Definite Descriptions in the Light of the Comprehension *vs.* Acceptance Distinction: Comparing Self-Paced Reading with Eye-Tracking Measures

Misha-Laura Müller * and Magali A. Mari*

Cognitive Science Center, Institute of Communication and Cognitive Sciences, University of Neuchâtel, Neuchâtel, Switzerland

This paper presents two experiments on the processing of informative definite descriptions in plausible vs. implausible contexts. Experiment 1 is a self-paced reading task (with French native speakers, *n* = 69), with sentences containing a definite *vs.* indefinite NP, each preceded by *plausible* or *implausible* contexts. Our study replicated Singh and colleagues' findings, namely that definite descriptions are significantly costlier when they occur in *implausible* contexts. The translation of the original stimuli from English to French did not affect the results, suggesting that the phenomenon applies cross-linguistically. Experiment 2 consists in an eye-tracking task, designed to measure the participants' (*n* = 44) gaze patterns on complete sentences with the same four conditions (definite *vs.* indefinite NP; implausible vs. implausible contexts). A mixed effect model analysis revealed that (a) the *total gaze duration* on target segments and (b) the *processing of the complete sentence* were significantly longer in implausible conditions. These results show that implausible contexts predict a marked increase in the offline processing costs of definite descriptions. However, no significant difference was found for online processing measures (i.e., *first fixation duration, first-pass reading time* and *regression path time* measures) across all experimental conditions. These results suggest that it is only once the sentence is fully processed that implausible contexts increase processing costs. Furthermore, these results raise methodological issues related to the study of the online processing of definite descriptions, to the extent that self-paced reading and eye-tracking methods in the present study lead to incompatible results. With respect to the eye-tracking results, we suggest that the contrast between online and offline processing is likely to reflect the fact that participants first adopt a stance of trust to understand utterances before filtering the information through their epistemic vigilance module.

Keywords: definite descriptions, implausible context, presupposition processing, comprehension, acceptance, epistemic vigilance, self-paced reading, eye-tracking

INTRODUCTION

Definite Descriptions: A Brief Review, From Frege to EEG

The presupposition effects of definite descriptions were first described by Frege's seminal paper, *Über Sinn und Bedeutung* (1892). One of the most emblematic examples to illustrate such effects can be found in the utterance "The king of France is bald", initially discussed by Russel. In this example, the definite description triggers an existential presupposition that can be paraphrased as "*there exists a king of France*". The existential component of the NP is lost when it contains an indefinite article as in "*A king of France is bald*"[1]. According to Frege, the presupposition of existence is not part of the thoughts expressed by a proposition (Frege, 1892) but rather serves as a precondition for the sentence to have a denotation. At that time, the discussion regarding definite descriptions focused essentially on assessing whether the truth of the presupposition is necessary for a sentence to have a truth value (cf. Russel 1905; Strawson 1950).

Beyond their existential component, definite descriptions trigger a presupposition of uniqueness or familiarity (cf. Heim 1982; Schwarz 2009; Roberts 2003). The familiarity property is divided between "weak familiarity," meaning that the definite description is globally familiar in the general culture, or "strong familiarity," meaning that the definite description has an anaphoric function (cf. Roberts 2003). Definite descriptions are said to have a strong anaphoric function when they directly refer to a previously stated indefinite description, as in "John bought a book and a magazine. *The book* was expensive" (Schwarz, 2013:535). Finally, in cognitive pragmatics, definite descriptions are defined as a linguistic mean, among other presupposition triggers, that allow the listener to spare cognitive efforts on information that is not relevant (Sperber and Wilson 1987; Saussure, 2012; Saussure, 2013; Vallauri et al., 2018).

Definite descriptions present the referent of the following noun as if it was mutually known by the speaker and the listener. However, it is not always the case that the content is actually shared by the listener at the time she hears the utterance. Thus, the presupposition of a definite description can either be satisfied (entailed by the context) or not satisfied (not entailed by the context), amounting–in the latter case–to a presupposition failure. In experimental settings, the satisfaction of a presupposition is generally controlled by context sentences that introduce the content which will then be presupposed in a target sentence. For instance, the presupposition "the graphic designer" is said to be satisfied in sentence (1) because it has previously been introduced in the context sentence (cf. "[…] a very bad-tempered *graphic designer*"). It is not the case in sentence (2) where the presupposition (i.e., the existence of a "familiar" graphic designer) is not explicitly stated in the context sentence:

1) In Paolo's office, there used to be a very bad-tempered **graphic designer**. […] Due to overstaffing problems, about a month ago **the graphic designer** was made redundant.

(*Presupposition satisfaction*, cf. Domaneschi et al., 2018:18)

2) In Paolo's office there are many employees. […] Due to overstaffing problems, about a month ago **the graphic designer** was made redundant.

(*Presupposition failure*, cf. Domaneschi et al., 2018:18)

When there is a presupposition failure, as in sentence (2) above, the content can be accommodated by the listener i.e., she can incorporate this information into her set of previous beliefs. In the case of utterance (2), the listener will infer an appropriate context in which there is an identifiable graphic designer. In psycholinguistics, accommodation is defined as a two-step process: First, the addressee recognizes the incompatibility between the presupposition and the context (i.e., the "presupposition failure"). In a second step, the addressee incorporates the information into his or her previous beliefs (Domaneschi and Di Paola, 2018:484). Let us note that accommodation tacitly implies the acceptance of the presupposed content (Von Fintel, 2008; Müller, 2018; Müller, 2021) and its memorization either in working, short-term or long-term memory. However, the distinction between working memory, short-term and long-term memory has not been thoroughly investigated (but see Clifton, 2013 and Schneider et.al., 2020 for studies on deep and superficial comprehension in accommodation processes). In the research field of memory illusions and definite descriptions, experimental investigations have focused on the incorporation of false content in short-term memory (Loftus and Guido, 1975; Bredart and Modolo, 1988; Barton and Sanford, 1993; Vallauri et al., 2018, *inter alia*).

Until now, psycholinguistics focused on the processing costs of definite descriptions in terms of reading times and number of fixations using self-paced reading and eye-tracking tasks. Definite descriptions have also been studied with EEG methods, focusing on language-related Event-Related brain Potentials. Across different methods, the main question that has been addressed is whether they are semantic or pragmatic content (cf. Schwarz, 2009; Schwarz, 2013; Domaneschi and Di Paola, 2018). Compared with other presupposition triggers, definite descriptions have been defined as mandatory triggers (as opposed to "optional triggers", cf. Glanzberg, 2005; Domaneschi et al., 2014). Furthermore, the processing costs of definite descriptions have been shown to significantly increase during accommodation (Domaneschi and Di Paola, 2018). This observation is also supported by EEG experiments that have shown that accommodating definite descriptions elicits a reaction on the N400[2], reflecting additional costs of the integration of a

---

[1]The uniqueness and familiarity components are also lost when comparing definite descriptions with indefinite ones. However, as it was pointed out by an anonymous reviewer, some indefinite descriptions can be presuppositional (see Heim (1982); Özge (2011), and also some of Grice's examples of generalized conversational implicatures, cf. (Grice, 1975: 56)).

[2]The N400 is language-related Event-Related brain Potential (ERP) component. A review of the functional role of the N400 (and P600) with respect to semantic processing can be found in Seyednozadi et al. (2021).

new discourse referent in a mental model where it was not previously introduced (Domaneschi et al., 2018:27).

Finally, informative definite descriptions have been shown to be significantly costlier in implausible contexts than in plausible contexts, leading participants to judge implausible definite description as "inappropriate" (Singh et al., 2016:617). This judgment was measured by a stop-making-sense task[3], during which participants were significantly more likely to end the task in implausible contexts (for both definite and indefinite articles). Singh et al.'s experiment involved stimuli composed by two sentences, which were divided into segments. The first sentence introduced a plausible (3a) vs. implausible (3b) context, and the following sentence began with a definite vs. indefinite NP:

3a) Gabriella went to a **concert** 2 weeks ago. **The/A** guitarist winked at her flirtatiously.

3b) Gabriella went to a **school** 2 weeks ago. **The/A** guitarist winked at her flirtatiously.

(Singh et al., 2016:631)

In both cases, the definite description is not entailed by the context, to the extent that there is no explicit mention of a guitarist in neither of the context sentences. Thus, both conditions lead to a presupposition failure that the listener will eventually seek to repair through accommodation[4]. In both contexts, accommodation would imply the inference of the relationship between the context sentence and the first NP of the second sentence. Such inference is best known as "bridging" (Clark, 1975), and can be further illustrated as follows:

4) John bought **a book**.

   The author is French.

5) John's hands were freezing as he was **driving** down the street.

The steering wheel was bitterly cold and he had forgotten his gloves.

(Schwarz, 2013:535)

In utterance (4), the listener infers that the "author" is the one of the "book" that John has bought and in utterance (5), the "steering wheel" is understood as being the one of the car that John was "driving". In both cases, the referent is implicit, yet highly accessible to the extent that it belongs to the same semantic domain as the focus object of the context sentence (i.e., there is a strong semantic relationship between "the author" and "the book", as well as between "the steering wheel" and "driving [a car]"). Semantic proximity has been shown to play a crucial role on the processing ease of definite descriptions that require a bridging inference to identify the referent (Haviland and Clark, 1974; Garrod and Anthony, 1977; Schwarz, 2019). According to Clifton (2013), definite descriptions containing a superordinate term, as in *the bird*, are easily processed as an anaphoric reference to a previously-introduced subordinate term. This applies specifically when the latter term is a prototypical member of the category (*the robin*).

Semantic proximity also plays a role on people's ability to identify false content. This was shown in the context of the Moses Illusion where participants made significantly more mistakes when the false lexical item was semantically close to the accurate one (Erickson and Mattson, 1981). For instance, more mistakes were made when the question was '*How many animals of each kind did **Moses** take on the Ark?*' as opposed to "*How many animals of each kind did **Kennedy** take on the Ark?*", because Moses is semantically closer to Noah (both being biblical figures)[5].

Considering the above, the study of definite descriptions appears as a saturated field, leaving only little space for further investigations. In spite of the numerous existing studies, we contend that there is more to say in light of the recent advances in cognitive pragmatics and psychology. The adopted perspective of this paper is to consider the relationship between definite descriptions and the listener's epistemic vigilance, in particular with respect to the comprehension vs. acceptance distinction (Sperber et al., 2010). According to the Epistemic Vigilance hypothesis, it is assumed that humans developed, in the course of their evolution, some mechanisms dedicated to the filtering of information. This competence can be found in the human ability to distinguish the comprehension of an utterance from its acceptance. It is assumed that humans first adopt a stance of trust (to comprehend utterances) before filtering the information through their epistemic vigilance module, so that they eventually accept the utterance. This ability would have evolved to meet our ancestors' needs, namely keeping the benefits of verbal communication, and reducing the risks of manipulation (Sperber et al., 2010:368). In the present study, we investigate the effect of context plausibility on the comprehension and acceptance of definite and indefinite descriptions.

