Deictic shift in the production of direct and indirect speech

Jianan Li, Katinka Dijkstra and Rolf A. Zwaan

Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, Rotterdam, The Netherlands
Corresponding author: Jianan Li; Email: lijianan37@gmail.com

Abstract

We investigated how deictic shifts and memory representation influence production difficulties in reported speech. In Experiment 1, participants read short stories, including a conversation between two protagonists. Participants then recalled the last sentence from the conversation in either direct or indirect speech. When participants had verbatim memory of the to-be-reported utterance, direct speech production was faster than indirect speech if indirect speech production required a deictic shift. If participants had verbatim memory and indirect speech production did not require a deictic shift, direct and indirect speech's speech latencies were similar. In Experiment 2, we investigated the production difficulties of reported speech when verbatim memory was unavailable to participants. First, participants read short stories. Next, half of the participants completed an intervening task, whereas the other half did not. Last, participants undertook a sentence completion task by filling in a missing word in direct or indirect speech. When participants did not have verbatim memory of to-be-reported utterances, direct speech production was faster than indirect speech if a deictic shift was needed to produce direct speech. These results suggest that production difficulties in reported speech depend on how to-be-reported utterances are represented in memory and whether deictic shifts are required.

Keywords: Direct speech; indirect speech; deictic shifts; memory representation

1. Introduction

Speakers face a considerable number of choices during language production. Given a communicative intention, they need to decide what information to express and in what order it should be arranged to express themselves clearly (Levelt, 1993). Speakers also have the flexibility to choose the sentence structure in which they convey information (Montag et al., 2017). For example, when reporting on what had been said previously, they could use direct speech (*Paul said, 'I am hungry'.*) or indirect speech (*Paul said that he was hungry*). The contrast between direct speech

and indirect speech has received substantial attention from both linguists and psychologists. This topic is intriguing because direct speech and indirect speech have almost the same semantic meaning but differ significantly from each other in surface forms. These two reporting methods lead people to represent, perceive, and comprehend them differently (Eerland et al., 2013; Eerland & Zwaan, 2018; Köder et al., 2015; Stites et al., 2013; Yao et al., 2011, 2012). For instance, comprehenders have a better memory of the exact wording of an utterance when it is presented in direct speech rather than indirect speech (Eerland et al., 2013). Furthermore, people need more time and are more prone to make mistakes when resolving the pronouns in direct speech than in indirect speech in an information transmission setting where the speaker is sharing information in the world around him or her (Köder et al., 2015; Köder & Maier, 2016). In this study, we investigate how grammatical features influence the production of direct and indirect speech in English. We explore whether the production difficulties of direct and indirect speech vary as a function of different surface structures.

The primary difference between direct speech and indirect speech lies in the reporter's perspective. In direct speech, the reported content is conveyed directly from the original speaker's perspective, while indirect speech involves reporting from the reporter's own perspective. Different perspectives in direct and indirect speech result in the use of different deictic terms. Consider the following examples[1]:

(1) *Mary said: 'I am going to visit my friend tomorrow'.*
(2) *Mary said that she was going to visit her friend the next day.*

As shown in examples (1) and (2), different pronouns (*I* versus *she*; *my* versus *her*), verbs (*am* versus *was*), and time references (*tomorrow* versus *the next day*) are used depending on which perspective the speaker has taken. Direct speech and indirect speech are also formulated into different sentence structures. Direct speech is constructed as a main clause (Banfield, 1973; De Vries, 2008). The quote is directly attached to the reporting verb (e.g., said) with no restrictions. Conversely, for indirect speech, a subordinate clause is used. The quote is introduced by the complementizer 'that'. A subordinate clause requires all obligatory constituents of a sentence (Mayes, 1990). In sum, direct speech and indirect speech generally differ from each other in terms of deictic expressions and sentence structures.

Previous studies have found that even a subtle difference in surface structures has a significant influence on production difficulty (e.g., the inclusion or omission of 'that' in the relative object clause) (Ferreira & Dell, 2000). As already illustrated, direct speech and indirect speech have several differences regarding surface sentence structures. A relevant question to ask given these characteristics would be: How do these distinctions influence the production difficulty of these two types of reporting? Even though no research to date has directly investigated this question, prior studies found a preference for direct speech over indirect speech among individuals with aphasia and children. People with aphasia and children are typically characterized by

---

[1]Please note that in this study, we distinguished between direct speech and indirect speech based on the English language. People should be cautious when generalizing results from this study to other languages because the differences between these two reporting styles may vary across different languages. For instance, direct speech and indirect speech are constructed as different sentence structures in English. However, in Mandarin Chinese, in most cases, they do not differ in sentence structure.

more limited language production competence compared with healthy adults (Lubinski, 1991; Spaccavento et al., 2014). The observed preference for direct speech has led researchers to assume that direct speech is potentially an easier communicative strategy (Li, 1986).

The first line of evidence suggesting that producing direct speech is easier than indirect speech comes from studies involving individuals with aphasia. Researchers have observed that direct speech is dominant in aphasic speakers' reported speech (Hand et al., 1979; Menn et al., 1995; Ulatowska et al., 2011). Another study further compared whether individuals with aphasia and non-brain-damaged individuals used the two reporting styles differently. The results show that speakers with aphasia use more direct speech than their healthy counterparts in picture description and personal narrative tasks (Groenewold et al., 2014). In particular, individuals with Broca's aphasia who have verb-finding difficulties used bare quotations (i.e., a quotation without a quotative verb) significantly more than other types of quotations. It is, therefore, argued that speakers with aphasia strategically used direct speech to deal with grammatical problems and difficulties in identifying the correct word (Groenewold et al., 2014).

