Disfluencies in conversation: a comparison of utterances with and without metaphors

Amy Han Qiu[1], Vanessa Vanzan[1], Chara Soupiona[2], and Christine Howes[1]

[1]Department of Philosophy, Linguistics and Theory of Science,
Faculty of Humanities, University of Gothenburg
[2]Department of Philology, Division of Linguistics, University of Crete

Abstract

Disfluencies are pervasive in conversations and commonly regarded as indicative of cognitive difficulties. However, they have rarely been examined in utterances with metaphors, which are considered to be more cognitively challenging than those without any metaphors. In this paper, we investigate the occurrence of filled pauses and self-repairs in conversational turns with and without metaphors, across various word counts. Results showed that metaphor presence and word count contributed significantly to the probabilities of filled pauses and self-repair. Notably, there was a significant interaction between metaphor presence and word count, highlighting the combined cognitive demands elicited by using metaphors and producing longer utterances as key factors influencing disfluencies in spontaneous conversations.

1 Introduction

Disfluencies are often characterised as disruptions or breaks in the flow of communication, such as hesitations, pauses, filled pauses, and self-repairs such as repetitions or reformulations. These "interruptions" occur commonly in everyday interactions and impact how language is conveyed and interpreted. Self-repairs and filled pauses reflect incremental processing, with real-time adjustments made word-by-word, as the speaker progresses through the utterance. Following the incremental view, disfluencies are natural byproducts of the dynamic processes involved in generating speech. Disfluencies may occur in different cases, for example cognitive difficulties (Levelt, 1983; Bortfeld et al., 2001; Clark and Tree, 2002), heightened attention of the ongoing communication (Cienki, 2020), and interactive issues (Goodwin, 1979).

Disfluencies have been extensively explored in various linguistic dimensions, for example word-related features like word class, utterance features like utterance types and sentence lengths, and conversational dynamics like speaker exchange (Shriberg, 1996). However, they were rarely examined for their relationships with the use of metaphors, which involves talking and potentially thinking about something in terms of something else (Semino, 2008).

Processing and producing metaphors are typically assumed to demand extra cognitive resources due to the inherent complexity in cross-domain mappings. The mapping of features between two distinct domains, as well as the wording, may lead to heightened cognitive pressure (Lakoff and Johnson, 1980; Steen, 2023), which makes metaphor use an intriguing yet unexplored avenue for exploring the conversational dynamics of disfluencies.

Below is an example of disfluencies in metaphor use, cited from a conversation about an ethical dilemma of sacrificing one person to save more people (metaphorical parts in bold type and disfluency markers italicised):

(1) "*Bu- bu- but* are you *s- saying* that *um uh* we need to **value** the sort of the **worth** of each person?"

In this example, the importance of a person is interpreted in terms of the financial worth of a property. The utterance is characterised by disfluencies, indicated by repetitions ("bu- bu-"), filled pauses ("um uh"), and self-repair ("s-saying"). They interrupt the flow of speech and may indicate uncertainty, interactive issues or even difficulty in articulating the intended message smoothly.

Despite the well-acknowledged link between cognitive pressure and disfluencies (Levelt, 1983; Bortfeld et al., 2001; Clark and Tree, 2002), whether the cognitive complexity associated with metaphor use contributes to the probabilities of disfluencies in an utterance remains an interesting research question.

Based on the transcripts of 19 triadic conversations, this study compares the probability of

Proceedings of the 28th Workshop on the Semantics and Pragmatics of Dialogue, September, 11–12, 2024, Trento, Italy.

filled pauses and self-repairs in utterances with and without metaphors. Disfluencies in metaphor use, which is a linguistic phenomenon characterized by inherent cognitive complexities, could provide insights into the interaction between metaphor, word count, and different types of disfluencies.

1.1 Filled pauses and self-repairs

Filled pauses and self-repairs are two of the most studied forms of disfluencies (Clark and Tree, 2002). Below is an example:

(2) *Mmm*, is there *any any* other line of thought, that we can think?