---

[3]In "stop-making-sense" tasks, participants are given the instruction to keep making words appear, segment by segment, as long as the sentence(s) make sense to them. As soon as an incoming word/phrase does not make sense in the context of the preceding words/phrases, participants are asked to end the task (cf. Singh et al., 2016:615).

[4]An anonymous reviewer argued that in implausible contexts, as in (3b), the definite description would not be accommodated by the listener because it is "impossible to understand". However, we contend that in a more natural setting i.e., when complete sentences are presented (as opposed to segment-by-segment), there is no reason to assume that participants would *not* seek to reconstruct a hypothetical context in which the definite description is relevant (cf. Sperber and Wilson's Relevance-guided comprehension heuristic, presented in Research question and hypotheses). Actually, the construction of an adequate context for implausible definite descriptions does not seem to be costly, as exemplified below: (3b') Gabriella went to a **school** 2 weeks ago. **The** guitarist [who was invited at the school to give a concert] winked at her flirtatiously. Furthermore, let us note that the implausibility of the noun can be considered as qualitatively different when preceded by a definite or an indefinite article. We conducted a pre-test (cf. *Materials*) to make sure that this component did not bias the interpretation of our results.

[5]Proper names, which are used in Moses Illusion experiments, are equivalent to definite descriptions (cf. Geurts 1997; Matushansky 2006).

Research Question and Hypotheses

The effect of context plausibility on the comprehension and acceptance of definite descriptions is addressed through a replication of (Singh et al., 2016) and the addition of an eye-tracking task. In their study, (Singh et al., 2016) designed a self-paced reading task coupled with a stop-making sense task, in which participants were asked to stop the task when they judged an utterance as inappropriate. In average, the reading pace in self-paced reading or stop-making sense tasks is about half less rapid than during normal reading because of task demands (Rayner, 1998). This is not the case in eye-tracking tasks, which benefit from valuable advantages compared to self-paced or stop-making sense tasks: they allow to assess moment-to-moment cognitive processes in natural reading settings, without additional processing costs due to task demands (Rayner, 1998). Furthermore, eye-tracking tasks allow to determine if a text region was fixated upon during the first reading (online processing) or if it was fixated later in reading (offline processing). In this way, it is possible "to make inferences about the time course of processing during text comprehension" (Liversedge et al., 1998:56). For this reason, it is worth comparing these methodologies and their effects on reading. Thus, the interest of this study is twofold: first, we evaluate whether Singh et al. (2016) results can be replicated cross-linguistically (from English to French); second, we compare these results with eye-tracking measures. This comparison allows, in a first stage, to test whether a change in methodology affects processing during reading. In a second stage (Experiment 2), the comparison between online and offline measures in the eye-tracking task provides insights regarding the comprehension vs. acceptance distinction, *as per* Sperber et al. (2010).

The framework adopted in this paper is Relevance theory, according to which utterance processing follows a relevance-guided comprehension heuristic, best described as follows:

Relevance-guided comprehension procedure:
Follow a path of least effort in computing cognitive effects:

a. Test interpretative hypotheses (disambiguations, reference resolutions, implicatures, etc.) in order of accessibility.
b. Stop when your expectations of relevance are satisfied.

(Wilson and Sperber, 2004:613)

Furthermore, we adopt the fundamental assumptions stemming from the Epistemic Vigilance hypothesis (Sperber et al., 2010): first, we assume that sentence processing is understood in the context of a massively modular mind (Sperber, 2005). That is to say, language comprehension can be associated with other domain-specific mechanisms, such as emotion reading, attributing mental states, social cognition and so on (Wilson, 2016). Second, language processing is conceived as being on a par with epistemic vigilance mechanisms, responsible for the assessment of the reliability of the communicated content. Importantly, as underlined by Mazzarella (2015), the role of epistemic vigilance mechanisms is to assess the believability of the interpretation *resulting from* utterance comprehension. In this sense, a behavioral output of the acceptance of an utterance is expected to be observed *after* the sentence is processed.

With respect to self-paced reading and eye-tracking measures, this paper holds two overarching assumptions: first, reading times and the length of time spent on an area of interest reflects the quantity of cognitive effort. In other words, the longer participants look at a certain region, the more costly the processing on this region is (Liversedge et al., 1998; Rayner, 1998; Rayner, 2009). In line with Relevance theory, we contend that the processing costs are due to the search for relevance, be it for the identification of a referent or for the comprehension or/and acceptance of an utterance. Second, both comprehension and acceptance can be assessed based on measures of online and offline processing respectively. Online processing is defined as the immediate processing of a word, or series of words, when that word is read for the first time. Traditionally, online processing measures are "sensitive to processing difficulty experienced immediately on reading [a] word" (Liversedge et al., 1998:58). Online processing can thus assess utterance comprehension, with both self-paced reading tasks (based on the time spent reading a segment) and eye-tracking tasks (based on measures of first fixation, first pass and regression durations). Offline processing corresponds to late processing of a word, or series of words, after having been read for the first time (Liversedge et al., 1998) and can thus be used to investigate utterance acceptance. An efficient way to assess offline processing is with eye-tracking tasks and measures of total gaze duration and processing time of sentences. Therefore, a mismatch between the comprehension and the acceptance of definite descriptions based on context plausibility should be observed in a difference in reading times between online and offline processing. More precisely, difficulties in comprehension should appear as costly online processing, whereas difficulty in the acceptance of definite descriptions should have an impact on offline processing measures. We bear in mind that offline processing costs are also imputable to pragmatic inferences such as implicatures[6] (Noveck and Andres, 2003; Noveck, 2004). However, implicatures are equally likely to be inferred across plausible and implausible conditions. Therefore, a significant difference between online and offline processing across plausible and implausible conditions can arguably be attributed to the cognitive costs dedicated to the *acceptance* of implausible sentences (as opposed inferences dedicated to their "comprehension," i.e., the inference of an implicature to identify the speaker's intended meaning). Finally, with respect to definite descriptions, we expect that their acceptance will be more costly than for indefinite descriptions, due the requirements to identify a salient referent.

In light of the above, specific hypotheses can be formulated for the two experiments: (1) in line with Singh and colleagues, we predict that implausible contexts compared to plausible contexts will increase the processing costs of definite and indefinite descriptions, (2) we predict that, within the implausible condition, the processing of definite article should be costlier

---

[6]One could argue that additional costs during offline processing could be attributed to bridging inferences. However, as shown by Burkhardt (2006), bridging inferences of definite descriptions (in new vs. given settings) are observed during online processing.

than of indefinite articles, due to a difficulty to identify a "familiar"/salient referent in an implausible context, but this difference should not be present for plausible conditions. Accordingly, context plausibility will affect more the processing of definite descriptions than indefinite descriptions. On the one hand, Experiment 1 investigates exclusively online processing (which is specific to self-paced reading tasks) and, therefore, tests the comprehension of definite descriptions while varying context plausibility. On the other hand, Experiment 2 assesses whether there is a distinction between comprehending and accepting definite descriptions in plausible and implausible contexts. Comprehension is tested with online processing measures (first fixation, first pass and regression durations), whereas acceptance is evaluated with offline processing measures (total gaze duration and processing time of sentences).

The remainder of this paper is structured as follows: *Experiment 1: Self-Paced Reading Task* presents the self-paced reading task (Experiment 1), using Singh et al., 2016 stimuli and experimental design. *Experiment 2: Eye-Tracking Task* presents the eye-tracking experiment (Experiment 2), testing the same stimuli as in Experiment 1. With these two experiments, we assess if a definite or indefinite description is first *comprehended*, with online processing measures, and then *accepted*, based on measures of offline processing. Finally, the general discussion compares the results in both experiments and evaluates whether the plausibility of context is likely to affect the acceptance of an utterance, as defined by Sperber et al. (2010).

EXPERIMENT 1: SELF-PACED READING TASK

Methods
Participants
For the first experiment, 81 participants were recruited in a swiss university (University of Neuchatel) and through advertisement by e-mail and on social networks. Only French-native speakers were eligible to take part in the experiment. Our experimental design is composed of two fixed effects and two random effects and follows the structure of a "counterbalanced design" as described by (Westfall et al., 2014:2026). This allowed to estimate the required sample size with Westfall and colleagues' website (https://jakewestfall.shinyapps.io/crossedpower/). The sample size estimation was conducted prior to data collection and based on the "standard case" values of variance components or VPCs proposed by (Westfall et al., 2014: 2025), with a power set at .90, a medium effect size of $d = 0.50$) and a number of 24 stimuli. The analysis revealed that 80.6 participants were required. This sample size of 81 participants was set before data collection. No additional participant was recruited once the pre-set sample size was reached. In line with (Singh et al., 2016), we excluded data from participants who had an accuracy rate for comprehension questions lower than 65%. This led to the exclusion of 12 participants. The final sample size included 69 participants i.e., 36 women and 33 men (mean age = 23.70 years, $SD = 4.56$). The study (Experiments 1 and 2) was reviewed and approved in advance by the university's ethics and research committee. Furthermore, each participant was asked to sign a written informed consent before starting the experiment in accordance with the Declaration of Helsinki standards. Finally, participants had the possibility to withdraw their consent at any time during or after the experiment.

Design and Hypotheses
The goal of this experiment was twofold: first, we sought to replicate (Singh et al., 2016) self-paced reading experiment, which assesses the impact of context plausibility on informative definite descriptions. Second, we translated Singh and colleagues' stimuli from English to French to see if there is a stable cross-linguistical effect of context plausibility on definite description comprehension. Let us note that this experiment, because of its methodology, investigates online processing and, therefore, tests exclusively the comprehension of definite descriptions. Unlike Singh and colleagues who tested two presupposition triggers (i.e., "the" and "too"), we focused on one trigger (i.e., "the"). However, we designed an additional task with an eye-tracking device to make further investigations on the online vs. offline processing of definite descriptions (cf. Experiment 2).

Singh and colleagues' experiment presented stimuli composed by two sentences, each of them presented segment-by-segment in a self-paced reading task (cf. *Materials*). The first phrase introduced the context and the second one introduced the critical region with either a definite NP (presupposition condition) or an indefinite NP (assertion condition). The first sentence was either plausible or implausible with respect to the critical region in the second sentence, as illustrated below:

6a) **Plausible context, definite/indefinite noun.**

Bill est allé dans un club vendredi soir. Le/Un videur s'est disputé avec lui pendant un moment.
'Bill went to a club Friday night. The/A bouncer argued with him for a while.'[7]

6b) **Implausible context, definite/indefinite noun.**

Bill est allé au cirque vendredi soir. Le/Un videur s'est disputé avec lui pendant un moment.
'Bill went to the circus Friday night. The/A bouncer argued with him for a while.'

