Metaphorically speaking

Conversational behavioural patterns in the use of linguistic metaphors

Amy Han Qiu, Christine Howes, and Vladislav Maraev
University of Gothenburg

While abundant studies have examined cognitive efforts required for metaphor comprehension, what happens during metaphor production remains underexplored. Based on 19 triadic conversations produced by 57 participants, this paper examines conversational behaviours associated with the use of metaphors and deliberate metaphors in particular. The data includes 2,631 conversational turns, of which 690 contained metaphors and 45 had deliberate metaphors. Four conversational behaviours were examined: turn duration, within-turn pause duration, between-turn gap duration, and co-speech gestures. Compared with turns without metaphors, those with metaphors lasted significantly longer and were more likely to co-occur with gestures, whereas differences in within-turn pause and between-turn gaps were not significant. The results suggest that metaphor production involves more cognitive efforts or a stronger awareness of the ongoing communication; however, metaphor processing by the listener does not necessarily take more time. The effect of metaphor deliberateness was not significant in any of the conversational behaviours.

Keywords: metaphor production, deliberate metaphors, conversational behaviours, cognitive processing, disfluencies

1. Introduction

Metaphor is the linguistic act of talking, and potentially thinking, about something in terms of something else (Semino, 2008). An example is "The researcher is quite far away from a breakthrough" (metaphorically used words italicised, the same hereafter), in which "far away from" is used metaphorically, exhibiting a clear meaning transfer from their 'basic' meaning of physical distance (i.e., the source domain) to a contextual meaning describing the researcher's chances of

https://doi.org/10.1075/rcl.00256.han | Published online: 24 April 2026
Review of Cognitive Linguistics ISSN 1877-9751 | E-ISSN 1877-976X

achieving a breakthrough (i.e., the target domain) (Lakoff & Johnson, 1980). In this example, the metaphor is most likely used in a non-deliberate way, which means the speaker talks about the target domain without paying attention to the source domain of the metaphorical utterance (Steen, 2011). Another example is "The researcher is twenty miles away from a breakthrough". Here, the effect of the utterance is for the listener to shift their attention temporarily from the target domain only to both the target and the source domain evoked by the metaphorical expression (Steen, 2015). Physical distance is presented as a source domain referent for inferring the chances of achieving academic success. Expressions like this, where the speaker intends to provide a new perspective on the topic, are referred to as deliberate metaphors (Steen, 2011). In contrast, a literal equivalent of the two metaphors, "The researcher is unlikely to make a breakthrough", expresses the same idea without establishing meaningful links between a concrete and an abstract domain.

How metaphors function in real-world contexts has long been considered an interesting question to cognitive linguists. Extensive research has examined the semiotic features of metaphor use (e.g., Kaal, 2012; Tay, 2016; Reijnierse et al., 2020; Ritchie, 2013) and the comprehension process (e.g., Brisard et al., 2001; Columbus et al., 2015; Olkoniemi et al., 2016). However, speakers' conversational behaviours in metaphor production, which may be non-linguistic and multimodal, remain unexplored. Do speakers behave differently when producing metaphors compared to literal utterances? What about when using deliberate compared to non-deliberate metaphors? What do the patterns say about the underlying cognitive mechanisms and/or the effects of being used in an ongoing interaction? These questions remain to be answered through an integration of linguistic and behavioural data.

In this paper, we focus on four conversational behaviours: turn duration, within-turn pause duration, between-turn gap duration, and co-speech gestures, comparing these measures in turns containing metaphors to those that do not. We first provide a review of existing literature, highlighting the need to adopt a multimodal approach in studying metaphor use in conversations, and then introduce the data and statistical methods. After that, we will present the descriptive statistics and statistical results, and discuss the implications, limitations, and future directions.

2. Literature review

2.1 Metaphors in conversation: Findings from the semiotic approach

Given the dialogic nature of conversation, the use of metaphors is seen as arising from the interaction among the linguistic, cognitive, affective, interactional, and sociocultural dimensions of the ongoing conversation (Kaal, 2012; Cameron, 2003; Cameron & Maslen, 2010; Müller & Tag, 2010).

A large body of research follows a semiotic approach, exploring the linguistic features and functions of metaphors based on textual analyses. According to Reijnierse (2017), about 7.3% of all lexical units in conversation data are metaphor-related. Metaphors are often used to clarify abstract ideas, shape perspectives, influence opinions, create common ground, and convey judgments, performing explanatory, persuasive, affective, and evaluative functions (e.g., Semino, 2008; Ritchie, 2013). According to the Conceptual Metaphor Theory (Lakoff & Johnson, 1980), metaphors create systematic mappings between abstract concepts and more concrete, bodily-related, and tangible experiences. Such mappings are expected to trigger embodied simulation Gallese and Lakoff (2005), which can be broadly defined as the co-activation of neural activities and bodily features (Cuccio et al., 2019). This process can facilitate the transmission of abstract ideas as well as the speakers' values and beliefs (Gallese & Lakoff, 2005), making metaphors particularly effective in influencing the audience's thoughts and reasoning (Lakoff & Johnson, 1980; Gibbs, 1994, 2005; Cuccio et al., 2019).

While certain metaphors appear to be more appealing or impressive than others, some researchers attributed the difference to the degree of deliberateness involved in metaphor use (Steen, 2011, 2015). According to Reijnierse et al. (2018), a metaphor is potentially deliberate "if the source domain is part of the referential meaning of the utterance in which it is used" (p. 136). Deliberate metaphors, like the example discussed earlier, are assumed to guide attention and prompt the listener to engage in cross-domain mapping. They are conceptually different from novel metaphors, which concerns the linguistic property of metaphors rather than communicative intent (Steen, 2023).[1] Corpus-based research shows that 1.3%

1. Importantly, deliberateness is not equated with novelty: Both novel and conventional metaphors can be used deliberately or non-deliberately depending on the communicative intent and context (Steen, 2011, 2023). Nevertheless, deliberate metaphors are indeed significantly more likely to involve the creation of novel expressions while non-deliberate metaphors are more likely to be conventional (Tay, 2016). Also, deliberateness is not necessarily the same as consciousness in metaphor use. According to Steen (2011), deliberate metaphors can be seen as metaphors in attention, which affords the possibility of conscious metaphorical cognition, but conscious cognition does not have to occur for a metaphor to be considered deliberate.

of metaphors in conversations are potentially used in a deliberate manner (Reijnierse, 2017). Among all types of metaphors, deliberate metaphors, are believed to best activate and shape the mental representation of the abstract idea due to the explicit emphasis they put on bodily-related information (Cuccio et al., 2019).

While the semiotic approach provides valuable insights into the linguistic characteristics of metaphors, it typically reduces conversation data to written text, focusing exclusively on verbal content while overlooking multimodal information in conversational behaviours. This tendency is noted by Linell (1982) as "the written-language bias", an oversight that disregards important multimodal information accompanying spoken language, such as disfluencies, pauses, and gestures. As a result, crucial conversational cues that co-occur with language use are often neglected, limiting our understanding of how metaphors, amongst other things, are used in spontaneous interactions. Another key limitation of the semiotic approach is its limited ability to capture cognitive activities associated with metaphor processing. While some studies argue that metaphors, particularly deliberate ones, may be more cognitively and affectively engaging than literal expressions, these conclusions are often based on inference rather than empirical evidence.

2.2 Experimental approaches to metaphor processing

Beyond identifying the communicative and affective functions of metaphor, recent theoretical work has sought to specify how different levels of cognitive effort arise during comprehension. Studies based on psycholinguistic experiments yielded mixed results, highlighting the nuanced demands of metaphor processing. Some suggest heightened cognitive demands for metaphors. For example, experiments on reaction time showed that understanding written metaphors takes longer than understanding non-metaphorical expressions with comparable meanings (Gibbs, 1990; Onishi & Murphy, 1993). Similarly, eye-tracking studies show that metaphorical sentences often result in longer reading times (Columbus et al., 2015; Olkoniemi et al., 2016). In a self-paced reading task, novel metaphors specifically require longer reaction times than literal expressions (Brisard et al., 2001). A more recent eye movement study further shows that deliberate metaphors are read significantly slower than non-deliberate metaphors, and both types of metaphors are read significantly slower than non-metaphorical words (Vries et al., 2018). Deliberate metaphors were also found to draw more attention to the speaker's intention (Gibbs, 2015; Thibodeau, 2017) or add to perceived effectiveness compared to non-deliberate cases (Reijnierse, 2017).