Filled pauses like *"mmm"* in example (2) can serve as markers of language processing, indicating moments of word retrieval, linguistic uncertainty and speech planning (Clark and Tree, 2002).

Self-repair in metaphorical dialogues reveals how speakers manage errors and refine their language in real-time communication. Unlike filled pauses, self-repairs like repetitions (*"any any"*) specifically involve the speaker interrupting their ongoing speech to correct or revise what they have just said. Self-repairs play a crucial role in maintaining shared understanding and mutual interpretation in effective communication (Clark, 1996). Additionally, self-repairs contribute to the negotiation of meaning between speakers, as they indicate the awareness of one speaker of the other's comprehension needs and their willingness to clarify or elaborate on their message.

Some studies showed that disfluencies can occur due to heightened cognitive pressure. Previous research has examined disfluencies in different utterance types (Oviatt, 1995; Shriberg, 1996; Lickley, 2001). Longer and more syntactically complex turns were found to have a higher frequency of repetition disfluencies, and giving instructions or expressing uncertainty when answering questions was associated with a greater use of filled pauses. Similar patterns have been found at the beginning of utterances, where cognitive pressure is assumed to be high due to speech planning. Some empirical studies found that disfluencies can serve as a compensatory cognitive strategy, aiding the speaker to manage the cognitive load in conversation (Brennan and Schober, 2001; Bailey and Ferreira, 2007; Howes et al., 2017).

Disfluencies could also arise for interactive reasons, assisting the interlocutors in adjusting their communicative comprehension strategies. For example, self-repair can reflect the speaker's intention to maintain their turn and to regulate the flow of conversation (Goodwin, 1979). When an interlocutor is puzzled or needs more time for speech planning, filled pauses can facilitate smoother communication by affording a longer time for accommodating these cognitive challenges.

There are also studies that linked disfluencies with the speaker's awareness of the ongoing communication. According to Cienki (2020), the occurrence of disfluencies can be a key signal of the speaker's awareness of the impact of their language use on the hearer, reflecting their "metacommunicative awareness (MCA)"[1] in conversations. The more effortful the speech is, the stronger the MCA can be (Cienki, 2020).

1.2 Disfluencies in metaphor use

An interesting phenomenon often overlooked in disfluency research is the use of metaphors. According to Kaal (2012), 2.9% to 10.1% of lexical units in conversations are metaphor related. Below is an example of conversational turns with metaphorical lexical units:

(3) He is quite **far away from** a breakthrough.

In example (3) the lexical units "**far**" "**away**" and "**from**" are metaphorically deployed to signify physical distance from achieving success, which may introduce a layer of cognitive complexity to the interpretation process. By contrast, utterances like "He is unlikely to achieve a breakthrough". is a direct account of the low probability of achieving success, without linking to a more basic meaning.

Utterances containing metaphors are typically regarded as more cognitively challenging than those without any metaphors (Gibbs, 1994). According to Lakoff and Johnson (1980); Steen (2023), processing metaphor presumably requires extra cognitive resources due to the need for inferential work and the mapping of complex relationships between domains. This argument finds support in a recent study, which showed that when word count is controlled for, speakers invest a longer time articulating an idea with metaphors compared to those without metaphors (Qiu et al., 2024, in progress). Neuroscientific research has also shown

[1] Other possible MCA signals include gestures, prosodic markers like stress, marked intonation, and use of pauses, verbal cues like modification, diversification, extension, literalisation, etc, (refer to Cienki, 2020 for more details).

Proceedings of the 28th Workshop on the Semantics and Pragmatics of Dialogue, September, 11–12, 2024, Trento, Italy.

that metaphor production, compared with the use of literal language, involves more intense cognitive work, and an increase of brain activation grows with the increase of creativity in the metaphors (Benedek et al., 2014).

As noted earlier, heightened cognitive pressure elicited by complicated linguistic tasks and producing long utterances may be major contributors to increased disfluencies in spontaneous speech (Levelt, 1983; Bortfeld et al., 2001; Clark and Tree, 2002). Focusing specifically on filled pauses and self-repairs, this paper explores whether the inherent cognitive complexity associated with metaphor use adds to the probability of disfluencies in conversation turns.

