



OPEN ACCESS

EDITED BY

Didier Maillat,
Université de Fribourg, Switzerland

REVIEWED BY

Caroline Jagoe,
Trinity College Dublin, Ireland
Johanna Miecznikowski,
University of Italian Switzerland, Switzerland

*CORRESPONDENCE

Magali A. Mari
✉ magali.mari@unine.ch

SPECIALTY SECTION

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

RECEIVED 03 November 2022

ACCEPTED 13 February 2023

PUBLISHED 03 March 2023

CITATION

Mari MA and Müller M-L (2023) Social cognition
and Relevance: How stereotypes impact the
processing of definite and indefinite
descriptions. *Front. Commun.* 8:1088861.
doi: 10.3389/fcomm.2023.1088861

COPYRIGHT

© 2023 Mari and Müller. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Social cognition and Relevance: How stereotypes impact the processing of definite and indefinite descriptions

Magali A. Mari* and Misha-Laura Müller

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

This paper focuses on the impact of social cognition on the processing of linguistic information. More specifically, it brings some insights to Relevance theory's construal of Meaning_{NN}, which seeks to account for non-propositional meanings. It shows, through two experiments, how gender and nationality-related stereotypes guide the processing of definite and indefinite descriptions. Experiment 1 consists of a self-paced reading task (with 59 French native speakers), introducing information confirming vs. violating gender stereotypes within a nominal phrase (NP). The NP (e.g., "chirurgien/chirurgienne", "surgeon_{male/female}") was itself introduced either by a definite article (presupposition) or an indefinite article (assertion). Results showed that information violating gender stereotypes was costlier to process than stereotype-congruent information. Moreover, when information violated gender stereotypes, definite descriptions became significantly costlier than indefinite ones, because they required the identification of a salient referent which contradicted stereotypical expectations. Experiment 2 tested the effects of definite vs. indefinite NP on processing nationality-related stereotypes in a self-paced reading task (with 49 French native speakers). Participants read definite vs. indefinite NPs referring to representatives of a country. The NP was subsequently paired with information that confirmed vs. contradicted nationality stereotypes. Results showed that information contradicting nationality stereotypes were significantly costlier to process than information confirming stereotypes. Furthermore, when information contradicted nationality stereotypes, indefinite descriptions (which promote a single occurrence reading) failed to facilitate information processing compared to definite descriptions (which promote a generalized representation of the social category). Overall, the present findings are consistent with research on stereotypes, in that they show that stereotype-incongruent information affect sentence processing. Importantly, while Experiment 1 revealed that stereotypes affected the processing of linguistic markers, Experiment 2 suggested that linguistic markers could not modulate the processing of stereotypes.

KEYWORDS

social cognition, gender stereotype, nationality stereotype, language processing, Relevance theory, massive modularity, definite description

1. Introduction

This paper lies at the intersection of social cognition and pragmatics. Using tools from the study of stereotypes, it contributes to the theoretical framework of Relevance theory (Sperber and Wilson, 1986, 2015). Relevance theory argues that meaning derivation is guided by a comprehension heuristic. When exposed to an ostensive verbal stimulus, the listener seeks for optimal relevance, minimizing processing costs, to obtain most cognitive effects through the acquisition, reinforcement, or revision of a belief (cf. Sperber and Wilson, 2015, p. 135).

Initially, Relevance theory developed Grice's theory of implicatures (Grice, 1957, 1975), providing a cognitive explanation for pragmatic inferences responsible for explicit and implicit meanings. However, in recent years, it focused more on argumentation and literary studies (Sperber et al., 2010; Mercier and Sperber, 2011; Cave and Wilson, 2018; Wharton and Strey, 2019). In this context, Relevance theory presented a new research agenda oriented toward a broader approach to ostensive communication (Sperber and Wilson, 2015): it is emphasized that an adequate theory of meaning should include not only "determinate propositions" conveyed by linguistic stimuli, but also non-propositional meanings conveyed by verbal and non-verbal cues. Among the examples mentioned, they present the following exchange, for which the levels of analysis are broader than those initially proposed in classical approaches in pragmatics:

(1) Rob: Do you live in London?