Studies involving children also seem to suggest that direct speech is easier to produce. For example, direct speech was observed to appear at an earlier age than indirect speech in several languages, including English, Swedish, Turkish, Dutch, and German (Ely & McCabe, 1993; Köder, 2013; Nordqvist, 2001; Özyürek, 1996). A significant preference for direct speech over indirect speech has been observed in children during various activities, such as the construction of personal narratives, dinner-time conversations with parents (Ely & McCabe, 1993), book-reading (Nordqvist, 2001), and make-believe play (Nordqvist, 2001). In one study comparing children with adults on the use of direct and indirect speech in a narrative production task, six-year-old children used direct speech more frequently than adults (Goodell & Sachs, 1992). In other words, direct speech occurs earlier and more frequently in children's language. Researchers argued that this is because producing direct speech is easier than indirect speech for reasons described as follows.

First, direct speech is a 'reproducing' and 'mimicking' of previous utterances, whereas indirect speech is an act of paraphrasing, which includes a level of interpretation (Li, 1986). Direct speech production requires less cognitive effort than indirect speech because 'mimicking' is easier than 'paraphrasing'. The second explanation proposed that the presence of paralinguistic and non-verbal information contributes to direct speech's ease of production. Studies indicate that people with aphasia make use of paralinguistic information to get around using verbs. They used prosody (i.e., an increase in pitch) to signal the use of direct speech instead of using reporting words such as 'say' and 'go' (Lind, 2002). Even though these two views have different focuses, they do not necessarily contrast each other. As a matter of fact, it is likely that both factors play a part in the production processes of direct and indirect speech. As of yet, neither of these two possibilities has been put to a direct experimental test, highlighting the need for further investigation.

As the first step toward evaluating these two possibilities empirically, this study aimed to test how the shift of perspectives (deictic terms) influences the production difficulties of direct and indirect speech. In this study, a perspective shift refers to the process of transforming deictic terms or adjusting pronouns to align the utterance with a new context or perspective. In two experiments, participants first read short written dialogues between two protagonists. Following this, participants were

prompted to answer questions related to the dialogue in either direct speech or indirect speech. By adopting this methodology, we compared the effect of perspective shifting on language production while controlling for the confounder of non-verbal information that is associated with direct speech.

As demonstrated earlier, direct speech and indirect speech differ in the use of deictic terms. To understand how deictic terms influence the production of direct and indirect speech, we first must consider how a previous utterance is represented. Since reported speech is defined as an utterance that refers to previous utterances[2], memory may play an important role in direct and indirect speech production. Prior studies suggest that sentences can be represented in either a verbatim way or a propositional way (Anderson, 1974; Fisher & Radvansky, 2018). The verbatim representation includes information about a sentence's surface properties (e.g., the exact wording), whereas the propositional representation is about the gist of a message. Usually, the verbatim representation is short-lived. Memory for the surface form is forgotten quickly after a sentence has been comprehended. In contrast, the gist of a sentence is retained for a more extended period (Sachs, 1967, 1974). For instance, one can easily remember reading about a car accident in the local news several days ago. It is more difficult to determine whether the original sentence that described this event was '*A man was hit by a car*' or '*A car hit a man*'.

We predict that direct and indirect speech production difficulties depend on which type of representation is accessed. When speakers rely on verbatim representation, producing direct speech will be less challenging than indirect speech. This is because direct speech shares the same speaking perspective and surface forms as the original utterance. The production of indirect speech, however, requires speakers to undertake a transformation of deictic terms. Reversely, when the production of reported speech relies on the gist representation, producing direct speech will become costlier than indirect speech because direct speech requires a narrative shift from the present context to a previously reported situation. Such a transformative process has been found to affect the comprehension of direct speech negatively. People take longer and make more errors when interpreting pronouns in direct speech (Köder et al., 2015; Köder & Maier, 2016). Therefore, when the verbatim memory is disrupted, the production of deictic terms in direct speech is expected to be more demanding than in indirect speech despite the similarities in surface forms between direct speech and the original utterance.

In this study, two experiments were conducted to test the effect of memory representation and deictic shifts on direct and indirect speech production difficulties. The effect of deictic shifts on direct and indirect speech production difficulties was examined when participants had verbatim memory of to-be-reported utterances (Experiment 1) and when the verbatim representation of those utterances was disturbed by an interfering task (Experiment 2). The hypothesis for each experiment is formulated as follows: ***Experiment 1: Direct speech production is faster when verbatim memory of the to-be-reported utterance is available, as compared to indirect speech production that involves transforming deictic expressions. Experiment 2: Direct speech production is slower when verbatim memory of the to-be-reported utterance is disrupted, as compared to indirect speech production.***

---

[2]We are interested in the reporting of utterances that were actually spoken out loud. Imagined direct speech (e.g., I was thinking: 'You idiot!') is beyond the scope of this study.

Jason came into Melanie's office because he could not reach her by phone.

Melanie was confused because she had been in her office the whole time.
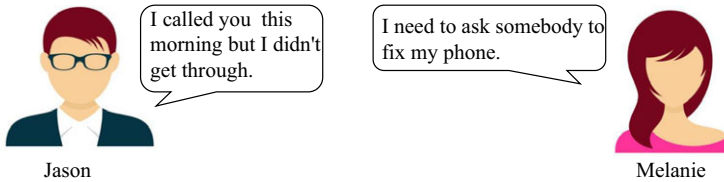

Jason: I called you this morning but I didn't get through.