2 Research questions

This paper aims to address the following research questions:

1. What is the difference between utterances with and without metaphors in the probabilities of containing filled pauses?

2. What is the difference between utterances with and without metaphors in the probability of self-repairs?

3. Does metaphor presence interact with word count in terms of the occurrence of filled pauses or self-repairs?

According to previous research on the inferential processes involved in metaphor use (Steen, 2023; Benedek et al., 2014) and research on disfluencies (Lavelle et al., 2012; Bortfeld et al., 2001; Oviatt, 1995; Shriberg, 1996), metaphor processing imposes greater cognitive demands and should thereby lead to more disfluencies, while longer sentences similarly increase cognitive load and disfluency rates (Bortfeld et al., 2001). We therefore hypothesise that:

1. Utterances containing metaphors are more likely to contain filled pauses compared to utterances not containing metaphors

2. Utterances containing metaphors are more likely to contain self-repairs compared to utterances not containing metaphors

3. When metaphors are present, longer turns will be associated with an increased frequency of self-repairs and filled pauses compared to turns without metaphors.

3 Methods

3.1 Data

The data consists of 19 face-to-face triadic conversations between 57 participants who were unfamiliar with each other. The conversations were video and audio recorded, lasting from 5 to 10 minutes each. The data were collected earlier as the control condition in an experiment comparing conversations among healthy people to those involving a patient with schizophrenia. The participants were instructed to discuss the balloon task, an ethical dilemma in which one of the four hot air balloon passengers should sacrifice themselves by jumping out to their certain death in order to save the other three. The data collection procedure and other details are reported in Lavelle et al. (2012) and Howes and Lavelle (2023).

This study examines conversation utterances at the level of turns, which refers to all sub-utterances, segmented by filled pauses, unfilled pauses, laughters, etc., produced by one speaker before the next speaker starts to talk. Turns may vary in length; some turns may comprise multiple sub-utterances and are thus longer than others.

The 19 conversations consisted of $3,785$ turns, among which 849 turns contained only laughter, cough, unclear utterances, or backchannels (e.g., "yeah", "ummm", "okay"). As including these turns may inflate the number of utterances without metaphors, they were filtered out from further analysis. $2,631$ turns were preserved, which include a total of $24,476$ words. The mean of total word count per conversations is 1288.21 (95%CI: 1060.85 - 1515.58). The mean word count of each turn is 9.28 (95%CI: 8.92 - 9.69).

3.2 Disfluencies Identification

Filled pauses were identified manually based on a find-and-replace operation on inconsistently spelt cases (see Howes et al. 2017 for more details).

Self-repairs were identified with STIR (STrongly Incremental Repair detection), an automatic incremental self-repair detection system (Hough and Purver, 2014). STIR was trained and initially tested on the Switchboard corpus of telephone conversations (Godfrey et al., 1992). The system has a high accuracy rate and high correlations with human coders in detecting self-repair rates (Howes et al., 2014).

Although the numbers of disfluencies detected

Proceedings of the 28th Workshop on the Semantics and Pragmatics of Dialogue, September, 11–12, 2024, Trento, Italy.

in each turn are available, convergence issues were found when running the statistical models. Therefore, filled pauses and self-repair were annotated as binary variables based on whether a disfluency marker of the relevant type was present in the turn. These annotations were taken directly from Howes et al. (2017).

3.3 Metaphor Identification

Metaphorically used lexical units were annotated manually following the Metaphor Identification Procedure VU (MIPVU; Steen et al., 2010). The criteria for identifying metaphoricity is whether the word has a more basic meaning that is "more concrete, body-related, more precise, or historically older" (Steen et al., 2010), and whether the contextual meaning contrasts with the basic meaning but can be understood in comparison with it.