Following (Singh et al., 2016) experiment, the predictions regarding the processing of presuppositions can be formulated as follows:

---

[7]In some cases, the translation of Singh et al. (2016) original stimuli required superficial modifications in order to seem natural for French native speakers, or to make each stimulus equally long. For instance, the first name in the original stimulus in (6a) was Carrie, which is unknown where the experiment was conducted. We changed it to a more common name (i.e., Mary). Also, the segment containing temporal information was also changed in order to have an equivalent number of syllables in the translated stimulus (e.g., in (6a), "Christmas eve" became "this fall"). However, the critical regions (i.e., the "context noun" in the context sentence as well as the critical noun in the second sentence were always faithfully translated).

Prediction 1: Reading times of definite and indefinite descriptions will be significantly longer in implausible conditions compared to plausible conditions (i.e., a main effect of context plausibility is expected).

Prediction 2: Within implausible conditions, reading times will be significantly longer for definite descriptions, compared to indefinite descriptions; within plausible conditions, no significant difference is expected between definite and indefinite descriptions (i.e., an interaction effect is expected between context plausibility and the type of description).

Prediction 1 goes in line with documented evidence that readers engage in lexical predictions while reading sentences (Kleinman et al., 2015; Brothers and Kuperberg 2020): nouns that are semantically distant from a preceding context correlate with longer reading times, due to a violation of expectation (see also Haviland and Clark 1974; Garrod and Anthony, 1977; Schwarz, 2019). For instance, in our experimental setting, the critical noun "A/The bouncer" is expected to be easier to process when the preceding context is located in a "club" than in a "circus".

Prediction 2 focuses on the processing of presuppositions within implausible contexts: given that definite descriptions presuppose the existence of a familiar/salient referent, they should be costlier to process in implausible contexts than indefinite descriptions, which do not have any requirements regarding the referent. In other words, a definite description in an implausible context encodes the instruction to identify a salient concept ("The bouncer") within a context that is incompatible with it (i.e., "in a circus", cf. Singh et al., 2016). We expect this to be a costly process, as it requires additional efforts to repair the context. This is not the case for an indefinite description in implausible contexts, because it does not require the identification of a familiar/salient category in the communicative context (i.e., "A bouncer" is presented as a new piece of information in the context of a circus).

Regarding prediction 2, let us note a difference between Singh and colleagues' experiment and the present one: when computing the reading time differences between definite and indefinite descriptions in implausible contexts, Singh and colleagues dismissed the participants who ended the experiment in their "Stop-making sense" task. This decision implied the exclusion of a significant number of participants who presented difficulties to process the definite description. We believe that the exclusion of such data might be responsible for the fact that no significant difference was found between definite descriptions in plausible vs. implausible contexts. Therefore, we maintained the prediction that, in implausible contexts, definite descriptions will be significantly costlier than indefinite descriptions.

Materials

We translated from English to French the final set of 24 stimuli selected by (Singh et al., 2016) for their first experiment (cf. **Supplementary Materials**). The stimuli are composed by sets of

two sentences: the first sentence provides a context and the second one–the target sentence–introduces information that is either plausible or implausible with respect to the context sentence. The context sentence introduces a location that can have a weak or a strong semantic relationship with the noun phrase of the target sentence, that introduces a specific agent. In addition, the NP of the target sentence was introduced either with a definite description (working as a presupposition trigger) or an indefinite description (i.e., working as an assertion). As a result, a total of 4 conditions were created for each stimulus: 1) Plausible-definite condition (*plausible context - definite description*); 2) Plausible-indefinite condition (*plausible context - indefinite description*); 3) Implausible-definite condition (*implausible context - definite description*); 4) Implausible-indefinite condition (*implausible context - indefinite description*). Furthermore, we conducted a norming study with 34 participants to ensure that the translated stimuli were equally divided into plausible vs. implausible categories, following (Singh et al., 2016: 613) procedure. In average, participants estimated that the probability of seeing one or more specific agents in a given implausible context was of 19.7%. Further stimuli examples are provided below (7a-d). Each of them was divided into 7 segments, marked by the vertical separation "|")[8].

7a) **Plausible-definite condition:**

Lucien | est allé | en prison | samedi soir. | **Le garde** | a parlé | avec lui pendant un moment.
*Lucien went to jail on Saturday night. **The guard** talked to him for a while.*

7b) **Plausible-indefinite condition:**

Lucien | est allé | en prison | samedi soir. | **Un garde** | a parlé | avec lui pendant un moment.
*Lucien went to jail on Saturday night. **A guard** talked to him for a while.*

7c) **Implausible-definite condition:**

Lucien | est allé | dans un café | samedi soir. | **Le garde** | a parlé | avec lui pendant un moment.
*Lucien went to a cafe on Saturday night. **The guard** talked to him for a while.*

7d) **Implausible-indefinite condition:**

Lucien | est allé | dans un café | samedi soir. | **Un garde** | a parlé | avec lui pendant un moment.
*Lucien went to a cafe on Saturday night. **A guard** talked to him for a while.*

The critical segment was in the 5th region[9], namely on the definite or indefinite NP that opened the second sentence (2–4

---

[8]For a complete list of the stimuli used, see in **Supplementary Material**.
[9]For definite descriptions, the triggering point corresponds to the article and the computation point corresponds to the head noun Domaneschi et al. (2018):16).

words). This specific noun phrase contains the presupposition trigger and the noun with a strong or weak semantic relationship with respect to the previously introduced context. In order to analyze if the experimental conditions affected the reading of the regions following the critical segment, the 6th and 7th regions were also considered as critical segments for the analyses.

Finally, we added 22 filler sentences[10] (cf. **Supplementary Materials**) with the same construction, resulting in a total number of 46 trials. The fillers all introduced a location or an activity in the first sentence, which had a highly plausible relationship with the noun introduced in the second sentence (*e.g.*, "Lea went to the spa on Monday. A beautician offered her a stone massage."). Participants had to answer yes/no comprehension questions about the filler sentences to ensure they read the sentences carefully. This task also allowed to eventually dismiss participants who did not adequately complete the task. In order to avoid cognitive overload, we did not add comprehension question to test whether participants accommodated the presupposition, *as per* Domaneschi and Di Paola (2018). Therefore, our study applies only to the processing of informative definite descriptions and not to their accommodation.

Procedure

The experiment was created with the E-Prime 2.0 software (Psychology Software Tools, Inc., 2012). Before beginning the experiment, participants were asked to indicate their age, gender and mother tongue. Once they provided this information, the experiment started with six practice trials and one comprehension question to familiarize participants with the task. A trial started with a white fixation cross on a black background, presented for 500 ms in the middle of the screen. Then, the first segment appeared on the screen. Each segment was presented with white letters on a black background, in a 16-point Arial font. Participants were instructed to read the sentences for comprehension and to press the space bar in order to display the segments consecutively. This procedure has the advantage of preventing participants from displaying the whole sentence before reading it.

Stimuli and fillers were presented randomly, and comprehension questions followed immediately their corresponding sentences. Participants saw the whole comprehension question (i.e., they were not divided in segments). They recorded their answers by pressing on the "E" or "I" keys on the keyboard, following the fixed location of the yes/no answer on the screen.

Participants saw only one version of each stimulus and as many stimuli from each of the four conditions, resulting in a within-subjects and within-stimuli design (Brauer and Curtin, 2018). Specifically, each participant read randomly 6 sentences with in the plausible-definite condition, 6 sentences in the plausible-indefinite condition, 6 sentences in the implausible-definite condition and 6 sentences in the implausible-indefinite condition. Sentences conditions were counterbalanced across subjects.

---

[10]The filler sentences were aimed at conducting another study on gender stereotypes.

Results
Data Analysis

The effects of context plausibility on definite and indefinite description comprehension were measured by reading times (i.e., the time spent reading a sentence segment). The critical segment for analysis is the NP composed by a presupposition trigger plus a noun with a strong or a weak semantic relationship with the previously described context (see example (8) below). Traditional analyses of self-paced reading tasks include the next two segments, following the critical segment, to assess spillover effects. According to Liversedge et al. (1998) spillover effect can be observed if the processing difficulty for the critical segment persists after the critical segments. For this reason, we also included the two segments following the critical region in our analyses to assess spillover effects. In the example below, we illustrate the segments that served as measuring references (i.e., critical segment, spillover 1, spillover 2).

8) Lucien est allé en prison samedi soir. [Le garde]$_{critical\ segment}$ [a parlé]$_{spillover\ segment\ 1}$ [avec lui pendant un moment] $_{spillover\ segment\ 2}$.
Lucien went to jail on Saturday night. [*The guard*] [*talked*] [*to him for a while*].

Just like Singh et al., 2016, reading times above 3,000 ms and below 100 ms have been excluded from the final dataset, resulting in the suppression of 5.4% of data and a dataset of 1566 datapoints (the dataset is available at https://osf.io/5peyh/). The data were logarithmically transformed to meet the assumptions of mixed effects model analyses (i.e., homoscedasticity, linearity and normality). Mixed effects modeling was conducted on RStudio (R Core Team, 2019, version 3.6.0) and implemented by the *lmer* function of the lme4 package (Bates et al., 2015a).

Model Selection

The model's fixed predictors are the context condition (*plausible* or *implausible*) and the article condition (*definite* or *indefinite*), in interaction. Subjects and stimuli were included in the models as random predictors with random intercepts. This choice was made because, in the present repeated measured design, both subjects and stimuli created nonindependence in the data. In addition, we included by-subjects and by-stimuli random slopes for the interaction of the fixed predictors (context condition and article condition), because these fixed predictors vary both within-subjects and within-stimuli. This choice of random slopes, based on the experimental design, is motivated by the recommendations of (Barr et al., 2013:263) and Brauer and Curtin (2018:402–403)[11]. Thus, reading time measures were assessed by the following maximal mixed effect models: model

---

[11]We acknowledge the fact that model selection, and specifically the choice of random effects structure, is subject to on-going debates among experts in the field. Most of them, however, agree on first determining the maximal model effect structure according to the experimental design (Brauer and Curtin, 2018).

TABLE 1 | Parsimonious models for reading measures on the three segments of Experiment 1.

| Reading measures segment	Parsimonious model
Critical segment	lmer (log critical segment ~ context * article + (article ‖ subjects) + (context + article ‖ stimuli))
First spillover segment	lmer (log spillover 1 ~ context * article + (context + article ‖ subjects) + (context + article ‖ stimuli)
Second spillover segment	lmer (log spillover 2 ~ context * article + (context ‖ subjects) + (article + context:article ‖ stimuli)

*Parsimonious models were selected after a random effect Principal Component Analysis, estimation of goodness of fit with likelihood ratio test and AIC criterion (Bates et al., 2015a; Matuschek et al., 2017). For all models, convergence was achieved with the built-in optimizer "bobyqa". Details of model selection are available in the **Supplementary Material**.*

< - lmer (log reading times ~ context * article + (context * article | subjects) + (context * article | stimuli).

