

The following paper was presented at The 9th Workshop on Disfluency in Spontaneous Speech (DiSS 2019) held at ELTE Eötvös Loránd University in Budapest, Hungary on 12–13 September, 2019.

Title: “Uh” is preferred by male speakers in informal presentations in American English

Author(s): Michiko Watanabe, Yusaku Korematsu and Yuma Shirahata

Abstract: This study investigates factors that are likely to be related to speakers' choice of filler type between *uh* and *um* in English, using an informal presentation speech corpus. The effects of the following factors on the probability of each filler type was examined: (1) immediately preceding clause boundary depth, (2) clause size measured as the number of words in the clause, (3) the number of quotation remarks in the clause, and (4) speaker's sex. The filler probabilities increased with the boundary depths. This trend was much stronger with *um* than with *uh*. *Ums* are more likely to appear clause-initially than *uhs*. Clause size had similar effect sizes on the two filler types. The number of quotation remarks had a stronger negative effect with *ums*. Speaker's sex had a significant effect only with *uhs*. *Uhs* are used more frequently by male speakers than by female speakers. The results indicate that speakers' choice of filler type is affected by the combination of multiple factors with various effect sizes.

DOI: <https://doi.org/10.21862/diss-09-012-wata-etal>

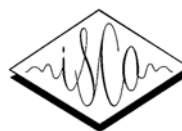
Citation (JIPA): Watanabe, Michiko, Yusaku Korematsu & Yuma Shirahata. 2019. “Uh” is preferred by male speakers in informal presentations in American English. In: R. L. Rose & R. Eklund (eds.), *Proceedings of DiSS 2019, The 9th Workshop on Disfluency in Spontaneous Speech*, 12–13 September, 2019, Budapest, Hungary, 43–46.

The complete proceedings for DiSS 2019 are available as follows.

ISBN: 978-963-489-063-8

DOI: <https://doi.org/10.21862/diss-09>

DiSS 2019 was sponsored by The Faculty of Humanities, ELTE Eötvös Loránd University and the International Speech Communication Association (ISCA).



“Uh” is preferred by male speakers in informal presentations in American English

Michiko Watanabe¹, Yusaku Korematsu² and Yuma Shirahata²

¹Center for Corpus Development, National Institute for Japanese Language and Linguistics, Tokyo, Japan

²School of Engineering, The University of Tokyo, Tokyo, Japan

Abstract

*This study investigates factors that are likely to be related to speakers' choice of filler type between *uh* and *um* in English, using an informal presentation speech corpus. The effects of the following factors on the probability of each filler type was examined: (1) immediately preceding clause boundary depth, (2) clause size measured as the number of words in the clause, (3) the number of quotation remarks in the clause, and (4) speaker's sex. The filler probabilities increased with the boundary depths. This trend was much stronger with *um* than with *uh*. *Ums* are more likely to appear clause-initially than *uhs*. Clause size had similar effect sizes on the two filler types. The number of quotation remarks had a stronger negative effect with *ums*. Speaker's sex had a significant effect only with *uhs*. *Uhs* are used more frequently by male speakers than by female speakers. The results indicate that speakers' choice of filler type is affected by the combination of multiple factors with various effect sizes.*

Introduction

When speakers need extra time to continue their speech, they are likely to slow down the current speech, pause longer, repeat syllables or words, and/or utter fillers to buy time for speech planning. *Uhs* and *ums* are the most common filler types in English. Are there any difference in the usage or functions between these two types? This study investigates factors that are likely to affect speakers' choice of fillers between the two types.

Clark and Fox Tree (2002) found that *uh* and *um* differ with respect to the length of following silent pauses. Long pauses are more likely to be preceded by *um* than by *uh*. From this finding, the authors argue that filler type informs listeners of the seriousness of upcoming problems. On the other hand, Brennan and Williams (1995) observed no difference in listeners' impression about the metacognitive states of speakers between the two types.