For the present study, each lexical unit was annotated based on its basic meaning and contextual meaning provided by three dictionaries, i.e., the Longman Dictionary, the Oxford English Dictionary, and WordNet. Annotations provided by the VUAMC[2] (Steen et al., 2010), the largest available corpus hand-annotated for metaphorical language use, were used as references to enhance inter-reliability. The presence of metaphors was annotated at the level of turns as a binary variable. Turns that contained at least one metaphorically used lexical unit were annotated as metaphorical, and those without any metaphorically used lexical units were annotated as non-metaphorical.

To test the inter-rater reliability, two annotators worked independently on 10% of randomly selected data. The annotations reached 97.1% agreement (Cohen's kappa = 0.88). More details about metaphor identification and inter-rater reliability checks are reported in Qiu et al. (2024, in progress).

Following this approach, 690 turns were identified as containing metaphorically used lexical units, and 1,941 as not containing any metaphors.

3.4 Statistical Methods

We ran a series of binomial Generalized Linear Mixed Models (GLMMs). A random intercept assigned for conversation groups was included in the models to account for potential correlation among observations within the same group. In this study, we compared models that took word count as an interaction term and those with word count as the co-variate. For cases where the interaction effect was significant, we compared the effect of word count on the two levels of metaphor presence with further stratified analyses.[3]

$P < 0.05$ was set as the threshold of statistical significance for all models. The analyses were run with the lmer function from the lme4 package of R.

4 Results and Discussion

4.1 Descriptive statistics

Among the 2,631 turns, 282 contained filled pauses and 551 contained self-repair. The overall rate of filled pause presence is 10.72%, and that of self-repair presence 20.94%. In utterances with and without metaphors, the proportion of filled pauses presence are 19.42% and 7.62%, and the proportion of self-repair presence are 30.43% and 17.57%. Both disfluency markers occurred more frequently in utterances with metaphors. Descriptive statistics are summarized in Table 1.

Metaphor	**Filled pauses**		**Self-repair**		**Total**
Presence	Yes	No	Yes	No	
Yes	134	556	210	480	690
No	148	1793	341	1600	1941
Total	282	2349	551	2080	2631

Table 1: Filled pauses and self-repair presence in the dataset

The probability of disfluency increases with word count. The Biserial correlation between word count and the presence of filled pauses is 0.3 (p<.01) and that between word count and the presence of self-repairs is 0.38 (p<.01).

Presence of filled pauses

The modelling results are summarised in Table 2. Both metaphor presence and word count have a significant main effect. In particular, when word count is held constant at 9.30, utterances with metaphors are more likely to contain filled pauses compared to utterances without metaphors. According to

[2]http://www.vismet.org/metcor/documentation/home.html (last accessed May 26, 2024).

[3]Including a random slope for word count and adding Participant ID as nested in Group caused singularity, which makes the options unfeasible.

the co-variate model, the predicted probabilities of filled pauses in the two utterance types are 12% ($95\%CI : 9\% - 16\%$) and 8% ($95\%CI 6\% - 10\%$), respectively.

The interaction effect between metaphor presence and word count on the presence of filled pauses is also significant. To compare the impact of word count on the two levels of metaphor presence, stratified analyses were performed (see Table 3). In both cases, word count has a significant effect on self-repair, with a more pronounced impact observed on utterances without metaphors. As word count increases by one unit, the probability of filled pauses increases more in utterances without metaphors (by approximately 0.075 units) than in utterances with metaphors (by about 0.048 units).

The predicted probabilities of filled pauses in the two levels of metaphor presence are plotted in Figure 1.

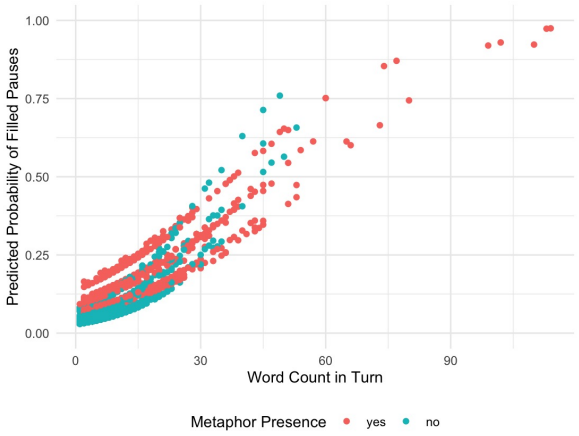

Figure 1: Predicted probabilities of filled pauses

When word count is below 30, utterances with metaphors are more likely to contain filled pauses compared to those without any metaphors. When word count is between 30 and 50, the probability of filled pauses in utterances without metaphors surpasses that in utterances with metaphors. When the word count goes beyond 50, the probability in utterances with metaphors continues to increased at a lower rate; however, no utterances without metaphors with comparable lengths were found in this range.