Jen: I live in Chelsea

(Sperber and Wilson, 2015, p. 144)

In the above, Jen implicitly answers Rob's question in the affirmative, given that Chelsea is a neighborhood in London. However, the relevance of the utterance will not only depend on determinate contents (on the level of explicature or implicature), but also on less determinate ones, triggered by the tone of voice or the social status of the speakers (Sperber and Wilson, 2015, p. 144). Here, Sperber and Wilson point out that when social status is manifest, they will guide inferences in different ways. For example, depending on their respective social status, Jen's utterance may express closeness because she shares more specific information about where she lives, or it may express a sense of social superiority that can be paraphrased as "I don't live in just any part of London".

According to Relevance theory, social status corresponds to "encyclopedic information", such as gender or nationality stereotypes in this study. Encyclopedic information is used by addressees to construct the context which guides them in making interpretive inferences. In response to an ostensive stimulus, the recipient constructs contextual hypotheses on the basis of information that is more or less salient, and, respectively, more or less easy to process. The construction of the context will allow the addressee to infer the premises leading to the derivation of an intentional explicature or implicature intended by the speaker (Sperber and Wilson, 1986, p. 37).

Furthermore, it should be noted that the most recent lines of research in Relevance theory argue that the comprehension heuristic should be conceived as a broader process than initially defined, accounting for less determinate meanings and

including non-verbal cues (Sperber and Wilson, 2015, p. 137). Following these new perspectives, Wilson (2016, p. 15) argues that linguistic markers may activate clusters of domain-specific modules of cognition, such as mindreading, emotion reading, or social cognition. The activation of these domain-specific modules is presumed to have an effect on the relevance-guided comprehension heuristic.

The present study aims to contribute to current discussions in Relevance theory by testing the impact of gender and nationality-related stereotypes on the processing of specific linguistic information, namely definite and indefinite descriptions.

1.1. The processing of stereotypes

While reading, one must not only visually process the written words but also understand their underlying meaning. To comprehend a text, readers draw on different sources of knowledge, namely linguistic, orthographic, and general world knowledge (Perfetti and Stafura, 2014; Kendeou et al., 2016). As pointed out by Relevance theory [cf. example (1)], making inferences lies at the core of comprehension. With respect to reading tasks: readers retrieve information from memory to construct a mental representation of a text (Graesser et al., 1994; Elbro and Buch-Iversen, 2013; Kendeou et al., 2016). The mental representation combines elements that are derived explicitly from the text, as well as elements that are implicit, coming from the readers' previously acquired knowledge (Gygax et al., 2021). As such, readers' world knowledge plays an essential role in reading comprehension.

While reading a word or a sentence, related concepts are automatically activated in semantic memory (Gerrig and McKoon, 1998; O'Brien et al., 1998; Rapp and van den Broek, 2005; Rubio-Fernández, 2013). For instance, upon reading sentence (2), concepts such as LAWYER*, LAW COURT*, or CRIMINAL* are likely to be activated and be more accessible in readers' memory:

(2) The judge sentenced a burglar to two years in prison.

Similarly, theoretical accounts of stereotyping propose that a given situation might increase the accessibility of stereotypic knowledge in memory (Gilbert and Hixon, 1991; Quadflieg and Macrae, 2011; Rees et al., 2020). For instance, upon reading sentence (2), stereotypical representations of "the judge" and "a burglar" will be activated in readers' memory, allowing readers to hold expectations about the likely traits, features, and behaviors of the two protagonists (Klein and Bernard, 2015; Beukeboom and Burgers, 2019). From this perspective, stereotypes function as heuristics as they guide expectations about members of a social category and are rapidly processed (Krieglmeyer and Sherman, 2012; Müller and Rothermund, 2014). When information violates a stereotype, more cognitive effort is required to access stereotype-incongruent information from associative memory, leading to increased processing difficulty (Banaji and Hardin, 1996; Kutas and Federmeie, 2000; Bartholow and Dickter, 2008; White et al., 2009). Importantly, previous research showed that regardless of personal opinions, people in the same context tend to be knowledgeable about the stereotypes in their culture (Devine, 1989; Lepore and Brown, 1997; Moskowitz et al., 1999; Quadflieg and Macrae, 2011;

Beukeboom and Burgers, 2019). As such, if a word or a sentence refers to a social category, readers within the same culture will spontaneously produce inferences about this social category and will most likely hold similar stereotypical expectations.