Melanie: I need to ask somebody to fix my phone.

Jason

Melanie

Figure 1. An example of the stories from the deictic shift condition in Experiment 1.

2. Experiment 1

2.1. Method

Design and Stimuli. Experiment 1 had a 2 (deictic shift: shift versus no shift) × 2 (speech type: direct speech versus indirect speech) within-subjects design. The dependent variable was speech latency (from the onset of the answer screen to the onset of the speech report). The independent variable deictic shift was manipulated to determine whether the deictic words in an original utterance needed to be changed when reporting indirectly. For example, imagine John wants to quote Linda's utterance, '*Mary had a pony when she was a little girl*'. The deictic words in direct and indirect speech are the same. Therefore, no deictic shift is needed. If Linda says, '*I had a pony when I was a little girl*', the deictic word '*I*' needs to be changed to '*She*' when John quotes Linda indirectly. For the experiment, 40 experimental stories were created as stimuli, among which 20 stories were for the deictic shift condition and the other 20 stories were for the no deictic shift condition. Each story consisted of four sentences in which the last two sentences were always presented as a dialogue between two people. All stories were presented in the form of a picture (Fig. 1).

Participants. We were interested in the interaction between deictic shift and speech type. The determination of the sample size was based on Brysbaert's (2019) power analysis investigation for a within-subjects design with two factors. 110 participants are needed to detect an interaction effect in a two-factor repeated-measures design with an effect size of $d = 0.4$ and a power of 80% (Brysbaert, 2019). We recruited 120 participants to have enough observations after removing invalid data. Six participants were excluded due to a program crash. This resulted in a final sample of 114 participants ($M_{age} = 20.21$; 79 females). All participants were either native English speakers or had a minimum level of 80 on the Test of English as a Foreign Language (TOEFL) or a 6.0 on the International English Language Testing System (IELTS). They received 1-hour research hour credit for participation. Informed consent was obtained from all participants before the experiment. This experiment has been approved by the Research Ethics Review Committee DPECS of the Erasmus University Rotterdam.
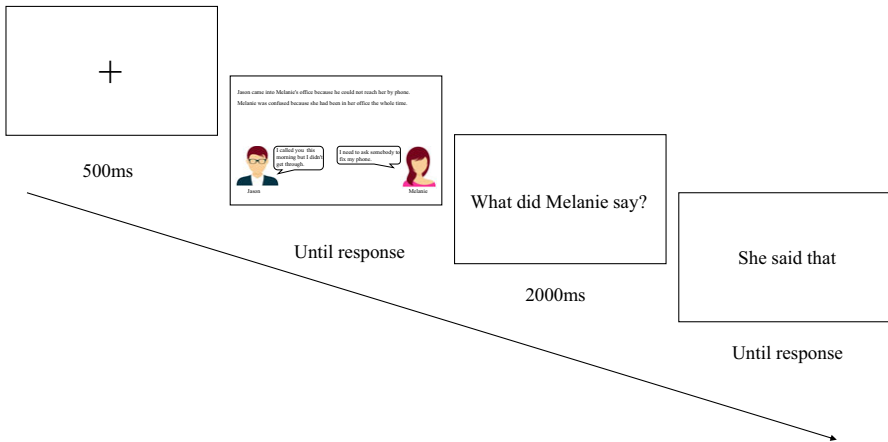
Figure 2. Experiment procedure for Experiment 1.

Procedure. Each participant was tested individually in a sound-attenuated room. First, participants read the instructions[3] and completed eight practice trials to become familiar with the procedure. The instructions indicated that they would be presented with a short story and thereafter needed to answer a question that corresponded to the story. The purpose of Experiment 1 was to test the production of direct and indirect speech when the verbatim memory of a to-be-reported utterance was accessible. Therefore, participants were asked to read the story carefully and memorize the last sentence of the story. In the deictic shift condition, the last sentence always contained at least two deictic terms that needed to be transformed when participants produced indirect speech.

Each trial began with a fixation cross appearing in the middle of the screen for 500 ms, followed by the story. After inspecting the story and memorizing the last sentence, participants pressed the SPACE bar to see the question. In total, 40 experimental questions were created to elicit direct or indirect speech (e.g., *What did Marie say?*). The questions were presented for 2,000 ms and then replaced by the answer screen. The first part of the answer was provided to prompt direct and indirect speech (Fig. 2). The participant's task was to complete the sentence according to the story they had just read. For example, when the question was '*What did Marie say?*' participants would be presented with either '*She said:*' to elicit direct speech or '*She said that*' to elicit indirect speech. Participants were asked to recall the sentence verbatim. This means that no word from the original sentence should be changed when producing direct speech, and only deictic terms were changed in indirect speech. To make the utterance planning more extensive, participants were instructed to speak as fluently as possible. Participants answered the questions by speaking into

---

[3]The instructions were as follows: 'You will be presented with a short story. Please read it carefully and memorize the last sentence of the story. After you have understood the story, please press the space bar to proceed to the answer screen. On the answer screen, you will be provided with the first part of the answer. Please think of your answer and formulate it beginning with the given words. Once you have formulated your answer, speak the sentence into the microphone as fluently as possible. While speaking, please avoid pauses and filler words (e.g., um)'.

a microphone. Speech latencies and audio responses were recorded for later analyses. Data and materials can be accessed via https://osf.io/zha4t/?view_only=e9e0b8ec491 d4af1bf1c0e6c605ca7c2.