Presence of self-repair

The modelling results on the presence of self-repair are summarized in Table 4. Both metaphor presence and word count have significant fixed effects. When word count is held constant at 9.30, utterances with metaphors are significantly more likely to contain self-repair compared to those without metaphors. According to the co-variate model, the predicted probabilities of self-repair are 11% ($95\%CI : 9\% - 15\%$) and 8% ($95\%CI, 6\% - 10\%$).

The interaction effect between word count and metaphor presence on the presence of self-repair is also significant. Results of stratified analyses are summarized in Table 5. We can see that word count has a significant effect on self-repair in both cases, and the impact is more pronounced on utterances without metaphors. As word count increases by one unit, the probability of self-repair in utterances without metaphors increases more (by approximately 0.133 units) than in utterances with metaphors (by about 0.075 units).

The predicted probabilities of self-repair in the two levels of metaphor presence are plotted in Figure 2.

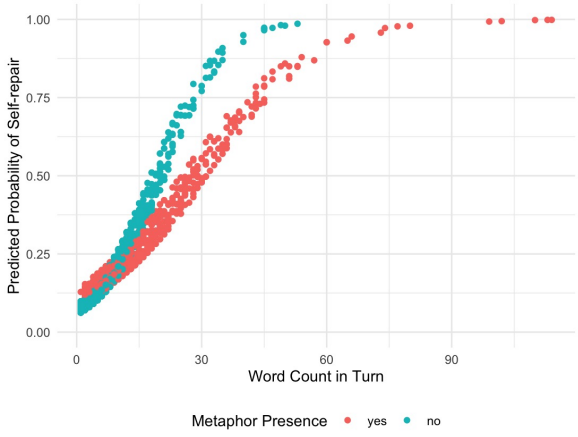

Figure 2: Predicted probabilities of self-repair

When the word count is below 10, utterances with metaphors generally have a higher probability of containing self-repairs than utterances without metaphors. When the word count exceeds 10, utterances without metaphors have a higher probability of containing self-repair.

Discussion

The significant main effect of metaphor presence on the probabilities of filled pauses and self-repair highlights its contributions to the occurrence of disfluencies. When word count is held constant, utterances with metaphors are associated with a heightened likelihood of filled pauses, which manifests as increased hesitation and interruptions in the speech flow. Utterances with metaphors also

Fixed Effects	Estimate	SE	z-value	p-value
(Intercept)	-2.391	0.205	-11.639	$< .001$
metaphor presence	-0.791	0.218	-3.625	0.01
word count	0.047	0.007	6.356	$< .001$
metaphor presence * word count	0.029	0.012	2.318	$< .001$

Table 2: Fixed effects of the interaction model on filled pauses

Metaphor presence	Fixed Effects	Estimate	SE	z-value	p-value
N	(Intercept)	-3.164	0.163	-19.362	$< .001$
	word count	0.075	0.010	7.344	$< .001$
Y	(Intercept)	-2.387	0.209	-11.40	$< .001$
	word count	0.049	0.008	6.44	$< .001$

Table 3: Stratified analysis of filled pauses in metaphor presence

have higher rates of self-repair, which plays out in the form of repetitions, substitutions, and deletions.

The occurrence of disfluency markers can provide insights into cognitive processing involved in metaphor production. Our results support the view that utterances with metaphors, compared to those without metaphors, may pose heightened cognitive demands on the speaker's end (Steen, 2023). As mentioned earlier, some previous studies take disfluencies as indicative of cognitive burdens or communication problems (Levelt, 1983; Colman and Healey, 2011), and some see disfluencies as a communicative solution to manage the cognitive pressure (Brennan and Schober, 2001; Bailey and Ferreira, 2007; Howes et al., 2017). Based on our results, it is plausible the increased cognitive demands associated with metaphor use requires more cognitive resources, potentially resulting in higher disfluencies rates.