Overall, the effects of stereotype information on reading are well-documented. An important line of research assessed how gender stereotypes affect anaphora resolution of personal or reflexive pronouns (see for e.g., Carreiras et al., 1996; Kennison and Trofe, 2003; Duffy and Keir, 2004; Irmen, 2007; Esaulova et al., 2014; Reali et al., 2015). These studies showed that reading times of anaphoric pronouns were longer when stereotypical expectations about role nouns did not match the gender of the pronoun (e.g., “The firefighter burned herself while rescuing victims from the building”, Duffy and Keir, 2004, p. 553). Another line of research tested whether readers make inferences about the gender of a person upon reading a role noun (and so, not only when required by the anaphora). For instance, Garnham et al. (2002) designed a study in which readers could make inferences about the gender of a character, without involving anaphora resolution (e.g., “The soldier drove to the playgroup after work, and picked up one of the children, who said ‘Look what I did today daddy!’”, Garnham et al., 2002, p. 442, see also Reynolds et al., 2006; Lassonde, 2015). Their findings suggest that readers automatically encode gender when they are exposed to role nouns, even though gender information is not crucial for comprehension (Gygax et al., 2021). Altogether, past studies revealed that, in various languages and cultures, reading is slowed down when gender is incongruent with stereotypical representations of role nouns (e.g., in English: Garnham et al., 2002; Reynolds et al., 2006; Lassonde, 2015; in Norwegian: Gabriel et al., 2017; in German: Irmen, 2007; Esaulova et al., 2014; in Italian: Cacciari et al., 1997; in Spanish: Carreiras et al., 1996). However, because gender is considered as a primary social category¹ (Brewer, 1988; Fiske, 1998), it is not clear whether information processing would be affected by stereotypical expectations about other social categories, that are less primary than gender, such as nationality stereotypes. For this reason, the present study compared, in a first experiment, the effects of well-studied primary stereotypes (i.e., professions associated with gender). In a second experiment, we assessed whether the observed effects also apply to less studied stereotypes, such as nationality stereotypes. Both experiments were designed to assess the extent to which stereotypes impact the processing of specific linguistic information (see next section below).

1.2. Research question and hypotheses

This study builds on Singh et al.’s (2016) experiment testing the impact of plausible vs. implausible contexts on the processing of definite and indefinite descriptions. While definite descriptions trigger a presupposition of a salient referent, indefinite descriptions

merely introduce a new referent (Singh et al., 2016, p. 619, but see also Sperber and Wilson, 1986, p. 706).

According to Relevance theory, presuppositions and assertions can be distinguished in terms of foreground and background implications. While asserted contents contribute to relevance by providing additional cognitive effects, presuppositions contribute to relevance by saving efforts (Sperber and Wilson, 1986, p. 706). With respect to indefinite descriptions, they will be responsible for generating more effort because they present a noun as a new referent to the reader. This is not the case with definite descriptions, which present the noun phrase as “familiar” in context (cf. Heim, 1982; Roberts, 2003; Schwarz, 2009).

In Experiment 1, definite descriptions occur in a context that requires a bridging inference (Clark, 1975). That is to say that the referent is not explicitly mentioned in the preceding context, thereby requiring the construction of a link between the context (e.g., a hospital) and the noun (e.g., a/the surgeon). However, each context sentence was designed to have a strong semantic proximity with the target definite description, which facilitates processing (Haviland and Clark, 1974; Garrod and Sanford, 1977; Clifton, 2013; Schwarz, 2019). Bridging inferences are even easier in Experiment 2, as they involve a context introducing a superordinate concept (i.e., the name of a country), followed by a noun for a subordinate concept (i.e., the inhabitants of the country).

With regard to Singh et al.’s (2016) study, they hypothesized that implausible contexts, as in (3) below, would lead to an increased processing difficulty upon reading the following sentence. Moreover, within the implausible condition in (3), the definite description was expected to be significantly more difficult to process than the indefinite one, as it requires the identification of a salient referent in an incompatible context.