2.2. Statistical analysis

Several types of responses were excluded from the final data analyses. First, trials with wrong responses (i.e., trials in which participants failed to recall verbatim or made errors in transforming deictic terms) were removed (8.3%). Second, utterances with disfluencies were removed (1.8%). Applying these exclusion criteria resulted in the removal of 10.1% of the trials in total. Data analyses were conducted with R version 4.2.2 (R Core Team, 2022) and lme4 version 1.1.31 (Bates et al., 2015) to perform a linear mixed-effects analysis of the influence of deictic shift and speech type on speech latencies of reported speech (Baayen et al., 2008). The model contained deictic shift and speech type and their interaction as fixed effects. As for random effects, we entered the participants and stimuli as varying intercepts, as well as by-participant and by-stimulus random slopes for the effect of deictic shift. P-values were obtained through likelihood ratio tests, comparing the full model including the specific variable to a reduced model that does not include the variable.

3. Results and discussion of Experiment 1

There was a significant interaction between shift and speech type ($\chi^2(1) = 32.61$, $p < 0.01$). In the shift condition, participants responded faster when they were asked to produce direct speech ($M = 1638$ ms, $SD = 727$ ms) than when they produced indirect speech ($M = 1795$ ms, $SD = 884$ ms) ($p < 0.01$). In the no-shift condition, there was no significant difference in speech latencies between direct speech ($M = 1831$ ms, $SD = 778$ ms) and indirect speech ($M = 1818$ ms, $SD = 808$ ms) ($p > 0.05$) (Fig. 3).
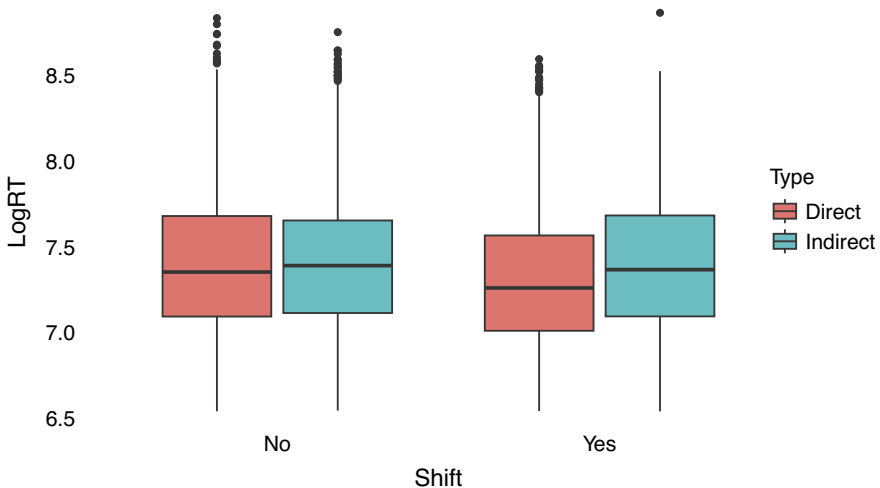

Figure 3. Log reaction time in each condition in Experiment 1.

The findings from Experiment 1 supported our predictions. When participants had verbatim memory, direct speech exhibited significantly shorter speech latencies than indirect speech when there was a deictic shift. Speech latencies for direct and indirect speech were not significantly different from each other when there was no deictic shift. These results are congruent with several previous studies regarding verbatim and gist memory in question answering. In those studies, participants were asked to answer questions about the texts they had just read. When there was no interval between the texts and the questions, participants responded faster and made fewer errors than when the questions were framed the same as the sentences in the texts. This matching advantage disappeared when there was an interval or an interference task between reading the texts and answering the questions (Anderson, 1974; Garrod & Trabasso, 1973; McKoon & Keenan, 1974; Wright, 1969). Even though direct speech may not always be a verbatim reproduction of previous utterances, it is an easier way of reporting if people have verbatim memory of to-be-reported utterances, as shown by Experiment 1.

If the existence of verbatim memory is the reason for the faster production of direct speech, the question is what will happen if the production of direct and indirect speech can only rely on the gist memory of original utterances? The purpose of Experiment 2 was to investigate the production of reported speech when the verbatim memory of the original utterance was hindered. To this end, we made several adjustments to the experimental procedures and stimuli from Experiment 1. First, we modified the procedures by adding an intervening task between reading the stories and producing reported speech. By adding a counting backward intervening task, we aimed to reduce participants' verbatim trace of the story. Second, Experiment 2 used a slightly different language production task. In Experiment 1, participants were presented with either '(S)he said: "" or '(S)he said that' and their task was to recall the whole quote. In Experiment 2, to control the content that would be produced, participants were presented with a sentence that had only one word (i.e., a deictic item) missing. Their task was to think of a word that could make this sentence complete according to the story they had just read. In addition, in all experimental trials, the missing word was always a personal pronoun. We chose to remove the personal pronouns because participants were the least likely to make mistakes in shifting personal pronouns compared with shifting verb tense and shifting time or space adverbs. Third, to keep the rest of the content of the sentence identical in the direct and indirect speech conditions, we created sentences with only one deictic item as stimuli. In Experiment 1, to maximize the effect of deictic shifting on the production of direct and indirect speech, we created sentences with at least two deictic words. One example of the sentences we used in Experiment 1 was '*I need to ask someone to fix my phone*'. Here, 'I', 'need', and 'my' must be transformed to report this sentence indirectly. In Experiment 2, we used sentences that had only one deictic term. For example, one experimental sentence is '*My friends would like to drink more, but the pub was already closed*'. The only difference between direct speech and indirect speech is the pronoun (my versus his). The participants' task was to think of the correct pronoun according to whether they were prompted to produce direct or indirect speech. We predicted that in the no-interval condition, producing direct speech would be faster than producing indirect speech. In the interval condition, producing direct speech would be slower than producing indirect speech.