It was observed that *ums* are more frequent at deeper discourse boundaries than *uhs* in Dutch monologues (Swerts, 1998). Shriberg (1994) observed that *um*'s rate is higher sentence-initially than sentence-medially, whereas *uh*'s rate is higher

sentence-medially than sentence-initially in “The Switchboard Corpus”, a corpus of informal telephone conversations in American English. These studies suggest that boundary depth is related to speaker's choice between the two types. *Um* has been reported to be more common at deeper boundaries than *uh*. We first examine whether *um* is more likely to appear clause-initially than *uh* by examining the ratios of *ums* and *uhs* at clause-initial and clause-medial positions in “The Corpus of Oral Presentations in English” (COPE). The details of the corpus are given in the method section.

Tottie (2011) observed that male speakers use *uhs* more frequently than *ums* whereas female speakers use roughly the same number of *uhs* and *ums* in British English. It was also observed that *ums* are more frequently used by younger speakers and by those with higher socio-economic status in British English. On the other hand, no sex difference or clear socio-economic effect was observed in American English (Tottie, 2014). These studies indicate that sociolinguistic factors affect speaker's choice of filler type. We examine sex effect on the probability of *uh* and *um* in the present study.

Watanabe and Korematsu (2019) investigated factors that are likely to be relevant to the clause-initial filler probability, using COPE. It was observed that the clause size—measured as the number of words in the clause—as well as the preceding clause boundary depth are related to the clause-initial filler probability. The filler probability increased with the clause boundary depth and with the number of words in the clause. It was also found that the number of quotation remarks in the clause negatively affects the clause-initial filler probability. Sex was also examined as a factor, but no significant effect was observed. Based on these findings, we included the clause size and the number of quotation remarks in the clause as possible predicting factors of the probability of *uh* and *um* at clause-initial and clause-medial positions in the present study.

Method

Speech material

We used COPE as material. COPE contains twenty informal English presentation speeches recorded in

Portland, Oregon and Los Angeles, California in 2013. An overview of COPE is given in Table 1.

Table 1. Overview of COPE.

Number of speakers	20
Total duration (minutes)	227
Total number of words	38,370
Total number of fillers	1,441
Clause initial fillers	941
Clause medial fillers	500

The speakers were university students or university graduates in their 20s and early 30s. They were given a topic, “the most memorable event in my life”, and instructed to give a talk for at least 10 minutes. They gave their talks in front of a small audience including their friends and the recording staff. Disfluencies were labeled, and clause boundaries were marked on the transcription. Only *uh* and *um* were identified as fillers. An excerpt from the transcription of COPE is given in Example (1) below.

Example (1)

- 1: <cb and when> and when we got to (I like) (r2 this little) this little food court area <ce and when>
 2: /cb/ he stopped me /ce/
 3: /cb and/ and he was like (qcb) so would you want (ncb to) to be my girlfriend (nce to) (qce) /ce and/
 4: /cb and/ and I literally jumped on him /ce and/
 5: /cb and/ and I have no idea (ncb why) why I did that (nce why) /ce and/
 (From losF09 in COPE)

Transcriptions of COPE are partitioned at coordinate or adverbial finite clause boundaries as in Example (1). Ankle brackets, <cb> and <ce> indicate finite adverbial clause beginning and end, respectively. Connectives are included within the brackets. “/cb/” and “/ce/” in line 2 mean coordinate clause beginning and end, respectively. “(qcb)” and “(qce)” in the third line indicate quotation clause beginning and end, respectively. “(ncb to)” and “(nce to)” in the third line indicate nominal infinitive clause beginning and end, respectively. Similarly, “(ncb why)” and “(nce why)” in the fifth line indicate nominal clause beginning with “why” and its end, respectively.

Procedure

The following factors were examined as those possible to predict the probability of *uh* and *um*.

(1) Clause boundary depth

Whether and how the effect size of clause boundary depth differs between *uh* and *um* was examined.

