite interesting that // (292) <u>uhm</u> <,> does it really // of course it doesn't // [...]
293	S1B-045:52	S2: Well I think uh yes I think that that's been done in The Garden / It's very complicated / I think one of the things about (293) <u>uhm</u> being openly gay in the cinema in this country meant that you had accepted certain limitations // [financially and <,>] / I I mean I I knew after I'd made Sebastiane that there wasn't much chance of going to Hollywood // even if I'd wanted to <,> //
294	S1B-045:75	S2: [...] I mean what might seem radical in cinema is in no way radical in theatre // (294) <u>uhm</u> the cinema in a way is like children's bedtime stories you kn- you know // and it always seemed that way to me // just simple really // whereas there was something much more complex // if you looked for instance at twentieth-century painting it got very very far away from the children's bedtime stories that you might equate with the pre-Raphaelites // you know the li- the (295) <u>uhm</u> Light of the World or The Howling Shepherd // it seemed to me that the cinema was stuck in a way in the nineteenth century // in an a very odd way //
295	S1B-045:79	
296	S1B-045:98.1	S1: Wh- wh- what what is collective activity on the Jarman set like // I mean clearly the same sort of rigid hierarch- hierarchies don't obtain // But (296) <u>uhm</u> are you very much the uh the the person in charge // or as I understand from some reports
297	S1B-045:98.2	

		about The Garden (297) <u>uhm</u> you've con- con- confessed that a lot of the film was shot by other people and indeed edited by other people //
298	S1B-045:128.1	S2: I have to say with that sort of film you're you're working on it all the time and you never know where it's going to go uh until it's finished // [...] // [and people come in and say <,>] / well (298) <u>uhm</u> Ian who is (299) <u>uhm</u> one of the assistants on the film said oh I've got a friend called Jessica Martin / she's a singer / she would love to be in the film / [...]
299	S1B-045:128.3	
300	S1B-046:44	S1: So you and you're your brother and two sisters enjoyed all that // did you // S2: Oh yes // [I think (300) <u>uhm</u> ] / I don't know what my mother would have done if we had not come out naturally bookish // but we did come out naturally bookish // [...]
301	S1B-046:51.1	S1: Do you do you make literary allusions as you go about the o- ordinary domestic business //
302	S1B-046:51.2	S2: I do indeed // and it annoys almost all of them // One of my daughters at a party opened the front door to the guests and said (301) <u>uhm</u> (302) <u>uhm</u> I am the youngest daughter of this house / I do not read books / there are too many books in this house / excuse me I will take your coat / [and then] <laugh> / (303) <u>Uhm</u> I think I'm a bit too intense about books //
303	S1B-046:56	
304	S1B-046:74	S1: You you've talked in in various articles over the years about her making you feel utterly inadequate and and horrible // how did she do that // S2: (304) <u>uhm</u> <,> really with noise / she had a sort of capacity for making a loud complaining noise / and anything that wasn't just as she had somehow envisioned it was going to be made her feel very nervous and hysterical / (305) <u>uhm</u> <,> and you knew when you came downstairs that she would have found something that you hadn't done <,> / which she didn't even really want you to do / but she would start to shout about it // (306) <u>uhm</u> <,> that isn't all that she was / I mean she did <,> give us a great many books / [and she did (307) <u>uhm</u> <,>] / there was sort of warmth in her as well // I've talked about this because of course you can't talk about it when people are alive //
305	S1B-046:77	
306	S1B-046:80	
307	S1B-046:81	
308	S1B-046:95	S1: Shall we have your second record now // S2: (308) <u>Uhm</u> the second record is <,> Reginald Kell laying the Adagio from Mozart's Clarinet Concerto // [...]
309	S1B-046:100	S1: I- if you suffered at home Antonia was there salvation for you at school where you could shine // S2: (309) <u>Uhm</u> I didn't like school // I think largely because I am a person who couldn't bear being at boarding school among other girls //
310	S1B-046:117	S1: Why didn't you speak to anybody outside class // S2: (310) <u>Uhm</u> I had no social graces // I had no <,> capacity to make friends I think // S1: But didn't the teachers take any pride in you : uh uh this this shining scholar // S2: It was it was a very good Quaker moral school and the one thing you mustn't do is shine // [...] / and that these things didn't really matter and that the girls who couldn't do the exams shouldn't care // (311) <u>uhm</u> and I see what she was trying to do and it was a good and kind thing // but it meant that the one thing I could do <,> was not valued //
311	S1B-046:122	
312	S1B-047:15	S1: [...] That's not a prospect which all scholars would relish <,> // S2: It isn't a prospect that I relish / but uh there's a Jewish tradition <,> of uh scholarship in which after a certain stage you just have to get stuck in to public life <,> // and (312) <u>uhm</u> that's the tradition which I reluctantly have to follow //
313	S1B-047:17	S1: Is it going to make serious inroads into your writing <,> // S2: Uh I should think so // but (313) <u>uhm</u> <,> in the end I think a religious thinker is judged <,> by what he does in life not just by the books that he writes // [...]