Despite this weight of evidence of increased processing demands of metaphors compared to non-metaphors, there are also studies that revealed only small or no difference between the processing of metaphorical and literal expressions. For example, Glucksberg (2008) found that metaphorical and non-metaphorical meanings can be processed equally quickly and in parallel, with neither having unconditional priority. Nevertheless, in cases where both are present, metaphorical expressions may be preferred to non-metaphorical ones. Similarly, Ortony et al. (1978) found that metaphor reading times exceed literal ones only when contextual information is lacking; when sufficient contextual information is provided, reading times of metaphorical and literal sentences did not differ significantly.

Apart from processing time, there are also studies that examined cognitive engagement in metaphor processing based on other factors, such as accuracy of judgement, perceived vividness, emotional engagement, and persuasiveness. A study on the speed-accuracy trade-off, for example, showed that although metaphors and their literal counterparts were processed at the same speed, the accuracy rate of metaphors can be lower than that of literal expressions (Bambini et al., 2021). This suggests that the derivation of metaphor is more complex than that of literal meanings. Some studies have also shown that metaphors across all types have a larger N400 effect, compared to literal expressions (see Lai et al., 2009 for a review), indicating greater processing demands or difficulty in semantic integration (Coulson, 2008). In addition, metaphorical expressions, both novel and conventional, are widely found to be more vivid or emotionally evocative than their literal counterparts (e.g., Bohrn et al., 2012; Reijnierse, 2017; Tay, 2024; Citron & Goldberg, 2014; Mon et al., 2021). Studies such as Thibodeau and Boroditsky (2011) also provide evidence that metaphors can influence the ways we reason about and act regarding complex abstract issues. These studies provide evidence that metaphors have greater potential in triggering Embodied Simulation and influencing audiences' opinions.

Overall, experimental evidence reveals that metaphor comprehension engages variable levels of cognitive effort depending on factors such as familiarity, context, and communicative intent. Drawing on Kahneman's (2011) dual-process model, Steen (2023) proposes that metaphor comprehension may shift between fast, automatic and slow, reflective processing, depending on the communicative context. Familiar metaphors in neutral contexts may be understood effortlessly, whereas novel or deliberate metaphors that extend across utterances require more conscious, reflective engagement. This dual-process account helps explain the mixed empirical results by linking variability in processing effort to contextual and communicative factors that influence metaphor processing.

While existing studies have provided valuable insights into the cognitive processing of metaphors, most have examined it only as an outcome of comprehension, leaving the process involved in metaphor production unexplored. Moreover, similar to studies that follow the semiotic approach, existing experimental studies also exhibit an apparent written-language bias. As noted by Kaal (2012), there are far more experimental studies on written metaphor comprehension than on spoken scenarios. As summarised above, these studies primarily use written materials displayed on screens, with metaphor comprehension measured through indicators such as mouse-clicking and eye movements. Even studies performed on spoken materials tend to present isolated audio to the experimental subjects and are thus also not interactive in nature. While precise measurements of cognitive indicators offer a deeper understanding of the processing mechanisms underlying metaphor comprehension, those indicators do not capture the dynamic, multimodal nature of spontaneous conversation.

2.3 Conversational behaviours

The study of multimodal behaviours in conversations provides an opportunity to gather indirect evidence about cognitive processing in metaphor production. Conversational features like repetition, prosody, and gestures, which are often neglected in previous research, may capture speakers' awareness of the form and/or content of their acts in the ongoing communication; they are thus referred to by Cienki (2020) as cues of "metacommunicative awareness (MCA)". According to Cienki (2020), these cues do not necessarily indicate that the speaker intended to speak or behave in a certain way; rather, they may be signals that the speaker is more or less aware of how they were expressing themselves. If the cues are presented in an effortful way, for example, deliberate use of gestures and elevated tone, they may play a role in the listeners' comprehension (Cienki, 2020). Apart from signalling the speakers' communicative intent, some cues, such as long utterances and disfluencies, may also indicate difficulties in formulating or expressing the idea or challenges in managing interpersonal dynamics (Bard et al., 2001). These cues can be a signal for listeners even if they were produced non-intentionally, influencing the conversation in implicit ways. In what follows, we introduce some specific conversational behaviours that may provide complementary insights into previous findings on metaphor use.

2.3.1 Duration of conversational turns

A key indicator of speakers' metacommunicative awareness and cognitive efforts in metaphor production is the duration of conversational turns. Longer turns, which involve more words or slower speech, may indicate slower information pro-

cessing, less coherent speech, and relatively higher cognitive load (Berthold & Jameson, 1999; Müller et al., 2001; utterances with a large number of words are unlikely to be fast, see Goldman-Eisler, 1954). In contrast, shorter turns, characterised by fewer words or faster speech, typically reflect quicker information processing, more coherent speech, and a lower cognitive load. From the perspective of communication, longer durations may arise when the speaker intends to emphasise a particular point or articulate the message with greater precision and clarity, or when the speaker experiences pressure when talking, whereas shorter turn duration may reflect a more spontaneous and natural flow of speech (Müller et al., 2001).

Turn duration and disfluencies

A behavioural indicator that can lead to increased turn duration is the presence of verbal disfluency markers such as self-repair, filled pauses (e.g. er, um), and repetition (Lickley, 2015). These markers may occur as a result of cognitive difficulties (Levelt, 1983; Bortfeld et al., 2001; Clark & Tree, 2002), attention to the ongoing communication (Cienki, 2020), and interactive issues Goodwin (1979). Previous research shows that longer and more cognitively demanding utterances typically involve more disfluencies (Lickley, 2001; Shriberg, 1996).

Qiu et al. (2024) found that turns with metaphors were significantly more likely to contain self-repair and filled pauses compared to turns without any metaphors. This suggests that the cognitive processes involved in producing metaphors may introduce additional planning or retrieval demands, or an increased awareness of the extra effort that the listener may need in comprehending the utterance, leading to more frequent disfluencies. The increased occurrence of self-repair and filled pauses in metaphor-rich turns indicates that speakers may need extra time to formulate and adjust their metaphorical expressions during the conversation (Qiu et al., 2024).

2.3.2 Within-turn pauses

Another conversation behaviour relevant to production difficulty, either cognitive or communicative, is the time lag between two conversational turns. In cases where the same speaker continues speaking after a short silence (i.e., they hold the turn), this time lag is referred to as the duration of within-turn pauses. Shorter within-turn pauses suggest more fluent or less deliberate speech, while longer pauses reflect less fluent speech or a stronger intent to emphasise the idea (Lickley, 2015; Cienki, 2020).

2.3.3 Between-turn gaps

When a different speaker takes the floor after a pause, the time lag is referred to as the duration of between-turn gaps. Shorter between-turn gaps indicate that the preceding turn may not be very much cognitively challenging to the listener (Holler et al., 2018; Roberts et al., 2015; Berthold & Jameson, 1999; Müller et al., 2001) or do not pose a challenge in formulating appropriate responses, whereas longer gaps may suggest greater difficulty in comprehension or that the turn is more difficult to formulate a response to.

2.3.4 Co-speech gestures

Gestures are widely acknowledged to play a crucial role in communication and thought. According to McNeill (1992), speech and gesture arise from the same "growth points" of thoughts, which are then "unpacked" into linguistic content and gestures during speech production. Some gestures convey visual, spatial, and kinesthetic information in iconic, metaphoric, and deictic ways (McNeill, 1992), and some are merely used to stress or emphasise certain parts of the utterance but do not carry substantive meanings (Clough & Duff, 2020).