The maximal mixed effect model failed to converge for reading time measures of the three segments analyzed (cf. **Supplementary Materials**), most likely because of the complexity of the random effects structure. The number of parameters estimates for the maximal model is 25, which might have been too high, given the number of datapoints (1566), to reach a stable maximum likelihood estimation within a reasonable number of iterations (Barr et al., 2013; Bates et al., 2015b; Brauer and Curtin 2018; Winter, 2019). Following Brauer and Curtin (2018: 404), we applied the "remedies" in hierarchical order to achieve convergence. The first effective remedy consisted in optimizing the number of iterations to reach a stable estimation. The models achieved convergence by using the built-in optimization procedure "bobyqa" of the lme4 package (Bates et al., 2015a). However, although this optimizer enabled to achieve convergence, the models still resulted in a singular fit (Bates et al., 2015b). According to Bates et al. (2015a), this is a hint that the models were overparametrized and that they should be reduced to arrive at parsimonious models in order to balance the Type I error rate and statistical power (Bates et al., 2015b; Matuschek et al., 2017). For this reason, we performed a random effect Principal Component Analysis with the *rePCA* function of the lme4 package (Bates et al., 2015a) and estimated goodness of fit with likelihood ratio test (LRT) and AIC criterion (Bates et al., 2015b; Matuschek et al., 2017). **Table 1** summaries the resulting parsimonious models for reading measures of the three segments. The details of model selection and comparison are available in the **Supplementary Materials**.

Reading Times on the Critical Segment

Reading times on the critical segments were expected to be longer in implausible conditions compared to plausible conditions for both definite and indefinite descriptions (cf. Prediction 1). We thus expected to find a main effect of context plausibility on reading times. This effect was indeed observed for the critical segment: reading times were longer, of about 91 ms, when the segments were implausible ($M = 1046.56$ ms, $SD = 556.33$) than when they were plausible ($M = 1055.38$ ms, $SD = 521.80$), $t (86.6) = -3.23$, $p = 0.002$ (see **Figure 1**; **Table 2**). Specifically, contrast analyses revealed that reading times in implausible-definite conditions ($M = 1149.27$ ms, $SD = 548.18$) were significantly longer than in plausible-definite conditions ($M = 1052.77$ ms, $SD = 519.15$), $t (86.6) = -3.23$, $p = 0.002$. These results replicate (Singh et al., 2016) findings and confirm Prediction 1 (i.e., reading times will be significantly longer in implausible conditions,
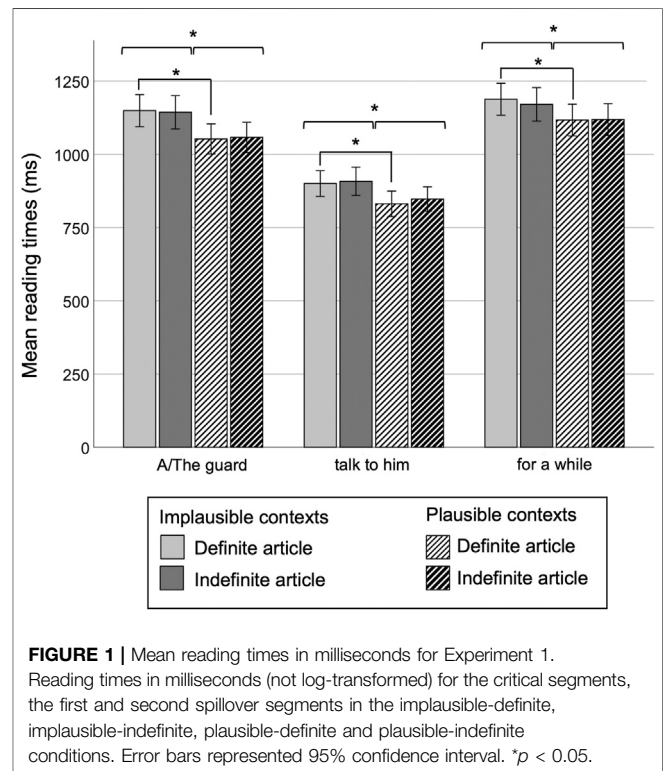
FIGURE 1 | Mean reading times in milliseconds for Experiment 1. Reading times in milliseconds (not log-transformed) for the critical segments, the first and second spillover segments in the implausible-definite, implausible-indefinite, plausible-definite and plausible-indefinite conditions. Error bars represented 95% confidence interval. *$p < 0.05$.

compared with plausible conditions, indicating increased processing costs).

However, no interaction effect was observed between context plausibility and the type of description, $t (1368) = -0.01$, $p = 0.995$ (see **Figure 1**; **Table 2**). Contrast analyses[12] were nonetheless conducted to investigate if definite compared to indefinite descriptions are more difficult to process in implausible context conditions (cf. Prediction 2). These analyses revealed that within the *implausible* conditions, segments containing a definite article were read slightly more slowly ($M = 1149.27$ ms, $SD = 548.18$) than the segments containing an indefinite article ($M = 1143.83$ ms, $SD = 565.15$). However, no significant effect was found, $t (80.5) = 0.12$, $p = 0.904$. The difference obtained is weaker than the one observed by Singh and colleagues, despite the fact that no participant was excluded due to a "stop-making-sense" task (i.e., for Singh and colleagues, a difference of 28 ms

---

[12]Contrast analyses were performed with the *difflsmeans* function of the lmerTest package (Kuznetsova et al., 2017).

TABLE 2 | Statistical results of the selected parsimonious models for Experiment 1.

Fixed effects							Random effects		
	Estimate	SE	CI (95%)	t-value	DF	p-value		Var.	SD
Critical segment									
(Intercept)	6.952	0.046	[6.86, 7.04]	152.01	99.90	<0.001	Subjects intercept	0.092	0.303
Context – Plausible	−0.085	0.026	[−0.14, −0.03]	−3.23	86.63	**0.002**	Subjects slope (article)	0.005	0.067
Article – Indefinite	−0.003	0.027	[−0.06, 0.05]	−0.12	80.49	0.904	Stimuli intercept	0.011	0.105
Interaction	−0.0002	0.035	[−0.07, 0.07]	−0.01	1368	0.995	Stimuli slope (context)	0.002	0.046
							Stimuli slope (article)	0.002	0.039
First spillover segment									
(Intercept)	6.713	0.044	[6.63, 6.80]	153.28	86.69	<0.001	Subjects intercept	0.105	0.324
Context – Plausible	−0.092	0.028	[−0.15, −0.04]	−3.30	68.10	**0.002**	Subjects slope (context)	0.008	0.089
Article – Indefinite	−0.005	0.026	[−0.06, 0.05]	−0.18	84.56	0.862	Subjects slope (article)	0.010	0.100
Interaction	0.030	0.032	[−0.03, 0.09]	0.93	1323	0.352	Stimuli intercept	0.003	0.057
							Stimuli slope (context)	0.004	0.060
							Stimuli slope (article)	0.001	0.027
Second spillover segment									
(Intercept)	6.991	0.054	[6.88, 7.10]	129.97	99.90	<0.001	Subjects intercept	0.084	0.289
Context – Plausible	−0.078	0.023	[−0.12, −0.03]	−3.33	86.63	**0.001**	Subjects slope (context)	0.0002	0.012
Article – Indefinite	−0.021	0.026	[−0.07, 0.03]	−0.81	80.49	0.423	Stimuli intercept	0.034	0.184
Interaction	0.027	0.035	[−0.04, 0.10]	0.80	1368	0.427	Stimuli slope (article)	0.003	0.053
							Stimuli slope (context*article)	0.002	0.048

DF stands for degrees of freedom, SE for standard error, CI for confidence interval, Var. for Variance and SD for standard deviation. Values in bold are significant at p < 0.05 (calculated using Satterthwaite's approximations). The selected mixed effects models are presented in **Table 1**.

was found between implausible-definite and implausible-indefinite conditions, cf. (Singh et al., 2016: 618). However, let us note that, just like Singh and colleagues, this trend was not observed either between plausible-definite ($M = 1052.77$ ms, $SD = 519.15$) and plausible-indefinite ($M = 1057.97$ ms, $SD = 525.05$) conditions, $t (76.3) = 0.13, p = 0.896$. These results do not confirm Prediction 2 (i.e., within implausible conditions, reading times will be significantly longer for implausible-definite condition, compared to implausible-indefinite condition).

Reading Times on the First and Second Spillover Segments

Reading times of the two spillover segments were analyzed to assess whether processing difficulty of the critical segment persisted (or emerged) after reading it. Similar to the critical segment, we predicted longer reading times in implausible conditions compared to plausible conditions for both definite and indefinite descriptions. A main effect of context plausibility was still observed on the first and second spillover segments. For the first spillover segment, reading times were slower, of about 65 ms, in implausible conditions (both with definite or indefinite articles) ($M = 904.19$ ms, $SD = 460.78$) than in plausible conditions ($M = 839.30$ ms, $SD = 433.01$), $t (68.1) = 3.30, p = 0.002$ (see **Figure 1**; **Table 2**). For the second spillover segment, reading times were also slower (of about 60 ms) in the implausible conditions ($M = 1176.96$ ms, $SD = 551.90$) compared to the plausible conditions ($M = 1117.13$ ms, $SD = 543.55$), $t (272.1) = 3.33, p = 0.001$. These results replicate Singh et al., 2016 findings and confirm Prediction 1.

Again, no interaction effect was observed between context plausibility and the type of descriptions for both, the first spillover segment, $t (1323) = 0.93, p = 0.352$, and the second spillover

segment, $t (1368) = 0.80, p = 0.427$. To assess a possible difference between definite and indefinite descriptions in implausible context conditions, we conducted contrast analyses. These analyses revealed no significant difference between segments in the implausible-definite condition (first spillover: $M = 900.50$ ms, $SD = 440.59$; second spillover: $M = 1187.90$ ms, $SD = 544.08$) and segments in the implausible-indefinite condition (first spillover: $M = 907.91$ ms, $SD = 480.88$; second spillover: $M = 1165.89$ ms, $SD = 560.19$), $t (84.6) = 0.17, p = 0.862$, for the first spillover, and $t (64.7) = 0.81, p = 0.423$, for the second spillover segment.

Discussion

Experiment 1 assessed the effects of context plausibility on the processing of informative definite articles. The results show a general effect of context plausibility on reading times for the critical segment and the two spillover segments that follow. This suggests that implausible segments trigger longer reading times because of increased processing costs. The increase of processing costs in implausible contexts were observed immediately when reading the critical segment and spilled over the following segments. These results confirm Prediction 1, namely that implausible contexts increase the processing costs of a new discourse referent when the preceding context is implausible. These findings provide a cross-linguistic support to Singh and colleagues' observations as well as those of other works on this topic (Gibson and Pearlmutter 1998; Tanenhaus and Trueswell 1995; Trueswell et al., 1994, *inter alia*). However, Prediction 2 is not confirmed by our findings: indeed, processing times in implausible-definite conditions were not significantly costlier than in implausible-indefinite conditions. These results persisted, despite the fact that we did not dismiss participants due to a "stop-making-sense" task (as it was the case for Singh and colleagues).