Disfluencies may also be related to the speaker's consciousness over language use in conversation. Cienki (2020) proposed the concept of meta-communicative awareness (MCA) to account for the speaker's degree of awareness of the form and/or content of their language use. Disfluency markers are recognised as key signals of MCA. When the signals are present, compared to cases with less effortful or no signals, the speaker is more likely to be aware of their ways of self-expression (Cienki, 2020). Following this line of thought, utterances with metaphors, given the increased disfluency rates, may be produced with higher degrees of MCA compared to those without metaphors. The heightened occurrence of disfluency markers, as exemplified above, may reflect the speaker's active engagement in shaping and refining their linguistic choices to effectively convey the complicated ideas.

Another interesting observation is that the disfluency markers are not necessarily attached to the metaphorical parts of the utterance. Rather, they may occur before or after, and sometimes quite far away from the metaphorical parts; example (1) presented earlier is illustrative. This suggests that the cognitive pressure may have arisen before the utterance is articulated, and may influenced the entire production process.

The findings offer clues regarding the interactive relationship between metaphor presence and word count in terms of the occurrence of disfluencies. Interestingly, the patterns differ across the two types of disfluencies. From Figure 1 we see that utterances containing metaphors have higher rates of filled pauses than those without metaphors. However, utterances without metaphors increase more sharply in filled pause rates than those with metaphors, especially when word count goes above 30. The presence of self-repair, as shown by Figure 2, exhibits a different pattern. Despite the fact that utterances without metaphors have higher

Proceedings of the 28th Workshop on the Semantics and Pragmatics of Dialogue, September, 11–12, 2024, Trento, Italy.

Fixed Effects	Estimate	SE	z-value	p-value
(Intercept)	-2.095	0.178	-11.787	$< .001$
metaphor presence	-0.512	0.200	-2.563	0.01
word count	0.073	0.008	8.788	$< .001$
metaphor presence *word count	0.059	0.013	4.527	$< .001$

Table 4: Fixed effects of the interaction model on self-repair

Metaphor presence	Fixed Effects	Estimate	SE	z-value	p-value
N	(Intercept)	-2.614	0.126	-20.72	$< .001$
	word count	0.133	0.010	12.88	$< .001$
Y	(Intercept)	-2.102	0.188	-11.198	$< .001$
	word count	0.075	0.009	8.61	$< .001$

Table 5: Stratified analyses on self-repair in the two levels of metaphor presence

self-repair rates when word count is held constant, the trend shifts when considering interactions between metaphor presence and word count. Notably, when word count is above 10 words, those without metaphors have generally higher rates of containing self-repair than utterances with metaphors. However, the majority of utterances without metaphors are shorter than 10 words ($1,579$ out of $1,942$) and have lower self-repair rates, which explains the main effect discussed earlier. This finding suggests that different linguistic variables may interact in shaping conversation behaviors, underscoring the need for disfluency research to consider the impact of word count, especially its interaction with other linguistic variables.

Consistent with previous findings (e.g., Oviatt, 1995; Bortfeld et al., 2001), the positive association between word count and the probabilities of disfluencies confirms that the cognitive effort involved in articulating longer utterances is higher, regardless of the presence of metaphors. More interestingly, our results also show that the impact of word count on disfluency rates is less prominent on longer utterances with metaphors. It is possible that the production of longer utterances with metaphors requires more deliberate planning and articulation, which leads to relatively lower disfluency rates. This might also be explained by the presence of compensatory cognitive strategies. A recent study (Qiu et al., 2024, in progress) showed that speakers may employ more compensatory cognitive strategies, such as gestures, in turns with metaphors than those without. These strategies were found to help to alleviate the speakers' cognitive pressure (Kita, 2000) and sustain mutual understanding (Healey et al., 2015). While these strategies may happen at a higher chance in longer utterances, it is possible that they mitigate the impact of cognitive difficulties, resulting in lower disfluency rates.