<b>314</b>	S1B-047:42	S1: [...] One commentator getting his his metaphors slightly mixed wrote about the the virtual canonization of Lord Jakobovits by the Prime Minister <,> // Now would you anticipate that (314) <u>uhm</u> in the fullness of time you would be similarly canonized // does orthodoxy in fact incline to the right politically <,> //
<b>315</b>	S1B-047:47	S2: No I think that's a uh misidentification // perhaps a half century ago the identification would have been made in completely the opposite direction <,> // then it was naturally assumed that Judaism was identical with socialism <,> // and (315) <u>uhm</u> had anyone challenged this Jews would have been surprised <,> // and therefore <,> those changes really represent sociological shifts in the Jewish community // [...]
<b>316</b>	S1B-047:100	S1: I was struck in in one of your books by (316) <u>uhm</u> a a question you pose / you called it a a tantalisingly simple question <,> / [...]
<b>317</b>	S1B-048:7	S1: What was it that inspired you to write poetry // [Was th-] / [you l-] / a lot of them are are subjects from nature : aren't they // [lots of birds and countryside <1> and </1>] = S2: = <1> I </1> honestly don't know //(317) <u>uhm</u> <,> people always ask me this about my novels too // what made you write them // [...]
<b>318</b>	S1B-048:24	S1: You always read poetry presumably // S2: Yes // always // I lectured in English at Durham University / [and uh <,>] / well I studied English all my life / so of course <,> I love poetry // poetry's lovely to read in bed at night // I think it (318) <u>uhm</u> <,> it quietens your mind and <,> flattens your spirit out //
<b>319</b>	S1B-048:49	S1: Does novel-writing not come easily <1> to you </1> // S2: <1> no // no </1> // it's very difficult indeed <,> // I I make (319) <u>uhm</u> <,> up to four drafts of of every book // and it takes me ages and ages //
<b>320</b>	S1B-048:60	S1: yes well we're we're actually talking a bit in the past now // [I'm taking life] / I'm sort of retired <,> // but when I was in full flow as it were of writing (320) <u>uhm</u> I had to discipline myself very severely so many hours a day //
<b>321</b>	S1B-048:81	S1: Is that the reason for carrying on doing it // it was does seem a very odd occupation <,> // carry on <1> flogging yourself th- through </1> S2: <1> it's an extremely odd </1> occupation // [it's it's (321) <u>uhm</u> ] / and probably every writer would say the same thing but it's an inner compulsion //
<b>322</b>	S1B-048:91	S2: [it was] <,> / the the compulsion to write stories was always there too // I used to write those oh from childhood <,> // (322) <u>uhm</u> what led me to write the first one <,> / I think it was wanting to re-create a a lovely hot place like Provence in the middle of an English winter <,> // and I found myself led into the story and realized that <,> well perhaps I'd finish one this time // and make a story of it // I still remember how I first started to write <,> / (323) <u>uhm</u> my husband is a very keen fisherman / and his friend used to call and they used to talk about fish <,> / so one cold January evening in Durham <,> I decided that I wasn't going to listen to them talking about fish anymore / [...]
<b>323</b>	S1B-048:95	
<b>324</b>	S1B-048:112	S2: All my books all the suspense novels the early ones are set in different places // in different countries // and (324) <u>uhm</u> it was a desire to re-create the places <,> and <,> perhaps explore the history of the place // I don't know // [...]
<b>325</b>	S1B-048:118	S1: Do you do a lot of travelling to find material <,> // or did it come the other way round // S2: No / it's the other way round / (325) <u>uhm</u> I used to go for you know fortnights abroad / holiday or something / [and if the place] / I'd go to a place because of its associations / usually historical or literary / [...]
<b>326</b>	S1B-048:128	S1: [...] I mean the bulk of your your writing was done at a time when I suppose the the whodunit the sort of Agatha Christie novel was at its most popular // but you've never touched that form : have you //