Overall, these results replicate Singh et al., 2016 findings and reveal that the observed phenomenon applies cross-linguistically. The results show that the plausibility of context weighs on processing costs of a noun, regardless of whether the following content is asserted or presupposed. However, let us underline that there is a slight tendency to process information more slowly in implausible-definite than in implausible-indefinite conditions. This goes in line with the view that the processing ease of definite descriptions depends on semantic proximity.

In the next section, we present a second experiment that aims at providing more insights regarding the comprehension vs. the acceptance of informative definite descriptions, in plausible and implausible conditions, with measures of online and offline processing.

EXPERIMENT 2: EYE-TRACKING TASK

Methods

Participants

Similar to Experiment 1, we conducted a sample size estimation for a "counterbalanced design" (Westfall et al., 2014:2026) with the "standard case" values of VPCs. However, because of the technical requirements of eye-tracking studies, we reduced the sample size while keeping a reasonably high power of 0.85 and a medium effect size of $d = 0.50$. Consequently, the sample size estimation was conducted prior to data collection, with a power set at 0.85, a medium effect size of $d = 0.50$ and with a number of 24 stimuli. The estimation revealed that 47.6 participants were required to achieve the pre-set power and effect size. We thus recruited a total of 48 participants with normal or corrected-to-normal vision, in a French-speaking university (University of Neuchatel, Switzerland) through advertisement by e-mail and on social networks. Again, only French-native speakers were eligible to take part in the experiment. We excluded data from participants who had an accuracy rate for comprehension questions lower than 65% ($n = 4$) like in Experiment 1 and similarly to Singh and colleagues' study (2016). The final sample size included 44 participants, comprising 24 women and 20 men (mean age = 23.30 years, $SD = 3.91$).

Design and Hypotheses

The self-paced reading task in Experiment 1 provided information that was limited to the online processing costs of definite descriptions. This type of data is not sufficiently informative to evaluate whether the participants produce inferences dedicated to the acceptance of the utterance, as opposed to its comprehension (cf. general assumptions in the Introduction).

Thus, in Experiment 2, we employed an eye-tracking task to assess the costs dedicated to the comprehension *and* the acceptance of definite descriptions compared to indefinite descriptions, while varying context plausibility. As argued previously (cf. *Introduction*), difficulty in the comprehension of an utterance should have an effect on online processing measures whereas difficulty in the acceptance of an utterance should have an impact on offline processing measures. The

overarching prediction is that context plausibility and the type of descriptions (definite or indefinite) will affect the gaze pattern of critical "areas of interest" (henceforth AOI).

For measures of online and offline processing, the stimuli were divided into two regions of analysis: (1) the noun phrase region of the target sentence (introducing a definite or indefinite description) and (2) the spillover region (see example (9) below). The spillover region assesses if a difficulty persists after the noun phrase region (Liversedge et al., 1998). For measures of offline processing only, we added a third area of analysis, region (3) which is the contextual phrase (see example (9) below). This specific AOI was not analyzed for measures of online processing. Indeed, this region was relevant only *after* the reader encountered the critical noun (i.e., AOI 1). The three AOIs can be found below, with their respective numerals written in subscript (i.e., noun phrase region$_1$; spillover region$_2$; context region$_3$):

9) Lucien est allé [en prison]$_3$ samedi soir. [Le garde]$_1$ [a parlé]$_2$ avec lui pendant un moment.

   Lucien went [to jail]$_3$ on Saturday night. [The guard]$_1$ [talked]$_2$ to him for a while.

With respect to online processing measures i.e., measures of first fixation duration, first-pass duration and regression path time, we expect to find similar results as in Experiment 1 on AOI 1 and AOI 2: (1) reading times of definite and indefinite descriptions will be significantly longer in implausible conditions compared to plausible conditions (i.e., a main effect of context plausibility is expected), and (2) within implausible conditions, reading times will be significantly longer for definite descriptions, compared to indefinite descriptions; within plausible conditions, no significant difference is expected between definite and indefinite descriptions (i.e., an interaction effect is expected between context plausibility and the type of description).

The two predictions above are related to the ones presented in Experiment 1. A replication of the effects with eye-tracking measures would corroborate the idea that plausibility is a significant variable weighing in the processing of definite descriptions. Furthermore, it would provide additional support for the idea that presuppositions are conventionally encoded in the lexical meaning, as opposed to being the result of a pragmatic inference (cf. Domaneschi and Di Paola 2018: 485). Finally, a replication of these results would also show the compatibility between two different methodologies for the study of this phenomenon i.e., self-paced reading and eye-tracking tasks.

The following predictions were formulated for offline processing measures, namely processing time of the whole sentences and total gaze duration on AOI 1, AOI 2, and AOI 3:

Prediction 3: Processing time of sentences and total gaze duration will be significantly longer in implausible conditions than in plausible conditions, suggesting that participants encountered more difficulties to accept

implausible contexts compared to plausible contexts (i.e., a main effect of context plausibility is predicted).

Prediction 4: Within implausible conditions, definite descriptions, unlike indefinite descriptions, should trigger the production of inferences dedicated to the acceptance of an utterance, resulting in an increase of offline processing costs. Within plausible conditions, this effect of definite descriptions should not be observed (i.e., an interaction is predicted between context plausibility and the type of description).

The measures of total gaze duration and processing time of sentences differ from the measures of online processing (of predictions 1 and 2), that focused on first fixation durations, first-pass durations and regression path times. The measures of total gaze duration and processing time of sentences provide insights regarding the specific processing patterns of participants after they read complete sentences. For instance, it can inform whether participants looked more at the context region (i.e., the location noun) *after* reading the implausible target sentence. A confirmation of predictions 3 and 4 would provide evidence in favor of the view that participants produce inferences to accept an utterance after they understood it.

Overall, the distinction between online and offline processing can provide some insights regarding the comprehension vs. acceptance distinction (Sperber et al., 2010): as argued earlier, to the extent that our stimuli involve a contrast between plausible (acceptable) vs. implausible (less acceptable) informative definite descriptions, the observation of additional processing costs is more likely to be due to the search of an appropriate context (to make the sentence acceptable) than to the computation of a conversational implicature (which, if it were the case, would also be found in plausible contexts). Furthermore, Experiment 2 should allow to see whether the inferences produced to accept an utterance are specific to the presupposed conditions (implausible-definite condition) or if it they are also produced for asserted content (implausible-indefinite condition).

Materials
The 24 stimuli and 22 fillers (cf. **Supplementary Materials**) were the same as the ones used in Experiment 1 and in Singh et al. (2016), with the exception that the context and target sentences were presented as a whole (not in segments). Similar to Experiment 1, participants had to answer yes/no comprehension questions to the filler sentences in order to encourage attentive reading.

Apparatus
Eye movements were recorded with a video-based eye-tracking device, SMI RED 5 (SensoMotoric Instruments, Teltow, Germany). To minimize head movements, participants were asked to place their head on a chinrest at a distance of 60 cm from a 22″ screen and the eye-tracking device. With these settings, 3 characters represented a visual angle of approximately 1°. Eye data were recorded binocularly and

non-invasively at a sampling rate of 500 Hz. A 5-point calibration was performed before stimuli presentation.

Procedure
The experiment was constructed with the SMI Experiment Center software (SensoMotoric Instruments, Teltow, Germany). Before beginning the experiment, participants were asked to indicate their age, gender and mother tongue. Similar to Experiment 1, a trial started with a fixation cross presented for 500 ms in the middle of the screen. Then, the context and target sentences appeared on the screen, written in black with a 32-point Arial font on a light gray background. The stimuli were presented on two lines, with the first line containing the context sentence and the second line containing the target sentence. The mean number of characters was 42 (±6) for the first line and 44 (±8) for the second line. Participants were instructed to read the sentences for comprehension and to press the space bar in order to display the next set of sentences.

The 24 stimuli and the 22 fillers were presented randomly. As in Experiment 1, participants saw only one version of each stimulus and as many stimuli i.e., six stimuli, from each of the four conditions, resulting in a within-subjects and within-stimuli design (Brauer and Curtin, 2018). Comprehension questions followed immediately their corresponding sentence. Participants recorded their answers by pressing on the "E" or "I" keys on the keyboard, following the location of the yes/no answer on the screen. Six practice trials and one comprehension question were used to familiarized participants with the task.

Results
Data Analysis
The effect of definite articles and context plausibility on information processing was investigated for both online and offline processing measures. In order to assess offline processing, measures of total gaze duration and the time to process the two presented sentences were computed. Total gaze duration consists in the sum of all fixation durations occurring on an AOI (cf. example (9) above). Processing time of sentences is the time participants spent reading and re-reading each stimulus until they moved on to a new stimulus. Online processing was assessed with measures of first fixation duration, first-pass reading time and regression path time. First fixation duration consists in the duration of the first fixation occurring on an AOI. First-pass reading time is the sum of all fixation durations from first entering a given AOI from the left until exiting it in any direction. Regression path time is the sum of all fixation durations on an AOI from entering it on the left and leaving it to the right, including any fixation made on previous regions.

The AOIs (cf. example (9) above) were delimited with the SMI software BeGaze Analysis (SensoMotoric Instruments, Teltow, Germany). Fixations were detected using the SMI built-in algorithm based on velocity, with a threshold set at 40°/s. Fixation shorter than 40 ms were discarded by the algorithm[13].

---

[13]According to Rayner (2009), typical minimal fixation duration during reading is around 50 ms.

Although recording of eye movement was binocular, only the data from the dominant eye were analyzed.

The data were logarithmically transformed to meet the assumptions of mixed effects model analyses. The final dataset was composed of 1056 datapoints (the data of Experiment 2 are available at https://osf.io/5peyh/). Mixed effects modeling was conducted on RStudio (R Core Team, 2019, version 3.6.0) and implemented by the *lmer* function of the lme4 package (Bates et al., 2015a).

Model Selection

Similar to Experiment 1, the model's fixed predictors are the context condition (*plausible* or *implausible*) and the article condition (*definite* or *indefinite*), in interaction. Random effects structure was selected based on the experimental design, as recommended by Barr et al. (2013):263) and Brauer and Curtin (2018): 402–403). Subjects and stimuli were thus included in the models as random predictors with random intercepts, because both subjects and stimuli created nonindependence in the data due to the experimental design (a repeated measures design). In addition, by-subjects and by-stimuli random slopes for the interaction of the fixed predictors were included, because these fixed predictors vary both within-subjects and within-stimuli (Brauer and Curtin, 2018). Eye movements measures were thus assessed by the following maximal mixed effect model: model < - lmer (log eye movement measure ~ context * article + (context * article | subjects) + (context * article | stimuli).

