



**HAL**  
open science

# Should 'uh' and 'um' be categorized as markers of disfluency? The use of fillers in a challenging conversational context

Loulou Kosmala, Aliyah Morgenstern

► **To cite this version:**

Loulou Kosmala, Aliyah Morgenstern. Should 'uh' and 'um' be categorized as markers of disfluency? The use of fillers in a challenging conversational context. *Fluency and Disfluency across Languages and Language Varieties*, 2018. hal-02360614

**HAL Id: hal-02360614**

**<https://hal.science/hal-02360614>**

Submitted on 12 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Kosmala, L. & Morgenstern, A. (2018). Should ‘uh’ and ‘um’ be categorized as markers of disfluency? The use of fillers in a challenging conversational context. In Liesbeth Degand, Gaëtanelle Gilquin, Laurence Meurant & Anne Catherine Simon (eds.) *Fluency and Disfluency across Languages and Language Varieties*. Corpora and Language in Use – Proceedings 4, Louvain-la-Neuve: Presses universitaires de Louvain, x-x.

## **Should ‘uh’ and ‘um’ be categorized as markers of disfluency? The use of fillers in a challenging conversational context**

Loulou Kosmala & Aliyah Morgenstern

Sorbonne Nouvelle University

### **Abstract**

This paper examines the use of ‘uh’ and ‘um’ in spontaneous speech. ‘Uh’ and ‘um’ have traditionally been labeled as markers of disfluency, and in question-answering, they are commonly said to reflect uncertainty, due to the unfamiliarity or difficulty of the questions. Our proposal is that fillers do not necessarily reflect difficulty and speech disruption. Our study consists of an experiment in which participants were asked questions about a film in a conversational setting. After the experiment, participants were invited to rate the “difficulty” of the questions they had been asked. Findings indicate that they did not often produce fillers when they found the questions “difficult” as little relation between the rate of fillers and the perceived difficulty of the questions was found. Fillers mainly occurred in initial position and served planning functions. This confirms (Tottie 2011) that fillers are predominantly used to buy time in conversation and reflect planning processes.

## 1. Introduction

When speakers are engaged in conversation, their speech is often filled with a number of disfluencies. A clear illustration of a disfluent utterance would be the following:

&uh (.) well you'd have [//] you wouldn't &hav [/] you wouldn't [//]  
you'd have to ask &uh  
other (..) people on that &uh I always seem to be you know growing  
up in &b +//.  
I think [/] I think in fact [/] I think that's my house over there.<sup>1</sup>

This transcription is taken from an interview with film director Tim Burton, who is answering a question previously asked by the interviewer (“How strange were you as a child?”). Although the speaker is a native American English speaker who was born and raised in the US, and who is used to public speaking and film interviews, he still produces a very disfluent utterance. Here disfluency is characterized by a series of truncated words (“hav”, “b”), self-repairs (“you'd have [//] you wouldn't have”; “you wouldn't [//] you'd have”) repetitions (“you wouldn't [/] you wouldn't”) silent pauses, and filled pauses (“uh”, repeated twice). In spontaneous speech, 5 to 6% of words are said to be affected by speech disfluency (Bortfeld *et al.* 2001; Fox Tree 1995; Shriberg 1994). A number of studies insist on the fact that disfluency represents a suspension, or an interruption in speech; disfluencies “interrupt the flow of speech and do not add propositional content to an utterance” (Fox Tree 1995: 709); conversely, they present a “deviation in speech from ideal delivery” (Ferreira & Bailey 2004: 231). This concept can also be opposed to the notion of “fluency”, characterized as “smoothness or continuity of speech” (Koponen & Riggenbach 2000: 8).

When people talk, they are aware that there is a conventional way of speaking, known as “the ideal delivery” (Clark 2006: 245), that is, “a single action with no suspensions- no silent pauses, no fillers, no repeats, no self-corrections, no delays except for those required by the syntax of the sentence”. In spontaneous speech, however, this delivery rarely happens, since conversations are most often unplanned. Speech disfluency is an inherent human phenomenon as speakers typically do not know in advance what they are going to say and how they are going to say it. They plan their utterances as they produce them. In fact,

---

1 The '&' is a conventional symbol used in CLAN transcriptions in CHAT format to indicate that the following characters have a special status so that they could be automatically searched. In our transcriptions, we used '&' before fillers and truncated words.

the course of human language can never be a continuous flow, and O’Connell & Kowal (2005: 457) argue that the notion of ideal delivery is inadequate for the following reasons:

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

Speakers may notice a problem before producing speech and it may be indicated by disfluencies (Merlo & Mansur 2004). Therefore, they may have to suspend their speech. Nooteboom (1983), suggested the “Main Interruption Rule”: “Stop the flow of speech immediately upon detecting trouble” (quoted by Levelt (1989: 478). Clark (2006: 245) called this detection the ‘detection of trouble’. He explained that after suspending speech, hiatuses occur. He described them as being: “the time interval between the point of suspension of fluent speech and the point of its resumption.” In this interval, speakers may insert a pause, a filler, or an editing expression. After the resumption occurs and the hiatus ends, speakers may resume in four alternative manners: (1) continuing where they left off; (2) restarting the last constituent; (3) correcting the error and replacing the constituent; (4) starting a new utterance.

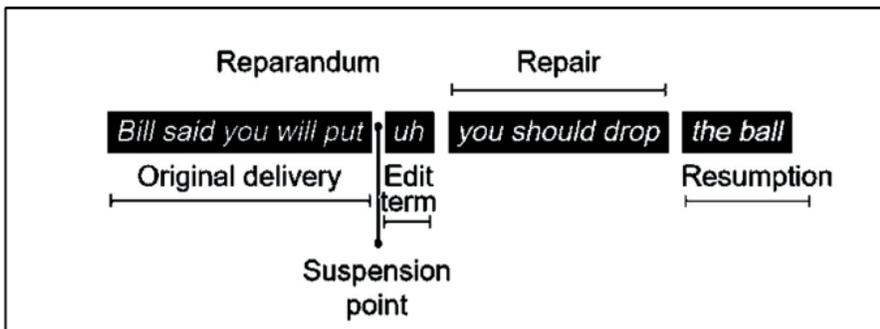


Figure 1. A disfluent utterance (Ferreira & Bailey 2004: 232).

