


ARTICLE

Discourse comprehension and referential processing: effects of contextual distance and semantic plausibility on presupposition processing

Susanne Dietrich , Verena C. Seibold and Bettina Rolke

Evolutionary Cognition, Department of Psychology, University of Tübingen, Tübingen, Germany

Corresponding author: Susanne Dietrich; Email: s.dietrich@uni-tuebingen.de

(Received 10 January 2024; Revised 27 July 2024; Accepted 30 July 2024)

Abstract

The present study aimed to investigate whether contextual factors influence how a reference is processed in discourse. We used intact and violated presuppositions (PSP), triggered by a definite or indefinite noun phrase, to monitor the reference process. In one sentence set, a contextual referent was explicitly mentioned close or far from the PSP-triggering noun phrase (*memory context*). In another sentence set, a referent was not explicitly mentioned in the context, but an inference to a referent was either plausible or implausible due to contextual semantic relations (*inference context*). Participants were asked to rate the coherence of the discourse after listening to it. Our results revealed a strong influence of the temporal distance of the contextual presentation of a referent. When the referent was far in the context (*memory context*), PSP violations were judged to be less severe than for close referents, suggesting that they are less clearly represented in memory. Furthermore, PSP violations seemed to play a subordinate role when the semantic context provided a basis for the plausible presence of a referent (*inference context*). Our results suggest that discourse comprehension involves referential processes whose importance may fade with distance in memory or may be obscured by semantic contextual content.

Keywords: bridging inference; coherence; discourse processing; plausibility; pragmatic processing; presupposition; spoken language comprehension

1. Introduction

Every thought expressed through language is interpreted within a specific context that determines the meaning of an utterance (e.g., Altmann & Steedman, 1988). Thus, to understand a discourse, it is necessary to establish a reasonable relationship between the utterance and the given context. This context can come from different sources, such as the discourse itself or the pragmatic circumstances in which the



discourse is uttered (Gundel et al., 1993). If necessary information is not mentioned in the context and a reference process cannot be carried out successfully, the information might be inferred from background knowledge (Stalnaker, 2002). The latter process is particularly appropriate when common knowledge between the discourse participants can be assumed and – in line with Grice’s maxims of optimized communication (Grice, 1975) – redundant information should be avoided.

1.1. Presuppositions

One way of referring to contextual information is by using presuppositions (PSPs). PSPs are triggered by specific words signaling that contextual information needs to be referred to. For example, the definite determiner as a modifier of an object noun phrase (NP) triggers the PSP that such an object must have been mentioned as existing in the context (existence PSP; Strawson, 1950). Furthermore, the indefinite determiner as a modifier of an object NP triggers the PSP that such an object has not previously existed in the context and can be regarded as new (novelty PSP, Anderson & Holcomb, 2005). In case of the definite determiner, the detection of a PSP trigger initiates a contextual search for a suitable referent, the so-called reference process (Spivey-Knowlton & Sedivy, 1995; Tiemann et al., 2011; Rolke et al., 2023). In case of the indefinite determiner, the detection of a PSP trigger might invoke the establishment of a new discourse protagonist. These discourse comprehension processes involve remembering relevant information that has been mentioned in the context or, if it has not been mentioned, inferring referential information by bridging inference (e.g., Burkhardt 2006, 2007). This means that discourse elements are related to each other, and sentences are semantically connected in a meaningful way to create a coherent discourse representation (Clark, 1975; Matsui, 2001).

Previous studies of PSP processing have shown that the detection and processing of PSP triggers within a discourse context is generally associated with longer reading times than the processing of non-PSP triggers (Tiemann et al., 2011), and that processing of a violated PSP takes longer than processing of an intact PSP (e.g., Domaneschi & Di Paolo, 2018; Schwarz, 2007). For instance, Domaneschi and Di Paolo (2018) conducted a self-paced reading study in which some context information was followed by a target sentence containing a change of state verb as a PSP trigger, such as *to give up* an action. The meaning of the change of state verbs directly encodes a precondition for the content of the context: If an action *is given up*, it must have existed before. In the intact PSP condition, the presupposed information was explicitly mentioned in the context

- (1) “Before her pregnancy G. *smoked* ten cigarettes per day. ... From the beginning she has *given up smoking*”

while in the violated PSP condition, the context did not satisfy the PSP of the target sentence, and thus the missing context must be inferred

- (2) “G. is at the third month of her first pregnancy. ... From the beginning she has *given up smoking*”

The authors observed longer reading times at the time point coinciding with the PSP trigger (“*given up*”) in the violated PSP condition than in the intact condition

(see also Altmann & Steedman, 1988; Singh et al., 2016). Furthermore it has been shown that the increase in reading time can occur quite early after the presentation of the PSP trigger (Kirsten et al., 2014; Rolke et al., 2023) and that the misuse of PSP triggers, that is, the use of the definite determiner in case of no existing referent or the use of the indefinite even if a referent had already been introduced in the discourse, reduces coherence judgments between the sentence containing the PSP trigger and the given discourse context (Dietrich et al., 2019; Rolke et al., 2023; Schneider et al., 2020). These results suggest that when reading times are prolonged by (i) a violated PSP or (ii) a PSP compared to a control without a PSP, some additional cognitive operations are required and are associated with increased cognitive load.

1.2. Cognitive processing of presuppositions

On the grounds of these observations, Rolke et al. (2023) proposed that the PSP trigger directly serves as a starting point within a sentence to initiate several cognitive processes: First, a reference process that is aimed at finding a referent for the PSP trigger in the established context through a memory search. Second, an evaluation process aimed at verifying whether the PSP is satisfied in the context. If the PSP can be verified in the course of the evaluation process, discourse comprehension can be completed quickly, resulting in a successful integration of the new content into the context (Van Berkum et al., 1999; Van Berkum et al., 2003) and the establishment of a coherent representation of the discourse (Garrod & Sanford, 1982). In contrast, if the PSP is violated, the evaluation process continues until the receiver decides that the presupposed information is missing and that the discourse is implausible. Third, as a repair mechanism in case of a violated presupposition, Rolke et al. (2023) suggested that an accommodation process (Beaver & Zeevat, 2007; von Stechow, 2008; Simons, 2003) might step in, in which missing information is inserted post hoc into the discourse context to save a sentence from being meaningless.

Besides the assumption of different processes being initiated by PSP triggers, several authors have argued that the processing of PSPs will directly affect cognitive load (Rolke et al., 2023; see also Altmann & Steedman, 1988). In general, cognitive load refers to the amount of working memory resources available to perform a given mental operation (see, e.g., Sweller et al., 1998). This concept is directly rooted in the assumption that working memory is a resource-limited system, so that only a limited amount of information can be processed or maintained at once (e.g., Baddeley & Hitch, 1974). With respect to PSP processing, Rolke et al. (2023) have proposed that reference and accommodation processes, in particular, increase cognitive load because these processes require contextual information to be maintained and (eventually) updated in working memory. Some empirical support for the idea that PSP processing is related to memory processes and, more specifically, to an increase in cognitive load comes from studies that measure brain activation or event-related potentials (ERP). For example, in an fMRI study, Dietrich et al. (2019; see also Dietrich et al., 2023) manipulated discourse coherence by using an existence PSP

(3) “Tina had *no swimming badge*. When Tina showed *the/a swimming badge*,”

or uniqueness PSP

(4) “Manuel saw ... a TV program about *a species of dolphin* He noticed that *a/the species of dolphin* was”

The authors reported increased brain activation in the inferior frontal gyrus, which is associated with working memory (e.g., Zhang et al., 2003), for PSPs falsely triggered by a definite or indefinite NP compared to intact PSPs. Moreover, in an ERP study, Anderson and Holcomb (2005; see also Kirsten et al. 2014) compared test sentences containing either a definite or an indefinite NP that referred either directly to a repeated critical NP or to a synonym of the critical NP in a preceding context. They reported an enhanced left anterior negativity (LAN) for the definite NP compared to the indefinite NP independent of word type (repetition vs. synonym), the latter being reflected in an N400 effect. As the LAN is assumed to mirror working memory demands (e.g., King & Kutas, 1995), the authors interpreted the enhanced LAN as an index of a more difficult referential assignment. Similar results were obtained in an ERP study by Burkhardt (2006). She varied the degree of availability of a referent for a definite NP within a target sentence, e.g.,

(5) “He said that *the conductor* was very impressive.”

She manipulated the context by (i) explicitly mentioning a referent in the context

(6) “Tobias visited *a conductor* in Berlin.”,

by (ii) inferring a referent.

(7) “Tobias visited *a concert* in Berlin.”,

or by (iii) presenting an unrelated context

(8) “Tobias talked to Nina.”