		S2: No // because I never liked it // I never liked reading it I mean // (326) <u>uhm</u> <,> I suppose it's a sort of guessing game <,> // [...]
327	S1B-048:175	S1: Your heroines are very much of a type : aren't they // they're very strong-willed <,> // (327) <u>uhm</u> they're usually attractive and feminine // S2: I think <,> not strong-willed but possibly with a very strong sense of right and wrong // [...]
328	S1B-049:4	S1: Sheila Mottley you've spent the past twenty-eight years caring for your children / and in particular your eldest daughter Jeanette who was born with no arms or legs <,> // S2: Yes // That is correct <,> // (328) <u>uhm</u> Jeanette <,> unfortunately as I found out later was the result of thalidomide // but there again she's not on her own //
329	S1B-049:9.1	S1: How soon was it that you knew there was something wrong with your baby <,> // S2: Well actually (329) <u>uhm</u> I was told when I was when I was pregnant <,> // much to my alarm / (330) <u>uhm</u> that I was carrying a child with no upper limbs <,> / and they didn't know to what extent the lower limbs had been damaged <,> / but not to sort of look forward to a great deal because they didn't think she'd live <,> // and (331) <u>uhm</u> they said well I'd just have to proceed with the pregnancy / and uh get it over and done with / and go and start again <,> on the next one // like a baby machine //
330	S1B-049:9.2	
331	S1B-049:10	
332	S1B-049:17	S1: It was a difficult birth // did they hand Jeanette to you immediately she was born // S2: (332) <u>Uhm</u> no // I didn't look at Jeanette to be totally honest with you until Jeanette was three days old <,> // and the reason for that before you ask me was that (333) <u>uhm</u> everybody was confusing my brain <,> // people were telling me to do one thing // [...]
333	S1B-049:18	
334	S1B-049:104	S1: What kind of child was Jeanette // S2: Very very demanding <,> // From the time I brought her out of hospital she never slept <,> // you've got twenty-four hours a day Jeanette took up twenty-two of them <,> // You know she's a bit like that now <,> // but (334) <u>uhm</u> <,> she was always wanting to do things // like I used to put her on the floor and she could roll <,> // [...]
335	S1B-049:167	S1: [...] He just said go out and be a prostitute // S2: No // no // it wasn't quite that easy actually / [...] / he dropped me off this road where all the girls worked / which is well-known road in Manchester / which I don't think I should mention cos they're trying to clear it up at the moment <,> / and (335) <u>uhm</u> <,> he left me there with this packet of Durex <,> / [...]

### Appendix 3: Position of *uh* and *uhm* in clauses

Note: the numbers highlighted in grey represent the sum of all instances occurring in the respective category