Most speech disfluencies can be divided into two parts, their *problem* and their *solution* (Clark 2002: 6). Figure 1 shows an example of a disfluent utterance in which the speaker corrected the previous sequence. The original ideal delivery was first uttered (“Bill said you will put”) and then the speaker detected trouble. The original delivery is therefore interrupted and the disfluency begins.

The filler ‘uh’ (which is here inserted into the hiatus) occurred at the “Suspension point”, and a repair replaced the original “reparandum”. The “resumption” refers to the return to fluent delivery. In this sense, it seems that disfluency mostly reflects planning difficulty; however, it is not always the case and a number of authors argue that disfluency markers have little to do with “problems” (Schegloff 2010; Tottie 2011).

### 1.1. Main disfluency markers

Even though many authors provided different categorizations for disfluency phenomena (see Lickley 2015 for a review) the main disfluency markers include:

1. Repeats: Repeated words or phrases which are not significant semantically, as in “I I saw a very very big boy” (only the pronoun ‘I’ is repeated as the repetition of ‘very’ is used as an intensifier) (Maclay & Osgood 1959: 24).
2. Self-repairs: Repairs made by the speaker and not by the interlocutor. Levelt (1989) listed three parts of the repair process: (1) the original utterance which contains the item to be repaired; (2) the moment of interruption and (3) the repair, which can be defined as “the correct version of what was wrong before” (1989: 44).
3. False starts: Incomplete or self-interrupted utterances (as defined by Maclay & Osgood 1959: 24).
4. Unfilled pauses: Pauses of unusual length (Maclay & Osgood 1959: 24), which are not used as an emphasis—this is the difference between “reactive” and “rhetorical” pauses as explained by Clark (2006). Drommel (1980) speaks of “D-pauses”, which are unintentional pauses likely to occur as a result of speech-planning.
5. Filled pauses, or vocalic fillers: non-lexical sounds such as ‘uh’ and ‘um’, defined as “the insertion at any moment within spontaneous speech of a long and stable vocalic segment” (Candea, Vasilescu & Adda-Decker 2005: 10).
6. Prolongations, or lengthenings: word prolongation, “disfluent lengthening” defined as “a marked prolongation of one or more phones, resulting in above-average syllable and word duration” (Betz & Wagner 2016: 1), that are distinguished from phrase-final lengthening, used as a cue for phrase boundaries.

This paper focuses on the use of ‘uh’ and ‘um’ in English spontaneous speech. The difference between those fillers and the other disfluency markers listed above is the fact that they are non-lexical sounds that are widely used in all languages (with some phonological variations), yet they do not derive from grammatical markers, have little phonetic weight, and are inserted in speech.

## 1.2. The so-called fillers ‘uh’ and ‘um’

One thing to point out when discussing ‘uh’ and ‘um’ is the fact that they are known under many different labels in the literature. Maclay & Osgood (1959) and Goldman-Eisler (1968) first referred to them as “filled pauses”, as opposed to “unfilled pauses”. Levelt (1989) spoke of “editing expressions” that he defined as sounds produced when speech is suspended or difficulty is encountered. We also find other terms such as “hesitation disfluencies” (Corley & Stewart 2008), simply “uh(m)s” (Schegloff 2010) and “planners” (Tottie 2011). Clark & Fox Tree (2002) chose to refer to them as “fillers” as they argued that they shared the same characteristics as words. Fox Tree (2007) in fact later referred to them as “filler words”. Tottie (2011) points out that “filler” is a negative term, which does not say anything about their discourse functions. However, Candea, Vasilescu & Adda-Decker (2005) used the term “autonomous fillers” to represent their vocalic features. In this paper, we chose to refer to them as ‘uh’ and ‘um’, and as “fillers” (short for “autonomous fillers”), although we will not carry out an acoustic analysis.

The position of fillers in the utterance and their form can reflect different problems in production, as Shriberg (1994: 54) argued: “The form ‘um’ may be used relatively more often during planning of larger units, while ‘uh’ may be relatively more likely to reflect local lexical decision-making.” In addition, Bortfeld *et al.* (2001) found that phrase-internal fillers were more likely to serve the functions of displaying word-finding problems and between-phrase fillers were more likely to serve the functions of beginning or ending turns.

According to Clark & Fox Tree, the main function that they serve is to signal a delay in speaking. ‘Uh’ is associated with a minor delay, while ‘um’ is associated with a major delay. Other researchers such as Corley & Stewart (2008) argued that even though it is true that fillers may occur in situations associated with uncertainty or trouble, this does not prove that they are specifically used by the speaker to signal a delay in speech. Moreover, Corley & Hartsuiker (2011) argued that delays facilitated word recognition, but they also demonstrated that any delay (whether it was ‘um’, or a silence) had this effect. Thus, fillers could

not only be viewed as signals. Schegloff (2010: 171) argued that although fillers were associated with delay, they did not “announce” a delay, but “embody” it.

A great many ‘uh(m)s’ are not followed by silence; and if it is argued that, in those instances, they themselves constitute the delay, then (a) that is “embodying,” not “announcing,” a delay and, in that case, (b) their claim cannot be falsified, as it is a self-fulfilling prophesy.

Other functions are mentioned in the literature: (1) Keeping control over the conversational ball (Maclay & Osgood 1959). (2) Displaying planning difficulty (Clark 2006; Fraundorf & Watson 2011; Bortfeld *et al.* 2001; Schnadt & Corley 2006); (3) Reflecting cognitive difficulty in selecting a name (Bortfeld *et al.* 2001; Finlayson & Corley 2012; Hartsuiker & Notebaert 2009). (4) Marking discourse structure (Swerts 1998; Tottie 2014) (5) Marking uncertainty (Smith & Clark 1993); (6) Signaling turn-taking (Kjellmer 2003).