Burkhardt (2006) reported differences in ERPs (N400, P600) depending on context condition. In the case of an unrelated context (vs. a context with a given referent), the definite phrase in the target sentence evoked an enhanced N400, whereas inference of a referent (vs. a context with a given referent) evoked a reduced N400 followed by an enhanced P600. This pattern of results suggests that a given referent and a semantic matching facilitates discourse processing – as indicated by the N400 – but that integrating an inferred referent into a mental discourse model involves additional processing costs – as indicated by the P600. The author concluded that the vaguer the context and thus the more difficult the referencing process, the greater the cognitive load associated with processing the definite determiner. Overall, processing of PSPs has been assumed to induce cognitively demanding processes, which presumably involve reference processes, evaluation processes, and under given circumstances, inferential as well as accommodation processes to establish a meaningful discourse representation. Moreover, these processes have been theoretically linked to increased cognitive load in working memory (e.g., Altmann & Steedman, 1988; Rolke et al., 2023). Based on these considerations and on the assumption that working memory has a limited processing capacity, the aim of the present study was to investigate whether the way PSPs are resolved varies as a function of different memory demands. Specifically, we reasoned that the way a PSP is processed and/or

resolved depends critically on the availability of the presupposed contextual information and the amount of resources available for PSP processing.

1.3. *Discourse processing*

The idea that the availability of discourse information plays a key role in how a discourse is mentally represented has been advocated in theories on discourse processing (Kintsch & van Dijk, 1978, Kintsch, 1988; Zwaan et al., 1995; Zwaan & Radvansky, 1998), and has been supported in empirical studies on discourse understanding. For instance, Kintsch and van Dijk (1978) proposed that the discourse content is represented in a hierarchical network of propositions (see also Van Dijk & Kintsch, 1983) with a longer retention in an active state for important high-level propositions like, e.g., frequently occurring protagonists or relevant objects, than for less relevant low-level propositions like, e.g., subordinate facts. In line with this assumption, several authors (Cirilo, 1981; Lesgold et al., 1979) have shown that high-level propositions in a written story are recalled more often than low-level propositions (Cirilo, 1981). Furthermore, it has been shown that the number of connections an antecedent has with several other discourse entities influences retrieval accuracy (Trabasso & Sperry, 1985; Trabasso & van den Broek, 1985; O'Brien & Myers, 1987), and that the association strength of an inferential relationship modulates ERP components. For instance, Burkardt (2007) found that a more demanding inference (e.g., “found dead” combined with “the pistol”) was associated with ERPs of larger amplitude as compared to a more direct relation (e.g., “shot” combined with “the pistol”), again indicating that the mental representation of the discourse influences referential memory processes.

These results are compatible with the finding that with assumed larger distance within the discourse representation, a larger amount of memory resources is required for finding a referent (Warren, 2001) and for integrating it into the discourse structure (Gibson, 2000). Moreover, the occurrence of a protagonist in the discourse seems to influence the discourse representation. For instance, O'Brien (1987) has shown that when one of two antecedents on the same semantic hierarchical level (e.g., “doll” and “pie”) appeared in two temporarily separated positions in a text passage, the search time to retrieve the later antecedent was shorter than that for the early antecedent (O'Brien, 1987). This result is in line with the findings of Streb et al. (2004), who also varied the temporal distance between anaphor and antecedent in text passages by inserting either a large, middle, or small number of words in between anaphor and antecedent. The authors reported that comprehension times, as well as ERP amplitudes, increased with higher distance. Similar results were obtained by a study of Fiebach et al. (2002). In this study, participants processed German subject/object WH questions with a short or long gap between a WH filler (“who_(subject) / who_(object)”) in sentences such as

- (9) “T. asks ..., who_(subject) [on Tuesday] the_(object) doctor called [has].”
- (10) “T. asks ..., who_(object) [on Tuesday] the_(subject) doctor called [has].”
- (11) “T. asks ..., who_(subject) [on Tuesday afternoon after the accident] the_(object) doctor called [has].”
- (12) “T. asks ..., who_(object) [on Tuesday afternoon after the accident] the_(subject) doctor called [has].”

The authors showed that when processing the interdependent far (vs. near) sentence constituents, individuals with low working memory capacity showed an increased number of errors in sentence comprehension compared to individuals with high working memory capacity. In addition, the authors reported that the amplitude of the electrophysiological LAN component measured at the time of the WH filler increased with increasing distance from the filler and was modulated by individual working memory capacity. Taken together, theoretical considerations and experimental findings suggest that the way discourse information is represented depends on the semantic distance or “hierarchical level” of propositions in a representational context (Burkhardt, 2007; Cirilo, 1981; Warren, 2001) as well as on the temporal distance of the presentation of the propositions (O’Brien, 1987; Streb et al., 2004).

1.4. Present study

Although the studies described above emphasize the involvement of working memory in discourse comprehension, previous research on the processing of PSPs has focused primarily on syntactic, semantic or theoretical aspects. In the present study, we adopt a more comprehensive approach, examining contextual factors on a more global level. Specifically, we investigated the interaction between working memory demands and PSP processing. To this end, we varied (i) the temporal distance between an NP trigger and its referent, and (ii) the semantic-pragmatic relations between an NP trigger and the context, and measured their influence on PSP processing, using the definite and indefinite determiner as PSP triggers. We generated spoken discourses consisting of several sentences establishing a context followed by a test sentence. The test sentence contained either a definite or indefinite NP, with the former requiring the existence of a referent and the latter introducing a novel discourse entity. To monitor PSP reference processes, we created a context for which the PSP in the test sentence was either intact or violated. Participants were asked to judge the coherence of the global discourse, including the test sentence combined with the discourse context. Whereas methods such as measuring reading times (e.g., Rolke et al., 2023) and event-related potentials (e.g., Kirsten et al., 2014) or monitoring eye movements (e.g., Schwarz & Tiemann, 2012) allow tracking incremental discourse reading and estimating cognitive load during understanding, coherence judgments provide insight into participants’ overall (subjective) comprehension and their interpretation of a discourse, making them a valuable measure for assessing the impact of contextual factors on PSP processing. In sentence set 1 (*memory context*), a referent for the definite determiner in the test sentence was explicitly mentioned as being existent (13) or inexistent (14), thereby fulfilling or violating the existence PSP of the definite determiner, respectively.

- (13) “... T. has a suitcase. When T. packs the suitcase,” (intact PSP, far distance)
- (14) “... T. has no suitcase. When T. packs the suitcase,” (violated PSP, far distance)

To manipulate temporal distance, we mentioned the referent in the context with a close or with a far distance to the definite NP trigger in the test sentence.

- (15) "... T. has a suitcase. When T. packs the suitcase," (intact PSP, close distance)

Because greater distance is associated with higher processing demands (see, e.g., O'Brien, 1987; Streb et al., 2004), and these demands should affect the amount of resources available for PSP processing, we expect a modulation of the PSP consistency effect by distance, with coherence judgments becoming less discrepant between violated and intact PSP conditions at greater distances. In sentence set 2 (*inference context*), in half of the test sentences, the definite determiner required an existing referent, whereas, in the other half, the indefinite was compatible with a situation in which a novel discourse entity was introduced. As the test sentence item was never mentioned in the context, the existence PSP of the definite was violated, whereas the novelty PSP of the indefinite was intact.

- (16) "... S. has a sofa. ... When S. moves a shelf," (intact PSP, plausible context)
- (17) "... S. has a sofa. ... When S. moves the shelf," (violated PSP, plausible context)

To measure the influence of semantic plausibility on PSP referencing, we manipulated the plausibility with which the test sentence item fits into the semantic-pragmatic content of the context, i.e., whether or not an item could be easily integrated via inferential bridging.

- (18) "S. goes hiking. ... When S. moves a shelf," (intact PSP, implausible context)

According to Burkhardt (2006), we assume that semantic bridging imposes high cognitive load and thus interferes with PSP referencing, thereby modulating the PSP effect. The way how PSP processing is influenced, however, might depend on the importance of semantic appropriateness compared to the consistency of PSP (cf. Cirilo, 1981). If participants perceive the PSP fit as primarily relevant, a consistency effect should occur in both semantically appropriate and semantically inappropriate contexts; if participants use the semantic fit as an interpretive basis for their discourse representation, the PSP consistency effect should be smaller in the inappropriate semantic case than in the appropriate case.

2. Experiment 1

2.1. Method

2.1.1. Participants

Twenty-eight volunteers participated in the experiment, recruited either via a recruiting platform of the University of Tübingen or via local social networks. Twenty-five participants were right-handed and three left-handed (8 male, 20 female), and they had a mean age of 23.8 years ($SD = 4.6$ years). According to their self-report, all participants were native speakers of German. The experiment was run in accordance with the World Medical Association's Declaration of Helsinki 2008, and all participants provided written informed consent prior to the experiment.

Participants received course credit or financial compensation in the form of a book voucher (€15) for their participation.

2.1.2. Apparatus

The experiment was programmed in PsychoPy3 (version 2020.1.2, <https://www.psychopy.org>; Peirce et al., 2019) and was conducted online via the platform Pavlovia (<https://pavlovia.org/>). Participants accessed the experiment via a link and used their own computer, and they provided demographic data and informed consent via e-mail. The behavioral data were initially stored anonymously on a server in the UK. Demographic data and informed consent were provided via email.