Singh et al. used two methods to test participants, namely a self-paced-reading task and a stop-making-sense task². In both methods, participants read a plausible vs. implausible context sentence (3), followed by a definite or indefinite noun phrase (henceforth NP). Implausible contexts were expected to make the target NPs significantly costlier than plausible ones. Furthermore, as mentioned above, definite NPs like (3b) were expected to be costlier than indefinite ones like (3a) within implausible contexts:

- (3) Mary went to the beach_{plausible} / office_{implausible} a few hours ago.
 (3a) A lifeguard warned her there about the weather.
 (3b) The lifeguard warned her there about the weather.

(Singh et al., 2016, p. 631)

Singh et al. observed an effect of context plausibility, where implausible contexts made the target NP (*A/The* lifeguard) significantly costlier to process than plausible ones. However, no significant difference was found when comparing definite NPs with indefinite ones. An effect was only found in the stop-making-sense task, when summing over all participants: In this

¹ Gender is considered as a primary social category because attention to gender emerges early (see for e.g., Quinn et al., 2002) and because children of 3–4 years of age are already aware of conventional gender stereotypes (see for e.g., Weinraub et al., 1984; Leinbach et al., 1997; Shutts et al., 2009).

² In the stop-making-sense task, participants were instructed to continue making words appear, segment by segment, as long as the sentences made sense. As soon as an incoming word or phrase did not make sense in the context of the preceding words/phrases, participants were asked to end the task (cf. Singh et al., 2016, p. 615).

case the proportion of dropouts was significantly higher in the presupposition condition (*The lifeguard*) than in the assertion condition (*A lifeguard*) (Singh et al., 2016, p. 617). Importantly, no effect was found between presupposition and assertion conditions in the self-paced reading task (Singh et al., 2016, p. 618). In a replication of Singh et al.'s study, using eye-tracking and self-paced-reading tasks (Müller and Mari, 2021), found significant results for plausibility effects, but no difference between definite and indefinite articles in the implausible condition, just like Singh et al.

The present study seeks to take these experiments further, using congruent vs. incongruent stereotypes instead of plausible vs. implausible contexts. The use of stereotypes, instead of context plausibility, is beneficial on two levels. First, it solves the problem of “context plausibility”, which involves effects from various possible sources (e.g., surprise, comprehension problems, or also typicality effects). Importantly, the stimuli in this experiment used only plausible contexts, thus allowing the critical variable to be isolated, excluding surprise effects or problems attributable to the comprehension of the utterance. Second, as presented in the previous section, stereotypes are widely studied and well-understood in terms of reading tasks.

Experiment 1 consisted in a self-paced-reading task, assessing the impact of gender stereotypes (i.e., a primary social category) on the processing of asserted vs. presupposed contents. More specifically, Experiment 1 aimed to replicate previous findings on the effects of gender stereotypes on reading times cross-linguistically (with French speaking Swiss participants) and sought to identify the specific time course of processing gender stereotypic information. To this end, Experiment 1 tested the following hypothesis:

Hypothesis 1: information violating gender stereotypes (4a) would be costlier to process than stereotype-congruent information (4b), within a compatible context (4).

- (4) Lucienne est allée à l'hôpital le mois dernier. (Context sentence)
 (4a) *La/Une chirurgienne* l'a opérée avec une grande précision. (Stereotype-incongruent)
 (4b) *Le/Un chirurgien* l'a opérée avec une grande précision. (Stereotype-congruent)

[Lucienne went to the hospital last month. (Context sentence)
The/A surgeon_{female} operated on her with great precision. (Stereotype-incongruent)
The/A surgeon_{male} operated on her with great precision. (Stereotype-congruent)]

Furthermore, and as in Singh et al. (2016) and Müller and Mari (2021), Experiment 1 tested whether definite NPs would lead to longer processing compared to indefinite NPs when the information contradicts a gender stereotype. In this case, the identification of a salient referent, required for definite NPs, is inconsistent with the encoding of stereotype-incongruent information.

Hypothesis 2: stereotype-incongruent NPs would be costlier to process when presupposed through a definite description (e.g., “la chirurgienne”; “the surgeon_{female}”) than when asserted through an indefinite description (e.g., “une chirurgienne”; “a surgeon_{female}”).