Within all the different functions, we find two opposite views on the role of ‘uh’ and ‘um’. One view is that they are typically associated with problems, trouble and difficulty. They occur when speakers detect trouble in processing; ‘uh’ is a symptom of recency of trouble indicating that the trouble is still present at the moment of interruption (Levelt 1983). They indicate the depth of the speakers’ retrieval problem (Smith & Clark 1993). They are also said to increase during question-answering. Respondents may often produce a ‘uh’ or an ‘um’ in order to save face (Goffman 1967; 1971) when they are uncertain of the answer, and therefore indicate little confidence, which ultimately delays their response. This is also known as the *Feeling of Knowing model* (Hart 1965). Swerts & Kraemer (2005) looked at how visual (eye gaze, facial expressions) and prosodic cues (including ‘uh’ and ‘um’) were used by speakers when answering factual questions and signaled uncertainty in their answers. The use of ‘uh’ and ‘um’ can therefore be linked to task difficulty. They are said to increase when the topic of the conversation is more abstract (Reynolds & Paivio 1968) or when the question asked is too difficult (Smith & Clark 1993). Other difficult tasks involving challenging choices, such as describing mazes with a varying number of alternate routes, as investigated by Christenfeld (1994), has shown that descriptions of mazes with more options lead to a higher rate of filled pauses. However, the same study also showed that in describing the simplest maze, the participants still produced a high number of filled pauses. Therefore, the other view is that ‘uh’ and ‘um’ have little to do with difficulty and problems as they may serve other functions such as “project[ing] further talk” (Schegloff 2010: 140), planning and structuring discourse (Tottie 2014), and other pragmatic functions such as turn-taking, turn-holding and highlighting

(Kjellmer 2003). In this view, ‘uh’ and ‘um’ do not necessarily signal difficulty or uncertainty, but rather the cognitive processes at play.

Moreover, a relation can also be drawn between the use of ‘uh’ and ‘um’, and eye gaze. Kendon (1967) points out that gaze direction tends to change at the beginning and ending of utterances; he suggests that by gazing away before a long utterance, speakers may be withdrawing their attention from their interlocutors and focusing on what they are planning to say. He argues that when speakers reach the end of an utterance, they are approaching a “choice-point”, defined as “a point in which subsequent action he might take will largely depend upon how his interlocutor is behaving” (p. 35). The processes Kendon describes (mainly speech planning and turn-taking) are also associated with functions served by fillers. Moreover, Goodwin & Goodwin (1986) found that speakers frequently gazed away from their interlocutor during word search; speakers tended to withdraw their gaze as the search began. The authors explain that gaze withdrawals usually occur near “perturbations in the talk displaying initiation of a word search” (p. 57). In later work, they argue that gaze withdrawal and thinking face may not be communicative, but rather “adjustments to the cognitive demands that a word search imposes (for example, ways of eliminating distracting visual information)” (p. 58). Swerts & Kraemer (2005) support this view, as they found that when answering a difficult question, speakers were more likely to produce fillers and to withdraw their gaze before producing their answer. They pointed out the “common observation” that it is usually easier to “think” when not looking at the interlocutor. Moreover, Glenberg *et al.* (1998) explained that in answers to general knowledge questions, speakers were more likely to withdraw their gaze before answering as a way to disengage from the environment. The frequency of gaze aversion was thus correlated with the difficulty of the question.

As the literature suggests, the terms “difficulty”, “problems” and “trouble” are often used to describe what fillers may reflect (Clark 2006; Finlayson & Corley 2012; Clark & Fox Tree 2002). The term “difficulty” will be used in this paper to refer to the perceived difficulty of the questions that the participants were asked. The other term suggested here will be “cognitive effort” in speaking/discourse planning (also used by Bortfeld *et al.* 2001) to refer the cognitive processes that are reflected by fillers.

In line with these issues, this paper aims to examine the role of ‘uh’ and ‘um’ in spoken discourse in what we call “challenging contexts” in order to test whether they are linked to notions of difficulty, and of cognitive effort. Visual cues such as eye gaze will also be analyzed.

## 2. Corpus and Methods

### 2.1. Experiment

A small experimental study was conducted, involving 16 native speakers of English, aged 18 to 23, coming from different regions (American English, Northern and Southern British English). The group included 8 males and 8 females. The experiment was based on a film, entitled *Big Fish* (2003), directed by Tim Burton. The starting point of our study was to analyze whether ‘uh’ and ‘um’ were associated with planning difficulty. The students were undergraduates, and none of them were film specialists or enrolled in a film course.

The film depicts a complex relationship between a man and his father, who is on his deathbed, and it recounts far-fetched stories told by the father. As there are constant shifts between fantasy elements and real-life events, it is difficult for the son – and the audience – to believe these stories. It is a challenging film, since the audience constantly has to make sense of the stories and question reality. It was selected for its visual complexity, as well as its loose plot.

8 of the 16 participants were asked to watch the film on their own at home, a few days prior to the experiment (no more than three days in order to avoid other factors such as long-term memory retrieval to be taken into account). The experiment was carried out in a series of eight sessions, during which the participants worked in pairs. All the paired participants knew each other fairly well. During each session, the participants who had seen the film (coded as Participant A1 to H1) answered a list of ten questions about the film. The questions were read by their partners (the other eight participants, coded as Participant A2 to H2) who had never seen the film. The video recordings were made in familiar settings (in their homes, or in student halls) and the participants were free to speak as much or as little as they wished. The participants were sitting opposite each other, and the camera was placed in the corner of the room. The participants answering the questions (A1-H1) were instructed to answer one question at a time, be as spontaneous as possible, and view the experiment as a casual conversation. The total duration of the corpus was 71 minutes and 63 seconds, with an average duration of 7 minutes and 9 seconds.