In English, it is difficult to tell whether a given clause boundary is also a sentence boundary or a boundary within a sentence, because no period or comma is used in speech. There is no syntactic or morphological cue to tell boundary depth in English. In order to evaluate clause boundary depth, we had three labellers add sentence boundary labels to clause boundaries which they judged to be deep boundaries. The labellers were instructed to mark sentence boundaries based on the content and the prosody of speech. We estimated boundary depth based on the number of labellers who marked the boundary as a sentence boundary. We regarded boundaries marked by three labellers as sentence boundaries to be the deepest, those marked by two labellers to be the second deepest, and so forth. Thus, clause boundaries were grouped into four types depending on perceived boundary depth. The number of boundaries in each group is given in Table 2.

Table 2. Number of boundaries in each group.

Type0	Boundaries marked by no labeler	719
Type1	Boundaries marked by one labeler	1521
Type2	Boundaries marked by two labelers	979
Type3	Boundaries marked by three labelers	806
Total		4025

(2) Clause size

The number of words in each clause was counted and regarded as an index of message complexity to be conveyed by the clause. Fillers and other disfluencies were not counted as words.

(3) Number of quotation clauses

Speech sometimes contains quoted remarks. The number of quotation clauses in a given clause was found to have a negative effect on the containing clause initial filler probability (Watanabe and Korematsu, 2019). The number of quotation clauses was included as a possible predicting factor of the probability of *uh* and *um*.

(4) Speaker’s sex

Speaker’s sex was included as a predicting factor, because Tottie (2011, 2014) reports that male speakers use more *uhs* than *ums* while female speakers use roughly the same number of the two types in British English, whereas no such difference was observed in American English.

We excluded clauses containing more than 30 words from analysis because samples were sparse for larger clauses.

We estimated the probability of *uh* and *um* at clause-initial and clause-medial positions separately using a generalized linear mixed model, with maximum likelihood estimation of variance components. Because the response variable was

binary, we conducted mixed-effects logistic regression. Fixed effects predictor variables were the factors (1) through (4). Speakers were treated as a random effects factor. We used `glmer` function in `lme4` package and `MuMIn` and `lmerTest` packages running under R version 3.5.1.

Results

The proportion of clause-initial and clause-medial *uh* and *um*

Table 3 shows the frequency and the ratio of *uh* and *um* at clause-initial and at clause-medial positions. *Ums* are far more frequent than *uhs* clause-initially, whereas the difference is small clause-internally. The ratio shows that 73% of *ums* are used clause-initially, while *uhs* are used roughly equally clause-initially and clause-medially. The results are in accordance with the results of previous studies that *ums* tend to be used at deeper boundaries than *uhs*.

Table 3. Frequencies and ratios of *ums* and *uhs* at clause-initial and clause-medial positions.

	Frequency			Ratio	
	Initial	Medial	Total	Initial	Medial
<i>uh</i>	245	259	504	0.49	0.51
<i>um</i>	683	254	937	0.73	0.27

Factors related to clause-initial *uhs* and *ums*

Table 4 shows the model results of clause-initial *uhs*, and Table 5 *ums*. Estimates of categorical variables are given as the relative values to one of its levels whose estimate is zero in R. The reference variables are included in the table. Odds ratio indicates the degree of effect size of each factor. When the ratio is close to 1.0, the effect size is small. The more distant the ratio from 1.0, the larger the effect size.

First, both clause-initial *uhs* and *ums* are significantly related to the boundary depth. The filler probabilities increase with the boundary depths. This

Table 4. Results of mixed-effects logistic regression for clause-initial *uh*.

Variable	Estimate	Std. Error	z value	Pr(> z)		Odds ratio
(Intercept)	-5.068	0.3868	-13.103	< 2e-16	***	0.006
Boundary depth	0.398	0.070	5.667	0.000	***	1.488
Number of words	0.039	0.012	3.176	1.49e-03	**	1.040
Number of quotation clauses	-0.092	0.146	-0.633	5.26e-01		0.912
Gender Female	0					
Male	1.655	0.456	3.628	0.000	***	5.233