		<i>uh</i> (N=142)		<i>uhm</i> (N=206)	
		N <sub>uh</sub>	% <sub>uh</sub>	N <sub>uhm</sub>	% <sub>uhm</sub>
<b>Category A: Prenuclear Positions</b>		<b>75</b>	<b>52.82</b>	<b>136</b>	<b>66.02</b>
<b>1</b>	<b>Initial / at Clause Boundary</b>	46	32.39	76	36.89
a	UH/M is situated at a clause boundary that is not introduced by a CCID	28	19.72	43	20.87
b	UH/M is situated before a clause-introductory CCID	18	12.68	33	16.02
<b>2</b>	<b>CCID # subject</b>	27	19.01	49	23.79
a	CCID # subject	19	13.38	29	14.08
b	CCID # adjunct + subject	2	1.41	7	3.40
c	CCID # <i>there</i>	3	2.11	3	1.46
d	CCID # CCID	3	2.11	10	4.85
<b>3</b>	<b>Adjunct # subject</b>	1	0.70	4	1.94
a	clausal adjunct	0	0.00	2	0.97
b	non-clausal adjunct	1	0.70	2	0.97
<b>4</b>	<b>CCID # adjunct + subject</b>	0	0.00	2	0.97
a	clausal adjunct	0	0.00	2	0.97
b	non-clausal adjunct	0	0.00	0	0.00
<b>5</b>	<b>CCID # verb (in clauses that lack a subject / imperatives / interrogations)</b>	1	0.70	5	2.43
<b>Category B: Middle Positions</b>		<b>25</b>	<b>17.61</b>	<b>40</b>	<b>19.42</b>
<b>1</b>	<b>Subject # verb</b>	4	2.82	4	1.94
a	subject # MVB	3	2.11	1	0.49
b	subject # copular verb	0	0.00	2	0.97
c	subject # auxiliary + MVB	1	0.70	1	0.49
<b>2</b>	<b>auxiliary verb # MVB</b>	5	3.52	11	5.34
a	auxiliary verb # MVB	2	1.41	9	4.37
b	modal verb # MVB	1	0.70	2	0.97
c	auxiliary # auxiliary + MVB	2	1.41	0	0.00
<b>3</b>	<b>MVB # non-clausal complementation</b>	11	7.75	14	6.80
a	MVB # object	4	2.82	6	2.91
b	MVB # subject complement	6	4.23	8	3.88
c	MVB # other complement	0	0.00	0	0.00
d	MVB + object # object complement	1	0.70	0	0.00
<b>4</b>	<b>MVB # clausal complementation</b>	4	2.82	10	4.85
a	MVB # object	4	2.82	4	1.94
b	MVB # subject complement	0	0.00	1	0.49
c	MVB # other complement	0	0.00	5	2.43
<b>5</b>	<b><i>there</i> + MVB # subject</b>	1	0.70	0	0.00
<b>6</b>	<b>auxiliary # MVB in interrogative clauses</b>	0	0.00	1	0.49
a	auxiliary + subject # MVB	0	0.00	1	0.49
b	auxiliary # subject + MVB	0	0.00	0	0.00
<b>Category C: Postnuclear Positions</b>		<b>12</b>	<b>8.45</b>	<b>4</b>	<b>1.94</b>
<b>1</b>	<b>MVB / Clause # adjunct (non-clausal)</b>	11	7.75	3	1.46
<b>2</b>	<b>MVB / Clause # adjunct (clausal)</b>	1	0.70	1	0.49
<b>3</b>	<b>MVB / Clause + adjunct (non-clausal) # adjunct (non-clausal)</b>	0	0.00	0	0.00
<b>4</b>	<b>MVB / Clause + adjunct (clausal) # adjunct (clausal)</b>	0	0.00	0	0.00
<b>5</b>	<b>MVB + object # object complement (non-clausal)</b>	0	0.00	0	0.00
<b>Category D: other positions</b>		<b>30</b>	<b>21.13</b>	<b>26</b>	<b>12.62</b>