4. Experiment 2

4.1. Method

Design and Stimuli. Experiment 2 used a 2 (interference: yes versus no) × 2 (speech type: direct speech versus indirect speech) mixed design. Interval was a between-subjects variable, with half of the participants needing to finish an interference task before the sentence completion task and the other half of participants only finishing the sentence completion task. Speech type was a within-subjects variable. A total of 20 stories were created as stimuli. Similar to Experiment 1, each story consisted of four sentences, and the last two sentences were presented as a dialogue between two people. The incomplete sentence was created by removing the first word of the last sentence in the dialogue.

Participants. According to Brysbaert (2019), for a two-way ANOVA with one within-groups factor and one between-groups factor, 67 participants for each group are needed to run a study with 80% power. To ensure a sufficient number of participants after removing invalid data, we recruited 150 university students as participants. Five participants from the interval condition were excluded due to a program crash, and the last five participants from the no-interval condition were removed to ensure equal numbers of participants in both conditions. The final sample size consisted of 140 participants (M_{age} = 22.15, 92 females). All participants were either native English speakers or had a minimum level of 80 on the Test of English as a Foreign Language (TOEFL) or a 6.0 on the International English Language Testing System (IELTS). They received 1-hour research hour credit for participation. Informed consent was obtained before the experiment. This experiment has been approved by the Research Ethics Review Committee DPECS of the Erasmus University Rotterdam.

Procedure. Participants were tested individually in a sound-attenuated room. After reading the instructions and indicating that they already understood the procedures, participants were asked to finish six practice trials to be familiarized with the task. The procedures began with a fixation at the center of the screen, reminding participants that the experiment would start soon. Following the fixation, participants were presented with a short story. They were instructed to read the story carefully and press the SPACE bar after they had understood it. After reading the story, half of the participants continued to complete a counting backward task. In the counting task, a 3-digit number (e.g., 469) appeared on the screen. Participants were instructed to first read out this number and then count backward in steps of 3. Participants kept counting for 20 seconds until they saw the word 'STOP' on the screen. Following the stop sign, a fixation appeared again on the screen for 200 ms, and an incomplete sentence was presented. Participants were asked to think of a word that could make this sentence complete based on the story they had just read. Once they had the answer, they read the whole sentence into a microphone as fluently as possible. Participants immediately finished the sentence completion task after reading the stories for the no-intervention condition. Speech latencies and audio responses were recorded for later analyses.

4.2. Statistical analysis

Responses with wrong answers (9.87%) or disfluencies (6.92%) were removed. Data analyses were conducted with R version 4.2.2 (R Core Team, 2022) and lme4 version 1.1.31 (Bates et al., 2015) to perform a linear mixed-effects analysis of the influence of

interval and speech type on speech latencies of reported speech(Baayen et al., 2008). The model contained interference and speech type and their interaction as fixed effects. In our analysis, we accounted for random effects by including participants and stimuli as varying intercepts. Additionally, we included by-participant and by-stimulus random slopes for the effects of speech type. *P*-values were obtained through likelihood ratio tests, comparing the full model including the specific variable to a reduced model that does not include the variable.

4.3. Results and discussion of Experiment 2

There was a significant interaction between interference and speech type (($\chi^2(1) = 24.86$, $p < 0.01$). A post-hoc analysis showed that, in the interval condition, speech latencies of indirect speech ($M = 1619$ ms, $SD = 609$ ms) were significantly shorter than speech latencies of direct speech ($M = 1886$ ms, $SD = 764$ ms) ($p < 0.01$). In the no-interval condition, speech latencies of indirect speech ($M = 1137$ ms, $SD = 311$ ms) did not significantly differ from speech latencies of direct speech ($M = 1096$ ms, $SD = 229$ ms) ($p > 0.05$) (Fig. 4). Our hypotheses were only partially supported by these results. In line with our prediction, when the verbatim trace of an utterance was impaired by an intervening task, recalling the utterance using indirect speech was faster than direct speech. To our surprise, Experiment 2 did not observe a difference in speech latencies for direct and indirect speech when verbatim memory was available. This result seems to be in contrast with the results from Experiment 1, where we observed that direct speech production was faster than indirect speech, but it can be explained by the differences in tasks. We will return to this discrepancy in the discussion section.