Table 5. Results of mixed-effects logistic regression for clause-initial *um*

Variable	Estimate	Std. Error	z value	Pr(> z)		Odds ratio
(Intercept)	-3.966	0.326	-12.184	< 2e-16	***	0.019
Boundary depth	1.059	0.054	19.440	< 2e-16	***	2.884
Number of words	0.045	0.009	4.794	1.63e-06	***	1.046
Number of quotation clauses	-0.708	0.184	-3.846	0.000	***	0.493
Gender Female	0					
Male	-0.192	0.416	-0.462	0.644		0.825

trend is much stronger with *um* than with *uh*. Second, the clause size has a significant effect on both filler types. The effect sizes are very close to each other. Third, the number of quotation clauses has a significant negative effect only with *um*. *Uh* is not affected by the factor. Fourth, sex effect is significant only with *uh*. Male speakers use significantly more *uhs* than their female counterparts.

We reanalyzed 90% of the data only with significant variables with speakers as a random effects factor, and estimated filler probability for the remaining data.

Figure 1 and Figure 2 illustrate estimated probabilities of clause-initial *uh* and *um* for each boundary type as a function of the number of words in the clause. As sex factor had a significant effect on *uh*, the probability of *uh* is shown separately for female and male speakers. Filled circles indicate observed boundaries with fillers and pluses indicate

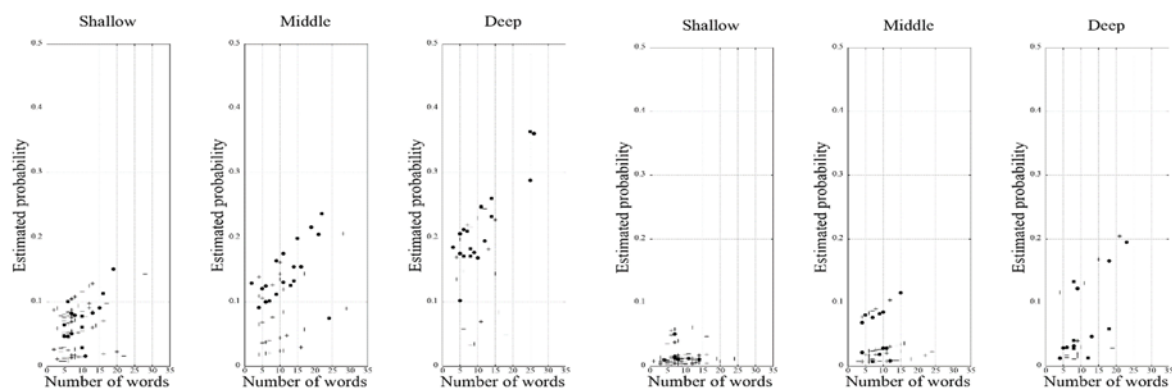


Figure 1. Estimated clause-initial probability of *uh* for each boundary type as a function of number of words in the clause: the upper figure for male speakers and the lower for female speakers. Filled circles indicate observed boundaries with fillers and pluses without fillers.

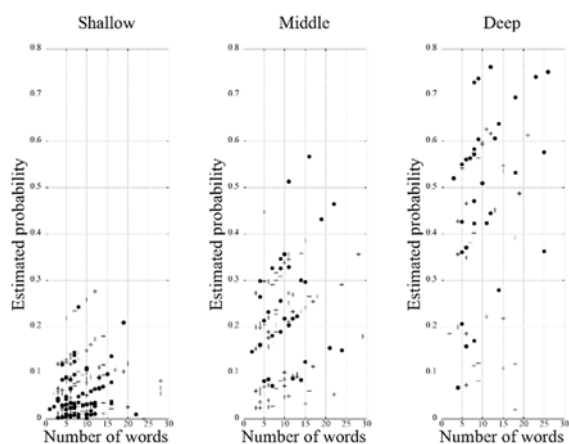


Figure 2. Estimated clause-initial probability of *um* for each boundary type as a function of number of words. Filled circles indicate observed boundaries with fillers and pluses without fillers.

observed boundaries without fillers. Type0 and Type1 boundaries in Table 2 are shown combined as shallow boundaries.