5 Conclusion

This study compared the probabilities of disfluencies in naturally produced conversational utterances with and without metaphors, taking the impact of word count into account. The findings offer insights into the conversational dynamics in metaphor use and the cognitive mechanisms underlying disfluencies. A strength of the study is that it captures how people talk in everyday life, which would be hard to replicate and control for in psycholinguistic experiments. We also have supportive evidence that the production of turns containing metaphors may pose greater cognitive challenges than those without metaphors.

Several key limitations need to be acknowledged. Firstly, even though utterances examined by this study are thematically consistent, it was not possible to control utterances in spontaneous conversations in terms of semantic content and lengths.

Proceedings of the 28th Workshop on the Semantics and Pragmatics of Dialogue, September, 11–12, 2024, Trento, Italy.

Future studies could consider using experimental designs to compare utterances with and without metaphors on the same topic and of similar word count.

Secondly, this paper focused exclusively on differences in the probabilities of disfluency markers. The placement of disfluency markers, especially in utterances with metaphors, remains to be explored by future research. Furthermore, we do not distinguish between different types of self-repair, for example, whether repetitions or reformulations are more associated with turns containing metaphors. Additionally, more fine-grained analysis distinguishing between, for example "forward-looking" and "backwards-looking" disfluencies (Ginzburg et al., 2014), remains for future work.

Thirdly, in this study, metaphor presence was annotated as a binary variable. In fact, there are some more fine-grained aspects of metaphors that may cause the utterance to be processed with different levels of ease, for example, the number of metaphorical lexical units, the degree of novelty/conventionality (Giora, 2002), and deliberateness of metaphor use (Reijnierse et al., 2018). Future research could explore how these features interact with disfluencies and other aspects of language use. This can be investigated either in spontaneous conversation, or with more controlled psycholinguistic methods like the tangram experiments in Clark and Wilkes-Gibbs (1986).

Despite these limitations, our results show that both word count and metaphor presence are significant factors contributing to the presence of disfluencies. Utterances with metaphors are generally more likely to contain filled pauses and self-repairs compared to those without metaphors. This may stem from heightened cognitive or communicative challenges associated with metaphor use, or potentially reflect the speaker's increased awareness of language use in the conversation (Cienki, 2020). Interestingly, the impact of word count on disfluencies varies between utterances with and without metaphors and across different disfluency markers, highlighting the combined influence of metaphor use and longer utterances on speech disfluencies.

Acknowledgements

This research was supported by ERC Starting Grant DivCon: Divergence and convergence in dialogue: The dynamic management of mismatches (101077927). Howes was further supported by the Swedish Research Council grant (VR project 2014-39) for the establishment of the Centre for Linguistic Theory and Studies in Probability (CLASP) at the University of Gothenburg. We are also thankful for the comments from the anonymous reviewers and the advice on statistics from Asad Sayeed.

References

Karl GD Bailey and Fernanda Ferreira. 2007. The processing of filled pause disfluencies in the visual world. In *Eye movements*, pages 487–502. Elsevier.

Mathias Benedek, Roger Beaty, Emanuel Jauk, Karl Koschutnig, Andreas Fink, Paul J Silvia, Beate Dunst, and Aljoscha C Neubauer. 2014. Creating metaphors: The neural basis of figurative language production. *NeuroImage*, 90:99–106.

Heather Bortfeld, Silvia D Leon, Jonathan E Bloom, Michael F Schober, and Susan E Brennan. 2001. Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and speech*, 44(2):123–147.

Susan E Brennan and Michael F Schober. 2001. How listeners compensate for disfluencies in spontaneous speech. *Journal of memory and language*, 44(2):274–296.

Alan Cienki. 2020. A multimodal perspective on mca: Cues of (possible) metacommunicative awareness. In *Drawing attention to metaphor: Case studies across time periods, cultures and modalities*, pages 63–92. John Benjamins Publishing Company.