According to the information packaging hypothesis proposed by Kita (2000), gestures help speakers organise and package visuo-spatial information into a linear, sequential format that can be readily verbalised. Speakers gesture more in cognitively challenging or linguistically complex tasks (Kita & Davies, 2009; Morsella & Krauss, 2004; Hostetter, 2011), when they perceive a lack of shared knowledge with their listeners (Campisi & Özyürek, 2013; Galati & Brennan, 2014; Hoetjes et al., 2015; Hilliard & Cook, 2016), and, in dialogue, when they are aware that their partner is having trouble understanding, for example in clarification sequences (Healey et al., 2015).

These conversational behaviours are well-established indicators of cognitive processing and communication dynamics; however, they have been underexplored in the context of spontaneous metaphor use. How these behavioural cues vary across turns with and without metaphors, and turns with deliberate and non-deliberate metaphors, remains an open question. Furthermore, it is unclear whether some behavioural cues, like turn duration, might be further influenced by other cues, such as within-turn pauses and verbal disfluency markers. The present study addresses this research gap by identifying conversation behaviours associated with the use of metaphors and deliberate metaphors in particular.

3. Data and methods

Our data consists of 19 face-to-face triadic conversations between 57 unfamiliar participants (26 females and 31 males). Their mean age was 31 ($SD = 10.09$, Min = 18, Max = 54), and the average age of completing training or education was 21 ($SD = 5.77$, Min = 13, Max = 43).

As shown in Figure 1, participants were seated close to one another in a triangular formation. They were video- and audio- recorded and motion-captured using markers positioned on their upper body, arms, and head. The dialogues were collected previously as the control condition in an experiment comparing these dialogues to others involving a patient with schizophrenia. The data collection methods and associated details are reported in Lavelle et al. (2012) and Howes and Lavelle (2023).

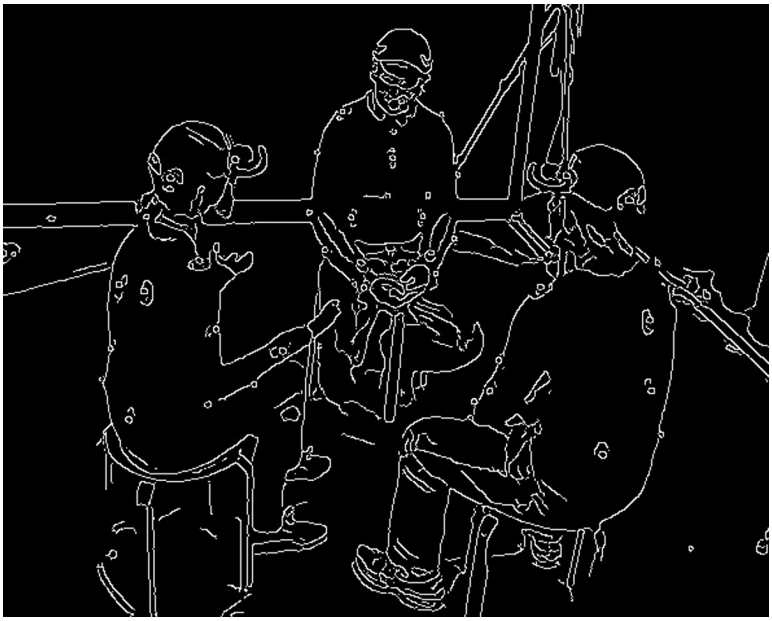


Figure 1. Participants in a triadic interaction

3.1 The balloon task

The context of the conversations is the balloon task. Participants were verbally presented with an ethical dilemma and informed that their task was to collaborate with their group to resolve it. The verbal instructions are presented in Table 1.

The balloon task was chosen as the conversational context primarily to control for topic variation: participants were instructed to go over all four passengers' cases, ensuring comparability across conversations. Moreover, moral dilemmas naturally engage higher-order cognitive and evaluative processes, such as weighing values, assessing risks, and negotiating perspectives, which are often expressed metaphorically, for example, through metaphorizing life, decision-making, or moral responsibility. Therefore, ethical dilemmas provide a structured and cognitively and socially rich context where metaphor use is both likely and meaningful.

Table 1. Verbal instructions for participants of the balloon task

Four people are in a hot air balloon. The balloon is losing height and about to crash into the mountains. Having thrown everything imaginable out of the balloon, including food, sandbags and parachutes, their only hope is for one of them to jump to their certain death to give the balloon the extra height to clear the mountains and save the other three. The four people are:

Dr Nick Rivers, a cancer research scientist who believes he is on the brink of discovering a cure for most common types of cancer. He is a good friend of Tom and Susie Derkins.

Mrs Susie Derkins a primary school teacher. She is over the moon because she is 7 months pregnant with her second child.

Mr Tom Derkins the balloon pilot. He is the husband of Susie, who he loves very much. He is also the only one with any balloon flying experience.

Carla Jenkins — a nine-year old musical prodigy, considered by many to be a "twenty-first century Mozart".

Your task is to debate the reasons for and against each individual being saved, and reach a mutual agreement on which of the four individuals should jump from the balloon.

3.2 Behavioural data

Time stamps for turn durations, within-turn pauses, and between-turn gaps were extracted automatically from the recordings using ELAN. Gestures during speech were recorded in terms of hand movement, regardless of semantic informativeness (see Howes & Lavelle, 2023). In this study, gesture presence was annotated as a binary variable at the level of turn (present/absent). Table 2 shows an example transcript together with the four conversation behaviours annotated at the level of turns. Turn durations, pauses, and gaps are in milliseconds. Descriptive statistics are summarised in Table 3.

Table 2. Example extract of a conversation

Turn No.	Spk	Duration (ms)	Pause (ms)	Gap (ms)	Transcript	Metaphor presence	Gesture presence	Disfl.
14	3	7781	926	−240	So, really, we've got to decide who, who is most likely to be able to save the greatest amount of life, if you like, isn't it, do you think? Do you agree?	yes, deliberate	no	no
15	3	1459	0	−450	So who is most [likely to]	no	no	no
16	3	1350	0	230	land the balloon safely.	no	yes	no
17	1	610	0	−528	[you see cause I-]	no	yes	no
18	3	3331	0	1307	[Because] otherwise, they the inc-the probability of them all dying is greater.	yes, non-deliberate	yes	yes
19	3	800	0	5	If, if you.	no	no	yes
20	1	5330	498	96	Bu- bu- but are you s-saying that um uh we need to value the sort of the worth of each person?, As they um uh, you know,	yes, deliberate	yes	yes
21	3	355	0	−101	Well.	no	no	no

Table 3. Descriptive statistics of behavioural variables

Variable	Descriptive statistics
Word count per turn	$M = 9.30$, $SD = 10.1$, 95%CI [8.92, 9.69]
Duration of turn (ms)	$M = 2319$, $SD = 2630$, 95%CI [2219, 2420]
Duration of within-turn pause (ms)	$M = 284$, $SD = 751$, 95%CI [256, 313]
Duration of between-turn gaps (ms)	$M = -78.1$, $SD = 766$, 95%CI [−107, −48.8]
Gesture presence	Yes: $N = 961$, No: $N = 1670$

3.3 Linguistic data

The transcription is illustrated in Table 2. A turn is defined as a speaker's contribution to the interaction between two points of transfer of speakership; a turn begins when a speaker starts talking and ends when they stop or when the next speaker begins. Backchannels and disfluencies (e.g., "yeah", "umm", "okay"), self-repairs, repetitions, and laughter were included. Overlapping speech was marked, and co-speech gestures were annotated where relevant. Unclear or inaudible segments are marked as.

The 19 transcripts, with an average length of 6.44 minutes, consisted of 3,784 turns. 1,153 turns that included only laughter, backchannels, or unclear utterances were excluded to avoid inflating the number of utterances without metaphors. 2,631 turns were kept for the analysis. The average word count per turn was 9.30 ($SD = 10.1$, 95%CI[8.92 −9.69]).

Metaphor identification

Metaphors were identified following the framework Metaphor Identification Procedure VU (MIPVU) (Steen et al., 2010). The criterion is whether the lexical unit has a more basic meaning, i.e., a meaning that is "more concrete, body-related, more precise, or historically older", and whether its contextual meaning contrasts with the basic meaning but can be understood in comparison with it. If these conditions are met, the basic meaning is interpreted as the source domain, the contextual meaning as the target domain, and the metaphorical meaning is evoked through cross-domain mappings between the two (Lakoff & Johnson, 1980).