2.1.3. Material

For both sentence sets (set 1, *memory context*; set 2, *inference context*), we created short discourse contexts that consisted of five spoken German sentences followed by a test sentence (Tab. 1). The first context sentence (1) introduced an everyday scenario such as hiking or tidying up. The second context sentence (2) contained a first noun phrase (NP1) which established either the existence of an object by means of an indefinite determiner (*a* – in German ‘ein’) or the non-existence of an object by negating the object’s presence (*no* – in German ‘kein’). The third (3) and fourth (4) context sentences were used to describe the contextual scenario in more detail. The fifth context sentence (5) contained a second noun phrase (NP2). This NP2 (i) established the existence of the second object if NP1 was negated or (ii) negated an object’s existence if NP1 was introduced as an existing object. The sixth sentence served as a test sentence (6).

In sentence set 1 (*memory context*) (see Table 1, left column), the test sentence described an action of a protagonist with the object mentioned in the context by either NP1 (in sentence 2) or by NP2 (in sentence 5). To monitor PSP reference processes, the PSP of the NP was either intact or violated (PSP CONSISTENCY). In more detail, as the PSP trigger in the test sentence of this set was always a definite NP, the PSP was intact when the definite NP in the test sentence referred to the existing NP in the context (intact existing PSP), and it was violated when the definite NP referred to the negated NP in the context (violated existing PSP). Memory requirements were varied by the distance between the PSP trigger in the test sentence and the position of the contextual NP (DISTANCE), which was presented either in context sentence 2 (far) or in context sentence 5 (close). Altogether, the two factors were combined for each context sentence set, resulting in four conditions for each context: There were two conditions with an intact existence PSP, i.e., the establishment of an object by the indefinite NP in the context at far and close distance combined with an indefinite NP in the test sentence, and two further conditions violating the existence PSP by negating the existence of an object at far and close distance also combined with an indefinite NP in the test sentence. These four conditions were realized for the two different discourse NP (NP1, NP2), resulting in eight trials per scenario. As we established eight different scenarios, 64 trials for sentence set 1 were presented.

In sentence set 2 (*inference context*) (see Table 1, right column), the object mentioned in the test sentence never occurred in the context scenario. Here, we manipulated PSP consistency by using either an indefinite or a definite NP in the test sentence. In this situation, the novelty PSP of the indefinite NP was intact as the

Table 1. Examples of sentence material originally presented in German. In sentence set 1 (*memory context*, left column), a referent was explicitly mentioned in the context by an indefinite NP (intact PSP) or by a negated NP (violated PSP) and was presented at close or far distances relative to the PSP trigger in the test sentence. In sentence set 2 (*inference context*, right column), the context was either plausible or implausible for a potential referent that was not explicitly mentioned. The PSP trigger in the test sentence could be either an indefinite NP (intact PSP) or a definite NP (violated PSP).

Sentence set 1: <i>Memory context</i>	Sentence set 2: <i>Inference context</i>
<p>Condition: intact/violated – far</p> <p>Context sentences:</p> <ol style="list-style-type: none"> (1) Tina is going to France. (2) Tina has a/no suitcase. (3) The travel is short. (4) She wants to relax for a few days in the countryside. (5) Tina has no/a duffel bag. <p>Test sentence:</p> <ol style="list-style-type: none"> (6) When Tina packs the suitcase, she notices a dark cloud. 	<p>Condition: intact/violated – plausible</p> <p>Context sentences:</p> <ol style="list-style-type: none"> (1) Simon rearranges the living room. (2) Simon has a sofa. (3) He puts his things in different places. (4) He does not like the arrangement anymore. (5) Simon has no TV. <p>Test sentence:</p> <ol style="list-style-type: none"> (6) When Simon moves a/the shelf, he suffers a painful strain.
<p>Condition: intact/violated – close</p> <p>Context sentences:</p> <ol style="list-style-type: none"> (1) Tina is going to France. (2) Tina has no/a duffel bag. (3) The travel is short. (4) She wants to relax for a few days in the countryside. (5) Tina has a/no suitcase. <p>Test sentence:</p> <ol style="list-style-type: none"> (6) When Tina packs the suitcase, she notices a dark cloud. 	<p>Condition: intact/violated – implausible</p> <p>Context sentences:</p> <ol style="list-style-type: none"> (1) Simon goes hiking in the mountains. (2) Simon has a tent. (3) He looks forward to the trip for a long time. (4) He will probably be on route for some days. (5) Simon has no rope. <p>Test sentence:</p> <ol style="list-style-type: none"> (6) When Simon moves a/the shelf, he suffers a painful strain.

object had not been mentioned, and a new object was introduced (intact novelty PSP). In contrast, the existence PSP of the definite was violated when the definite NP referred to a previously unmentioned object (violated existence PSP, PSP CONSISTENCY). Memory requirements in this set were modulated by establishing either a plausible or an implausible semantic context for the object mentioned in the test sentence (PLAUSIBILITY). For this sentence set 2 (*inference context*), we built 16 new scenarios containing a test sentence with either a semantically plausible or implausible object with respect to the context. The combination of context sentences with the plausible test sentence object of another scenario from sentence set 2 served as implausible discourses. Each test sentence object in this sentence set occurred four times: with a definite determiner, with an indefinite determiner, and in a plausible and implausible semantic context, resulting in 64 trials for the 16 different scenarios. In addition, eight practice trials were constructed that contained different discourse scenarios than those in the experimental trials, but the same discourse structure as in the sentence set 1 (*memory context*) and set 2 (*inference context*).

Each discourse, including the test sentence, was spoken on tape condition as one unit by a female actor with consistent prosodic modulation to maintain a natural flow of speech without discontinuity effects.

2.1.4. Procedure

Participants were instructed to complete the experiment on a computer. They were asked to use the computer's built-in speakers and to sit in front of the computer. After a short on-screen briefing on the experimental procedure, eight practice trials similar to the experimental ones were presented to familiarize participants with the experimental procedure. The time course of a single trial is shown in Figure 1.

Participants were asked to look at a red fixation cross in the center of the screen while listening attentively to the discourse consisting of five context sentences and one test sentence. 2,500 ms after the end of the discourse, the fixation cross turned green, and participants rated the coherence of the test sentence relative to the context sentences ("How well does the test sentence match the context?") on a Likert scale from 1 to 4 (1 = poor, 2 = fairly poor, 3 = fairly good, 4 = good). It was made clear that "coherence" refers to the logical and coherent connection of information between the discourse sentences. To ensure that the participants understood the concept of coherence, they were given examples of coherent and incoherent discourses. Participants answered by pressing keys on their computer keyboard with their index and middle fingers of the two hands: Q key with left middle finger = good (4), W key with left index finger = fairly good (3), O key with right index finger = fairly poor (2), P key with right middle finger = poor (1). For half of the participants, the left/right assignment to the rating scale was reversed. If no response was made within 6,500 ms, a prompt ("Please press a key") appeared on the screen, which disappeared when one of the response keys was pressed. Only when a response was made, the trial continued, the green fixation cross disappeared, and a new trial began after 2,000 ms.

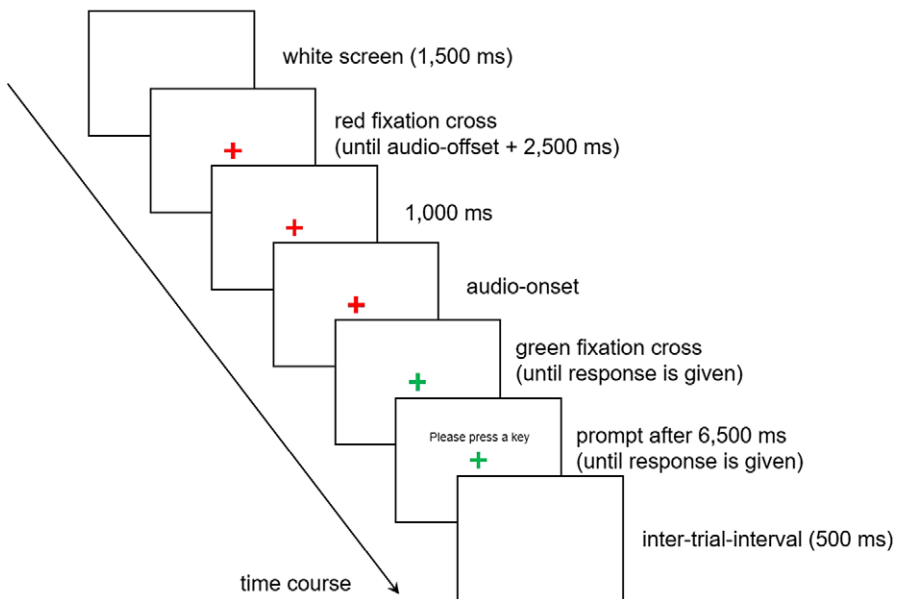


Figure 1. Illustration of an exemplary trial. A red fixation cross indicated the start of a trial and was visible for the entire duration of the audio presentation. After a further 2,500 ms, the red fixation cross was replaced by a green fixation cross, signaling participants to judge the coherence of the discourse. The green fixation cross disappeared when the response was given and a new trial began after 2,000 ms. If no response was made within 6,500 ms, a prompt ("Please press a key") appeared on the screen, which disappeared when one of the response keys was pressed.