Experiment 2 focused on the processing of nationality-related stereotypes, i.e., a secondary social category, and their interaction with definite and indefinite descriptions. To our knowledge, only two papers have studied the processing of secondary social categories. Dickinson (2011) focused on stereotypical inferences regarding heterosexuality during reading tasks, and Lassonde (2015) assessed stereotypical expectations regarding the behaviors of social groups³. Whereas, Lassonde (2015) found that reading times were longer for information that violated stereotypical expectations about social groups, Dickinson (2011) failed to reach conclusive results. Thus, given the limited information available on secondary social categories, it is worth providing new investigations.

In Experiment 2, participants first read a context sentence introducing the name of a country. Two countries were alternatively presented, for example Italy vs. Japan, as presented below (5). The second sentence introduced a redundant NP (“A/The Italian/s” vs. “A/The Japanese”), followed by an attribute (“great seducer/s”) which was congruent (5a) or incongruent (5b) with a stereotype:

- (5) Mathilde est allée en Italie/ au Japon le week-end dernier. (Context sentence)
 (5a) Un/Les italien/s a/ont joué au/x grand/s séducteur/s durant tout le séjour. (Stereotype-congruent)
 (5b) Un/Les japonais a/ont joué au/x grand/s séducteur/s durant tout le séjour. (Stereotype-incongruent)
 [Mathilde went to Italy/Japan last weekend. (Context sentence)
 An/The Italian/s played the great seducer/s during the whole stay. (Stereotype-congruent)
 A/The Japanese played the great seducer/s during the whole stay. (Stereotype-incongruent)]

Theoretical perspectives on stereotyping propose that any kind of stereotype-incongruent information should be difficult to process because it requires more cognitive effort to access this information from associative memory (see for e.g., Banaji and Hardin, 1996; Kutas and Federmeie, 2000; Bartholow and Dickter, 2008; White et al., 2009). Drawing from this perspective, the following hypothesis was tested:

³ In Lassonde's (2015) study, the stereotype-incongruent information was introduced by a whole sentence (e.g., “The nuns said there was not enough alcohol” vs. “The rockers said there was not enough alcohol”, Lassonde, 2015, p. 161). In Dickinson's (2011) study, the stereotype-incongruent information was initiated by anaphora resolution (e.g., “Last night, in the packed movie theatre Hannah screamed loudly until her wife held her close”, Dickinson, 2011, p. 457).

Hypothesis 3: information violating nationality stereotypes should elicit longer reading times than stereotype-congruent information.

As illustrated above, the noun introducing the inhabitants of the country was preceded either by a plural definite or by an indefinite article. It should be noted that in French (in which language the study was conducted), plural definites invite a generic reading (Robinson, 2005, p. 18), thereby favoring a generalized and taxonomic representation of the social category described. However, in the present experimental setting, plural definites remain referential, thus fulfilling the condition of a presupposition (i.e., referring to a salient referent in the context)⁴. As for indefinite NPs, they favor a single occurrence reading, thus presenting information about the social category as singular in the provided context⁵.

Following Sperber and Wilson (1986, p. 706), Experiment 2 tested whether readers would save processing efforts for presupposed contents, as opposed to asserted ones:

Hypothesis 4: definite articles would be read more quickly than indefinite articles because they presuppose a referent which is highly salient (redundant in the context).

Finally, we conducted exploratory analyses to evaluate whether stereotype-incongruent information would be easier to process when introduced by an indefinite article (single occurrence reading) than by a definite one (generalized and definitional representation of the social category). These exploratory analyses aimed to evaluate whether stereotype-incongruent information was easier to process when it is under the scope of an indefinite description, as it promotes the reading of only one occurrence of an unexpected representation.

2. Experiment 1

Experiment 1 aimed to further assess the specific processing time course of gender stereotypes and to replicate previous findings (i.e., that gender stereotype-incongruent information is costly to process) cross-linguistically with French speaking Swiss participants (Hypothesis 1). Experiment 1 also investigated whether stereotype-incongruent information is costlier to process

4 Robinson (2005, p. 18) points out that generic readings in French can be encoded either by singular or plural definite descriptions. One test allowing to claim the presence of a generic reading is to see if the predicate cannot apply to an individual (*Paul est rare [Paul is rare]). In the present experimental setting, predicates can apply to an individual (e.g., Paul a joué au grand séducteur [Paul played the great seducer.]). This speaks in favor of a non-generic reading of the stimuli.