Following the MIPVU, lexical units were operationalised as single words and multi-word expressions, which were identified using the BNC List of Multiwords.[2] All lexical units were annotated based on the meanings provided by three dictionaries, i.e., the Longman Dictionary, the Oxford English Dictionary, and WordNet.[3] To enhance interrater reliability, the VUAMC (Steen et al., 2010),[4] the largest available corpus hand-annotated for metaphorical language use, was used to assist

2. Accessed from https://ucrel.lancs.ac.uk/bnc2/multiwd.htm (June 19, 2025).

3. Dictionary meanings provide a baseline for identifying metaphorical uses, allowing annotators to detect departures from conventional usage that may invite conceptual processing of the source domain. More specifically, this allows annotators to distinguish between purely literal senses, conventionalized metaphorical uses, which are typically listed in dictionary as metaphors, and more novel or deliberate metaphors, which are not captured by dictionaries or extended creatively based on conventional metaphors. This approach is used in established metaphor identification procedures such as MIP (Praglejaz Group, 2007), MIPVU (Steen et al., 2010), and DMIP (Reijnierse et al., 2018).

4. Accessed from http://www.vismet.org/metcor/documentation/home.html (June 19, 2025).

with decision-making. To account for the fact that metaphors in real-world data can span across words (Cameron & Maslen, 2010) and to facilitate statistical analyses, for this study, metaphor use was annotated as a binary variable (present/absent).

Table 2 shows examples of turns with and without metaphors. Turns 14, 18 and 20 were marked as metaphorical because each contains at least one metaphorically used word or phrase that maps abstract moral or ethical concepts onto concrete, quantifiable domains. The most notable metaphors are "the *greatest amount* of life" and "*greater* probability of dying", which conceptualize human survival as a measurable substance or quantity. Likewise, Turn 20 describes moral worth in terms of economic or social *value*, framing ethical judgment through the metaphor of evaluation. In contrast, other turns in this extract talk about the possibility of a safe landing without establishing obvious metaphorical links; they were therefore labeled as not containing metaphors.

Deliberate metaphor identification

Deliberate metaphors were identified following the framework Deliberate Metaphor Identification Procedure (DMIP) (Reijnierse et al., 2018), also converted to a binary variable at the level of turns. Focusing on the communicative function of the metaphor, DMIP builds on MIPVU by examining whether the metaphor is intended to be noticed and conceptually processed as a metaphor by the addressee. According to DMIP, a metaphor is considered deliberate if the source domain is not merely a conventional part of the metaphorical expression but plays an active role in constructing the referential meaning of the utterance (Reijnierse et al., 2018). Cues of potentially deliberate metaphors include metaphor markers (e.g., *like, as if, as though*), comparisons or similes, novel and extended metaphors, and paralinguistic features like intonation and stress (Reijnierse et al., 2018).

Turns 14 and 20 in Table 2 were marked as deliberate metaphors because the speakers explicitly guided attention to the metaphorical construal of moral decision-making in quantitative terms. Turn 14 was classified as a deliberate metaphor because "*amount*" evokes a figurative mapping between life and a measurable quantity, and the use of the superlative "*greatest*" further highlights this metaphorical construal. This is further supported by the fact that the metaphor was taken up and extended by another speaker in Turn 20 as "*to value the sort of the worth of each person*", demonstrating successful highlighting and shared awareness of the metaphorical meaning. As noted by Cienki (2020), the presence of tuning devices such as "*sort of*" in Turn 20 signals reflective processing and awareness of the figurative expression; therefore, this turn was also classified as containing a deliberate metaphor. In contrast, Turn 18 was marked as non-deliberate, because the adjective "*greater*" has been conventionalised for

expressing quantitative comparison in terms of probability, and no other word or phrase in this turn further highlights its metaphorical origin.

Interrater reliability checks

Interrater reliability checks were conducted with a novice researcher. According to Bolognesi et al. (2017), a linguistic annotation scheme has a greater degree of replicability if the annotations provided by a trained rater are comparable to those provided by a novice rater with no prior experience in metaphor identification. After the pre-annotation training, the two annotators worked together on 10% of data, randomly selected, to calibrate their understanding of metaphor use in the data.[5] Project notes for dealing with frequently occurring words like prepositions and context-dependent expressions were created for reference in independent annotation.

The two annotators worked independently on 25% of the data. The annotation of turns with and without metaphors reached 97.1% agreement (Cohen's kappa = 0.88), and that of turns with deliberate and non-deliberate metaphors reached 97% agreement (Cohen's kappa = 0.75). After the interrater reliability checks, the two annotators resolved inconsistent cases in their annotations. The remainder of the dataset was annotated by the first author of the paper.

3.4 Research questions and methods

In this paper, we address the following questions:

RQ1 How does metaphor presence interact with turn duration, and what is the role played by disfluencies?
RQ2 How does metaphor presence interact with within-turn pause duration, between-turn gap duration, and co-speech gestures?
RQ3 How does metaphor deliberateness interact with each conversational behaviour?

A series of mixed effect models were run to investigate the effect of linguistic metaphors and metaphor deliberateness on the behavioural variables. Given the nature of variables and data distribution, different models were applied for the variables of interest. Generalised Linear Mixed Models (GLMMs) with a Gamma distribution and log link were chosen for the model on turn durations, and the presence of disfluency markers, a key factor influencing turn durations, was included as an interaction term. Zero-inflated models with a Gaussian distribu-

5. The training materials included Steen et al. (2010); Steen (2011, 2015, 2023); Reijnierse (2017); Reijnierse et al. (2018, 2020).

tion for the conditional part were run for within-turn pauses, given the large number of zeros and negative values in the data. Linear Mixed Models (LMMs) were used for the duration of between-turn gaps. The binomial family of GLMM with a logit link was chosen for the models on gesture presence. For all models, $p < .05$ was set as the threshold of statistical significance.

Due to multicollinearity when including word count as a fixed covariate, all models except the two on between-turn gaps were specified with word count as a random slope, allowing its effect to vary across groups. Conversation group was included as a random intercept. For the models on between-turn gaps, including word count as a random slope led to multicollinearity; therefore, word count was instead included as a fixed covariate. Including Participant as nested in Group caused multicollinearity for most models, indicating minimal within-group variability; therefore, only Group was included as a random effect. The models were run with R packages lmerTest and glmmTMB.

4. Results and discussion

4.1 Descriptive statistics

Among the 2631 turns, 1941 turns did not contain any metaphor-related words. 690 contained at least one metaphor-related words. Among these, 45 were identified as deliberate metaphors and 645 as non-deliberate metaphors, corresponding to 26.23% of all turns containing metaphor-related words and 6.5% of all metaphor-containing turns being potentially deliberate. Consistent with previous findings that such metaphors are inherently infrequent in spontaneous conversation (e.g., Reijnierse et al., 2018; Steen, 2023), deliberate metaphors were rare in our dataset, as speakers generally rely on conventionalised expressions to maintain communicative efficiency.

720 of the 2631 turns included disfluency markers, either self-repair or filled pauses, and 1911 were fluent utterances. 276 of the disfluent cases contained metaphors, including 23 deliberate and 253 non-deliberate, and 444 were without metaphors. The dataset had 1911 fluent utterances, among which 414 contained metaphors and 1497 were without.

Descriptive statistics of each utterance type and associated behavioural measures are summarised in Table 4, and those for fluent and disfluent utterances are shown in Table 5.