The experiment lasted approximately 70 minutes. The total sentence material of 128 trials was divided into eight blocks of 16 discourse scenarios each, to allow participants to take a short break after each block. For half of the participants, the first four blocks contained only trials of sentence set 1 (*memory context*), and the following four blocks contained only trials of sentence set 2 (*inference context*) and this order was reversed for the other half of participants. Within a block, trials were presented in a pseudo-random order that assured that each condition was presented four times and that all scenarios were different from each other.

2.1.5. Data analyses

Statistical analyses of the coherence judgment data were performed using R 4.4.1 (R Core Team 2024). We performed separate statistical analyses for sentence set 1 (*memory context*) and set 2 (*inference context*). Coherence judgments were analyzed by means of cumulative link mixed models (CLMMs) with a logistic linking function – a method appropriate for repeated measures data resulting from a Likert scale (Stroup, 2012, Taylor et al., 2023). We used the *clmm* ordinal and *emmeans* package in R as suggested by Christensen (2023). In addition to the estimated marginal means (EMM) and standard error (SE), we also reported the original means (*M*) and standard deviation (*SD*) without plotting them to illustrate the extent to which participants used the extreme ends of the scale. Our fixed effects for sentence set 1 (*memory context*) were DISTANCE and PSP CONSISTENCY, with an interaction between them. Our fixed effects for sentence set 2 (*inference context*) were PLAUSIBILITY and PSP CONSISTENCY, with an interaction between them. Furthermore, for both sentence sets, we included both participants and discourse scenarios (items) as random effects because we expected differences in the extent to which they are affected by our experimental manipulations. To derive an appropriate random effects structure with comparable complexity across sentence set, we took the following approach: We fitted the CLMMs following Brown (2021), so that the models did not show convergence warnings or singular fits, both of which indicate problems with parameter estimation. This involved fitting the most complex model first and gradually removing terms until convergence was reached. Accordingly, for sentence set 1, we started with fitting a CLMM with (i) correlated by-participant random intercepts and slopes for DISTANCE and PSP CONSISTENCY and (ii) correlated by-item random intercepts without slopes. For sentence set 2, we started with fitting a CLMM with (i) correlated by-participants random intercepts and slopes for PLAUSIBILITY and PSP CONSISTENCY, and (ii) correlated by-item random intercepts without slopes. For sentence set 1 (*memory context*), a model with interaction between the by-participant random slopes for the two factors showed a better fit (vs. without interaction).¹ For sentence set 2 (*inference context*), however, a model without interaction between the by-participant random slopes for the two factors is sufficient to describe the data.² To make the models comparable, we chose the model with the more parsimonious random effects structure, that is, without interaction

¹Results of the likelihood ratio test to compare the goodness of fit of the model with and without the interaction between the by-participant random slopes for the two factors: In Experiment 1, sentence set 1 (*memory context*) the comparison was significant ($\chi^2(17) = 10.1, p < .05$), suggesting a better fit of the model with the interaction.

²In Experiment 1, sentence set 2 (*inference context*) the comparison was not significant ($\chi^2(17) = 2.2, p = .693$), suggesting no difference in fitting the models.

between the by-participant random slopes for the two factors for both sentence sets. Removing the interaction for set 1 (*memory context*) makes no difference to the fixed effects. In case of significant interaction effects, we conducted post-hoc tests based on contrasts using the Wald-z-statistic, for which we report Bonferroni corrected p -values for four tests each.

2.2. Results

Figure 2 shows the estimated marginal means (EMM) and standard error (SE) for the coherence judgments for sentence set 1 (left panel) and set 2 (right panel). In sentence set 1 (*memory context*), intact PSPs were judged to be more coherent ($EMM = 2.66$, $SE = 0.28$) than violated PSPs ($EMM = -2.36$, $SE = 0.31$), $\beta = -5.81$, $SE = 0.58$, $z = -10.09$, $p < .001$ (PSP CONSISTENCY). Moreover, the effect of DISTANCE was not significant, $\beta = -0.27$, $SE = 0.30$, $z = -0.90$, $p = .371$ (close distance: $EMM = -0.1150$, $SE = 0.17$; far distance: $EMM = 0.41$, $SE = 0.16$). The interaction PSP CONSISTENCY \times DISTANCE, $\beta = 1.57$, $SE = 0.32$, $z = 4.88$, $p < .001$, however, shows that distance modulated the consistency effect. Violated PSPs were judged to be less incoherent if they had a far referent ($EMM = -1.71$, $SE = 0.32$) than if they had a close referent ($EMM = -3.01$, $SE = 0.36$), $\beta = -5.81$, $SE = 0.58$, $z = -10.09$, $p_{Bonf} < .001$, whereas for intact PSPs there was no difference between close referents ($EMM = 2.80$, $SE = 0.31$) and far referents ($EMM = 2.53$, $SE = 0.32$), $\beta = -0.27$, $SE = 0.30$, $z = -0.90$, $p_{Bonf} = .371$. The mean (M) and standard deviation (SD) of the original Likert scale coherence ratings (1 = poor and 4 = good) for sentence set 1 were as follows: intact PSPs with close distance, $M = 3.62$, $SD = 0.49$; intact PSPs with far distance, $M = 3.56$, $SD = 0.50$;

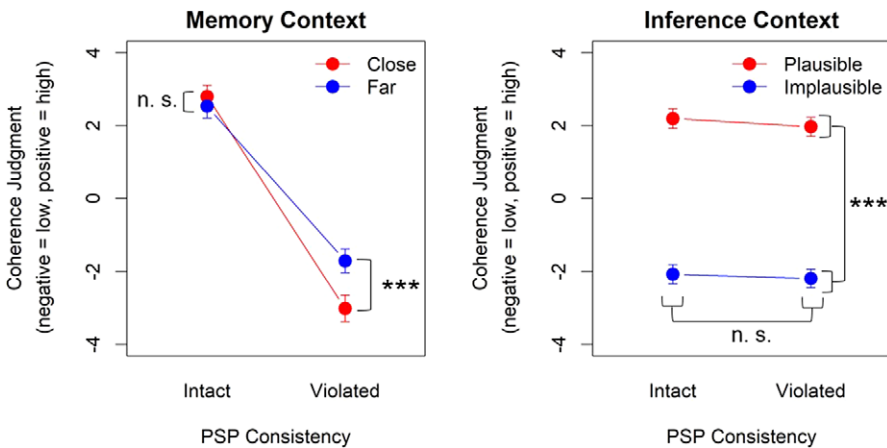


Figure 2. Estimated marginal means of coherence judgments as a function of PSP CONSISTENCY (intact vs. violated). In sentence set 1 (*memory context*, left panel), the reference information was presented at close or far distance (DISTANCE), whereas in sentence set 2 (*inference context*, right panel), the context was either plausible or implausible for a potential referent (PLAUSIBILITY). Error bars represent the standard error based on the variance of the residuals and the structure of the model. It takes into account the uncertainty in the estimates resulting from the data and the model. Asterisks indicate threshold significance *** $< .001$, n. s. = not significant for (i) four post hoc tests (intact PSP: close vs. far; violated PSP: close vs. far; close distance: intact vs. violated; far distance: intact vs. violated) using the Wald-z-statistic (set 1, left panel) and (ii) the two main effects for PSP CONSISTENCY and PLAUSIBILITY (set 2, right panel).

violated PSPs with close distance, $M = 1.37$, $SD = 0.51$; violated PSPs with far distance, $M = 1.80$, $SD = 0.60$.

In sentence set 2 (*inference context*), there was no significant effect of PSP CONSISTENCY, $\beta = -0.17$, $SE = 0.19$, $z = -0.89$, $p = .372$ (intact PSPs: $EMM = 0.06$, $SE = 0.19$; violated PSPs: $EMM = -0.11$, $SE = 0.18$). Moreover, plausible relationships ($EMM = 2.08$, $SE = 0.23$) were in general judged to be more coherent than implausible ones ($EMM = -2.13$, $SE = 0.22$), $\beta = -4.21$, $SE = 0.31$, $z = -13.54$, $p < .001$ (PLAUSIBILITY). There was no significant interaction between PSP CONSISTENCY and PLAUSIBILITY, $\beta = 0.11$, $SE = 0.37$, $z = 0.29$, $p = .773$. The mean (M) and standard deviation (SD) of the original Likert scale coherence ratings (1 = poor and 4 = good) for sentence set 2 were as follows: intact PSPs with plausible relationships, $M = 3.35$, $SD = 0.29$; intact PSPs with implausible relationships, $M = 1.68$, $SD = 0.28$; violated PSPs with plausible relationships, $M = 3.27$, $SD = 0.37$; violated PSPs with implausible relationships, $M = 1.64$, $SD = 0.33$.