#### Appendix 4: Analysis of the precise syntactic positions of phrase-internal *uhs* and *uhms*

phrase type	position	N <sub>uh</sub>	% <sub>uh</sub>	N <sub>uhm</sub>	% <sub>uhm</sub>
<b>inside noun phrases</b>		<b>8</b>	<b>29.63</b>	<b>7</b>	<b>25.93</b>
<b>NP</b>	DT # NPHD	3	11.11	5	14.29
	DT + ADJ # NPHD	1	3.70	0	0.00
	CJ # CJ	4	14.81	2	7.41
<b>inside verb phrases</b>		<b>5</b>	<b>18.52</b>	<b>12</b>	<b>34.29</b>
<b>VP</b>	AUX # MVB	2	7.41	10	28.57
	AUX # AUX	2	7.41	0	0.00
	MOD # MVB	0	0.00	1	2.86
	MOD # MOD	1	3.70	0	0.00
	MVB # MVB <sub>new</sub>	0	0.00	1	3.70
<b>inside preposition phrases</b>		<b>14</b>	<b>51.85</b>	<b>16</b>	<b>45.71</b>
<b>PP</b>	P # NP	14	51.85	16	45.71
<b>inside adjective phrases</b>		<b>0</b>	<b>0.00</b>	<b>0</b>	<b>0.00</b>
<b>Sum</b>		<b>27</b>	<b>100.00</b>	<b>35</b>	<b>100.00</b>

## Appendix 5: Lexical/pausal environment of *uh* and *uhm*

#...point of insertion

s...short silent pause

w... for word

ss...long silent pause

	Position	N <sub>uh</sub>	% <sub>uh</sub>	N <sub>uhm</sub>	% <sub>uhm</sub>
no delay	w#w	123	63.40	164	48.96
	- #w	11	5.67	22	6.57
	w# -	6	3.09	14	4.18
short delay	- #s	0	0.00	1	0.30
	s# -	1	0.52	1	0.30
	w#s	17	8.76	57	17.01
	s#w	16	8.25	20	5.97
medium delay	s#s	1	0.52	6	1.79
long delay	ss#ss	0	0.00	4	1.19
	- #ss	3	1.55	3	0.90
	ss#-	0	0.00	0	0.00
	s#ss	1	0.52	5	1.49
	ss#s	0	0.00	2	0.60
	w#ss	4	2.06	28	8.36
	ss#w	10	5.15	3	0.90
other	- # -	0	0.00	3	0.90
	laugh#w	0	0.00	1	0.30
	w#laugh	1	0.52	1	0.30
Sum	---	194	100.00	335	100.00

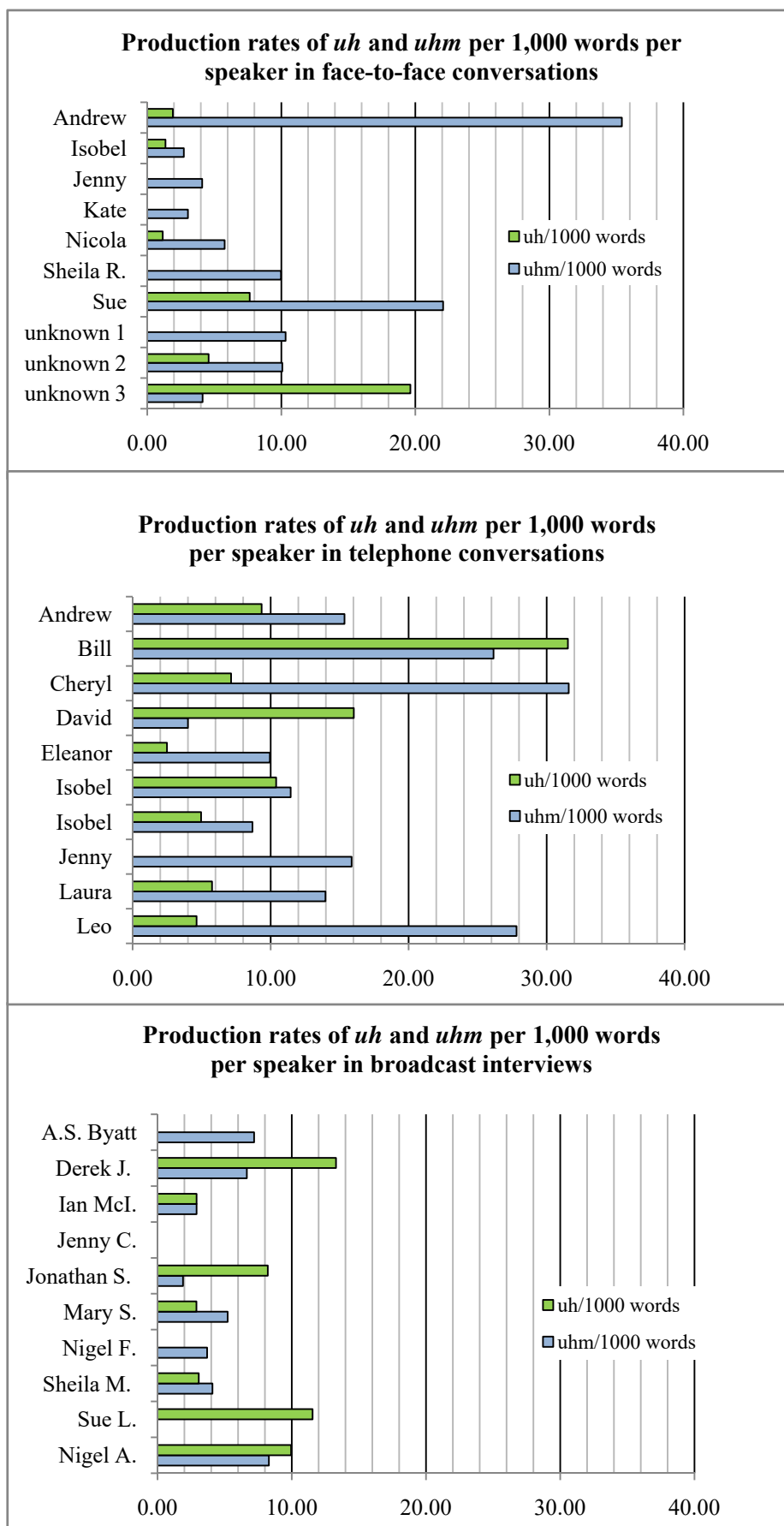
## Appendix 6: Position of *uh* and *uhm* with regard to self-interruptions