Our goal was to make some of the questions “difficult” and challenging to answer. Questions can be difficult to answer when respondents have difficulty: (1) understanding the question, (2) retrieving the requested information, and (3) formulating their response. (Smith & Clark 1993: 26). We wanted to test whether fillers were related to the difficulty of the questions, and therefore to

conceptual and planning difficulty, or whether they served other pragmatic functions. The questions are listed below.

*Table 1. Questions about the film*

No.	Questions
1	What is the movie about?
2	Who are the main characters? Describe them.
3	Describe two key scenes from the film.
4	Describe the father and son relationship.
5	How would you define the notion of ‘reality’ in the film?
6	Did you enjoy the film, why?
7	Which scene was your favorite, why?
8	Did you pay any attention to the way it was directed? (the music, sets, visual effects etc.)
9	How would you interpret the ending of the film?
10	Would you like your parents to keep on telling you stories that are not necessarily true? Would you do it to your own children?

A questionnaire was sent to all the participants who had answered the questions during the experiment. They were asked to rate the questions from “very easy to answer” to “very difficult to answer”. Note that this is similar to the “feeling of knowing” model (Hart 1965), although here we evaluated their judgment on the question itself, rather than on their ability to answer.

Since task difficulty has been associated with the rate of fillers, our hypothesis was that questions that were judged “difficult to answer” by the respondents would be more likely to be answered with utterances containing a higher rate of fillers than the ones judged “easy to answer”.

## 2.2. Transcription and coding

The 8 recordings were transcribed in detail using the transcription software CLAN; and included fillers (transcribed as &uh and &um) unfilled pauses (transcribed as (.) (..) and (...)<sup>2</sup>; repetitions (transcribed as [/]), self-repairs (transcribed as [//], and self-interruptions (transcribed as +//). Only fillers were analyzed for this study and were coded according to (1) their form— ‘uh’ or ‘um’;

<sup>2</sup> Length of pauses was coded by ear in the transcriptions and was not used for this specific study. In future studies, their length will be measured more accurately.

(2) their position in the utterance— ‘beginning’, ‘middle’ ‘end’ and ‘isolated’ (surrounded by more than three seconds of silence); (3) their functions<sup>3</sup>: (a) planning—planning made at the micro- or macro-level, *i.e.* plan a whole utterance (at the beginning), plan the continuation of the utterance (in the middle), or plan a specific lexical item (before a noun phrase); (b) reformulating—cases in which the utterance is reformulated because the speaker detected a problem; here ‘uh’ and ‘um’ typically co-occured with other disfluency markers such as repairs, repeats and restarts. What follows ‘uh’ and ‘um’ was reformulated again or differently; (c) marking uncertainty—the speakers used ‘uh’ and ‘um’ as a signal that they were uncertain about the answer; it typically co-occured with non-answers. The second author checked the whole coding and the few disagreements were solved together. We also coded whether speakers tended to avert their gaze when they produced the fillers.

Table 2. Example of coding

Utterance	Part.	Question	Form	Position	Function	Gaze withdrawal
&um the movie’s about (..) a father.	A2	1)	Um	Beginning	Planning	Yes
so:o he and his wife <go to> &um [/] go to see him and stuff.	C2	1)	Um	Middle	Reformulating	Yes
(...) &um (...) that’s a tough question.	B2	5)	Um	Beginning	Marking uncertainty	Yes

### 3. Results

347 occurrences of ‘uh’ and ‘um’ were found, but 40 were excluded from the analysis as they were produced by the participants who read the questions—they all produced them before reading the questions and therefore their use did not apply to our analysis.

3 Position and function are related factors especially in the case of the sub-categories of the planning function, but we coded them separately in order to obtain clear results for the three functions (planning, reformulating and marking uncertainty).

### 3.1. Overall results

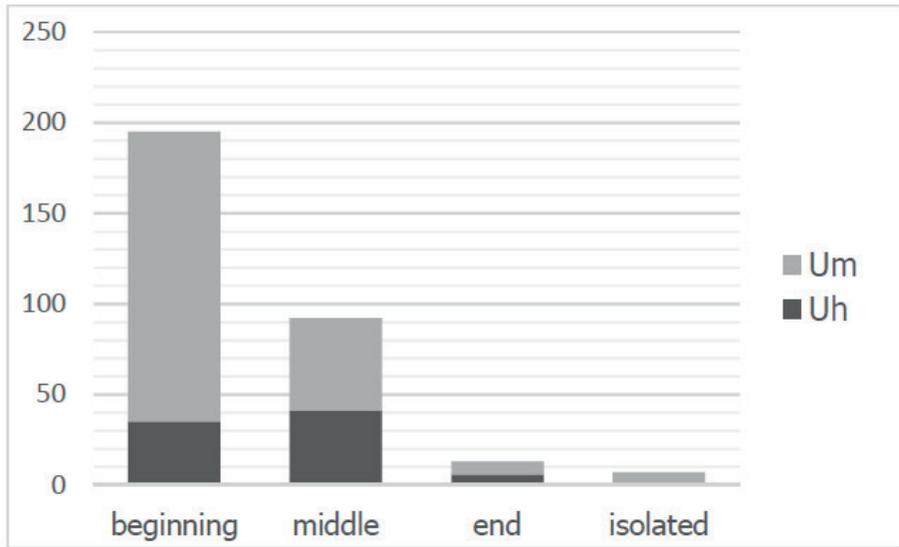
Results indicate that speakers produced 4.2 fillers per minute on average, although we did find great individual differences across speakers. Some produced a great number of fillers (Participant H2 and A2) while others produced a low number (Participant E2 and D2), but these differences are linked with the length of their speech. Our results also show that ‘um’ (224 out of 307) was much more frequently used than ‘uh’ (83 out of 307). This is not consistent with Shriberg (1994) and Clark & Fox Tree (2002), among others. These results show a large variation in ‘uh’/‘um’ ratios.