2.3. Discussion

In this experiment, we investigated whether PSP reference processing, which is important for establishing discourse coherence, is affected by different contextual situations. For sentence set 1 (*memory context*), we obtained an effect of PSP CONSISTENCY indicating that participants noticed the violation of the PSP entailed by the definite NP. The effect of DISTANCE was not significant. The main result for sentence set 1 is, however, that distance modulated the consistency effect. Specifically, for violated PSPs, coherence judgments were higher for far distances than for close ones. This result suggests that when participants have a vaguer representation of the referents because they occurred in an earlier part of the discourse, the referencing processes result in a fuzzier mental model of the discourse. We come back to this point in the General Discussion.

The effect of discourse plausibility in sentence set 2 (*inference context*) shows that the semantic relations between the context sentences are crucial for coherence judgments when a potential referent is not explicitly mentioned. We assume that through bridging inference, participants were able to create a mentally coherent representation when the context information was plausible, whereas they failed to do so when the context information was implausible. The result that PSP consistency did not influence the coherence judgments supports the idea that semantic-pragmatic plausibility overrules PSP processing and that participants no longer cared about violated PSP. This finding is surprising as we would have expected that, at least in the plausible context, PSP consistency should play a role (as it did in sentence set 1). One possible explanation for why we did not find any indication of PSP violation within the ratings for sentence set 2 might be that we presented the two different sets block-wise. Although the discourse structure was the same in the two sentence sets, because of the block-wise presentation, participants knew in advance that in sentence set 2, no possible referent occurred in the context. For this reason, the PSP triggering function of the definite NP might have decreased, and consequently, the violation of PSP consistency might have been entirely ignored during the coherence task. To investigate this possibility, we intermixed sentences from the different sets in Experiment 2.

3. Experiment 2

In this experiment, we tested the hypothesis that *a priori* knowledge about discourse structure had prevented processing of PSP consistency in sentence set 2. To this end, we intermixed the two sentence sets so that participants could not foresee which type of sentence set was presented.

3.1. Method

3.1.1. Participants

Twenty-eight participants were recruited in the same way as in Experiment 1. The sample consisted of 24 right-handed and four left-handed participants (7 male, 21 female) with a mean age of 22.9 years ($SD = 5.6$ years). According to their self-report, all participants were native speakers of German. Again, all participants provided written informed consent prior to the experiment in accordance with the World Medical Association's Declaration of Helsinki 2008. They received course credit or financial compensation in the form of a book voucher (€15) for their participation.

3.1.2. Material, procedure, and data analyses

The identical sentence material from Experiment 1 was used, but we now intermixed the two sentence sets. The entire procedure was analogous to Experiment 1, except that there were eight trials of sentence set 1 (*memory context*) and eight trials of sentence set 2 (*inference context*) in each of the eight blocks. Within a block, trials were presented in a pseudo-random order to ensure that each condition was presented twice and that all scenarios were different from each other. The statistical analysis was conducted analogously to Experiment 1. To make the models comparable, we chose the model with the more parsimonious random effects structure, that is, without interaction between the by-participant random slopes for the two factors for both sentence set.³

3.2. Results

Figure 3 shows the estimated marginal means (EMMs) for coherence judgments for sentence set 1 (left panel) and set 2 (right panel). In sentence set 1 (*memory context*), intact PSPs were judged to be more coherent ($EMM = 2.86$, $SE = 0.28$) than violated PSPs ($EMM = -2.29$, $SE = 0.35$), $\beta = -6.07$, $SE = 0.57$, $z = -10.56$, $p < .001$ (PSP CONSISTENCY). Moreover, close distances ($EMM = 0.15$, $SE = 0.19$) were, in general, judged to be less coherent than far ones ($EMM = 0.42$, $SE = 0.17$), $\beta = -0.65$, $SE = 0.24$, $z = -2.74$, $p = .006$ (DISTANCE). The interaction PSP CONSISTENCY \times DISTANCE, $\beta = 1.83$, $SE = 0.31$, $z = 5.84$, $p < .001$, however, shows that distance modulated the

³Results of the likelihood ratio test to compare the goodness of fit of the model with and without the interaction between the by-participant random slopes for the two factors: In Experiment 2, sentence set 1 (*memory context*) the comparison was significant ($X^2(17) = 28.8$, $p < .001$), suggesting a better fit of the model with the interaction. In Experiment 2, sentence set 2 (*inference context*) the comparison was not significant ($X^2(17) = 4.1$, $p = .392$), suggesting no difference in fitting the models. Removing the interaction for set 1 (*memory context*) makes no differences to the fixed effects.

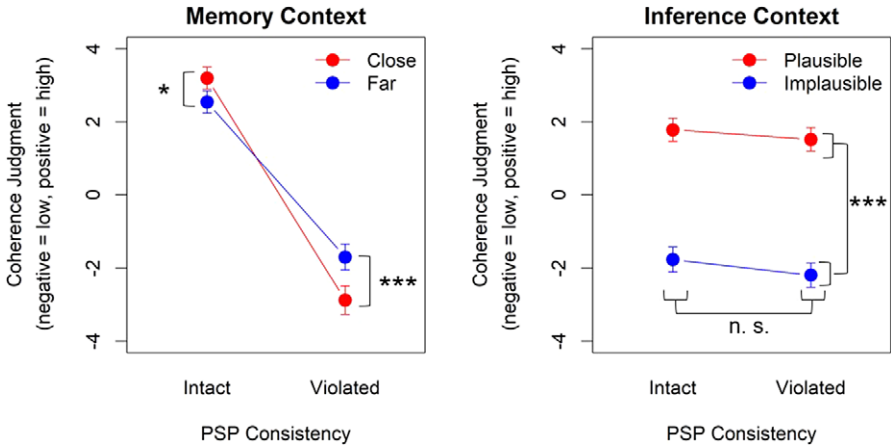


Figure 3. Estimated marginal means of coherence judgments as a function of PSP CONSISTENCY (intact vs. violated). In sentence set 1 (*memory context*, left panel), the reference information was presented at close or far distance (DISTANCE), whereas in sentence set 2 (*inference context*, right panel), the context was either plausible or implausible for a potential referent (PLAUSIBILITY). Error bars represent the standard error based on the variance of the residuals and the structure of the model. It takes into account the uncertainty in the estimates resulting from the data and the model. Asterisks indicate significance of threshold $* < .05$, $*** < .001$, n. s. = not significant for (i) four post hoc tests (intact PSP: close vs. far; violated PSP: close vs. far; close distance: intact vs. violated; far distance: intact vs. violated) using the Wald-z-statistic (set 1, left panel) and (ii) the two main effects for PSP CONSISTENCY and PLAUSIBILITY (set 2, right panel).

consistency effect. Violated PSPs were judged to be less incoherent if they had a far referent ($EMM = -1.70$, $SE = 0.35$) than when they had a close referent ($EMM = -2.88$, $SE = 0.39$), $\beta = -1.18$, $SE = 0.23$, $z = -5.19$, $p_{Bonf.} < .001$. A reversed pattern was observed for the intact PSPs, which were judged to be less coherent if they had a far referent ($EMM = 2.54$, $SE = 0.30$) than when they had a close referent ($EMM = 3.19$, $SE = 0.30$), $\beta = 0.65$, $SE = 0.24$, $z = 2.74$, $p_{Bonf.} = .025$. The mean (M) and standard deviation (SD) of the original Likert scale coherence ratings (1 = poor and 4 = good) for sentence set 1 were as follows: intact PSPs with close distance, $M = 3.69$, $SD = 0.55$; intact PSPs with far distance, $M = 3.53$, $SD = 0.50$; violated PSPs with close distance, $M = 1.45$, $SD = 0.59$; violated PSPs with far distance, $M = 1.86$, $SD = 0.53$.

In sentence set 2 (*inference context*), there was no significant effect of PSP CONSISTENCY, $\beta = -0.26$, $SE = 0.29$, $z = -0.90$, $p = .370$ (intact PSPs: $EMM = 0.01$, $SE = 0.22$; violated PSPs: $EMM = -0.33$, $SE = 0.22$), mirroring the absence of a significant PSP CONSISTENCY effect in Experiment 1. As in Experiment 1, plausible relationships ($EMM = 1.65$, $SE = 0.28$) were judged to be more coherent than implausible ones ($EMM = -1.98$, $SE = 0.30$), $\beta = -3.55$, $SE = 0.49$, $z = -7.29$, $p < .001$ (PLAUSIBILITY). However, there was no interaction between PSP CONSISTENCY and PLAUSIBILITY, $\beta = -0.17$, $SE = 0.41$, $z = -0.40$, $p = .689$. The mean (M) and standard deviation (SD) of the original Likert scale coherence ratings (1 = poor and 4 = good) for sentence set 2 were as follows: intact PSPs with plausible relationships, $M = 3.16$, $SD = 0.44$; intact PSPs with implausible relationships, $M = 1.84$, $SD = 0.49$; violated PSPs with plausible relationships, $M = 3.07$, $SD = 0.45$; violated PSPs with implausible relationships, $M = 1.72$, $SD = 0.42$.