5 Grice (1975, p. 56) provides examples with indefinite articles to illustrate the phenomenon of generalized conversational implicatures [e.g., "X is meeting a woman this evening."; "X went into a house yesterday and found a tortoise (...)]. He explains that the use of the indefinite article promotes the inference that the item is unfamiliar. In the present experimental design, the use of the indefinite article includes the notion of unfamiliarity. However, it also promotes a single occurrence reading.

when it is presupposed through a definite description compared to when it is asserted through an indefinite description (Hypothesis 2).

2.1. Methods

2.1.1. Participants

For Experiment 1, 59 French speaking participants were recruited from a university in Switzerland. Only native French speakers were selected to participate in the experiment. The total sample size was set before data collection and based on the sample size estimation for "counterbalanced designs" developed by Westfall et al. (2014: 2026). The sample size estimation was conducted on Westfall and colleagues' website (<https://jakewestfall.shinyapps.io/crossedpower/>). We used the "standard case" values of variance components (VPCs; Westfall et al., 2014, p. 2025), with a power set at 0.85, a medium effect size of $d = 0.50$, and a number of 22 stimuli. The sample size estimation revealed that 58.8 participants were required. No additional participant was recruited once the pre-set sample size of 59 participants was reached. Following Singh et al. (2016) and Müller and Mari (2021), which employed the same experimental design as the current study, we excluded data from participants who had an accuracy rate for comprehension questions lower than 65%. This led to the exclusion of two participants. The final sample size resulted in 57 participants (31 women and 26 men; with an age mean of 23.87 years old, $SD = 4.29$).

2.1.2. Materials

The materials were constructed following a 2×2 design, manipulating (a) information about the social category, which either confirmed or violated stereotypical expectations, and (b) the NP introducing the social category, either with a definite article "le/la", "the" (presupposition condition), or with an indefinite article "un/une", "a/an" (assertion condition). The stimuli were created from the same model as those employed in Singh et al. (2016) and Müller and Mari (2021). Namely, the stimuli consisted in sets of two sentences written in French. The first sentence introduced a context, which was then followed by a target sentence matching or violating a gender stereotype. The target sentence introduced a specific agent marked grammatically by gender (e.g., chirurgien/chirurgienne, surgeon_{male/female}). The NP of the target sentence, i.e., the NP containing the social category concept, was introduced either with a definite article (working as a presupposition trigger) or an indefinite article (working as an assertion). In the end, each stimulus varied across four conditions which manipulated the effect of stereotypes and the article preceding the NP: (1) *stereotype-congruent and definite NP*, (2) *stereotype-congruent and indefinite NP*, (3) *stereotype-incongruent and definite NP* and (4) *stereotype-incongruent and indefinite NP* (see Table 1).

Gender stereotypes were based on a selection of role nouns tested in Misersky et al. (2014) as well as additional role nouns commonly found in French speaking Switzerland. A list of 50 role nouns were pre-tested on another sample of 36 subjects (50% self-identified as women) from the same population as the

TABLE 1 Example of a stimulus of Experiment 1 in the four experimental conditions.

Condition	Context sentence	Target sentence
Stereotype-congruent and definite NP	Lucienne est allée à l'hôpital le mois dernier.	Le chirurgien l'a opérée avec une grande précision.
	Lucienne went to the hospital last month.	The surgeon _{male} operated on her with great precision.
Stereotype-congruent and indefinite NP	Lucienne est allée à l'hôpital le mois dernier.	Un chirurgien l'a opérée avec une grande précision.
	Lucienne went to the hospital last month.	A surgeon _{male} operated on her with great precision.
Stereotype-incongruent and definite NP	Lucienne est allée à l'hôpital le mois dernier.	La chirurgienne l'a opérée avec une grande précision.
	Lucienne went to the hospital last month.	The surgeon _{female} operated on her with great precision.
Stereotype-incongruent and indefinite NP	Lucienne est allée à l'hôpital le mois dernier.	Une chirurgienne l'a opérée avec une grande précision.
	Lucienne went to the hospital last month.	A surgeon _{female} operated on her with great precision.