The maximal mixed effect models failed to converge for almost all measures (except from measure of total gaze duration on AOI 1, regression path time on AOI 1 and AOI 2, cf. **Supplementary Material**). Convergence could not be reached most likely because of the complexity of the random effects structure, that includes 25 parameters estimates. This high number of parameter estimates, given the number of datapoints (1056) prevented the model from reaching a stable maximum likelihood estimation within a reasonable number of iterations (Barr et al., 2013; Bates et al., 2015b; Brauer and Curtin, 2018; Winter 2019). Following Brauer and Curtin (2018), we applied the "remedies" in hierarchical order to achieve convergence. The first effective remedy consisted in optimizing the number of iterations to reach a stable estimation. The models achieved convergence by using the built-in optimization procedure "bobyqa" or "nlminbwrap" of the lme4 package (Bates et al., 2015a). In addition, all maximal models resulted in a singular fit, indicating that the models were over parametrized and should be reduced to arrive at parsimonious models in order to balance Type I error rates and statistical power (Bates et al., 2015b; Matuschek et al., 2017). For this reason, we performed for each model a random effect Principal Component Analysis with the *rePCA* function of the lme4 package (Bates et al., 2015a) and estimated goodness of fit with likelihood ratio test (LRT) and AIC criterion (Bates et al., 2015b; Matuschek et al., 2017). **Table 3** present the parsimonious models selected from this procedure for each eye movement measure. The details of model selection and comparison are available in the **Supplementary Material**.

Online Processing Measures
First Fixation Duration
No main effect of context plausibility, nor of the interaction between context plausibility and the type of description was observed on the critical and on the spillover regions (AOI 1 and AOI 2). Contrast analyses revealed no significant differences between conditions of the two AOIs (for AOI 1, the critical region, all $ps > 0.38$; for AOI 2, the spillover region, all $ps > 0.60$). These results suggest that Predictions 1 and 2 are wrong, i.e., first fixation durations were expected to be significantly longer in implausible conditions (compared with plausible conditions), and within implausible conditions, first fixation durations were expected to be significantly longer for definite descriptions (compared to indefinite descriptions).

First-Pass Duration
No main effect of context plausibility, nor of the interaction between context plausibility and the type of description was observed in the duration of the first passage on the critical and spillover regions. Contrast analyses revealed no significant differences between conditions of the two AOIs (for AOI 1, all $ps > 0.075$; for AOI 2, all $ps > 0.12$). These results do not support Predictions 1 and 2, i.e., first-pass durations were expected to be significantly longer in implausible conditions (compared with plausible conditions) and within implausible conditions first-pass durations were expected to be significantly longer for definite descriptions (compared to indefinite descriptions).

Regression Path Time
Again, no main effect of context plausibility, nor of the interaction between context plausibility and the type of description was observed on the critical and on the spillover regions for the duration of regressions made from these regions. Contrast analyses revealed no significant differences between conditions of the two AOIs (for AOI 1, all $ps > 0.39$; for AOI 2, all $ps > 0.57$). These results do not support Predictions 1 and 2, i.e., regression path time was expected to be significantly longer in implausible conditions (compared to plausible conditions) and within implausible conditions regression path time was expected to be significantly longer for definite descriptions (compared to indefinite descriptions).

Offline Processing Measures
Processing Time of Sentences
Processing times were affected by the plausibility of the context. Processing times were in general longer for sentences in implausible conditions ($M = 6765.32$ ms, $SD = 3382.30$) than in plausible conditions ($M = 6208.57$ ms, $SD = 3027.26$), $t(91.5) = 3.33$, $p = 0.001$. These results support Prediction 3, namely that processing times of sentences are longer in implausible conditions compared to plausible conditions.

However, no effect of the interaction between context plausibility and the type of description was observed on processing times (see **Table 4**). To further investigate implausible context conditions, we conducted contrast analyses. The implausible-definite condition tended to be

TABLE 3 | Parsimonious models selected for eye movements measures on the AOIs of Experiment 2.

Eye movement measure	Parsimonious model
Processing time (whole sentence)	lmer (log processing time ~ context * article + (article ‖ subjects) + (context ‖ stimuli))
Total gaze duration	
AOI 1 (noun phrase)	lmer (log gaze duration AOI 1 ~ context * article + (context ‖ subjects) + (article ‖ stimuli))
AOI 2 (spillover)	lmer (log gaze duration AOI 2 ~ context * article + (context + article ‖ subjects) + (article ‖ stimuli))
AOI 3 (contextual phrase)	lmer (log gaze duration AOI 3 ~ context * article + (1 | subjects) + (context | stimuli))
First fixation duration	
AOI 1 (noun phrase)	lmer (log first fixation AOI 1 ~ context * article + (context:article ‖ subjects) + (0 + context + article ‖ stimuli))
AOI 2 (spillover)	lmer (log first fixation AOI 2 ~ context * article + (article ‖ subjects) + (article ‖ stimuli))
First-pass duration	
AOI 1 (noun phrase)	lmer (log first-pass AOI 1 ~ context * article + (1 | subjects) + (1 | stimuli))
AOI 2 (spillover)	lmer (log first-pass AOI 2 ~ context * article + (article ‖ subjects) + (1 | stimuli))
Regression path time	
AOI 1 (noun phrase)	lmer (log regression path time AOI 1 ~ context * article + (1 | subjects) + (1 | stimuli))
AOI 2 (spillover)	lmer (log regression path time AOI 2 ~ context * article + (1 | subjects) + (context:article ‖ stimuli))

*Parsimonious models were selected after a random effect Principal Component Analysis, estimation of goodness of fit with likelihood ratio test and AIC criterion (Bates et al., 2015b; Matuschek et al., 2017). The built-in optimizer "bobyqa" enabled to achieve convergence for all models except for the model of total gaze duration on AOI 1, for which no optimizer was necessary, and for the model of first fixation duration on AOI 1, for which the built-in optimizer "nlminbwrap" was used. Details of model selection are available in the **Supplementary Materials**.*

TABLE 4 | Statistical results of the selected parsimonious models for processing time and total gaze duration in Experiment 2.

Fixed effects							Random effects			
	Estimate	SE	CI (95%)	*t*-value	DF	*p*-value		Var	SD	Corr
Processing time										
(Intercept)	8.716	0.056	[8.61, 8.83]	156.98	70.48	<0.001	Subjects intercept	0.091	0.301	
Context – Plausible	−0.095	0.029	[−0.15, −0.04]	−3.33	91.47	**0.001**	Subjects slope (article)	0.013	0.112	
Article – Indefinite	−0.011	0.033	[−0.08, 0.05]	−0.34	110.65	0.739	Stimuli intercept	0.015	0.124	
Interaction	0.029	0.039	[−0.05, 0.11]	0.73	932.58	0.464	Stimuli slope (context)	0.001	0.035	
Total gaze duration – Critical region (AOI 1)										
(Intercept)	6.287	0.091	[6.11, 6.47]	69.32	65.04	<0.001	Subjects intercept	0.128	0.358	
Context – Plausible	−0.203	0.054	[−0.31, −0.10]	−3.74	167.57	**<0.001**	Subjects slope (context)	0.008	0.091	
Article – Indefinite	0.007	0.054	[−0.10, 0.11]	0.13	92.42	0.894	Stimuli intercept	0.094	0.307	
Interaction	0.099	0.074	[−0.05, 0.24]	1.33	940.66	0.183	Stimuli slope (article)	0.004	0.602	
Total gaze duration – Spillover region (AOI 2)										
(Intercept)	6.217	0.069	[6.08, 6.35]	90.09	74.85	<0.001	Subjects intercept	0.097	0.311	
Context – Plausible	−0.143	0.055	[−0.25, −0.04]	−2.61	139.36	**0.010**	Subjects slope (context)	0.023	0.153	
Article – Indefinite	−0.075	0.055	[−0.18, 0.03]	−1.38	71.37	0.173	Subjects slope (article)	0.015	0.124	
Interaction	0.069	0.070	[−0.07, 0.21]	0.99	888.24	0.324	Stimuli intercept	0.033	0.180	
							Stimuli slope (article)	0.005	0.068	
Total gaze duration – Context region (AOI 3)										
(Intercept)	6.559	0.088	[6.38, 6.74]	74.59	58.69	<0.001	Subjects intercept	0.124	0.351	
Context – Plausible	−0.085	0.075	[−0.24, 0.07]	−1.13	38.81	0.264	Stimuli intercept	0.089	0.299	
Article – Indefinite	−0.0003	0.049	[−0.10, 0.10]	−0.01	963.4	0.996	Stimuli slope (context)	0.078	0.279	−0.54
Interaction	0.005	0.070	[−0.13, 0.14]	0.07	961	0.945				

*DF stands for degrees of freedom, SE for standard error, CI for confidence interval, Var. for Variance, SD for standard deviation and Corr for correlation. Values in bold are significant at p < 0.05 (calculated with Satterthwaite's approximations). The selected mixed effects models are presented in **Table 3**.*

processed more slowly ($M = 6812.46$ ms, $SD = 3552.64$) than the implausible-indefinite condition ($M = 6718.18$ ms, $SD = 3209.01$). However, no significant difference was found, $t (110.6) = 0.33$, $p = 0.739$.

With respect to plausible conditions, no significant difference was observed between definite ($M = 6180.55$ ms, $SD = 3085.75$) and indefinite descriptions ($M = 6236.59$ ms, $SD = 2973.23$), $t (110.7) = −0.55$, $p = 0.583$. These results suggest that the plausibility of context is the sole variable contributing to the acceptance of an utterance and disqualifies the hypothesis that the

definite article requires more processing than indefinite articles for the acceptance of an utterance.

Total Gaze Duration

***Noun phrase region* (*AOI 1*).** For the NP region (i.e., the critical AOI), a main effect of context plausibility was observed. Total gaze duration was longer in implausible conditions ($M = 714.77$ ms, $SD = 585.22$) than in plausible conditions ($M = 612.79$ ms, $SD = 508.55$), $t (167.6) = −3.74$, $p < 0.001$ (see **Table 4**). This result supports Prediction 3, namely that total

gaze duration should be significantly longer in implausible conditions for this specific region.

However, no main effect of the interaction between context plausibility and the type of description was observed (see **Table 4**). Contrast analyses were nonetheless performed to assess Prediction 4, namely that total gaze duration should be longer in the implausible-definite than in the implausible-indefinite condition. Contrast analyses revealed that total gaze duration in the implausible-definite condition ($M =$ 721.96 ms, $SD = $ 633.99) was not significantly longer than in the implausible-indefinite condition ($M = $ 707.58 ms, $SD = $ 533.12), t (92.4) = −0.13, $p = $ 0.894. Thus, these results do not support Prediction 4: according to our results, within implausible conditions, definite articles are not more costly than indefinite articles.

With respect to plausible conditions, we found a marginally significant difference between plausible-definite ($M = $ 598.41 ms, $SD = $ 537.26) and plausible-indefinite ($M = $ 627.16 ms, $SD = $ 478.72) conditions, t (92.4) = −1.97, $p = $ 0.052, suggesting that when the context is plausible, definite description have a tendency to be less costly.