3.3. Discussion

In this experiment, we investigated whether the absence of an effect of PSP consistency in sentence set 2 had been due to the block-wise presentation in Experiment 1, which might have led participants to emphasize plausibility for this sentence set and thus might have masked PSP consistency effects. The results of Experiment 2 showed a replication of the results of Experiment 1 for sentence set 1, namely an effect of PSP CONSISTENCY and a modulation of PSP CONSISTENCY by DISTANCE. The effect for DISTANCE shows that in this mixed design, the distance between the test item and the potential referent in the discourse context became more prominent as compared to Experiment 1. For sentence set 2, we obtained a strong effect of PLAUSIBILITY. Contrary to our hypothesis, we found no clear evidence for PSP consistency processing in this sentence set. It thus seems that participants weighted the plausibility of the semantic context much more strongly than PSP consistency, even though in this unpredictable mixed experimental context. Overall, this experiment shows that the global embedding of sentence material can evoke specific processing strategies that can influence the focus of different discourse processing processes.

4. General discussion

In this study, we investigated whether the way in which PSPs are processed varies according to different memory demands and depends on the availability of the presupposed contextual information and the amount of resources available for PSP processing. We modulated contextual factors by (i) referring to a referent mentioned early or late in the context, existing or non-existing, in sentence set 1 (*memory context*) or (ii) inferring a referent in sentence set 2 that was more or less plausible due to semantic-pragmatic relations between an NP as PSP trigger and the context. We induced PSP processing in this sentence set by using the definite determiner, whose triggered reference processes should fail in any case, and by using the indefinite determiner, which should allow semantic-pragmatic inference depending on the context content (*inference context*). The two-sentence sets were presented in separate blocks in Experiment 1, whereas they were presented intermixed in Experiment 2 to avoid anticipation of the respective context. In sentence set 1 (*memory context*), the temporal distance of a potential context referent to the definite determiner in the test sentence influenced coherence judgments. Specifically, we observed that violated PSPs were judged to be more coherent when the potential referent was mentioned early in the context compared to a referent mentioned late in the context (Experiments 1 and 2). In Experiment 2, distance had an effect on both violated and intact PSPs. In sentence set 2 (*inference context*), no PSP consistency effect was observed, whereas the pragmatic-semantic fit of the test sentence NP exerted a very strong influence on coherence judgments independent of the presentation mode of sentence sets (blocked vs. intermixed).

The finding that coherence judgments depend on the temporal distance of a potential referent in the discourse (*memory context*) is noteworthy because it suggests that the reference process triggered by the definite determiner was modulated by the different temporal mention of the referent in the discourse. This finding may be explained in terms of increased memory demands for retrieving temporally distant information (e.g., Streb et al., 2004), which may interfere with PSP reference processes. For example, some authors report that memory demands for referent

integration increase when referring to relevant information mentioned early in the discourse (Gibson, 2000; Warren, 2001). Since increased cognitive demands occur generally when processing PSP vs. non-PSP sentence structures (e.g., Tiemann et al., 2011) and when processing violated vs. intact PSP (e.g., Rolke et al., 2023; Schneider et al., 2020), an interaction could take place here in that more difficult memory retrieval leaves fewer resources for the PSP reference processes (e.g., Fiebach et al., 2002). This, in turn, could lead to the observed outcome of less clear recall of whether a potential referent was mentioned as negated and non-existent or as existent, and thus less clear classification of distant referents.

Interestingly, the results showed a more robust distance effect for negated referents (violated PSP consistency) than for non-negated referents (intact PSP consistency) in both Experiments. In our sentence material, the existence PSP was violated by negating the referential object in the context, whereas the existence PSP was intact by contextually introducing an indefinite NP as a referent (e.g., "...Tina has *no/a suitcase*. When Tina packs *the suitcase*, ..."). As the violation of PSP consistency was realized by a negation in our study, processing negation might have contributed to our results. Dudschig et al. (2021a) suggest that more cognitive resources are required for processing a negation than for processing the non-negated opposition. They demonstrated this experimentally by testing the use of negation within directional instructions (e.g., "not left", "now left", "not right", "now right") in children, and the response to negative instructions was slower than to affirmative instructions (Dudschig et al., 2021b). The authors concluded that semantic understanding of negation requires inhibitory control processes in which the affirmation must first be represented and then inhibited. Thus, the retrieval of negated information, in addition to the already cognitively demanding processing of temporal distance, could strain the capacity of working memory, making its processing resources scarce for the retrieval and integration of the referent and explaining unclear coherence judgments, especially in the case of distant and violated PSP consistency.

One further possibility to explain the influence of the temporal distance of the referent from its trigger on the current coherence judgment would be a different way of representing the discourse elements in far and close contexts. Empirical evidence shows that in progressive reading or listening, the lexical or syntactic information mentioned in the text remains in memory only until the meaning content has been semantically interpreted and a long-term representation has been created (Jarvella, 1971; Sachs, 1967). Van Dijk and Kintsch (1983) summarized in their tripartite model that lexical input provides the basis for initial surface processing, which is only briefly retained in memory. In a subsequent stage of processing, propositions and concepts are formed, which contribute to a basic semantic understanding of the discourse content and are retained in memory. Finally, situational and individual experiences are linked to the discourse content (see also Sparks & Rapp, 2010 for an overview). It is possible that concrete discourse information, e.g. whether a mentioned referent was present in the context or not, gets lost in the increasingly abstract and semantic discourse representation. Since this abstraction is presumably more advanced for distant information than for briefly mentioned content, the probability of losing concrete information is higher for distant content than for close content. Such a distant and more abstract representation would then contain less details when being accessed later, and this might impair the ability to judge discourse coherence. In addition to a possible change in representation mode, some authors suggest that contextual information that is mentioned late (e.g., O'Brien, 1987), perceived as

important (e.g., Cirilo, 1981), or highly related (e.g., Burkhardt, 2007) is highlighted in its activation level, while the activation of other contextual information is suppressed (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983). To minimize memory load, early presented representations may be deactivated once a higher-level referential representation has been formed (Just & Carpenter, 1992). In the current sentence material, a referent mentioned early in the context would be lexically processed but would receive a lower level of activation, making it more difficult to recall and thus more difficult to judge coherence. Activation differences between distant and close possible referents of the discourse representation also provide a possible basis for explaining the more moderate consistency effect at longer discourse distances.

The results obtained for sentence set 2 (*inference context*) also provide some interesting new insights into discourse processing. The results of Experiment 1 (blocked presentation of the two sentence sets) and Experiment 2 (mixed presentation of the sentence sets) show that in the case of apparent semantic fit vs. non-fit, the semantic context is considered essential for coherence judgments, whereas PSP consistency seems to play no or only a very minor role. It was surprising to us that PSP fit had no effect because, according to previous findings (Rolke et al., 2023), we would have expected the incorrect use of the definite determiner without a referent to be a strong violation of communication rules, which should lead to comprehension difficulties and thus to a judgment of lower coherence. However, Ferstl and von Cramon (2001) have already shown that general discourse comprehension or global discourse coherence processing was not impaired when cohesive ties, such as the pronoun *ihr* ('her'), falsely suggested a relationship between sentences that did not exist. In sentence set 2, we can assume that an absent consideration of PSP consistency does not jeopardize global discourse comprehension and that PSP consistency seems to play a subordinate function in global discourse comprehension, while the semantic-pragmatic plausibility of the test sentence NP in relation to the context is in focus.

This idea is in line with several theoretical accounts of discourse processing. For instance, Sanford and Garrod (1989) have argued in their immediacy hypothesis that anaphora resolution need not be the key process in all discourse structures, but that other processes may take precedence and occur immediately. According to the authors, narrative discourse structures are characterized by a causal chain of events: Sequences of events lead to subsequent events and can be causally traced back to earlier sequences of events (Trabasso & van den Broek, 1985; van den Broek, 1990) or can create expectations about subsequent discourse content and facilitate discourse comprehension (Zwaan & Rapp, 2006). The discourse used in the present experiments had a narrative character: The action of a protagonist (e.g., "Simon goes hiking") was described in an action space characterized by the presence or absence of objects (e.g., "a rope", "no tent"). In the test sentence, which was introduced by a temporal adverb ("when"), another action of the same protagonist ("Simon") took place with regard to an object that has already been mentioned (sentence set 1) or had not yet been mentioned (sentence set 2). Because of this narrative discourse structure, it is possible that for discourse understanding participants weighted certain discourse features, such as the semantic-pragmatic causality of events, more heavily than others, possibly saving cognitive resources for processing less weighted discourse features. Furthermore, according to Kintsch's model of the construction and integration of verbal input (Kintsch, 1988; Kintsch & van Dijk, 1978), the semantic

meaning of a constituent is activated in a construction phase, and its integration is pragmatically supported by the meaning of the context (Zwaan & Rapp, 2006). In the construction phase, events are indexed and linked in memory along five dimensions: time, space, protagonist, causality, and intentionality (Zwaan et al., 1995). When events are combined, certain dimensions can be brought into focus (De Vega et al., 2004), while dimensions such as time and protagonist are observed independently of instructions (Therriault et al., 2006). In sentence set 2, a new object appears that was not previously mentioned in the context, which expands the discourse space and is therefore indexed and combined as a new event. According to Zwaan and Rapp (2006), anaphoric resolution has no event-indexing relevance, but determines the way in which information must be integrated into a context. In this respect, our results show that anaphoric information might, in principle, modulate discourse coherence judgments (sentence set 1), but does not have to (sentence set 2), thus revealing the possibility of weighting event-indexing and -combining operations to construct a situational discourse model. Assuming that working memory capacity is limited, this weighting range could be explained by strategies to minimize memory load by focusing on the most prominent hint to establish a coherent discourse model.