Vertical bars (|) indicate the separation between each segment.

participants of Experiment 1. The pre-test was run on Qualtrics (Provo, UT) and followed a procedure similar to Misersky et al. (2014). Participants had to indicate on a 5-point Likert scale their opinion about the extent to which role nouns consisted of women or men⁶. Response options included “mostly women,” “more women,” “as much women as men,” “more men,” “mostly men” (coded as 1 for “mostly women” and 5 for “mostly men”). Role nouns that obtained the smallest scores ($M = 2.23, SD = 0.32$) were selected as female stereotypes and roles nouns that obtained highest scores ($M = 3.79, SD = 0.38$) were used as male stereotypes. In total, 22 stimuli were used, half related to female role nouns and half related to male role nouns. An additional set of 24 filler sentences was used to mask the purpose of the experiment. The complete list of stimuli and fillers is available at <https://osf.io/b8h5q/>.

Stimuli were also pre-tested in terms of plausibility. A total of 34 raters indicated, via Qualtrics (Provo, UT), the probability to encounter a specific social agent in a given situation (e.g., seeing surgeons in a hospital). The questions were asked in the following form: “Si Marie va dans un hôpital, il est probable qu'elle rencontre ... chirurgien.ne(s)”, “If Mary went to the hospital, it is likely that she encounters ... surgeon(s)”. Raters could choose between “zero,” “only one,” “one or more,” “necessarily more than one” to replace the dots. For the selected stimuli, 76.4% of the raters chose “one or more”⁷, assuring that the stimuli were considered as plausible.

2.1.3. Procedure

The experiment was created with and ran on E-Prime 2.0 software (Psychology Software Tools, Inc., 2012). We masked the purpose of the study from participants by informing them that they would participate in a study that investigated the links between causal information and its effects on the perception of narrativity in a reading task. Participants were instructed to read the sentences

for comprehension. At the end of the study, the real purpose of the study was revealed.

Before running the experiment, participants were asked to indicate their age, gender, and mother tongue. The stimuli and fillers were then presented in sentence segments of 2–3 words (see Table 1), written in white 16-point Arial font on a black background. Each trial started with a white fixation cross on a black background, presented for 500 ms in the middle of the screen. The first segment then appeared on the screen. Participants would then press the spacebar to display the segments consecutively. This procedure prevented participants from displaying the whole sentence before reading it. Participants read only one condition 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). Stimuli and fillers were presented randomly. Comprehension questions were used to assess whether participants remained attentive during the whole task. Comprehension questions were asked about the filler sentences only, and directly followed the corresponding filler. Participants answered yes or no by pressing on the “E” or “I” keys on the keyboard, according to the location of the yes/no answers on the screen. The experiment started with six practice trials, including one comprehension question, to familiarize participants with the task.

2.2. Results

2.2.1. Data analysis

The effects of stereotype-congruent vs. incongruent information and the article preceding the NP on information processing were measured by reading times, i.e., the time spent reading a sentence segment before clicking on the space bar to make a new segment appear. Three segments are considered for the analysis: (a) the critical segment consisting in the stereotype-congruent/incongruent information and the definite/indefinite NP, (b) the first spillover segment that follows the critical segment, and (c) the second spillover segment [see example (6); vertical bars separate the sentence segments]. The two segments following the critical segment are traditionally included in the analysis of self-paced reading measures. In this way, it is possible to assess

6 For example, *Veillez indiquer si vous trouvez que plus de femmes ou d'hommes occupent la profession de chirurgiens/chirurgiennes*, [Please indicate whether you find that more women or men work as surgeon_{male/female}].

7 21.5% of the raters chose “necessarily more than one”, 2.1% of the raters chose “only one”, and none of the raters chose “zero”.

potential processing difficulties that emerged or persisted after reading the critical segment (Liversedge et al., 1998).

- (6) Lucienne | est allée | à l'hôpital | le mois dernier. | La chirurgienne critical segment | l'a opérée spillover 1 | avec grande précision spillover 2.
[Lucienne | went to | the hospital | last month. | The surgeon_{female} critical segment | operated on her spillover 1 | with great precision spillover 2.]