Spillover region (***AOI 2***). For the spillover region (i.e., the region following the critical AOI), a main effect of context plausibility was observed, with longer total gaze duration for implausible conditions ($M = $ 605.87 ms, $SD = $ 426.49) than for plausible conditions ($M = $ 547.41 ms, $SD = $ 407.37), t (139.4) = 2.61, $p = $ 0.010. These results support Prediction 3 (i.e., implausible conditions trigger longer gaze durations than plausible conditions).

No main effect of the interaction between context plausibility and the type of description was found (see **Table 4**). Contrast analyses were nonetheless performed to investigate Prediction 4 and revealed that total gaze duration in the implausible-definite condition ($M = $ 623.59 ms, $SD = $ 426.46) was not significantly longer than in the implausible-indefinite, ($M = $ 588.07 ms, $SD = $ 426.61), t (71.4) = 1.38, $p = $ 0.173.

Within plausible condition, we found no significant difference between plausible-definite ($M = $ 545.38 ms, $SD = $ 403.58) and plausible-indefinite conditions ($M = $ 549.38 ms, $SD = $ 411.80), t (72.6) = 0.11, $p = $ 0.912. These results seem to falsify Prediction 4: there is no evidence that the implausible-definite condition is costlier than the implausible-indefinite condition. Overall, the results obtained for the spillover region are similar to the ones obtained for the noun phrase region (AOI 1).

Context region (***AOI 3***). The AOI introducing the context was analyzed to determine if this context region was likely to increase total gaze duration when the critical AOI that followed was implausible, as opposed to plausible. Surprisingly, no main effect of context plausibility was observed, although total gaze duration tended to be longer for implausible contexts ($M = $ 896.20 ms, $SD = $ 631.29) than for plausible contexts ($M = $ 835.91 ms, $SD = $ 616.29), t (38.1) = 1.13, $p = $ 0.264. These results disconfirm Prediction 3 for this specific region (i.e., total gaze duration should be significantly longer in implausible conditions).

Furthermore, no effect of the interaction between context plausibility and the type of description was observed (see **Table 4**). Contrast analyses revealed no significant differences between implausible-definite ($M = $ 891.94 ms, $SD = $ 626.10) and implausible-indefinite ($M = $ 900.47 ms, $SD = $ 637.62) conditions, t (963.4) = 0.01, $p = $ 0.996.

Regarding plausible conditions, again, no significant difference was observed between plausible-definite ($M = $ 836.70 ms, $SD = $ 595.18) and plausible-indefinite conditions ($M = $ 835.11 ms, $SD = $ 638), t (963.4) = −0.09, $p = $ 0.927. Altogether, these findings falsify Prediction 4 for this specific region. Overall, the results of the context region (AOI 3) are distinct from the ones of the critical and spillover regions (AOI 1, AOI 2) to the extent that they both seem to falsify Predictions 3 and 4.

Discussion

Experiment 2 aimed at investigating the effects of context plausibility on the comprehension and acceptance of informative definite descriptions, based on measures of online and offline processing, respectively. The experimental setting allowed to test our stimuli in more natural reading display than in Experiment 1. On the basis of the above results, we can discuss several findings.

First, a main effect of context plausibility was observed on offline processing measures but not on online processing measures. More precisely, the effect of context plausibility was observed for measures of processing times of sentences and measures of total gaze duration (on the noun phrase region and the spillover region but *not* on the context region). This lends support to Prediction 3 (i.e., participants will encounter more difficulties to *accept* implausible contexts compared to plausible contexts), but not to Prediction 1 (i.e., implausible contexts should increase processing costs during utterance *comprehension* compared to plausible contexts). Second, we found no evidence that implausible-definite condition was costlier than implausible-indefinite condition, for both online and offline processing measures, i.e., utterance comprehension and acceptance, respectively. The present results cannot confirm Prediction 2 (related to utterance comprehension) and Prediction 4 (related to utterance acceptance).

These results do not replicate the ones obtained in Experiment 1 with the self-paced reading task. The absence of significant effect of context plausibility on online processing measures is problematic, to the extent that they suggest an incompatibility between the two methodologies considered, namely self-paced reading (Experiment 1) and eye-tracking (Experiment 2). Let us note in addition that, for all measures of online processing, we did not find any tendency of longer reading times for implausible contexts compared to plausible contexts. However, we can draw some conclusions from the results of offline processing measures: context plausibility affects the offline processing of presuppositions. Specifically, it is only once the sentence is fully processed that implausible contexts increase the processing costs. Following the general assumptions presented earlier, this suggests that participants produce inferences for the acceptance of an utterance.

GENERAL DISCUSSION

A self-paced reading task and an eye-tracking reading task were designed to assess whether the plausibility of the context can affect the processing of informative definite descriptions. The self-paced reading task provided a cross-linguistic replication of Singh et al. (2016) findings, namely that the plausibility of the context has an effect on the processing of definite and indefinite descriptions. Specifically, when a definite description introduced a weak semantic relationship with a previously established context, this implied longer reading times for the critical region (definite NP) and the following segments (spillover regions). Experiment 1 showed that the processing of definite and indefinite descriptions slows down when the context is implausible. However, we found no evidence supporting that definite descriptions are more difficult to process than indefinite descriptions in implausible contexts.

The eye-tracking task was designed to bring more insights regarding the online and offline processing of presuppositions in a more natural reading setting. This experiment revealed no significant difference across our four conditions for online processing measures i.e., measures of utterance comprehension. We found however a significant effect of context plausibility for offline processing measures i.e., measures of utterance acceptance. In a nutshell, the plausibility of the context is, in this experiment, the sole variable weighing on the processing costs of the acceptance of an utterance. Surprisingly, these results are incompatible both with Experiment 1 (Singh et al.'s self-paced reading task) and with previous findings in psycholinguistics, which suggest that the attempt to repair a presupposition failure occurs during online processing (cf. Domaneschi and Di Paola, 2018).

Importantly, implausible conditions influenced reading times only after the sentence was fully processed. The offline processing measures of our three areas of interest (noun phrase region, spillover region and context region) suggest that listeners spend more time looking at the NP and the spillover region when the context is implausible. However, they do not spend more time looking at the previous context when it is implausible. Nonetheless, readers do spend significant cognitive efforts to make a sentence acceptable when it involves an implausible discourse referent (preceded both by a definite or an indefinite articles).

In our perspective, the crucial question that is raised here is whether the assumptions related to the slow-downs in self-paced reading and eye-tracking tasks are correct. As underlined by Miller (2015), the foundational assumptions on which self-paced reading and eye-tracking methods are based on are rarely questioned in research papers. For instance, a slow-down in self-paced reading tasks is generally attributed to a high level of engagement from the participant, implying deep processing and metacognitive strategies to evaluate and reconstruct understanding of the material in concert with previous beliefs (Miller, 2015:33). However, longer reading times in self-paced reading task could also reveal a surprise effect due to unmeant expectations of lexical predictions (Kleinman et al., 2015; Brothers and Kuperberg, 2020). Importantly, with respect to online processing measures, Experiment 2 showed no significant results on regression measures (AOI) across all conditions. As underlined by Miller (2015):37), regression measures are *par excellence* the ones that reveal extra effort to repair errors or reconsider information (see also Hyönä et al., 2002). Thus, the fact that we did not obtain any significant result for regression measures could suggest that readers understand the sentences during online processing, but present difficulties to accept them, as revealed by offline processing measures. More precisely, we propose that reading times in self-paced reading tasks assess both utterance comprehension and utterance acceptance, at the same time. It is possible that readers completing the self-paced reading task comprehended *and* accepted each segment before moving to the next one. In contrast, with eye-tracking tasks, utterance comprehension and acceptance can be investigated separately. Future studies could specifically address this possibility and compare systematically the discrepancy between the two methodologies employed in the present study, namely self-paced reading and eye-tracking methods.

CONCLUSION

The main contributions of the present paper are the following: first, we provided a cross-linguistic replication of Singh et al. (2016) findings, showing that informative definite and indefinite descriptions can be costly when it is preceded by an implausible context. However, just like Singh and colleagues, our results were not able to show an interaction between the plausibility of context and the article preceding the head noun.

Furthermore, this paper revealed a crucial incompatibility between self-paced reading and eye-tracking methodologies. The fact that the online processing results were not replicated in the eye-tracking tasks requires further investigations. The best interpretation we could provide is that the slowdowns in self-paced reading are due to a surprise effect in the course of lexical predictions or to the computation of two tasks, namely the comprehension *and* the acceptance of the segment. The fact that no significant difference was found for regression path times could support the hypothesis that participants first adopted a stance of trust to understand utterances, before filtering them in their epistemic vigilance module (Sperber et al., 2010:368). In other words, it is only once the sentence was fully processed that participants may have produced inferences in order to accept the sentence.

More investigations are required to better draw the threshold between the comprehension vs. acceptance of an utterance and assess whether presuppositions participate to either or both of these categories. This type of research would bring significant insights regarding the deceptive uses of presuppositions, in the context of commercial advertisements or political discourse. Furthermore, with respect to the "massive modular mind" hypothesis (cf. Sperber, 2005), it is worth investigating the interaction between definite descriptions and other psychological variables that are likely to affect the acceptance of an utterance, such as stereotype effects or validity effects.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://osf.io/5peyh/.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by ethics committee, University of Neuchatel (Switzerland). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2021.634362/full#supplementary-material

REFERENCES

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random Effects Structure for Confirmatory Hypothesis Testing: Keep it Maximal. *J. Mem. Lang.* 68 (3), 255–278. doi:10.1016/j.jml.2012.11.001

Barton, S. B., and Sanford, A. J. (1993). A Case Study of Anomaly Detection: Shallow Semantic Processing and Cohesion Establishment. *Mem. Cogn.* 21 (4), 477–487. doi:10.3758/bf03197179

Bates, D., Mächler, M., Ben, B., and Walker, S. (2015a). Fitting Linear Mixed-Effects Models Using Lme4. *J. Stat. Softw.* 67 (1). doi:10.18637/jss.v067.i01

Bates, D., Kliegl, R., Vasishth, S., and Baayen, H.. (2015b). 'Parsimonious Mixed Models'. ArXiv Preprint ArXiv:1506.04967.