The above conclusion assumes that the processing of PSPs depends on the context (i.e., the sentence sets) in which the PSPs are embedded. When comparing the two sentence sets, however, it is essential to consider an additional factor. The two-sentence sets differ not only concerning their context, but they additionally differ in terms of the PSP that is triggered in using the definite determiner in sentence set 1 (*memory context*) and using the definite and indefinite determiner in sentence set 2 (*inference context*). While in sentence set 1, the definite determiner contains a violation of the existence PSP via the negation of a potential referent, a violation of the existence PSP in sentence set 2 is given by the fact that a potential referent for the definite determiner has not been mentioned. The different results concerning the PSP consistency effect between the sentence sets could therefore also be explained by a differential degree of PSP violation. In this sense, the use of negated referents in the memory context is probably a stronger violation than in a neutral context in which the referent was simply not mentioned. In the latter case, it is possible to perceive the PSP violation rather as a small blur, accommodating the unmentioned referent as existing in an imagined contextual space, whereas it should be hard to accommodate an explicitly negated referent.

Various studies actually suggest that the same PSP trigger (here, the definite determiner) might contain different PSPs, which might be furthermore processed in a different manner. As described in the introduction, several studies (Domaneschi & Di Paola, 2018; Haviland & Clark, 1974; Kirsten et al., 2014; Schneider et al., 2020) show that processing an unfulfilled PSP triggered by the definite determiner comes along with enhanced processing costs and with lowered ratings for coherence (see Rolke et al., 2023). These enhanced processing costs occur irrespective of whether the context lacks an adequate referent for the definite determiner (Clifton, 2013; Garnham et al., 1997) or whether the context is ambiguous, thereby impeding the smooth performance of the reference process (see, e.g., Altmann & Steedman, 1988). Moreover, Rolke et al. (2023) report that, depending on the context of the subsequent sentence, unfulfilled PSPs triggered by the definite determiner can either immediately result in a slowdown in reading-time or cause possible reference processes to be

postponed to a later point in the sentence. The authors assumed that the way a reference process is solved might depend on the actual context in which the trigger occurred and on the potential of the context to supplement absent information. Rolke et al. (2023) conceptualized this potential accommodation process as occurring at the end of a cognitive PSP processing sequence. Given that the present study solicited assessments of coherence, it is probable that such accommodation processes occurred prior to the evaluation, and this in turn would explain the absence of a PSP consistency effect in sentence set 2.

In addition to other studies, the present study shows that PSP processing is not an all-or-nothing process that always takes place to the same extent and in a stereotypical manner. On the contrary, we show that depending on the memory load and the global context, the specific PSP violation is perceived differently by the cognitive system and violations might possibly be accommodated. Given the finding that working memory load affects PSP processing, future research could examine how individual differences in discourse processing strategies correlate with variations in working memory capacity. Furthermore, our results suggest that accommodation is likely related to coherence judgments. This opens new avenues for investigating the underlying mechanisms of coherence judgments and their interaction with accommodation. For example, examining brain activation patterns during coherence judgment tasks associated with the accommodation process could shed light on the cognitive processes involved and provide valuable insights and possibly identify neurophysiological markers that signal when accommodation is occurring. Finally, based on the findings by Carter and Nieuwland (2022), who observed ERP evidence for expectation-based modulation of reference processes for the definite and the indefinite determiner, it would be interesting to see whether these expectancy-based processes might be triggered by different discourse contexts or by knowledge of communication partners, which in turn might influence reference processing and accommodation behavior.

5. Conclusion

The aim of the present study was to investigate whether PSP reference processing depends on contextual factors of a discourse. We showed that reference processing depends on the temporal distance of the contextual presentation of a referent (*memory context*). Given that increased cognitive demands in storing early contextual information are accompanied by cognitive demands in referencing that may exceed the capacity of our working memory, we suggest that resources for memory and referencing processes are shared in working memory. In discourse contexts in which potential referents were not mentioned but had to be inferred from semantic-pragmatic contextual relations (*context inference*), there seems to be a stronger weighting of event-indicating and -combining processes over integrative operations for anaphoric resolution.

Acknowledgements. The authors thank the anonymous reviewers for their thoughtful and thorough comments on the manuscript. We acknowledge support by Open Access Publishing Fund of University of Tübingen. The study was supported by the German Research Foundation (DFG Project DI 2246/1-3).

Data availability statement. Data and syntax are available at <https://osf.io/w5nx9>.

Competing interest. The authors declare none.

References

- Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191–238. [https://doi.org/10.1016/0010-0277\(88\)90020-0](https://doi.org/10.1016/0010-0277(88)90020-0)
- Anderson, J. E., & Holcomb, P. J. (2005). An electrophysiological investigation of the effects of coreference on word repetition and synonymy. *Brain and Language*, 94, 200–216. <https://doi.org/10.1016/j.bandl.2005.01.001>
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–89). Academic Press
- Beaver, D., & Zeevat H. (2007). Accommodation. In G. Ramchand & C. Reiss (Eds.), *The Oxford handbook of Linguistic Interfaces* (pp. 503–538). Oxford University Press.
- Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, 4(1). <https://doi.org/10.1177/2515245920960351>
- Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language*, 98, 159–168. <https://doi.org/10.1016/j.bandl.2006.04.005>
- Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *Cognitive Neuroscience and Neuropsychology*, 18(17), 1851–1854. <https://doi.org/10.1097/WNR.0b013e3282f1a999>
- Carter, G. A., & Nieuwland, M. S. (2022). Predicting definite and indefinite referents during discourse comprehension: evidence from event-related potentials. *Cognitive Science*, 46(2), e13092. <https://doi.org/10.1111/cogs.13092>
- Christensen, R. H. B. (2023). Ordinal – regression models for ordinal data. <https://CRAN.R-project.org/package=ordinal>
- Cirilo, R. K. (1981). Referential coherence and text structure in story comprehension. *Journal of Verbal Learning and Verbal Behavior*, 20(3), 358–367. [https://doi.org/10.1016/S0022-5371\(81\)90505-3](https://doi.org/10.1016/S0022-5371(81)90505-3)
- Clark, H. H. (1975) Bridging. In R. C. Schank & B. L. Nash-Webber, *Theoretical issues in natural language processing* (pp. 169–174). Association for Computing Machinery.
- Clifton Jr, C. (2013). Situational context affects definiteness preferences: Accommodation of presuppositions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(2), 487.
- De Vega, M., Robertson, D. A., Glenberg, A. M., Kaschak, M. P., & Rinck, M. (2004). On doing two things at once: Temporal constraints on actions in language comprehension. *Memory & cognition*, 32(7), 1033–1043.
- Dietrich, S., Hertrich, I., Seibold, V. C., & Rolke, B. (2019). Discourse management during speech perception: A functional magnetic resonance imaging (fMRI) study. *NeuroImage*, 202, 116047. <https://doi.org/10.1016/j.neuroimage.2019.116047>
- Dietrich, S., Hertrich, I., Blum, C., Seibold, V. C., & Rolke, B. (2023). The cortical dynamics of context-dependent language processing. *Language, Cognition and Neuroscience*, 38(7), 903–924. <https://doi.org/10.1080/23273798.2023.216512>
- Domaneschi, F., & Di Paola, S. (2018). The processing costs of presupposition accommodation. *Journal of Psycholinguistic Research*, 47, 483–503. <https://doi.org/10.1007/s10936-017-9534-7>
- Dudschig, C., Kaup, B., Liu M., & Schwab, J. (2021a). The processing of negation and polarity: an overview. *Journal of Psycholinguistic Research*, 50, 1199–1213. <https://doi.org/10.1007/s10936-021-09817-9>
- Dudschig, C., Kaup, B., Svaldi, J., & Gulewitsch, M. D. (2021b). Negation processing in children with ADHD: The generic problem of using negation in instructions. *Journal of Psycholinguistic Research*, 50(6), 1309–1320. <https://doi.org/10.1007/s10936-021-09789-w>
- Ferstl, E. C., & von Cramon, D. Y. (2001). The role of coherence and cohesion in text comprehension an event-related fMRI study. *Cognitive Brain Research*, 11, 325–340. [https://doi.org/10.1016/S0926-6410\(01\)00007-6](https://doi.org/10.1016/S0926-6410(01)00007-6)
- Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory and Language*, 47(2), 250–272. [https://doi.org/10.1016/S0749-596X\(02\)00004-9](https://doi.org/10.1016/S0749-596X(02)00004-9)
- Garnham, A., Oakhill, J., & Cain, K. (1997). The interpretation of anaphoric noun phrases time course, and effects of overspecificity. *The Quarterly Journal of Experimental Psychology Section A*, 50(1), 149–162.
- Garrod, S.C., & Sanford, A.J. (1982). The mental representation of discourse in a focused memory system: implications for the interpretation of anaphoric noun phrases. *Journal of Semantics*, 1, 21–41. <https://doi.org/10.1093/jos/1.1.21>

- Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). MIT Press.
- Grice, H. P., 1975. Logic and conversation. In P. Cole, & J. L. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts* (pp. 41–58). Academic Press.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69(2), 274–307. <https://doi.org/10.2307/416535>
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of verbal learning and verbal behavior*, 13(5), 512–521.
- Jarvella, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, 10(4), 409–416. [https://doi.org/10.1016/S0022-5371\(71\)80040-3](https://doi.org/10.1016/S0022-5371(71)80040-3)
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, 99(1), 122–149. <https://doi.org/10.1037/0033-295X.99.1.122>
- King, W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7(2), 376–395. <https://doi.org/10.1162/jocn.1995.7.3.376>
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review*, 95(2), 163–182. <https://doi.org/10.1037/0033-295X.95.2.163>
- Kintsch, W. & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363–394. <https://doi.org/10.1037/0033-295X.85.5.363>
- Kirsten, M., Tiemann, S., Seibold, V. C., Hertrich, I., Beck, S., & Rolke, B. (2014). When the polar bear encounters many polar bears: Event-related potential context effects evoked by uniqueness failure. *Language, Cognition and Neuroscience*, 29(9), 1147–1162. <https://doi.org/10.1080/23273798.2014.899378>
- Lesgold, A. M., Roth, S. F., & Curtis, M. E. (1979). Foregrounding effects in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior*, 18(3), 291–308. [https://doi.org/10.1016/S0022-5371\(79\)90164-6](https://doi.org/10.1016/S0022-5371(79)90164-6)
- Matsui, T. (2001). Experimental pragmatics: towards testing relevance-based predictions about anaphoric bridging inferences. In V. Akman, P. Bouquet, R. Thomason, & R. A. Young (Eds.), *Modeling and using context: Third international and interdisciplinary conference on modeling and using context* (pp. 248–260). Springer Verlag.
- O'Brien, E. J. (1987). Antecedent search processes and the structure of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 278–290. <https://doi.org/10.1037/0278-7393.13.2.278>
- O'Brien, E. J., & Myers, J. (1987). The role of causal connections in the retrieval of text. *Memory & Cognition*, 15(5), 419–427. <https://doi.org/10.3758/BF03197731>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberg, R., Sogo, H., Kastman, E., & Lindelov, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51, 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rolke, B., Kirsten, M., Seibold, V. C., Dietrich, S., & Hertrich, I. (2023). Processing references in context: when the polar bear does not meet a polar bear. *Cognitive Processing*, 24, 497–520. <https://doi.org/10.1007/s10339-023-01150-4>
- Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, 2(9), 437–442. <https://doi.org/10.3758/BF03208784>
- Sanford, A. J., & Garrod, S. C. (1989). What, when, and how?: Questions of immediacy in anaphoric reference resolution. *Language and Cognitive Processes*, 4(3–4), SI235–SI262. <https://doi.org/10.1080/01690968908406369>
- Schneider, C., Bade, N., Janczik, M. (2020). Automatic or capacity-limited? A combination of the PRP approach with a self-paced reading task. *Journal of Psycholinguistic Research*, 49, 247–273. <https://doi.org/10.1007/s10936-019-09686-3>
- Schwarz, F. (2007). Processing presupposed content. *Journal of Semantics*, 24, 373–416. <https://doi.org/10.1093/jos/ffm011>
- Schwarz, F., & Tiemann, S. (2012). Presupposition processing—The case of German wieder. In *Logic, Language and Meaning: 18th Amsterdam Colloquium, Amsterdam, The Netherlands, December 19–21, 2011, Revised Selected Papers* (pp. 200–209). Springer Berlin Heidelberg. https://florianschwarz.net/wp-content/uploads/papers/schwarz_tiemann.pdf
- Simons, M. (2003). Presupposition and accommodation: understanding the Stalnakerian picture. *Philos. Stud.* 112, 251–278. <https://doi.org/10.1023/A:1023004203043>

- Singh, R., Fedorenko, E., Mahowald, K., Gibson, E. (2016). Accommodating presuppositions is inappropriate in implausible contexts. *Cognitive Science*, 40, 607–634. <https://doi.org/10.1111/cogs.12260>
- Sparks, J. R., & Rapp, D. N. (2010). Discourse processing—examining our everyday language experiences. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(3), 371–381. <https://doi.org/10.1002/wics.11>
- Spivey-Knowlton, M., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, 55(3), 227–267. [https://doi.org/10.1016/0010-0277\(94\)00647-4](https://doi.org/10.1016/0010-0277(94)00647-4)
- Stalnaker, R. (2002). Common Ground. *Linguistics and Philosophy*, 25(5/6), 701–721. <https://doi.org/10.1023/A:1020867916902>
- Strawson, P. F. (1950). On referring. *Mind*, 59, 320–344. <https://doi.org/10.1093/mind/LIX.235.320>
- Streb, J., Hennighausen, E., & Rösler, F. (2004). Different anaphoric expressions are investigated by event-related brain potentials. *Journal of Psycholinguistic Research*, 33, 175–201. <https://doi.org/10.1023/B:JOPR.0000027961.12577.d8>
- Stroup, W. W. (2012). *Generalized linear mixed models: modern concepts, methods and applications*. CRC press.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251–296. <https://doi.org/10.1023/A:1022193728205>
- Taylor, J. E., Rousselet, G. A., Scheepers, C., & Sereno, S. C. (2023). Rating norms should be calculated from cumulative link mixed effects models. *Behavior Research Methods*, 55(5), 2175–2196.
- Tiemann, S., Schmid, M., Bade, N., Rolke, B., Hertrich, I., Ackermann, H., Knapp, J., & Beck, S. (2011). Psycholinguistic evidence for presuppositions: On-line and of-line data. In I. Reich, E. Horch, & D. Pauly (Eds.), *Proceedings of Sinn & Bedeutung* (Vol. 15, pp. 581–595). Universaar - Saarland University Press: Saarbrücken, Germany.
- Therriault, D. J., Rinck, M., & Zwaan, R. A. (2006). Assessing the influence of dimensional focus during situation model construction. *Memory & Cognition*, 34(1), 78–89. <https://doi.org/10.3758/BF03193388>
- Trabasso, T., & Sperry, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language*, 24, 595–611. [https://doi.org/10.1016/0749-596X\(85\)90048-8](https://doi.org/10.1016/0749-596X(85)90048-8)
- Trabasso, T., & Van Den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24(5), 612–630. [https://doi.org/10.1016/0749-596X\(85\)90049-X](https://doi.org/10.1016/0749-596X(85)90049-X)
- Van Berkum, J. J. A., Zwitserlood, P., Hagoort, P., Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, 17, 701–718. [https://doi.org/10.1016/S0926-6410\(03\)00196-4](https://doi.org/10.1016/S0926-6410(03)00196-4)
- Van Berkum, J. J. A., Hagoort, P., Brown, C. M. (1999). Semantic integration in sentence and discourse: evidence from the N400. *Journal of Cognitive Neuroscience*, 11, 657–671. <https://doi.org/10.1162/089892999563724>
- Van den Broek, P. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension Processes in Reading* (pp. 423–445). Routledge.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.
- Von Fintel, K. (2008). What is presupposition accommodation, again? *Philosophical Perspectives*, 22, 137–170. <https://doi.org/10.1111/j.1520-8583.2008.00144.x>
- Warren, T. C. (2001). *Understanding the role of referential processing in sentence complexity* [Doctoral dissertation]. Massachusetts Institute of Technology.
- Zhang, J. X., Leung, H.-C., Johnson, M. K. (2003). Frontal activations associated with accessing and evaluating information in working memory: an fMRI study. *Neuroimage*, 20, 1531–1539. <https://doi.org/10.1016/j.neuroimage.2003.07.016>
- Zwaan, R., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, 6(5), 292–297. <https://doi.org/10.1111/j.1467-9280.1995.tb00513.x>
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162–185. <https://doi.org/10.1037/0033-2909.123.2.162>
- Zwaan, R. A., & Rapp, D. N. (2006). Discourse comprehension. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 725–764). Academic Press.

Cite this article: Dietrich, S., Seibold, V. C., & Rolke, B. (2024). Discourse comprehension and referential processing: effects of contextual distance and semantic plausibility on presupposition processing. *Language and Cognition* 16: 2032–2054. <https://doi.org/10.1017/langcog.2024.45>