4.4. General discussion

The current study examined whether and how memory representation and deictic shifts influenced direct and indirect speech production difficulties. We compared
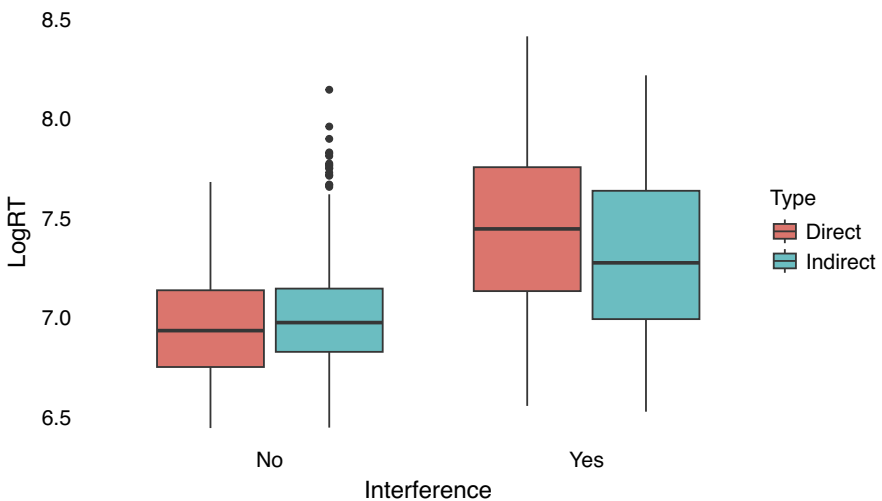

Figure 4. Log reaction time in each condition of Experiment 2.

production difficulties of direct and indirect speech in conditions where either a deictic shift was or was not required to produce reported speech. In both experiments, participants read short dialogues between two interlocutors and were prompted to answer questions using either direct or indirect speech. Experiment 1 investigated the effect of deictic shifts on production difficulties when participants had verbatim memory of the to-be-reported utterances. The results showed that deictic shifts significantly influenced speech latencies. When indirect speech production required a deictic shift, participants took a longer time to initiate indirect speech than direct speech. When no deictic shifts were needed, direct and indirect speech latencies were similar. In Experiment 2, we continued to investigate the effect of deictic shifts on production difficulties when participants' verbatim memory of to-be-reported utterances was interfered by an intervening task. The results showed that when verbatim memory was hindered, producing direct speech was slower than indirect speech. Unlike Experiment 1, we did not observe a difference in speech latencies between direct speech and indirect speech when verbatim memory was available.

This study was motivated by several arguments and observations, suggesting that direct speech is an easier mode of reporting than indirect speech (Groenewold et al., 2014; Li, 1986). The results of our study, however, painted a more complicated picture. First, we argued that the memory representation of the to-be-reported utterances should play a role in the production of reported speech because reported speech involves a recollection of previously said content. Even though direct speech may not necessarily be a verbatim replication of to-be-reported content, the retention of surface structures plays an essential role in language production. Verbatim memory can be used in the immediate output (Garrod & Trabasso, 1973). Experiment 1 tentatively supported the role of memory in reported speech production, showing that direct speech production was faster than indirect speech when verbatim memory was readily accessible to participants. We predict that the reason for the longer speech latency for indirect speech is that it requires a transformation of deictic words, whereas direct speech does not. When direct and indirect speech shared the same deictic expressions (e.g., *She said: 'John was late'.* versus *She said that John was late.*), the production of these two reporting styles was equally fast. This result again supports the argument that when verbatim memory was accessible to participants, the difference in speech latencies was caused by deictic shifts. When no deictic shift was required, there was no difference between direct speech and indirect speech in terms of speech latencies. In sum, the findings from Experiment 1 demonstrate two points: First, the retention of verbatim memory facilitates the production of direct speech. Second, verbatim retention only has a facilitating effect when the production of indirect speech requires a transformation of deictic expressions. When the deictic terms are the same in direct and indirect speech, there is no difference in speech latencies.

Experiment 1 helps to clarify the existing view arguing that direct speech is easier to produce because it is a verbatim replication of previous utterances (Li, 1986). While this belief has been criticized because people do not always report verbatim (Wade & Clark, 1993), we cannot deny that, between direct speech and indirect speech, only direct speech allows people to recite in a verbatim manner. Our findings from Experiment 1 modify this view by suggesting that, when people have verbatim memory of previous utterances, direct speech is easier than indirect speech because no deictic shift is needed for direct speech production.

Previous studies repeatedly show that the retention of an utterance's surface form decays quickly and is prone to interference (Fisher & Radvansky, 2018; Sachs, 1967). In most daily communication circumstances, the production of direct and indirect speech can only rely on the gist memory of previously said utterances. In Experiment 2, we compared speech latencies of direct and indirect speech when verbatim memory was not disturbed by an interfering task. We introduced a counting backward task that came in between the dialogue reading and the speech production task. We expected the production of direct speech to be slower than indirect speech when the verbatim trace was disrupted. We observed that when participants' verbatim memory of an utterance was hindered, indirect speech production was faster than direct speech. This aligns with previous findings from reported speech comprehension studies showing that participants spent more time interpreting pronouns in direct speech than in indirect speech (Köder et al., 2015; Köder & Maier, 2016). Resolving pronouns in direct speech requires a shift from the third-person perspective to the first-person perspective, resulting in a longer resolution time (Köder et al., 2015; Köder & Maier, 2016). Similarly, in Experiment 2, the production of direct speech also required a shift from the current situation to the time and space in which the original utterance was produced. This shift leads to longer reaction times for direct speech production than for indirect speech. The finding from Experiment 2, together with prior reported speech comprehension studies, suggests that perspective shift contributes to processing difficulty differences between direct speech and indirect speech.

Unlike Experiment 1, we did not observe significant differences in speech latencies when there was verbatim memory available. This is unexpected, considering that the production of direct speech was significantly faster than indirect speech in Experiment 1 when participants had access to verbatim memory. We speculate that this could be because different tasks were used in Experiment 1 and Experiment 2. In Experiment 1, participants recalled the whole sentence. Each sentence contained at least two deictic terms that needed to be changed to produce indirect speech. In contrast, in Experiment 2, participants only needed to paraphrase one pronoun. Presumably, the transformation of only one deictic term is not very resource-consuming and therefore can happen at a fast pace. Because the task demand of Experiment 2 was less taxing than that of Experiment 1, no difference in speech latencies was detected when verbatim memory was available in Experiment 2.