*Table 3. Distribution of fillers across participants*

<b>Participant</b>	<b>Uh</b>	<b>Um</b>	<b>Total</b>	<b>Total Per Minute</b>	<b>Duration</b>
A2	17	31	48	4.5	10:44 mn
B2	5	17	22	4.0	05:44 mn
C2	3	25	28	3.7	07:39 mn
D2	9	22	31	3.2	09:49 mn
E2	3	15	18	3.3	05:40 mn
F2	1	35	36	3.9	09:22 mn
G2	4	36	40	4.3	09:20 mn
H2	41	43	84	5.7	14:51 mn
<b>Total</b>	<b>83</b>	<b>224</b>	<b>307</b>	<b>4.2</b>	<b>73:13 mn</b>

### 3.2. The position of fillers in the utterance

Fillers were used at the beginning of utterances in 63% of the cases, 29% in the middle, 4% at the end, and 4% occurred in isolation. ‘Uh’ occurred equally at the beginning (35 occurrences) and in the middle (41 occurrences), while ‘um’ occurred much more at the beginning (160 occurrences) than in the middle (51 occurrences). This is consistent with Shriberg’s view (1994: 154) that ‘um’ is more frequently used during planning of larger units.



*Figure 2. Location of fillers*

### 3.3. The functions of fillers

Figure 3 indicates that 89% of the fillers served planning functions, which is significantly more than the other functions. Figure 4 shows that fillers served planning functions at the macro-level more frequently (67% of the time) than at the micro-level (43% of the time) as we find more occurrences of them in initial position (185 out of 307). Our results support the view that fillers are used to buy time to plan the utterance (Jucker 2015; Holmes 1988; Fehring & Fry 2007; Tottie 2014; among others). Therefore, contrarily to our initial hypothesis that followed Smith & Clark (1993), they seem to be more closely linked to speech planning than to production difficulties. ‘Um’ (205/275) was much more used for planning than ‘uh’ (70/275); and ‘uh’ (13/20) occurred more frequently when used for reformulating than ‘um’ (7/20). We also observed no occurrences of ‘uh’ corresponding to a display of uncertainty.

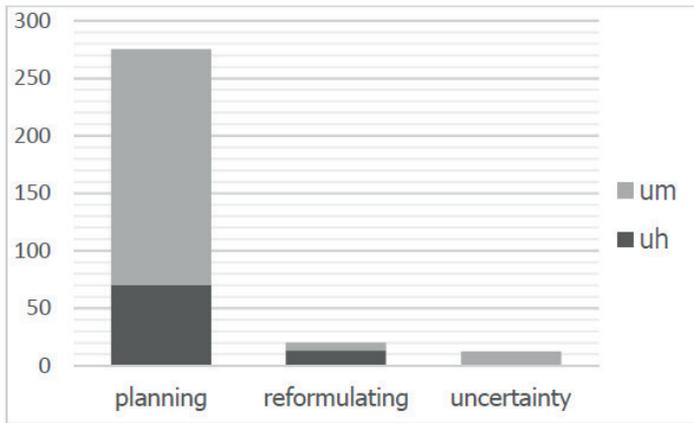


Figure 3. Distribution of fillers across functions

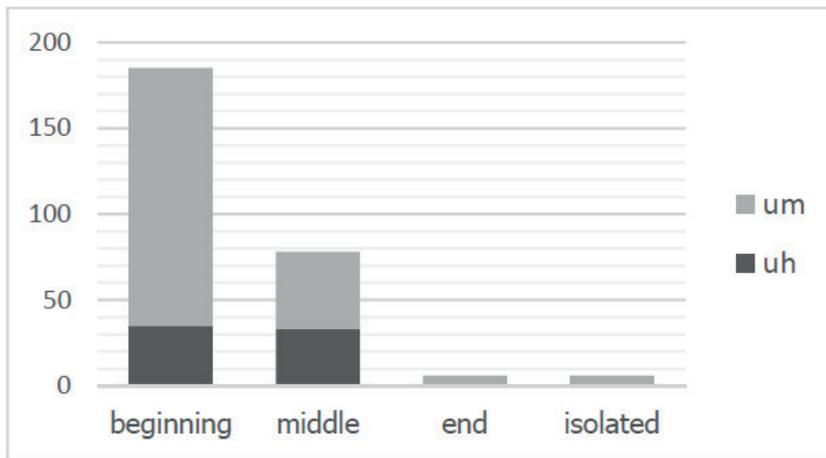


Figure 4. Position of the fillers serving planning functions

### 3.4. Gaze withdrawal

Figure 5 shows that fillers were very often accompanied by gaze withdrawal (in 88% of the cases). Speakers rarely looked at their interlocutors while producing fillers. This is consistent with the view that speakers disengage from interaction when producing fillers as a way to suppress control of the environment over cognition (Glenberg *et al.* 1998).

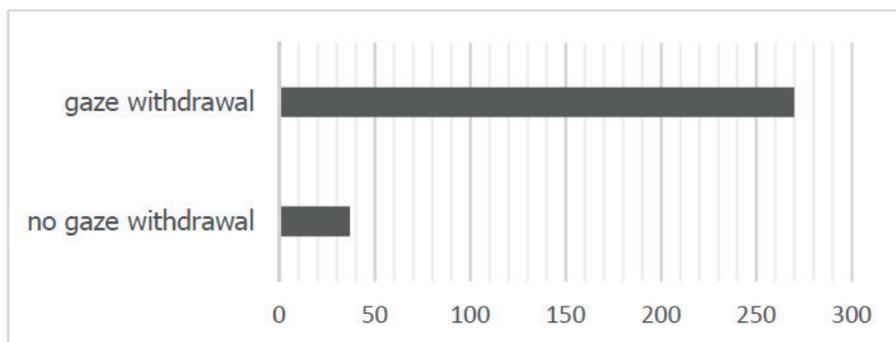


Figure 5. Number of fillers accompanied with gaze withdrawal

### 3.5. Rate of difficulty

Table 4 summarizes the distribution of fillers (per 100 words) per answer to each question, and Table 5 shows the total rate of difficulty per question (according to the judgment made by each participant). The average is 30.7 fillers per answer (3%), with only answers to questions 2) and 3) quite above average (6% and 5%), whereas our prediction was that answers to questions judged difficult to answer (such as 7, 8, 9 and 10, see Table 5) would contain the highest rate of fillers.