Table 4. Descriptive statistics for linguistic variables and associated behavioural measures

Linguistic variable	Turn duration	Within-turn pause duration	Between-turn gap duration	Gesture presence
Turns with metaphors (N=690)	M=4027, SD=4026, 95% CI [3726, 4327]	M=535, SD=988, 95%. CI [461, 609]	M=93.6, SD=784, 95% CI [−152, −35]	X=319 (46.23% out of 690)
deliberate metaphors (N=55)	M=6703, SD=8677, 95% CI [4096, 9310]	M=845. SD=1680, 95% CI [340, 1349]	M=−105, SD=676, 95% CI [−308, 98.2]	N=20 (36.36% of 55)
non-deliberate metaphors (N=645)	M=3840, SD=3415, 95% CI [3576, 4104]	M=513, SD=919, 95%. CI [442, 585]	M=−92.8, SD=791, 95% CI [−154, −31.6]	N=299 (46.35% of 645)
Turns without metaphors (N=l,941)	M=1712, SD=1488, 95% CI [1646, 1779]	M=195, SD=622, 95% CI [168, 223]	M=−72.6, SD=759, 95% CI [−106, −38.3]	N=642 (33.08% out of 1,941)

Table 5. Duration of fluent and disfluent utterances

Linguistic variable	N	Fluent turn duration, ms	N	Disfluent turn duration, ms
Turns with metaphors	414	M=2624, SD=2163, 95%. CI [2415, 2833]	276	M=6130, SD=5117, 95%, CI [5524, 6737]
deliberate metaphors	22	M=3503, SD=2903, 95%, CI [2216, 4719]	23	M=9763, SD=11068, 95%, CI [4977, 14549]
non-deliberate metaphors	392	M=2575, SD=2107. 95% CI [2365, 2784]	253	M=5800, SD=4071. 95% CI [5296, 6304]
Turns without metaphors	1497	M=1410. SD=1120, 95%, CI [1353, 1466]	444	M=2733, SD=2027, 95%, CI [2544, 2922]

4.2 Statistical modelling results

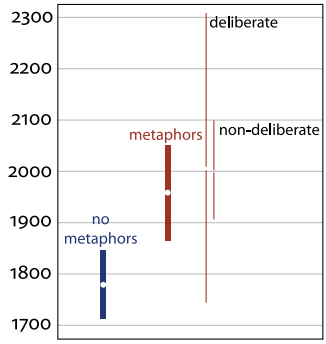
The effects of metaphors and metaphor deliberateness on the four behavioural variables are illustrated in Figures 2a to 2d.

4.2.1 Turn duration

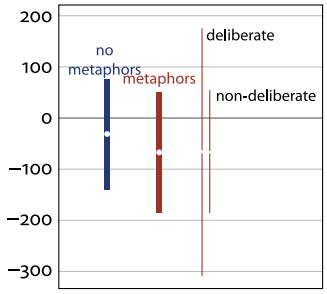
Descriptive statistics shown in Table 4 suggest that turns with metaphors generally lasted longer than those without metaphors after accounting for group-level variation in word count. Our statistical modelling results show that the main effect of metaphor presence on turn duration was significant (t=−5.57, p<.001),

with minimal variability in the effect of word count across groups. The predicted durations were 1956.21 (95%CI[1865.88, 2050.91]) for turns with metaphors and 1780.38 (95%CI[1715.44, 1847.77]) for those without.

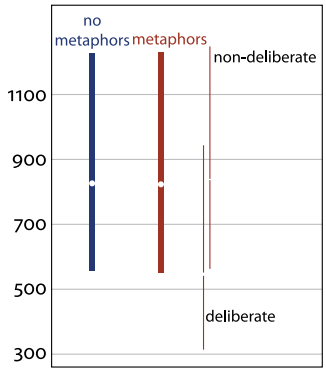
When the influence of word count, also minimal in this model, was accounted for at the group level, the effect of metaphor deliberateness was not significant. Predicted durations were similar for turns with deliberate metaphors ($M = 2005.54$, 95%CI[1744.92, 2305.07]) and those with non-deliberate metaphors ($M = 2001.58$, 95%CI[1907.78, 2099.78]); both were longer than turns without metaphors. The difference was only significant between turns with non-deliberate and no metaphors ($z = 3.80$, $p < .001$).



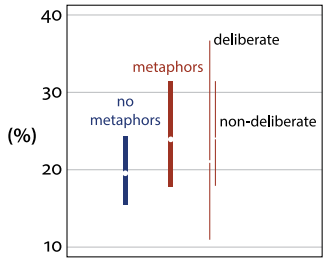
a. Turn duration (ms)



b. Between-turn pause duration (ms)



c. Within-turn pause duration (ms)



d. Gesture presence (ms)

Figure 2. Effects of metaphor presence and deliberateness on the four behavioural variables

Turn duration and disfluencies

Unsurprisingly, as can be seen in Table 4, disfluent utterances, which contain either filled pauses or self-repairs, lasted longer than fluent ones. As shown in Figure 3, disfluent turns had a significantly longer predicted mean duration ($M = 2166.23$, 95%CI [2025.87, 2316.31]) compared to fluent turns ($M = 1882.47$, 95%CI [1786.06, 1984.10]), after accounting for group-level variation in word count ($t = 4.00$, $p < .001$), which was minimal in this case. Among both fluent and disfluent utterances, turns with metaphors tended to be longer than those without. However, a significant interaction between metaphor presence and disfluency ($t = 3.98$, $p < .001$) showed that the difference was only significant in fluent speech ($z = 4.00$, $p < .001$), with turns with metaphors lasting longer ($M = 1882.47$, 95%CI [1786.06, 1984.10]) than those without ($M = 1637.68$, 95%CI [1574.65, 1703.24]). Possibly due to the limited sample size, the interaction model on metaphor deliberateness failed to converge.

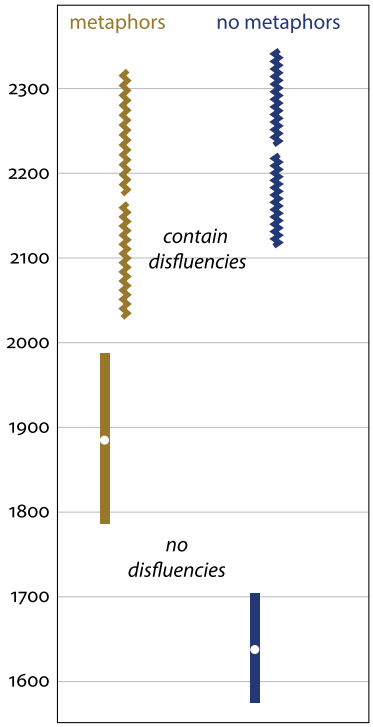


Figure 3. Metaphor × Disfluencies interaction on turn duration

4.2.2 Within-turn pause duration

Turns with metaphors generally had longer within-turn pauses compared to those without metaphors. However, the difference was not significant after accounting for group-level variation in word count ($z = -0.035$, $p = .97$), which was also minimal in this model. The predicted within-turn pauses following turns with and without metaphors were 828.99 (95%CI[556.70, 1234.47]) and 831.34 (95%CI[562.19, 1229.35]), respectively.

Descriptive statistics showed that turns with deliberate metaphors had generally longer within-turn pauses compared to turns with non-deliberate metaphors and those without metaphors. When the influence of word count was accounted for at the group level, which was again minimal, within-turn pauses after turns with deliberate metaphors turned out to be shorter ($M = 552.39$, 95%CI[320.80, 951.17]) than in the other two utterance types ($M = 843.08$, 95%CI[568.92, 1249.37], $M = 828.18$, 95%CI[562.36, 1219.66], although the differences were not significant (p-values from .09 to .97).

4.2.3 Between-turn gap duration

Turns with metaphors were followed by shorter gaps compared to those without metaphors. The negative values suggest that participants often started responding to the utterance before the preceding speaker finished speaking, a pattern widely observed by previous studies (e.g., Jefferson, 2004). After accounting for the influence of word count, which did not have a significant effect ($t = 0.77$, $p = 0.44$), turns with metaphors ($M = -61.12$, 95%CI[−180.36, 58.13]) were predicted to have shorter between-turn gaps compared to those without metaphors ($M = -25.37$, 95%CI[−133.70, 82.96]), but the difference was not significant ($t = 0.99$, $p = .32$).