Brauer, M., and Curtin, J. J. (2018). Linear Mixed-Effects Models and the Analysis of Nonindependent Data: A Unified Framework to Analyze Categorical and Continuous Independent Variables that Vary within-Subjects and/or within-Items. *Psychol. Methods* 23 (3), 389–411. doi:10.1037/met0000159

Bredart, S., and Modolo, K. (1988). Moses Strikes Again: Focalization Effect on a Semantic Illusion. *Acta Psychologica* 67 (2), 135–144. doi:10.1016/0001-6918(88)90009-1

Brothers, T., and Kuperberg, G. R. (2020). Word Predictability Effects Are Linear, Not Logarithmic: Implications for Probabilistic Models of Sentence Comprehension. *J. Mem. Lang.* 116, 104174. doi:10.1016/j.jml.2020.104174

Burkhardt, P. (2006). Inferential Bridging Relations Reveal Distinct Neural Mechanisms: Evidence from Event-Related Brain Potentials. *Brain Lang.* 98 (2), 159–168. doi:10.1016/j.bandl.2006.04.005

Clark, H. H. (1975). "Bridging," in Proceedings of the 1975 Workshop on Theoretical Issues in Natural Language Processing, 169–174. doi:10.3115/980190.980237

Clifton, C., Jr (2013). Situational Context Affects Definiteness Preferences: Accommodation of Presuppositions. *J. Exp. Psychol. Learn. Mem. Cogn.* 39 (2), 487–501. doi:10.1037/a0028975

Domaneschi, F., Carrea, E., Penco, C., and Greco, A. (2014). The Cognitive Load of Presupposition Triggers: Mandatory and Optional Repairs in Presupposition Failure. *Lang. Cogn. Neurosci.* 29 (1), 136–146. doi:10.1080/01690965.2013.830185

Domaneschi, F., Canal, P., Masia, V., Lombardi Vallauri, E., and Bambini, V. (2018). N400 and P600 Modulation in Presupposition Accommodation: The Effect of Different Trigger Types. *J. Neurolinguist.* 45, 13–35. doi:10.1016/j.jneuroling.2017.08.002

Domaneschi, F., and Di Paola, S. (2018). The Processing Costs of Presupposition Accommodation. *J. Psycholinguist Res.* 47 (3), 483–503. doi:10.1007/s10936-017-9534-7

Erickson, T. D., and Mattson, M. E. (1981). From Words to Meaning: A Semantic Illusion. *J. Verbal Learn. Verbal Behav.* 20 (5), 540–551. doi:10.1016/s0022-5371(81)90165-1

Frege, G. (1892). "On sense and reference," in *The philosophy of language*. Editors A. P. Martinich and D. Sosa (Oxford: Oxford University Press), 35–47.

Garrod, S., and Sanford, A. (1977). Interpreting Anaphoric Relations: The Integration of Semantic Information while Reading. *J. Verbal Learn. Verbal Behav.* 16 (1), 77–90. doi:10.1016/s0022-5371(77)80009-1

Geurts, B. (1997). Good News about the Description Theory of Names. *J. Semantics.* 14 (4), 319–348. doi:10.1093/jos/14.4.319

Gibson, E., and Pearlmutter, N. J. (1998). Constraints on Sentence Comprehension. *Trends Cogn. Sci.* 2 (7), 262–268. doi:10.1016/s1364-6613(98)01187-5

Glanzberg, M. (2005). Presuppositions, Truth Values, and Expressing Propositions. *Contextualism Philos.*, 349–396.

Grice, P. H. (1975). "Logic and Conversation," in *Syntax and Semantics 3: Speech Acts*, Editors P. Cole and J. L. Morgan (New York: Academic Press), 45–47.

Haviland, S. E., and Clark, H. H. (1974). What's New? Acquiring New Information as a Process in Comprehension. *J. Verbal Learn. Verbal Behav.* 13 (5), 512–521. doi:10.1016/s0022-5371(74)80003-4

Heim, I. (1982). The Semantics of Definite and Indefinite Noun Phrases. PhD dissertation. Amherst: GLSA, University of Massachussetts.

Hyönä, J., Lorch, R. F., Jr, and Kaakinen, J. K. (2002). Individual Differences in Reading to Summarize Expository Text: Evidence from Eye Fixation Patterns. *J. Educ. Psychol.* 94 (1), 44–55. doi:10.1037/0022-0663.94.1.44

Kleinman, D., Runnqvist, E., and Ferreira, V. S. (2015). Single-Word Predictions of Upcoming Language during Comprehension: Evidence from the Cumulative Semantic Interference Task. *Cogn. Psychol.* 79, 68–101. doi:10.1016/j.cogpsych.2015.04.001

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). LmerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* 82 (13), 1–26. doi:10.18637/jss.v082.i13

Liversedge, S. P., Paterson, K. B., and Pickering, M. J. (1998). *Eye Movements and Measures of Reading Time. Eye Guidance in Reading and Scene Perception.* Amsterdam: Elsevier, 55–75.

Loftus, E. F., and Zanni, G. (1975). Eyewitness Testimony: The Influence of the Wording of a Question. *Bull. Psychon. Soc.* 5 (1), 86–88. doi:10.3758/bf03336715

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., and Bates, D. (2017). Balancing Type I Error and Power in Linear Mixed Models. *J. Mem. Lang.* 94, 305–315. doi:10.1016/j.jml.2017.01.001

Matushansky, O. (2006). Why Rose Is the Rose: On the Use of Definite Articles in Proper Names. *Empirical Issues in Syntax and Semantics* 6, 285–307.

Mazzarella, D. (2015). Pragmatics and Epistemic Vigilance: The Deployment of Sophisticated Interpretative Strategies. *Croat. J. Philos.* 15 (44), 183–199.

Miller, B. W. (2015). Using Reading Times and Eye-Movements to Measure Cognitive Engagement. *Educ. Psychol.* 50 (1), 31–42. doi:10.1080/00461520.2015.1004068

Müller, M. (2018). Accommodation: A Cognitive Heuristic for Background Information. *Anglophonia. French J. English Linguistics*, 25. doi:10.4000/anglophonia.1491

Müller, M. (2021). "Presupposition Effects: Beyond and within Speaker's Meaning," in *Beyond Meaning. Pragmatics and Beyond*. Editors E. Ifantidou, L. de Saussure, and T. Wharton (Cambridge: Cambridge University Press).

Noveck, I. A., and Posada, A. (2003). Characterizing the Time Course of an Implicature: An Evoked Potentials Study. *Brain Lang.* 85 (2), 203–210. doi:10.1016/s0093-934x(03)00053-1

Noveck, I. A. (2004). "Pragmatic Inferences Related to Logical Terms," in *Experimental Pragmatics* (United Kingdom: Springer), 301–321.

Özge, U. (2011). "What Does it Mean for an Indefinite to Be Presuppositional?," in *International Tbilisi Symposium on Logic, Language, and Computation* (Berlin: Springer), 138–154.

Psychology Software Tools, Inc (2012). Pittsburgh, PA. [E-Prime 2.0].

R Core Team (2019). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychol. Bull.* 124 (3), 372–422. doi:10.1037/0033-2909.124.3.372

Roberts, C. (2003). Uniqueness in Definite Noun Phrases. *Linguistics Philos.* 26 (3), 287–350. doi:10.1023/a:1024157132393

Rayner, K. (2009). Eye Movements in Reading: Models and Data. *J. Eye Mov. Res.* 2 (5), 1–10. doi:10.16910/jemr.2.5.2

Russel, B. (1905). "On denoting," in *The philosophy of language*. Editors A. P Martinich and D. Sosa (Oxford: Oxford University Press), 105–113.

Saussure, L. de. (2012). Cognitive Pragmatic Ways into Discourse Analysis. The Case of Discursive Presuppositions. *Lodz Pap. Pragmatics* 8 (1), 37–59. doi:10.1515/lpp-2012-0004

Saussure, L. de. (2013). Background Relevance. *J. Pragmatics, Biases Constraints Commun. Argumentation, Persuasion Manipulation* 59, 178–189. doi:10.1016/j.pragma.2013.08.009

Schneider, C., Bade, N., and Janczyk, M. (2020). Is Immediate Processing of Presupposition Triggers Automatic or Capacity-Limited? A Combination of the PRP Approach with a Self-Paced Reading Task. *J. Psycholinguist Res.* 49, 247–273. doi:10.1007/s10936-019-09686-3

Schwarz, F. (2009). 'Two Types of Definites in Natural Language' Open Access Dissertations, September. https://scholarworks.umass.edu/open_access_dissertations/122.

Schwarz, F. (2013). Two Kinds of Definites Cross-Linguistically. *Lang. Linguistics Compass* 7 (10), 534–559. doi:10.1111/lnc3.12048

Schwarz, F. (2019). "Definites, Domain Restriction, and Discourse Structure in Online Processing," in *Grammatical Approaches to Language Processing*. Editors C. J. Clifton, K. Carlson, and J. D. Fodor (Springer), 187–208.

Seyednozadi, Z., Pishghadam, R., and Pishghadam, M. (2021). Functional Role of the N400 and the P600 in Language-Related ERP Studies with Respect to Semantic Processing: An Overview. Archives of Neuropsychiatry, January. http://profdoc.um.ac.ir/paper-abstract-1083599.html.dissertation.

Singh, R., Fedorenko, E., Mahowald, K., and Gibson, E. (2016). Accommodating Presuppositions Is Inappropriate in Implausible Contexts. *Cogn. Sci.* 40 (3), 607–634. doi:10.1111/cogs.12260

Sperber, D., and Wilson, D. (1987). *Relevance: Communication and Cognition.* 2nd Edn. Oxford; Cambridge, MA: Blackwell Publishers.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., et al. (2010). Epistemic Vigilance. *Mind Lang.* 25 (4), 359–393. doi:10.1111/j.1468-0017.2010.01394.x

Sperber, D. (2005). Modularity and Relevance. *The Innate Mind: Struct. Contents* 1, 53–68. doi:10.1093/acprof:oso/9780195179675.003.0004

Strawson, P. F. (1950). Iii.-on Referring. *Mind* LIX (235), 320–344. doi:10.1093/mind/lix.235.320

Tanenhaus, M. K., and Trueswell, J. C. (1995). "Sentence Comprehension," in *Speech, Language, and Communication*. Editors J. L. Miller and P. D. Eimas (San Diego, CA: Academic Press), 217–262. doi:10.1016/b978-012497770-9.50009-1

Trueswell, J. C., Tanenhaus, M. K., and Garnsey, S. M. (1994). Semantic Influences on Parsing: Use of Thematic Role Information in Syntactic Ambiguity Resolution. *J. Mem. Lang.* 33 (3), 285–318. doi:10.1006/jmla.1994.1014

Vallauri, E. L., Masia, V., and Masia, V. (2018). Facilitating Automation in Sentence Processing: The Emergence of Topic and Presupposition in Human Communication. *Topoi* 37 (2), 343–354. doi:10.1007/s11245-016-9417-9

Von Fintel, K. (2008). What Is Presupposition Accommodation, Again?*. *Philos. Perspect.* 22, 137–170. doi:10.1111/j.1520-8583.2008.00144.x

Westfall, J., Kenny, D. A., and Judd, C. M. (2014). Statistical Power and Optimal Design in Experiments in Which Samples of Participants Respond to Samples of Stimuli. *J. Exp. Psychol. Gen.* 143 (5), 2020–2045. doi:10.1037/xge0000014

Wilson, D. (2016). Reassessing the Conceptual-Procedural Distinction. *Lingua* 175–176, 5–19. doi:10.1016/j.lingua.2015.12.005

Wilson, D., and Sperber, D. (2004). "Relevance Theory," in *Handbook of Pragmatics*. Editors L. R. Horn and G. L. Ward (Oxford).

Winter, B. (2019). *Statistics for Linguists: An Introduction Using R*. New York: Routledge. CrossRef Full Text