Table 4. Distribution of fillers per answer to each question

Participants	1)	2)	3)	4)	5)	6)	7)	8)	9)	10)
A2	2%	8%	1%	5%	4%	3%	5%	0%	3%	8%
B2	3%	3%	5%	0%	3%	2%	1%	3%	1%	1%
C2	4%	5%	6%	2%	2%	1%	4%	2%	1%	1%
D2	4%	7%	0%	1%	3%	1%	4%	3%	1%	7%
E2	2%	1%	5%	4%	3%	1%	1%	1%	0%	0%
F2	0%	1%	5%	0%	4%	3%	2%	8%	3%	1%
G2	7%	6%	8%	2%	2%	2%	11%	1%	0%	1%
H2	2%	2%	12%	8%	4%	5%	9%	9%	3%	12%
<b>Total</b>	<b>2%</b>	<b>6%</b>	<b>5%</b>	<b>2%</b>	<b>3%</b>	<b>2%</b>	<b>4%</b>	<b>3%</b>	<b>1%</b>	<b>3%</b>

Table 5. Total rate of difficulty per answer to the questions

Question	Very easy	Easy	Difficult	Very difficult
1)	3	4	1	0
2)	5	3	0	0
3)	2	4	1	1
4)	0	4	4	0
5)	0	1	7	0
6)	4	2	1	1
7)	1	4	3	0
8)	0	2	5	1
9)	0	1	4	3
10)	0	2	5	1

We also observed that the answer to the question that was judged the easiest (5 out of 8 judged it “very easy”) had the highest average rate of fillers (6%). This result indicates that speakers do not necessarily produce fillers as a result of uncertainty or topic difficulty, while they may produce other types of disfluencies (*e.g.* repairs, repetitions etc.) when they are uncertain, or find the question difficult (Bortfeld *et al.* 2001; Smith & Clark 1993; Brennan & Williams 1995).

#### 4. Discussion

We are aware of the limitations of our analysis as we did not look at all the disfluency phenomena but paid particular attention to fillers. Our analysis is also limited as it lacks evidence from “listeners”; the participants who read questions and listened to answers did not intervene much in the conversation. Therefore, our analysis was only conducted on participants who answered the questions.

Our findings show that ‘uh’ and ‘um’ are used frequently during conversation, and the first thing to note is that they are not linked to conceptual and planning difficulty, and do not reflect speakers’ uncertainty as we found little relation between their use and the participants’ perception of the difficulty of the questions. Although it does not mean that speakers produced fillers because the topic was “not difficult”, it does show that topic difficulty is not necessarily a factor of filler use. Let us consider the following examples:

\*INV: brilliant &ah awesome &uh and &uh [%act:reads] would you like your parents to keep on telling you stories that are not necessarily true? •

\*INV: and [/] and would you do it to your own children in the future ?

•

\*PAR: &oh that's a tough one &=smiles &um [%gaze withdrawal] +/. •

\*INV: +< it's quite a deep question &=smiles. •

\*PAR: I think it depends on the context of that. •

\*INV: yeah [/] yeah. •

\*PAR: I think if it's (...) +//. •

\*PAR: I don't know (.) I think I [//] if I was like a little kindegarden. •

\*PAR: yeah it's quite nice to have like (...) yeah really interesting stories. •

\*INV: +< yeah. •

\*PAR: funny stories told to you by your parents. •

\*INV: +> definitely yeah. •

\*PAR: but then if they were telling me stories to like cover up the fact that they we:ere [%INV laughs] &=shakes his head I don't know criminals or something. •

\*INV: +< yeah [/] yeah &=smiles. •

\*PAR: then no obviously &=laughs. •

\*INV: that'd be out of context. •

\*PAR: &=laughs yeah it's like [/] it's like a good fun thing I think. •

\*INV: sure. •

\*PAR: but in that case like (...) if the son was right and the father is like &=shakes his head (...) a complete fraud and he has just lied all his life. •

\*INV: &m &m. •

\*PAR: I don't think that's (.) •

\*INV: no [/] no. •

\*PAR: +< that's great. •

\*INV: +< but it teaches him a moral lesson at some point\_ •

\*PAR: yeah I think so yeah I think there's [/] there's a balance and I think the father probably went too far in a way &=laughs. •

\*INV: brilliant. •

\*PAR: yeah •

\*INV: fine good stuff. •

21oct15[E|CHAT] 1

Figure 6. Chat transcription, Session H, Question 10

Figure 6 shows Participant H2's answer to question 10, which he later rated "very difficult to answer" according to the questionnaire. Even before uttering his answer, he comments on the difficulty of the question "oh that's a tough one", followed by an 'um', which we coded as marking uncertainty in this case. Yet, even though it seems clear that he is experiencing "difficulty" while formulating his answer at first, he does not produce any other fillers in the rest of his answer. We find other disfluency markers such as repetitions (in "I think there's [/] there's a balance") repairs (in "I think I [//] if I was a little kindergarten") and self-interruptions ("I think if it's +// I don't know"), but no other fillers.

However, if we look at his answer to the first question, which he rated "very easy to answer" (Fig. 7) the amount of fillers is much higher.