Between-turn gaps following turns with deliberate metaphors were generally shorter than those following turns with non-deliberate metaphors and turns without metaphors. After accounting for the influence of word count, which had no significant impact ($t = 0.76$, $p = 0.45$), the two utterance types were predicted to have similar mean between-turn gap duration (deliberate: $M = -61.12$, 95%CI[−303.74, 181.50]; non-deliberate: $M = -61.12$, 95%CI[−180.97, 58.73]). The duration was shorter than gaps after turns without metaphors, but the differences were not significant ($p = .95$ and .59, respectively).

4.2.4 Co-speech gesture

Participants gestured more in turns with metaphors compared to those without (see Table 3). After accounting for group-level variance in the influence of word count, which was minimal in this model, the predicted probability in the former

was 0.24 (95%CI[0.18, 0.32]), significantly higher than that in the latter (0.20, 95%CI[0.16, 0.25]; $z = -2.27$, $p = 0.02$)

Descriptive statistics show that a slightly lower proportion of turns with deliberate metaphors were gestured (44.44%) compared to turns with non-deliberate metaphors (46.35%), and both were more frequently gestured than turns without metaphors. The modelling results revealed the same pattern: the predicted probabilities of gesture occurrence in the three utterance types were 0.21 (95%CI[0.11, 0.37]), 0.24 (95%CI[0.18, 0.32]), and 0.20 (95%CI[0.16, 0.25]), respectively, and the influence of word count was still minimal. Turns with non-deliberate metaphors were significantly more likely gestured compared to those without metaphors ($z = 2.32$, $p = .05$), but not significantly different from turns with deliberate metaphors ($z = -0.47$, $p = .89$).

4.3 Discussion

The results show that speakers generally spend a longer time producing turns with metaphors than producing turns without any metaphors. The fact that the effect was significant even after controlling for word count indicates that the longer duration is not simply due to these turns containing more words but may reflect cognitive or processing factors associated with metaphor use. In other words, the speakers are more likely to slow down when using metaphors. This might reflect the complexity or expressive nature of metaphor, which requires more cognitive and verbal adjustments from the speaker. Across models, the effect of word count remained consistently small and non-significant, suggesting that sentence lengths did not play a major role in shaping the conversation behaviours under examination.

In actual conversations, extended turn duration can manifest as more frequent disfluencies, longer within-turn pauses, or slower pronunciation of each word. While our previous study shows that turns with metaphors were indeed more likely to be disfluent than turns without metaphors (Qiu et al., 2024), the interaction between metaphor presence and disfluency markers suggests that prolonged duration in turns with metaphors should not be attributed to increased disfluencies. Quite the contrary, fluent utterances with metaphors lasted significantly longer than those without metaphors, while no significant difference was found among disfluent utterances.

The lack of a significant difference in within-turn pause duration between turns with and without metaphors further suggests that within-turn pause is not a key factor leading to increased turn duration. The prolonged duration observed for turns with metaphors was most likely influenced by the time attributed to articulating each word in the turn. Compared to the presence of disfluencies and

within-turn pauses, slowed articulation in turns with metaphors is not merely a result of cognitive strain but more likely a pragmatic strategy to assist self-expression. This pattern reflects heightened meta-communicative awareness in the spontaneous production of metaphors.

Descriptive statistics suggest that metaphor presence in a turn may enhance other speakers' engagement in the conversation, potentially prompting quicker responses. However, between-turn gap durations were predicted to be similar across utterance types, suggesting that the next speaker spent comparable amounts of time responding, regardless of metaphor use or deliberateness. In other words, responding to turns with metaphors and turns with deliberate metaphors did not substantially increase the cognitive load on the other speakers. It is worth noting, however, that between-turn gap duration offers only a general measure of response time in conversation; it does not capture the full scope of cognitive processes that the listeners may engage in when interpreting metaphorical language. Additional measures would be necessary to understand the complexities of listeners' processing efforts and the effects of interacting.

Another significant difference was observed for co-speech gestures. Turns with metaphors were significantly more likely to be gestured compared to those without. A possible reason is that metaphors can invoke concrete, tangible, or bodily-related concepts (Citron et al., 2019; Desai, 2022), which can or need to be emphasised through non-verbal resources. In this process, gestures may thus serve as a communicative aid, enriching the metaphorical expression through visual-spatial dimensions. While this study leaves a detailed analysis of gesture types for future work, we anecdotally observe that some gestures and paralinguistic cues are aligned with the metaphorical part of the utterance. An example is provided below in Example (1). The gestures generated by the speaker are illustrated by Figure 4 and the linguistic example below (the letters refer to aligned gestures shown in the pictures).

A. **B.** **C.**

Figure 4. Gestures when using metaphors

(1) Well apparently *on the verge* of a breakthrough so surely he's quite close to it.
A B C

As shown in Figure 4A, the speaker on the right has raised his right hand. The movement began whilst he says the word "apparently", and the gesture hold occurs during his articulation of the word "verge". This co-speech gesture is accompanied by increased volume and a slower speech rate. The speaker then lowers his hand, sitting on it whilst saying "breakthrough" (4B). Subsequently, he raises his right hand again, in an open-palmed gesture aligned with production of the word "surely" (4C). The increased probabilities of gesture presence can be a reflection of cognitive engagement, which helps to alleviate the emerging cognitive load. It could also be viewed as an indicator of heightened meta-communicative awareness, through which speakers consciously adjust their non-verbal behaviours to support or clarify their metaphor use. While our current analysis does not classify gestures by type or examine their overlap with specific speech content, how different types of gestures support metaphor production represents a promising avenue for future research.

An interesting finding for metaphor researchers is that the effect of metaphor deliberateness was much less pronounced than that of metaphor presence across the four conversational behaviours. The occurrence of deliberate metaphors, in most cases, did not have a significant impact on conversational behaviours compared to the use of non-deliberate metaphors. The only significant difference was observed for gesture presence in turns with non-deliberate metaphors and those without any metaphors, which suggests that gestures may arise spontaneously in response to metaphor use in real time rather than being employed to reinforce particular types of metaphor use.

Given the overall lack of significant differences between deliberate and non-deliberate metaphors, it is tempting to conclude that whether a metaphor is intentional or incidental does not substantially influence how long the speakers spend articulating their ideas, how long they pause during their own and other speakers' utterances, and how often they gesture in speech. It is also possible that non-deliberate metaphors, although not produced with the explicit intention of highlighting their figurative nature, nonetheless involve underlying metaphorical meaning-making processes that engage embodied associative processes in a way that is distinct from literal language. However, it is worth noting that turns with deliberate metaphors exhibited larger variability in all four conversational behaviours, and the predicted data also had large confidence intervals, indicating considerable uncertainty in the estimated probabilities. A likely explanation is the relatively small number of deliberate metaphors within each category. Because the limited sample size may reduce statistical power, conclusions drawn from this comparison should be interpreted with caution. This variability also represents an interesting avenue for future research, particularly with larger datasets or experimental designs that can elicit more instances of deliberate metaphor use. An alter-

native explanation for the unbalanced categories is that the distinctions between deliberate and non-deliberate metaphor use in conversation may be more fluid than current linguistic identification methods can fully capture. As DMIP primarily focuses on linguistic features but less on paralinguistic cues that signal speakers' communicative awareness, some metaphors classified as non-deliberate on linguistic grounds may still involve degrees of intentional highlighting conveyed through non-verbal channels. Future research could therefore integrate linguistic, paralinguistic, and multimodal indicators to better account for varying degrees of communicative awareness in metaphor use.

5. Conclusions and future work

In this paper, we examined 2631 conversational turns, produced by 57 participants, to see whether the presence of metaphors and deliberate metaphors has an impact on conversation behvaviours like turn duration, within-turn pause, between-turn gaps, and co-speech gesture. By linking linguistic and behavioural data, our study provides empirical evidence for how metaphors are produced and received in spontaneous conversation, revealing subtle and previously neglected ways in which speakers manage timing, pacing, and gesture to construct and convey metaphorical meanings.

