Understanding Fillers May Facilitate Automatic Sarcasm Comprehension: A Structural Analysis of Twitter Data and a Participant Study

Fatemeh S. Tarighat
UKRI CDT in NLP
University of Edinburgh
f.samadzadeh-tarighat@sms.ed.ac.uk

Walid Magdy
School of Informatics
University of Edinburgh
wmagdy@inf.ed.ac.uk

Martin Corley
School of PPLS
University of Edinburgh
martin.corley@ed.ac.uk

Abstract

Sarcasm detection models are often built based on self-annotated tagged data. However, fillers (e.g., *um* and *hmm*), deliberate use of which may indicate sarcasm, do not get enough attention in these models. We analyze five fillers in different categories of untagged tweets. We also present participant ratings of sarcasm, offensive language, language formality, and basic emotions in tweets with and without *um* and *hmm*. Our evidence, albeit weak, points to the importance of linguistic features such as these fillers in determining sarcastic meaning.

1 Introduction

Transcribed spoken language and user-generated online text are two of the main sources of training data for language models. Traditionally, fillers have been dismissed as noise in transcription of spoken language. However, the importance of understanding fillers and disfluencies of natural language has been emphasized in human-computer interaction research (e.g, Bates et al., 1993; Oviatt, 1995; Wigdor et al., 2016) with focused studies on real-time dialogue systems (Passali et al., 2022), question answering systems (Gupta et al., 2021), and autonomous vehicles (Large et al., 2017) in recent years. Moreover, sarcasm detection models do not account for linguistic details of their training resources and mainly rely on user-generated tags and indicators of sarcasm to flag remarks as sarcastic or non-sarcastic (Oprea and Magdy, 2020). This is all the more important as written language online is adopting elements of spoken language, e.g., the deliberate inclusion of fillers such as *uh* and *um* (also known as filled pauses and discourse markers). Whether spoken or written, fillers can convey sarcasm, among other things (D'Arcey et al., 2019).

With this in mind, we investigated 5 fillers and their potential sarcastic meanings on Twitter. Taking into account the type of tweets, we hypothesized that [1] users have position preferences when using fillers online and fillers appear more in the middle if the tweet is a stand-alone one and not in response to another tweet. [2] In contextually self-sufficient tweets, fillers are often perceived to deliver sarcasm. [3] Contextually independent tweets with filler somewhere in the middle get rated as sarcastic more than structurally similar tweets with filler appearing at the beginning or at the end.

2 Data Collection and Processing

We studied over 1.4 million English tweets containing *um*, *uh*, *hmm*, *erm*, *er*, and *#sarcasm* collected through the Twitter Application Programming Interface[1] using `twitter_collector`[2] over the span of 23 days. We excluded *#sarcasm* data from the study because most tweets including this tag did not include the fillers under investigation.

Our investigation focused on tweets that were classified as stand-alone, which could contain mentions (*@username*) or media but were not quotes, replies, or retweets. We reviewed random samples of these tweets to look for context-independent content to be used in our participant study. To ensure context-independence of the language in the tweets, we divided our sample into two groups; SELF-CONTAINED: tweets that only include text and emojis and MEDIA-URL: tweets that contain a form of media (e.g., image, GIF, video) and/or URLs (Table 1). 10% of the tweets analyzed included mentions.

	um	*uh*	*hmm*	*erm*	*er*
MEDIA-URL	33439	40934	42373	3302	19274
SELF-CONTAINED	106104	136066	167242	8006	56098
	139543	177000	209615	11308	75372

Table 1: Stand-alone tweets including each type of filler, with media content or links, or fully self-contained.

[1] https://developer.twitter.com/en/docs/twitter-api

[2] https://github.com/yalhariri/twitter_collector

Proceedings of the 26th Workshop on the Semantics and Pragmatics of Dialogue, August, 22-24, 2022, Dublin.

	um	*uh*	*hmm*	*erm*	*er*
Beginning	0.34	0.27	0.38	0.38	0.06
Middle	0.62	0.69	0.45	0.54	0.89
End	0.04	0.04	0.17	0.08	0.06

Table 2: Proportions of tweets in each position, by matched filler in automatically selected database (612,838 tweets).

Table 2 shows the proportions of tweets which matched the search criteria containing each of the fillers under investigation at the beginning, in the middle, and at the end. As can be seen, there is a tendency for fillers to occur in the middle of tweets.

3 Participant Study

We created a pool of 2300 SELF-CONTAINED tweets by randomly selecting 10 tweets per day. We applied several rounds of filtering, e.g., to remove false positives such as 'ER' for 'emergency room'. We manually selected 48 tweets, 24 containing *um* and 24 containing *hmm*. Two independent NLP researchers conducted context sufficiency checks for us. Each set of 24 tweets included 8 with the filler at the beginning, 8 with the filler in the middle (defined as any word except the first or last), and 8 with the filler at the end. To investigate the specific role played by the fillers, we controlled for content, by creating versions of each of the 48 tweets which were identical in every respect, but had the fillers removed. The resulting 96 tweets were counterbalanced into two lists of 48, each including equal numbers of examples of each filler at tweet beginning, middle, and end, as well as a matched number of tweets with fillers excised from the same positions.