\*INV: [%act:reads] what is the movie about? •  
 \*PAR: (.) well ok it's about &um [%gaze withdrawal] (.) a guy. •  
 \*PAR: a:and he:e's getting married. •  
 \*PAR: a:and &uh [%gaze withdrawal] he ends up telling (.) a story of his  
 (..) dad. •  
 \*PAR: &um [%gaze withdrawal] (be)cause basically the dad is at the wedding  
 and everything •  
 \*INV: right. •  
 \*PAR: and &um [%gaze withdrawal] his dad+/. •  
 \*INV: +< xxx good. •  
 \*PAR: &=laughs yeah it's always good isn't it? •  
 \*INV: yeah &=laughs. -  
 \*PAR: <his dad &um> [%said while laughing] ends up telling everybody a  
 story. •  
 \*PAR: and he [//] the [/] the son gets really annoyed (be)cause &um  
 [%gaze withdrawal] he's like &oh my dad is always telling stories  
 and they're all rubbish and made up and he [//] I don't know who my  
 dad really is. •  
 \*PAR: &um [%gaze withdrawal] and the movie just kind of follows (.) like  
 (.) all his dad's stories. •  
 \*INV: &=nods ok. •  
 \*PAR: but he never kind of really finds out the true of +//. •  
 \*PAR: like you get an idea but you never really know. •

Figure 7. Chat transcription, Session H, Question 1

7 occurrences of fillers are found (6 ‘um’ and 1 ‘uh’); only one of which occurs in co-production with another disfluency marker, and four that are placed in initial position.

Planning speech, searching from memory, or conceptualizing an answer is not necessarily a “difficult” task in the strict sense, but it requires a *cognitive effort* from the speaker. Planning an utterance requires thinking ahead, as speakers do not know in advance what they are specifically going to say. The fact that speakers tend to gaze away while they are producing fillers shows that speakers momentarily retreat from the interaction to think about what they are going to say next and how they are going to say it.

Fillers did not often occur in the middle of utterances and did not co-occur with other disfluency markers, which suggests that they do not necessarily display production problems. ‘Uh’ and ‘um’ are usually seen as reflecting lack of confidence and as the “hallmarks of youth” (Fox Tree 2007: 297). Yet, they represent a major part of spontaneous speech. Fillers are universal and found across many languages (Clark & Fox Tree 2002).

Our results show that they mainly occur at the beginning of utterances and serve planning functions at the macro-level. Consider the following utterances:

- (1) **&um** (...) they come at one point that the father didn’t see much himself in the son. (Session A, Participant A2)
- (2) **&um** there is some point at which (.) you’re like ok this is real life. (Session A, Participant A2)
- (3) **&um** so basically he go:oes (.) back to America (be)cause he lives in Paris with his French wife. (Session D, Participant D2)
- (4) **&um** but for me once we got to the ending (.) I was like (.) yeah I feel like I somewhat get this now and get what’s it’s all been about. (Session H, Participant H2)
- (5) **&um** and other than that I’d say the end scene just because it brought everything together for me and (.) made me realize what the film was actually doing. (Session H, Participant H2)
- (6) and **&um** she:e is kind of (.) more sympathetic towards the dad and likes talking to him. (Session G, Participant G2)
- (7) because **&uh** he’s not actually told me lies and all of the characters and people he’s spoken about are real. (Session H, Participant H2)

(8) (.) &um (.) we:ell I think the film would say there is no such thing as reality.  
(Session H, Participant H2)

The utterances above present a certain level of complexity as they are at least made of two clauses. They are also in relation with what was previously said. Levelt (1989: 3) explains that in planning an utterance, speakers need to decide on a purpose for their “next move”. In (3), (4), (5), (6) and (7) fillers are either preceded or followed by conjunctions (‘but’, ‘and’, ‘so’ and ‘because’), which connect the new utterance with previous ones, but which can also “project further talk” (Schegloff 2010: 140). Therefore, fillers allow the speakers to have more time to plan their utterance, and structure upcoming ideas. In these cases, fillers may help speakers manage their speech, and give them enough time to deliver the right message; this idea is consistent with Clark’s (2002) view that fillers are not problems in speaking, but *solutions* in speaking.

## 5. Conclusion

Our hypothesis was that the use of fillers would be linked to the perceived difficulty of the questions we asked our participants (Smith & Clark 1993). However, our findings showed that ‘uh’ and ‘um’ contributed to the planning and structuring of speech (Tottie 2014; Swerts 1998). In this sense, labeling ‘uh’ and ‘um’ as “disfluencies” could be problematic. Tottie (2014: 26) argues:

I would thus like to argue that the stigmatization of uhm as deplorable disfluency or ‘verbal blundering’ is inappropriate. It is a question of balance: Unless it is overused or used in the wrong context, uhm gives listeners time to take in the message that the speaker wishes to transmit in a way that more elegant delivery often does not. Indeed, 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.

Indeed, other authors agree with this view. When describing the notion of ‘fluency’, Hieke (1981: 150) describes filled pauses as being “devices used by speakers to produce more error-free, high quality speech”. Götz (2013: 134) in her book on fluency in native and nonnative speech explains that filled pauses (among other markers) are part of performance phenomena and can be used by L2 learners to sound more native like and less disruptive. Rose (2008) argues in favor of filled pauses as contributing to fluency as they can be seen as strategic devices that language teachers could teach L2 learners to be perceived and used in foreign speech.

If fillers are considered to be markers of disfluency associated with cognitive effort, then one should view disfluency not as a speech disruption phenomenon, but rather as an ambivalent phenomenon. Fillers can reflect planning difficulty, or on the contrary they can be used for planning purposes, depending on the context. Fluency and disfluency could thus be seen as “two sides of the same coin” (Crible, Degand & Gilquin 2017: 71), following Götz’s (2013) componential approach to “(dis)fluency”.

This study has only focused on the specific uses of ‘uh’ and ‘um’ in question-answering, which ultimately limits our conclusions. Further research must thus be carried out, taking into account this componential approach. More attention could also be paid to the pragmatic functions of fillers and their production could be compared to other pragmatic markers. An extended study should also be conducted on the other disfluency markers described in this paper (repairs, repetitions, unfilled pauses, prolongations, false starts) by analyzing their distribution in speech, their co-production, their relation to gaze, and their relation to task difficulty.