We recognize several limitations of our study. Firstly, we note that deictic shifts might only account for part of the production difficulty differences between direct speech and indirect speech. This study did not test whether other syntactic differences between these two reporting styles would influence their production difficulties. Some sentence structures have been demonstrated to be more difficult to produce than others, as is reflected by longer speech latencies or a larger pupil size (Altmann & Kemper, 2006; Sevilla et al., 2014). For example, producing passive and object-dislocated sentences evokes larger pupil dilation than active–canonical subject–verb–object sentences, suggesting that passive–noncanonical structures are more difficult to produce and require more cognitive effort (Sevilla et al., 2014). As discussed before, direct speech and indirect speech have different sentence structures, with direct speech constructed as a main clause and indirect speech as a subordinate

clause. In a main clause, to-be-reported content is directly attached to the reporting verbs (e.g., say). Consequently, direct speech can convey all syntactic structures, whereas indirect speech can only convey utterances that are grammatically correct in a subordinate clause. Observational and empirical evidence showed that people tend to use direct speech more frequently when the to-be-reported utterances are perceived as grammatically less acceptable in indirect speech (Mayes, 1990; Li et al., 2022). The use of direct speech sometimes involves much less syntactic paraphrasing than indirect speech. For instance, suppose *A* would like to quote B's utterance 'Say what?'. *A* can say, '*B* said, "Say what?"' in direct speech. *A* can also use indirect speech and say, '*B* asked if I could repeat what I said'. One would expect that when indirect speech requires much syntactic paraphrasing, it will be more difficult to produce compared with direct speech.

Second, we did not test the influence of non-verbal information on the production of direct and indirect speech. As demonstrated by prior studies, the use of non-verbal information might facilitate speech production, especially for people with limited language competence (Groenewold et al., 2014; Ulatowska et al., 2011). The non-verbal information could facilitate the production of direct speech by either replacing a difficult word with gestures or helping to recall the content. Despite this limitation, the results of our study remain informative because people often encounter direct speech without non-verbal information involved, such as producing direct speech in a written form.

Finally, this study did not consider reported speech production difficulties in various discourse types. Prior studies show that discourse types might influence the difficulty of shifting perspectives. In particular, people make more mistakes and have longer reaction times when interpreting pronouns in direct speech than in indirect speech because comprehending direct speech requires a deictic shift. However, people's abilities to interpret pronouns in direct speech are significantly improved when direct speech is embedded in a narrative setting compared with an information transmission setting (Köder et al., 2015; Köder & Maier, 2016). The narrative setting makes it easier to take the first-person setting and therefore reduces the time for perspective shifting (Köder & Maier, 2016). Future research can compare whether different reporting contexts will also influence the difficulty of perspective shifting in reported speech production.

As discussed in the limitation section, the comparison of production difficulties is an extremely complicated topic. The relative production difficulty of direct and indirect speech may depend on factors such as the memory representation of the utterances, the properties of the original utterances, or the contexts in which reported speech is produced. Despite the above-mentioned limitations, this study contributes to further our knowledge and understanding of direct and indirect speech. The most important theoretical contribution is that we provided empirical evidence for the role of memory representation in the production of direct and indirect speech. We can conclude that the availability of verbatim memory facilitates the production of direct speech, reflected by shorter speech latencies compared with indirect speech. When the verbatim representation is hindered, the production of indirect speech is faster instead.

References

Altmann, L. J., & Kemper, S. (2006). Effects of age, animacy and activation order on sentence production. *Language and Cognitive Processes*, 21(1–3), 322–354. https://doi.org/10.1080/0169096054400006

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 6(4), 451–474. https://doi.org/10.1016/0010-0285(74)90021-8

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Banfield, A. (1973). Narrative style and the grammar of direct and indirect speech. *Foundations of Language*, 10(1), 1–39.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.48550/arXiv.1406.5823

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, 2(1), 16. https://doi.org/10.5334/joc.72

De Vries, M. (2008). The representation of language within language: A syntactico-pragmatic typology of direct speech. *Studia Linguistica*, 62(1), 39–77. https://doi.org/10.1111/j.1467-9582.2007.00142.x

Eerland, A., Engelen, J. A., & Zwaan, R. A. (2013). The influence of direct and indirect speech on mental representations. *PLoS One*, 8(6), e65480. https://doi.org/10.1371/journal.pone.0065480

Eerland, A., & Zwaan, R. A. (2018). The influence of direct and indirect speech on source memory. *Collabra: Psychology*, 4(1), 5. https://doi.org/10.1525/collabra.123

Ely, R., & McCabe, A. (1993). Remembered voices. *Journal of Child Language*, 20(3), 671–696. https://doi.org/10.1017/S0305000900008539

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40(4), 296–340. https://doi.org/10.1006/cogp.1999.0730

Fisher, J. S., & Radvansky, G. A. (2018). Patterns of forgetting. *Journal of Memory and Language*, 102, 130–141. https://doi.org/10.1016/j.jml.2018.05.008

Garrod, S., & Trabasso, T. (1973). A dual-memory information processing interpretation of sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 12(2), 155–167. https://doi.org/10.1016/S0022-5371(73)80005-2

Goodell, E. W., & Sachs, J. (1992). Direct and indirect speech in English-speaking children's retold narratives. *Discourse Processes*, 15(4), 395–422. https://doi.org/10.1080/01638539209544820

Groenewold, R., Bastiaanse, R., Nickels, L., Wieling, M., & Huiskes, M. (2014). The effects of direct and indirect speech on discourse comprehension in Dutch listeners with and without aphasia. *Aphasiology*, 28(7), 862–884. https://doi.org/10.1080/02687038.2014.902916

Hand, C. R., Tonkovich, J. D., & Aitchison, J. (1979). Some idiosyncratic strategies utilized by a chronic Broca's aphasic. *Cortex*, 8, 191–212.