Our results show that after accounting for the influence of word count, turns with metaphors had significantly longer durations compared to turns without metaphors, which means people tend to speak more slowly and use more gestures when producing metaphors. However, turns with and without metaphors had roughly similar within-turn pause durations, which suggests that the use of metaphors does not appear to pose substantial cognitive challenges for the listener. The fact that speakers articulated each word more slowly with relatively shorter within-turn pauses suggests that the production of metaphors more likely reflects a stronger meta-communicative awareness than cognitive difficulties. While metaphors may create subtle shifts in conversational pacing, they do not uniformly accelerate or delay responses. In addition, deliberateness of metaphor use does not seem to be a key factor in shaping speakers' conversational behaviours.

Some limitations should be acknowledged. Firstly, the number of turns with deliberate metaphors was small. While the low frequency is consistent with previous studies (e.g., Cameron & Maslen, 2010; Kaal, 2012; Reijnierse et al., 2018), findings from the statistical analyses should be interpreted with caution. Future research could consider using a larger dataset that contains more deliberate metaphors. Secondly, only a limited number of conversational behaviours were

examined. Paralinguistic cues like intonation, pitch, gaze, and facial expressions were not taken into account due to constraints in data collection. Thirdly, gestures in this study were identified based on index of hand movement, and the forms of gestures were not taken into account. Future studies can use more comprehensive data collection methods to capture a broader range of conversational behaviours. Lastly, because this study examines utterances in spontaneous conversation tasks, the specific utterances being compared were not controlled in terms of semantic content, and responses from other speakers did not always address the preceding turn. Follow-up studies may use controlled stimuli to elicit metaphorical and non-metaphorical utterances on the same topic to explore the connection between topic and utterance types and the relationship between gestures and the content of speech. This can be achieved, for example, by applying picture-prompted discussions of abstract topics (see Tay et al. (2020) for an example of non-interactive tasks) in interactive contexts.

As an initial and quantitative exploration, this study only focused on a limited range of conversational behaviours, although many other features may also be associated with metaphor use, for example, stress, intonation, and facial expressions. Future research could examine a wider range of behaviours following our approach or explore whether different types of utterances are accompanied by distinct combinations of conversational behaviours. Based on our results, future research could also consider incorporating conversational behaviours into the identification of deliberate metaphors. Another interesting direction is conversational behaviours in online communicative contexts such as text-based chats and video-based chats. Text-based chats can be studied in terms of a different range of conversational behaviours, for example, typing time, correction before sending, and the use of emojis and emoticons, and the study of video-based chats could provide a new perspective on our knowledge about facial expression, gaze, and gestures in conversation. Studies in the two directions are expected to bring deeper insights into the interaction of multimodal behaviours in spontaneous conversations and the design of AI chatbots that have more natural and expressive communication abilities.

Funding

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 GU. Maraev was supported by an international postdoc grant from the Swedish Research Council (VR 2023-00358).
Open Access publication of this article was funded through a Transformative Agreement with University of Gothenburg.

References

Bambini, Valentina, Bott, Lewis, & Schumacher, Petra B. (2021). It is not always a matter of time: Addressing the costs of metaphor and metonymy through a speed-accuracy trade-off study. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale*, *75*(2), 189–196.

Bard, Ellen G., Lickley, Robin J., & Aylett, Matthew P. (2001). Is disfluency just difficulty? In *Proceedings of the international workshop on disfluency in spontaneous speech (DiSS'01)* (pp. 97–100). International Speech Communication Association.

Berthold, Andre & Jameson, Anthony. (1999). Interpreting symptoms of cognitive load in speech input. In *Proceedings of the seventh international conference* (pp. 235–244). Springer.

Bohrn, Isabel C., Altmann, Ulrike, & Jacobs, Arthur M. (2012). Looking at the brains behind figurative language — a quantitative meta–analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia*, *50*(11), 2669–2683.

Bolognesi, Marianna, Pilgram, Roosmaryn, & van den Heerik, Romy. (2017). Reliability in content analysis: The case of semantic feature norms classification. *Behavior Research Methods*, *49*, 1984–2001.

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

Brisard, Frank, Frisson, Steven, & Sandra, Dominiek. (2001). Processing unfamiliar metaphors in a self-paced reading task. *Metaphor and Symbol*, *16*(1–2), 87–108.

Cameron, Lynne. (2003). *Metaphor in educational discourse*. Continuum.

Cameron, Lynne & Maslen, Robert. (Eds.). (2010). *Metaphor analysis: Research practice in applied linguistics, social sciences and the humanities*. University of Toronto Press.

Campisi, Emanuela & Özyürek, Asli. (2013). Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. *Journal of Pragmatics*, *47*(1), 14–27.

Cienki, Alan. (2020). A multimodal perspective on MCA: Cues of (possible) metacommunicative awareness. In Camila Di Biase-Dyson & Markus Egg (Eds.), *Drawing Attention to Metaphor: Case studies across time periods, cultures and modalities* (pp. 63–91). John Benjamins.

Citron, Francesca M., & Goldberg, Adele E. (2014). Metaphorical sentences are more emotionally engaging than their literal counterparts. *Journal of Cognitive Neuroscience*, *26*(11), 2585–2595.

Citron, Francesca M., Cacciari, Cristina, Funcke, Jakob M., Hsu, Chun-Ting, & Jacobs, Arthur M. (2019). Idiomatic expressions evoke stronger emotional responses in the brain than literal sentences. *Neuropsychologia*, *131*, 233–248.

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

Clough, Sharice & Duff, Melissa C. (2020). The role of gesture in communication and cognition: Implications for understanding and treating neurogenic communication disorders. *Frontiers in Human Neuroscience*, *14*, 323.

Columbus, Georgie, Sheikh, Naveed A., Côté-Lecaldare, Marilena, Häuser, Katja, Baum, Shari R., & Titone, Debra. (2015). Individual differences in executive control relate to metaphor processing: An eye movement study of sentence reading. *Frontiers in Human Neuroscience*, *8*, 1057.

Coulson, Seana. (2008). Metaphor comprehension and the brain. In Raymond W. Gibbs. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 177–194). Cambridge University Press.

Cuccio, Valentina & Steen, Gerard. (2019). Deliberate metaphors and embodied simulation. In Ignasi Navarro i Ferrando (Ed.), *Current approaches to metaphor analysis in discourse* (pp. 185–204). De Gruyter Mouton.

Desai, Rutvik H. (2022). Are metaphors embodied? The neural evidence. *Psychological Research*, *86*(8), 2417–2433.

de Vries, Clarissa, Reijnierse, Gudrun, & Willems, Roel M. (2018). Eye movements reveal readers' sensitivity to deliberate metaphors during narrative reading. *Scientific Study of Literature*, *8*(1), 135–164.

Galati, Alexia & Brennan, Susan E. (2014). Speakers adapt gestures to addressees' knowledge: Implications for models of co-speech gesture. *Language, Cognition and Neuroscience*, *29*(4), 435–451.

Gallese, Vittorio & Lakoff, George. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3), 455–479.

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

Gibbs, Raymond W. (2005). *Embodiment and cognitive science*. Cambridge University Press.

Gibbs, Raymond W. (2015). Do pragmatic signals affect conventional metaphor understanding? A failed test of deliberate metaphor theory. *Journal of Pragmatics*, *90*, 77–87.

Gibbs, Raymond W. (1990). Comprehending figurative referential descriptions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(1), 56–66.

Glucksberg, Sam. (2008). How metaphors create categories-quickly. In Raymond W. Gibbs. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 67–83). Cambridge University Press.

Goldman-Eisler, Frieda. (1954). On the variability of the speed of talking and on its relation to the length of utterances in conversations. *British Journal of Psychology*, *45*(2), 94–107.

Goodwin, Charles. (1979). The interactive construction of a sentence in natural conversation. In George Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 97–121). Irvington.

Healey, Patrik G. T., Plant, Nicola, Howes, Christine, & Lavelle, Mary. (2015). When words fail: Collaborative gestures during clarification dialogues. In *AAAI Spring Symposium: Technical report* (pp. 23–29). AAAI Press.