The calculation of percentages was based on the respective total number of *uhs/uhms* occurring at points of self-interruption.

	N <sub>uh</sub>	% <sub>uh</sub>	N <sub>uhm</sub>	% <sub>uhm</sub>
self-interruption	14	66.67	29	78.38
self-correction / reformulation	4	19.05	8	21.62
mid-word interruption	2	9.52	0	0.00
miscellaneous	1	4.76	0	0.00
Sum	21	100.00	37	100.00



**Appendix 7: Individual speakers' production rates of *uh* and *uhm* per 1,000 words (sorted by genre)**



## Appendix 8: Abstract (English)

If *uh* and *uhm* had formerly been dismissed as undesirable, disruptive elements hindering the development of fluent speech, there is now a vast body of literature suggesting that they are in fact valuable communicative resources that serve numerous functions, such as the verbal representation of one's cognitive processes or the signaling of one's turn-keeping intentions. However, since most of these publications are generic studies of *uh* and *uhm*, it remains unknown up to this point whether or not the two variants fulfill different functions in spoken discourse. Until this day, only few have examined this question, and it is now often assumed that *uh* and *uhm* signal either different lengths of delay (Clark and Fox Tree 2002: 84), or different planning processes associated with speech production (Shriberg 1994: 154). The present thesis seeks to contribute to the bridging of this research gap by examining the following three questions: (1) whether genre produces different effects on *uh* than on *uhm*; (2) whether *uh* and *uhm* display different syntactic tendencies; and (3) whether *uh* and *uhm* fulfill different functions in spoken discourse. For the purpose of answering these three questions, this paper includes a corpus-based study of *uh* and *uhm* as observed across three different genres – private face-to-face conversations, telephone calls, and broadcast interviews – in ICE-GB. The results obtained in this study suggest that *uh* and *uhm* do not seem to fulfill different functions in spoken discourse, but that they appear to differ from each other with regard to their contextual and speaker-specific determinants. This finding is embedded in the theoretical part of this thesis – in which it is argued that both *uh* and *uhm* are multi-functional collateral signals that serve both the speaker and the listener in attaining their shared goal of efficient interpersonal communication – to construct a theoretical model of *uh* and *uhm* that seeks to outline the mechanisms underlying their production. It is suggested that the production of *uh* and *uhm* is determined by four different factors: (1) a speaker's need to buy time for an ongoing cognitive process, (2) a speaker's need or desire to verbally signal this process to the listener, (3) the linguistic and extra-linguistic context of a given discourse situation, and (4) the speaker as individual who demonstrates distinct idiosyncratic mannerisms based on his personal preferences and experiences. In suggesting this model, this thesis proposes a potential framework for future research on the study of the differences between *uh* and *uhm*.