## References

- Betz, S. & Wagner, P. (2016). Disfluent Lengthening in Spontaneous Speech. *Elektronische Sprachsignalverarbeitung (ESSV)*.
- Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F. & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech* 44(2), 123-147.
- Brennan, S. E. & Williams, M. (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(3), 383.
- Candea, M., Vasilescu, I. & Adda-Decker, M. (2005). Inter-and intra-language acoustic analysis of autonomous fillers. In *DISS 05, Disfluency in Spontaneous Speech Workshop*, 47-52.
- Christenfeld, N. (1994). Options and Ums. *Journal of Language and Social Psychology* 13(2), 192-199.
- Clark, H. H. (2002). Speaking in time. *Speech Communication* 36(1), 5-13.
- Clark, H. H. (2006). Pauses and hesitations: Psycholinguistic approach. In K. Brown (ed.), *Encyclopedia of Language and Linguistics (2nd ed.)*. Oxford: Elsevier, 244-248.

- Clark, H. H. & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition* 84(1), 73-111.
- Corley, M. & Hartsuiker, R. J. (2011). Why um helps auditory word recognition: The temporal delay hypothesis. *PLoS One* 6(5), e19792.
- Corley, M. & Stewart, O. (2008). Hesitation disfluencies in spontaneous speech: the meaning of *um*. *Language and Linguistic Compass* 2(4), 589-602.
- Crible, L., Degand, L. & Gilquin, G. (2017). The clustering of discourse markers and filled pauses. *Languages in Contrast* 17(1), 69-95.
- Drommel, R. H. (1980). Towards a subcategorization of speech pauses. In H.W. Dechert & M. Raupach (Eds.). *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*. The Hague: Mouton, 227-238.
- Ferreira, F. & Bailey, K. G. (2004). Disfluencies and human language comprehension. *Trends in Cognitive Sciences* 8(5), 231-237.
- Finlayson, I. R. & Corley, M. (2012). Disfluency in dialogue: an intentional signal from the speaker? *Psychonomic Bulletin & Review* 19(5), 921-928.
- Fox Tree, J. E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language* 34(6), 709.
- Fox Tree, J. E. (2007). Folk notions of um and uh, you know, and like. *Text & Talk-An Interdisciplinary Journal of Language, Discourse Communication Studies* 27(3), 297-314.
- Fraundorf, S. H. & Watson, D. G. (2011). The disfluent discourse: Effects of filled pauses on recall. *Journal of Memory and Language* 65(2), 161-175.
- Glenberg, A. M., Schroeder, J. L. & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition* 26(4), 651-658.
- Goffman, E. (1967). *Interaction ritual: Essays on face to face behavior*. Garden City, NY: Double Day.
- Goffman, E. (1971). *Relations in public*. New York: Harper and Row.
- Goldman-Eisler, F. (1968). *Psycholinguistics: experiments in spontaneous speech*. London: Academic Press.
- Goodwin, M. & Goodwin, C. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica* 62(1-2), 51-76.

- Götz, S. (2013). *Fluency in native and nonnative English speech* (Vol. 53). Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology* 56(4), 208.
- Hartsuiker, R. J. & Notebaert, L. (2009). Lexical access problems lead to disfluencies in speech. *Experimental Psychology*.
- Hieke, A. E. (1981). A content-processing view of hesitation phenomena. *Language and Speech* 24(2), 147-160.
- Holmes, V. M. (1988). Hesitations and sentence planning. *Language and cognitive processes* 3(4), 323-361.
- Jucker, A.H. (2015). Pragmatics of fiction: Literary uses of *uh* and *um*. *Journal of Pragmatics* 86, 63-67.
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica* 26, 22-63.
- Kjellmer, G. (2003). Hesitation. In defence of *er* and *erm*. *English Studies* 84(2), 170-198.
- Koponen, M. & Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In *Perspectives on fluency* (pp. 5-24). University of Michigan.
- Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition* 14(1), 41-104.
- Levelt, W. J. (1989). *Speaking. From intention to articulation*. Cambridge, MA: MIT Press Series in Natural-Language processing.
- Lickley, R. J. (2015). Fluency and disfluency. In M. Redford (ed.), *The handbook of speech production* (pp. 445-469). Wiley-Blackwell.
- Maclay, H. & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word* 15(1), 19-44.
- Merlo, S. & Mansur, L. L. (2004). Descriptive discourse: topic familiarity and disfluencies. *Journal of Communication Disorders* 37(6), 489-503.
- O'Connell, D. C. & Kowal, S. (2005). Uh and um revisited: Are they interjections for signaling delay? *Journal of Psycholinguistic Research* 34(6), 555-576.
- Reynolds, A. & Paivio, A. (1968). Cognitive and emotional determinants of speech. *Canadian Journal of Experimental Psychology* 22, 164.

- Rose, R. (2008). Filled pauses in language teaching: Why and how. *Bulletin of Gunma Prefectural Women's University* 29, 47-64.
- Schegloff, E. A. (2010). Some Other "Uh(m)"s. *Discourse Processes* 47(2), 130-174.
- Schnadt, M. J. & Corley, M. (2006). The influence of lexical, conceptual and planning based factors on disfluency production. In *Proceedings of the twenty-eighth meeting of the Cognitive Science Society* (pp. 8-13).
- Shriberg, E. E. (1994). *Preliminaries to a theory of speech disfluencies*. Berkeley: University of California.
- Smith, V. L. & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language* 32(1), 25.
- Swerts, M. (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30(4), 485-496.
- Swerts, M. & Kraemer, E. (2005). Audiovisual prosody and feeling of knowing. *Journal of Memory and Language* 53(1), 81-94.
- Tottie, G. (2011). Uh and um as sociolinguistic markers in British English. *International Journal of Corpus Linguistics* 16(2), 173-197.
- Tottie, G. (2014). On the use of uh and um in American English. *Functions of Language* 21(1), 6-29.