Hilliard, Caitlin & Cook, Susan Wagner. (2016). Bridging gaps in common ground: Speakers design their gestures for their listeners. *Journal of Experimental Psychology*, *42*(1), 91–103.

doi Hoetjes, Marieke, Krahmer, Emiel, & Swerts, Marc. (2015). On what happens in gesture when communication is unsuccessful. *Speech Communication*, *72*, 160–175.

doi Holler, Judith, Kendrick, Kobin H., & Levinson, Stephen C. (2018). Processing language in face-to-face conversation: Questions with gestures get faster responses. *Psychonomic Bulletin & Review*, *25*(5), 1900–1908.

doi Hostetter, Autumn B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*(2), 297–315.

doi Howes, Christine & Lavelle, Mary. (2023). Quirky conversations: How people with a diagnosis of schizophrenia do dialogue differently. *Philosophical Transactions of the Royal Society B. 378*(1875), 20210480.

doi Jefferson, Gail. (2004). A sketch of some orderly aspects of overlap in natural conversation. In Gene H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 43–59). John Benjamins Publishing Company.

Kaal, Anna Albertha. (2012). Metaphor in conversation [Doctoral dissertation, Vrije Universiteit Amsterdam].

Kahneman, Daniel. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

doi Kita, Sotaro. (2000). How representational gestures help speaking. In David McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge University Press.

doi Kita, Sotaro & Davies, Thomas Stephen. (2009). Competing conceptual representations trigger co-speech representational gestures. *Language and Cognitive Processes*, *24*(5), 761–775.

doi Lai, Vicky Tzuyin, Curran, Tim, & Menn, Lise. (2009). Comprehending conventional and novel metaphors: An ERP study. *Brain Research*, *1284*, 145–155.

Lakoff, George & Johnson, Mark. (1980). *Metaphors we live by*. University of Chicago Press.

doi Lavelle, Mary, Healey, Patrik G. T., & McCabe, Rose. (2012). Is nonverbal communication disrupted in interactions involving patients with schizophrenia? *Schizophrenia Bulletin*, *39*(5), 1150–1158.

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

Lickley, Robin. (2001). Dialogue moves and disfluency rates. In *Proceedings of the workshop on disfluency in spontaneous speech (DiSS'01)* (pp. 93–96). International Speech Communication Association.

doi Lickley, Robin. (2015). Fluency and disfluency. In Melissa A. Redford (Ed.), *The handbook of speech production* (pp. 445–474). Wiley.

Linell, Per. (1982). *The written language bias in linguistics*. Linköping University.

McNeill, David. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

Mon, Serena K., Nencheva, Mira, Citron, Francesca M., Lew-Williams, Casey, & Goldberg, Adele E. (2021). Conventional metaphors elicit greater real-time engagement than literal paraphrases or concrete sentences. *Journal of Memory and Language*, *121*, 104285.

Morsella, Ezequiel & Krauss, Robert M. (2004). The role of gestures in spatial working memory and speech. *The American Journal of Psychology*, *117*(3), 411–424.

Müller, Christian, Großmann-Hutter, Barbara, Jameson, Anthony, Rummer, Ralf, & Wittig, Frank. (2001). Recognizing time pressure and cognitive load on the basis of speech: An experimental study. In *Proceedings of the 8th international conference on user modeling (UM 2001)* (pp. 24–33). Springer.

Müller, Cornelia & Tag, Susanne. (2010). The dynamics of metaphor: Foregrounding and activating metaphoricity in conversational interaction. *Cognitive Semiotics*, *6*(s1), 85–120.

Olkoniemi, Henri, Ranta, Henri, & Kaakinen, Johanna K. (2016). Individual differences in the processing of written sarcasm and metaphor: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(3), 433–450.

Onishi, Kristine H., & Murphy, Gregory L. (1993). Metaphoric reference: When metaphors are not understood as easily as literal expressions. *Memory & Cognition*, *21*, 763–772.

Ortony, Andrew, Schallert, Diane L., Reynolds, Ralph E., & Antos, Stephen J. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior*, *17*(4), 465–477.

Pragglejaz Group (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, *22*(1), 1–39.

Qiu, Amy Han, Vanzan, Vanessa, Soupiona, Chara, & Howes, Christine. (2024). Disfluencies in conversation: a comparison of utterances with and without metaphors. In *Proceedings of the 28th workshop on the semantics and pragmatics of dialogue* (Article 13). SemDial.

Reijnierse, Gudrun. (2017). The Value of Deliberate Metaphor. [PhD-Thesis — Research and graduation external, Universiteit van Amsterdam].

Reijnierse, Gudrun, Burgers, Christian, Krennmayr, Tina, & Steen, Gerard. (2020). The role of co-text in the analysis of potentially deliberate metaphor. In Camila Di Biase-Dyson & Markus Egg (Eds.), *Drawing attention to metaphor: Case studies across time periods, cultures and modalities* (pp. 15–38). John Benjamins.

Reijnierse, Gudrun, Burgers, Christian, Krennmayr, Tina, & Steen, Gerard. (2018). DMIP: A method for identifying potentially deliberate metaphor in language use. *Corpus Pragmatics*, *2*, 129–147.

Ritchie, David. (2013). *Metaphor*. Cambridge University Press.

Roberts, Sean G., Torreira, Francisco, & Levinson, Stephen C. (2015). The effects of processing and sequence organization on the timing of turn taking: A corpus study. *Frontiers in Psychology*, *6*, 509.

Semino, Elena. (2008). *Metaphor in discourse*. Cambridge University Press.

Shriberg, Elizabeth. (1996). Disfluency in Switchboard. In *Proceedings of the international conference on spoken language processing* (pp. 11–14). International Speech Communication Association.

Steen, Gerard. (2011). From three dimensions to five steps: The value of deliberate metaphor. *Metaphorik.de*, *21*, 83–110.

Steen, Gerard. (2015). Developing, testing and interpreting deliberate metaphor theory. *Journal of Pragmatics*, *90*, 67–72.

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

Steen, Gerard, Dorst, Aletta, Herrmann, J. Berenike, Kaal, Anna, Krennmayr, Tina, & Pasma, Tryntje. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. John Benjamins.

Tay, Dennis. (2016). A variational approach to deliberate metaphors. *Cognitive Linguistic Studies*, *3*(2), 277–298.

Tay, Dennis. (2024). Affective engagement with metaphor in introductory statistics education. *The Journal of Experimental Education*, *94*(1), 105–119.

Tay, Dennis, Huang, Jin, & Zeng, Huiheng. (2020). Prompting strategies and outcomes in picture-based counseling. In Bernadette Watson & Janice Kreiger (Eds.), *Expanding horizons in health communication: An Asian perspective* (pp. 29–46). Springer.

Thibodeau, Paul. (2017). The function of metaphor framing, deliberate or otherwise, in a social world. *Metaphor and the Social World*, *7*(2), 270–290.

Thibodeau, Paul & Boroditsky, Lera. (2011). Metaphors we think with: The role of metaphor in reasoning. *PLoS One*, *6*(2), e16782.

Address for correspondence

Amy Han Qiu
Faculty of Humanities
University of Gothenburg
C580
405 30 Gothenburg
Sweden
amy.han.qiu@gu.se

Biographical notes

Amy Han Qiu is a postdoctoral researcher at the Department of Philosophy, Linguistics and Theory of Science at the University of Gothenburg (FLoV, GU), Sweden. Her research focuses on the use and understanding of figurative language in dialogue and how mental health disorders influence or manifest in language use. She received her PhD in professional communication, specializing in mental health communication, from the Hong Kong Polytechnic University.

https://orcid.org/0000-0003-2658-2312

Christine Howes is a Professor of Computational Linguistics at FLoV, GU. Her main interest is linguistic interaction, specifically how people use the resources provided by language in conversation to coordinate their understanding.

https://orcid.org/0000-0002-2794-1586

Vladislav Maraev is a postdoctoral researcher in computational linguistics with a background in network engineering and cognitive science. He builds computational models that reproduce human behaviour, aimed at understanding how our cognition works. He received his PhD from GU on the topic of laughter in interaction and ways of enhancing spoken dialogue systems with a capability to understand the communicative functions of laughter.

https://orcid.org/0000-0001-9209-286X

Publication history

Date received: 23 June 2025
Date accepted: 24 December 2025
Published online: 24 April 2026