An experiment was administered via Prolific[3]. Participants were asked to rate tweets for sarcasm (SARCASM), offensive language (OFFENSE), language formality (FORMALITY), and emotions associated with the tweets (Ekman's six basic emotions, not discussed further here) in 5-point Likert and slider question formats. We assumed that self-contained tweets containing fillers should get rated as more sarcastic, more offensive, and less formal compared to their without-filler counterparts. We also wanted to know whether any effect of presence/absence of fillers was moderated by their positions in the tweets.

[3]https://www.prolific.co/

um			*hmm*		
Question	Position	Mean	Question	Position	Mean
Sarcasm	Beginning	3.15	Sarcasm	Beginning	3.59
	Middle	3.50		Middle	3.51
	End	2.67		End	2.85
Offense	Beginning	2.63	Offense	Beginning	2.89
	Middle	2.63		Middle	2.92
	End	1.99		End	2.35
Formality	Beginning	2.12	Formality	Beginning	2.27
	Middle	2.05		Middle	2.37
	End	1.95		End	2.41

Table 3: Mean ratings (0–5) for tweets with *um* or *hmm* present/absent in three positions, for SARCASM, OFFENSE, and FORMALITY.

96 participants took part in the study. We found weak evidence supporting our claims (Table 3). [1] For both *um* and *hmm*, SARCASM scores are slightly higher when fillers are present. [2] For OFFENSE, fillers seem to contribute to offensive tone with the highest contrast in *um* beginning and *hmm* middle. Also, *um* middle and end are the only instances where offensive language scores are slightly lower when the filler is present. Thus, they are the only instances that seem to slightly take the sting away from remarks. [3] Tweets with both fillers in all positions get rated as less formal when fillers are present. [4] Surprise is the only emotion that gets rated more when the filler is present.

4 Discussion

The present study is limited in scope and shows only weak evidence in support of its hypotheses. However, the numerical indication that inclusion of fillers increases the perception of sarcasm suggests that a larger-scale study is warranted. As our next step, we will study MEDIA-URL tweets along with quotes and replies in our data set in a similar fashion. We can then investigate fillers in self-annotated sarcastic tweets to check whether tweets are perceived sarcastic regardless of the filler in them. A better understanding of linguistic features such as fillers would allow us to train language understanding, prediction, and detection models with more accuracy.

Acknowledgments

This work was supported in part by the UKRI Centre for Doctoral Training in Natural Language Processing, funded by the UKRI (grant EP/S022481/1) and the University of Edinburgh, School of Infor-

matics, and School of Philosophy, Psychology & Language Sciences. Ethical approval for data collection and participant study was granted by the University of Edinburgh, School of Informatics Ethics Committee according to the Informatics Research Ethics Process, RT number 2021/29989.

References

Madeleine Bates, Robert Bobrow, Pascale Fung, Robert Ingria, Francis Kubala, John Makhoul, Long Nguyen, Richard Schwartz, and David Stallard. 1993. The bbn/harc spoken language understanding system. In *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 111–114. IEEE.

J Trevor D'Arcey, Shereen Oraby, and Jean E Fox Tree. 2019. Wait signals predict sarcasm in online debates. *Dialogue & Discourse*, 10(2):56–78.

Aditya Gupta, Jiacheng Xu, Shyam Upadhyay, Diyi Yang, and Manaal Faruqui. 2021. Disfl-qa: A benchmark dataset for understanding disfluencies in question answering. *arXiv preprint arXiv:2106.04016*.

David R Large, Leigh Clark, Annie Quandt, Gary Burnett, and Lee Skrypchuk. 2017. Steering the conversation: a linguistic exploration of natural language interactions with a digital assistant during simulated driving. *Applied ergonomics*, 63:53–61.

Silviu Vlad Oprea and Walid Magdy. 2020. The effect of sociocultural variables on sarcasm communication online. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1):1–22.

Sharon Oviatt. 1995. Predicting spoken disfluencies during human-computer interaction. *Computer Speech and Language*, 9(1):19–36.

Tatiana Passali, Thanassis Mavropoulos, Grigorios Tsoumakas, Georgios Meditskos, and Stefanos Vrochidis. 2022. Lard: Large-scale artificial disfluency generation. *arXiv preprint arXiv:2201.05041*.

Noel Wigdor, Joachim de Greeff, Rosemarijn Looije, and Mark A Neerincx. 2016. How to improve human-robot interaction with conversational fillers. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 219–224. IEEE.