Herbert H Clark. 1996. *Using language*. Cambridge university press.

Herbert H Clark and Jean E Fox Tree. 2002. Using uh and um in spontaneous speaking. *Cognition*, 84(1):73–111.

Herbert H Clark and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition*, 22(1):1–39.

M. Colman and P. G. T. Healey. 2011. The distribution of repair in dialogue. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, pages 1563–1568, Boston, MA.

Raymond W Gibbs. 1994. *The poetics of mind: Figurative thought, language, and understanding*. Cambridge University Press.

Jonathan Ginzburg, Raquel M Fernández, and Schlangen David. 2014. Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics*, 7(9):64.

Rachel Giora. 2002. Literal vs. figurative language: Different or equal? *Journal of pragmatics*, 34(4):487–506.

John J Godfrey, Edward C Holliman, and Jane McDaniel. 1992. Switchboard: Telephone speech corpus for research and development. 1:517–520.

Charles Goodwin. 1979. The interactive construction of a sentence in natural conversation. *Everyday language: Studies in ethnomethodology*, 97:101–121.

Patrick Healey, Nicola Plant, and Mary Howes, Christine amd Lavelle. 2015. When words fail: Collaborative gestures during clarification dialogues. In *Turn-Taking and Coordination in Human-Machine Interaction: Papers from the 2015 AAAI Spring Symposium*, Austin, Texas, USA.

Julian Hough and Matthew Purver. 2014. Strongly incremental repair detection. *arXiv preprint arXiv:1408.6788*.

Christine Howes, Julian Hough, Matthew Purver, and Rose McCabe. 2014. Helping, I mean assessing psychiatric communication: An application of incremental self-repair detection.

Christine Howes and Mary Lavelle. 2023. Quirky conversations: How people with a diagnosis of schizophrenia do dialogue differently. *Philosophical Transactions of the Royal Society B*. In press.

Christine Howes, Mary Lavelle, Patrick G. T. Healey, Julian Hough, and Rose McCabe. 2017. Disfluencies in dialogues with patients with schizophrenia. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, London, UK.

Anna Albertha Kaal. 2012. *Metaphor in conversation*. Phd-thesis - research and graduation internal, Vrije Universiteit Amsterdam.

Sotaro Kita. 2000. How representational gestures help speaking. *Language and gesture*, 1:162–185.

George Lakoff and Mark Johnson. 1980. *Metaphors We Live By*. University of Chicago Press.

Mary Lavelle, Patrick G. T. Healey, and Rose McCabe. 2012. Is nonverbal communication disrupted in interactions involving patients with schizophrenia? *Schizophrenia Bulletin*.

Willem JM Levelt. 1983. Monitoring and self-repair in speech. *Cognition*, 14(1):41–104.

Robin J Lickley. 2001. Dialogue moves and disfluency rates. In *ISCA tutorial and research workshop (ITRW) on disfluency in spontaneous speech*.

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

Amy Han Qiu, Christine Howes, and Vladislav Maraev. 2024. Metaphoricity, deliberateness and conversation behaviours: A case study of behavioral patterns in the use of linguistic metaphors. Work in Progress.

W Gudrun Reijnierse, Christian Burgers, Tina Krennmayr, and Gerard J Steen. 2018. Dmip: A method for identifying potentially deliberate metaphor in language use. *Corpus Pragmatics*, 2:129–147.

Elena Semino. 2008. *Metaphor in discourse*. Cambridge University Press Cambridge.

Elizabeth Shriberg. 1996. Disfluencies in switchboard. In *Proceedings of international conference on spoken language processing*, volume 96, pages 11–14. Citeseer.

Gerard J Steen. 2023. Thinking by metaphor, fast and slow: Deliberate metaphor theory offers a new model for metaphor and its comprehension. *Frontiers in Psychology*, 14:1242888.

Gerard J Steen, Aletta G Dorst, J Berenike Herrmann, Anna A Kaal, Tina Krennmayr, Tryntje Pasma, et al. 2010. *A method for linguistic metaphor identification*. John Benjamins Publishing Company Amsterdam.