Köder, F. (2013). How children acquire reported speech in German and Dutch: A corpus study. In B. Sonnenhauser, C. Trautmann, & P. N. A. Hanna (Eds.), *Perspektiven Diskussionsforum Linguistik in Bayern/Bavarian Working Papers in Linguistics* (vol. 2, pp. 15–28). München: Bamberg.

Köder, F., & Maier, E. (2016). Children mix direct and indirect speech: Evidence from pronoun comprehension. *Journal of Child Language*, 43(4), 843–866. https://doi.org/10.1017/S0305000915000318

Köder, F., Maier, E., & Hendriks, P. (2015). Perspective shift increases processing effort of pronouns: A comparison between direct and indirect speech. *Language, Cognition and Neuroscience*, 30(8), 940–946. https://doi.org/10.1080/23273798.2015.1047460

Levelt, W. J. (1993). *Speaking: From intention to articulation*. MIT Press.

Li, C. (1986). Direct and indirect speech: A functional study. In C. Coulmas (Ed.), *Direct and indirect speech* (pp. 29–45). De Gruyter Mouton.

Li, J., Jongerling, J., Dijkstra, K., & Zwaan, R. (2022). The influence of utterance-related factors on the use of direct and indirect speech. *Collabra: Psychology*, 8(1), 33631. https://doi.org/10.1525/collabra.33631

Lind, M. (2002). The use of prosody in interaction: Observations from a case study of a Norwegian speaker with a non-fluent type of aphasia. In F. Windsor, M. L. Kelly, & N. Hewlett (Eds.), *Investigations in clinical phonetics and linguistics*. Lawrence Erlbaum Associates Inc.

Lubinski, R. (1991). *Dementia and communication*. B. C. Decker, Inc.

Mayes, P. (1990). Quotation in spoken English. *Studies in Language*, 14(2), 325–363. https://doi.org/10.1075/sl.14.2.04may

McKoon, G., & Keenan, J. M. (1974). Response latencies to explicit and implicit statements as a function of the delay between reading and test. In W. Kintsch (Ed.), *The representation of meaning in memory* (pp. 166–176). Erlbaum.

Menn, L., O'Connor, M., Obler, L., & Holland, A. (1995). *Nonfluent aphasia in a multilingual world*. John Benjamins. https://doi.org/10.1075/sspcl.5

Montag, J. L., Matsuki, K., Kim, J. Y., & MacDonald, M. C. (2017). Language specific and language general motivations of production choices: A multi-clause and multi-language investigation. *Collabra: Psychology*, 3(1), 1–22. https://doi.org/10.1525/collabra.94

Nordqvist, A. (2001). The use of direct and indirect speech by 1½-to 4-year-olds. *Psychology of Language and Communication*, 5(1), 57–66.

Özyürek, A. (1996). How children talk about a conversation. *Journal of Child Language*, 23(3), 693–714. https://doi.org/10.1017/S0305000900009004

R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, 2(9), 437–442.

Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory & Cognition*, 2(1), 95–100.

Sevilla, Y., Maldonado, M., & Shalóm, D. E. (2014). Pupillary dynamics reveal computational cost in sentence planning. *Quarterly Journal of Experimental Psychology*, 66(6), 1041–1052. https://doi.org/10.1080/17470218.2014.9119

Spaccavento, S., Craca, A., Del Prete, M., Falcone, R., Colucci, A., Di Palma, A., & Loverre, A. (2014). Quality of life measurement and outcome in aphasia. *Neuropsychiatric Disease and Treatment*, 10, 27–37. https://doi.org/10.2147/NDT.S52357

Stites, M. C., Luke, S. G., & Christianson, K. (2013). The psychologist said quickly, "Dialogue descriptions modulate reading speed!". *Memory & Cognition*, 41(1), 137–151. https://doi.org/10.3758/s13421-012-0248-7

Ulatowska, H. K., Reyes, B. A., Santos, T. O., & Worle, C. (2011). Stroke narratives in aphasia: The role of reported speech. *Aphasiology*, 25(1), 93–105. https://doi.org/10.1080/02687031003714418

Wade, E., & Clark, H. H. (1993). Reproduction and demonstration in quotations. *Journal of Memory and Language*, 32(6), 805–819. https://doi.org/10.1006/jmla.1993.1040

Wright, P. (1969). Transformations and the understanding of sentences. *Language and Speech*, 12(3), 156–166. https://doi.org/10.1177/002383096901200302

Yao, B., Belin, P., & Scheepers, C. (2011). Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience*, 23(10), 3146–3152. https://doi.org/10.1162/jocn_a_00022

Yao, B., Belin, P., & Scheepers, C. (2012). Brain 'talks over' boring quotes: Top-down activation of voice-selective areas while listening to monotonous direct speech quotations. *NeuroImage*, 60(3), 1832–1842. https://doi.org/10.1016/j.neuroimage.2012.01.111