## Appendix 9: Abstract (German)

Lange Zeit wurden *uh* und *uhm* – die englischen Äquivalente zum deutschen *äh* und *ähm* – als unerwünschte, störende Verlegenheitslaute angesehen. In den letzten Jahren wandten sich allerdings immer mehr Forschende von dieser Betrachtungsweise ab, weshalb heute verstärkt davon ausgegangen wird, dass *uh* und *uhm* einen wichtigen Beitrag zur Entwicklung effizienter zwischenmenschlicher Kommunikation leisten, indem sie Sprechern und Sprecherinnen ermöglichen, andauernde kognitiven Prozesse sowie den Wunsch, das Rederecht zu behalten, zu signalisieren. Was allerdings bis heute weitgehend ungeklärt bleibt, ist die Frage, ob die beiden Varianten *uh* und *uhm* unterschiedliche Funktionen erfüllen – eine Frage mit der sich die vorliegende Diplomarbeit beschäftigt. Bis heute haben sich nur wenige ForscherInnen mit dieser Fragestellung beschäftigt; die wenigen vorhandenen Arbeiten zu diesem Thema gehen allerdings davon aus, dass es tatsächlich einen beobachtbaren Unterschied zwischen den beiden Varianten gibt – sei es, dass sie unterschiedliche Verzögerungslängen signalisieren (Clark und Fox Tree 2002: 84) oder dass sie unterschiedliche Planungsprozesse widerspiegeln (Shriberg 1994: 154). Um diese Thematik näher zu erforschen, wurde anhand von Texten aus der britischen Komponente des International Corpus of English (ICE-GB) eine korpusbasierte Studie von *uh* und *uhm* durchgeführt. Untersucht wurden hierbei die folgenden drei Forschungsfragen: (1) Hat das Genre einer gegebenen Sprechsituation andere Auswirkungen auf *uh* als auf *uhm*?; (2) Gibt es syntaktische Unterschiede zwischen *uh* und *uhm*?; und (3) erfüllen *uh* und *uhm* unterschiedliche Funktionen im gesprochenen britischen Englisch? Die Ergebnisse dieser Studie legen den Schluss nahe, dass es keine systematischen funktionalen Unterschiede zwischen *uh* und *uhm* gibt, ihre Produktion aber von unterschiedlichen Faktoren abhängig zu sein scheint. Basierend auf einer ausführlichen theoretischen Betrachtung verschiedener relevanter Studien wird dieses Ergebnis dazu genutzt, ein theoretisches Modell von *uh* und *uhm* zu konstruieren, welches versucht, jene Mechanismen darzustellen, die der Produktion von *uh* und *uhm* scheinbar zu Grunde liegen. Es wird die Hypothese aufgestellt, dass die Produktion von *uh* und *uhm* von vier Faktoren abhängig ist: (1) die Notwendigkeit, Zeit für die Durchführung eines kognitiven Prozesses zu gewinnen, (2) der Wunsch, diesen Prozess dem Zuhörer zu signalisieren, (3) der linguistische und extralinguistische Kontext einer gegebenen Sprechsituation, und (4) der Sprecher als Individuum, der idiosynkratische Eigenheiten aufweist, welche sich auf persönlichen Vorlieben und Erfahrungen begründen. Die vorliegende Arbeit leistet sowohl durch die empirische Untersuchung syntaktischer und funktionaler Unterschiede zwischen *uh* und *uhm*, als auch durch die Erstellung eines

theoretischen Modells von  $uh$  und  $uhm$  einen wertvollen Beitrag zum Forschungsstand auf diesem Gebiet.