A comparison of Figure 1 and Figure 2 demonstrates difference in effect size of boundary depth. A comparison of the effects on female and male speakers in Figure 1 indicates that sex effect is not small with *uh*. Male speakers use *uh* at a much higher rate than female speakers.

Factors related to clause-medial *ums* and *uhs*

Table 6 shows the results of clause-medial *uhs*, and Table 7 *ums*. Boundary effects with clause-medial fillers are weaker than those with clause-initial fillers for both *uh* and *um*. The effects are marginally significant for both types. Effect size of number of words with clause-medial fillers are larger than that with clause-initial fillers for both types. Effects of number of quotation clauses are significant for both types. Sex effect is significant only with *uh*, as is the case with clause-initial fillers.

Discussion

The results are in accordance with those of previous research that *um* is preferred clause-initially than *uh* (Shriberg, 1994; Swerts, 1989). Message units between deep boundaries are generally larger than those between shallow boundaries. It is inferred that *um* reflects speaker’s cognitive load of planning larger units than *uh*.

Clause size measured as the number of words had similar effect sizes on *uh* and *um*. It is likely that *uh* and *um* reflect the message complexity to be conveyed in the clause to a similar degree.

Table 6. Results of mixed-effects logistic regression for clause-medial *uh*.

Variable	Estimate	Std. Error	z value	Pr(> z)		Odds ratio
(Intercept)	-5.594	0.3829	-14.611	<2e-16	***	0.004
Boundary depth	0.133	0.076	1.759	0.079	.	1.142
Number of words	0.133	0.012	10.946	<2e-16	***	1.142
Number of quotation clauses	-0.288	0.146	-1.979	4.78e-02	*	0.750
Gender Female	0					
Male	1.341	0.447	3	0.003	**	3.823

Table 7. Results of mixed-effects logistic regression for clause-medial *um*.

Variable	Estimate	Std. Error	z value	Pr(> z)		Odds ratio
(Intercept)	-5.261	0.334	-15.75	< 2e-16	***	0.005
Boundary depth	0.153	0.078	1.959	0.050	.	1.165
Number of words	0.185	0.013	14.531	< 2e-16	***	1.203
Number of quotation clauses	-0.687	0.198	-3.472	0.001	***	0.503
Gender Female	0					
Male	-0.368	0.383	-0.962	0.336		0.692

A sex effect was observed with *uh*. It is likely that sociolinguistic factors play a role in the choice of filler type not only in British English but also in American English.

Acknowledgements

This research is supported by JSPS KAKENHI, Grant Numbers 15K02553 and 18K00559.

References

Brennan, S. E. & M. Williams. 1995. The feeling of another’s knowing: prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and language* 34(3): 383–398. <https://doi.org/10.1006/jmla.1995.1017>

Clark, H. H. & J. E. Fox Tree. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84(1): 73–111. [https://doi.org/10.1016/S0010-0277\(02\)00017-3](https://doi.org/10.1016/S0010-0277(02)00017-3)

Shriberg, E. 1994. *Preliminaries to a theory of speech disfluencies*. Ph.D. dissertation, University of California at Berkeley.

Swerts, M. 1998. Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30(4): 485–496. [https://doi.org/10.1016/S0378-2166\(98\)00014-9](https://doi.org/10.1016/S0378-2166(98)00014-9)

Tottie, G. 2014. On the use of *uh* and *um* in American English. *Functions of Language* 21(1):6–29. <https://doi.org/10.1075/fo1.21.1.02tot>

Tottie, G. 2011. *Uh* and *Um* as sociolinguistic markers in British English. *International Journal of Corpus Linguistics* 16(2): 173–197. <https://doi.org/10.1075/ijcl.16.2.02tot>

Watanabe, M. & Y. Korematu. 2019. Comparison of factors related to clause-initial filler probabilities in English and Japanese, *Proceedings of ICPHS 2019, the International Congress of Phonetic Sciences*, 4–10 August 2019, Melbourne, Australia, 2440–2444.