Reading times below 100 ms and above 4,000 ms were excluded from the final dataset, leading to the suppression of 1.4% of data and a final dataset of 1,238 datapoints (the dataset is available at <https://osf.io/b8h5q/>). The data were logarithmically transformed to meet the assumptions of mixed effects model analyses (i.e., homoscedasticity, linearity, and normality). Data analysis was conducted on RStudio (R Core Team, 2019, version 3.6.0), using the lme4 package (Bates et al., 2015b).

2.2.2. Model selection

Model specification was driven by the experimental design, as recommended by experts in the field (Barr et al., 2013; Winter and Wieling, 2016; Brauer and Curtin, 2018). Fixed predictors are composed of the interaction between the stereotype condition (stereotype-congruent or incongruent information) and the NP condition (definite or indefinite article). Due to the repeated measures design, both subjects and stimuli created non-independence in the data and were thus included as by-subjects and by-stimuli random effects (Brauer and Curtin, 2018, p. 401). According to Barr et al. (2013), each fixed predictor that vary within-unit should include a random slope, as well as interactions when all factors vary within-units. In the present study, the stereotype condition and the NP condition varied both within-subjects and within-stimuli. Consequently, reading times were assessed with the following maximal mixed effect model: model <- lmer (log reading times ~ stereotype * NP + (stereotype + NP + stereotype*NP | subjects) + (stereotype + NP + stereotype*NP | stimuli)).

The maximal mixed effect model for the three analyzed segments converged. For the first spillover segment convergence was reached by using the built-in optimization procedure “bobyqa” of the lme4 package (Bates et al., 2015b). This procedure has been acknowledged as one of the “remedies” that should be used to achieve convergence⁸ (Brauer and Curtin, 2018, p. 404). The maximal mixed effect models for the three segments analyzed resulted however in a singular fit. Singular fits are indicators that the models are overparametrized and that they should be reduced to parsimonious models, balancing at the same time the Type I error rate and statistical power (Bates et al., 2015a,b; Matuschek et al., 2017). We thus conducted a random effect Principal Component Analysis, using the rePCA function of the lme4 package (Bates et al., 2015b). Goodness of fit was estimated with the likelihood ratio

⁸ Failures of convergence are often due to the complexity of the random effect structure required by the experimental design. For the present study, the number of parameters estimates was 25, which might have been too high to reach a stable maximum likelihood estimation given the 1,238 datapoints (Barr et al., 2013; Bates et al., 2015a; Brauer and Curtin, 2018; Winter, 2019).

TABLE 2 Resulting parsimonious models for reading times on the three analyzed segments of Experiment 1.

Segment analyzed	Final parsimonious model
Critical segment	lmer (log critical segment ~ stereotype * NP + (stereotype*NP subjects) + (stereotype stimuli))
First spillover	lmer (log spillover1 ~ stereotype * NP + (stereotype + stereotype*NP subjects) + (stereotype stimuli))
Second spillover	lmer (log spillover2 ~ stereotype * NP + (stereotype + NP subjects) + (stereotype stimuli))

Parsimonious models were selected after a random effect principal component analysis, estimation of goodness of fit with likelihood ratio test, AIC, and BIC criteria (Bates et al., 2015a,b; Matuschek et al., 2017). Details of model selection are available at <https://osf.io/b8h5q/>.

test (LRT) and AIC/BIC criteria (Bates et al., 2015a; Matuschek et al., 2017). The resulting models for reading times of the three segments are displayed in Table 2. The details of model selection and comparison are available at <https://osf.io/b8h5q/>. We also ran models including participants’ gender to assess potential differences between self-identified male and female participants. For all three analyzed segments, we found no effect of gender. Gender was thus not included as a fixed predictor in the final models.

2.2.3. Reading times for the critical segment

The effect of stereotype-congruent and incongruent information on reading times was first assessed. The analysis revealed that there was no main effect of stereotype on reading times of the critical segment. Although reading times of stereotype-incongruent information ($M = 1,106.44$ ms, $SD = 604.44$) were longer than reading times of stereotype-congruent information ($M = 1,046.3$ ms, $SD = 571.2$), this difference was not significant, $t(115.5) = -1.64, p = 0.103$ (see Table 4).

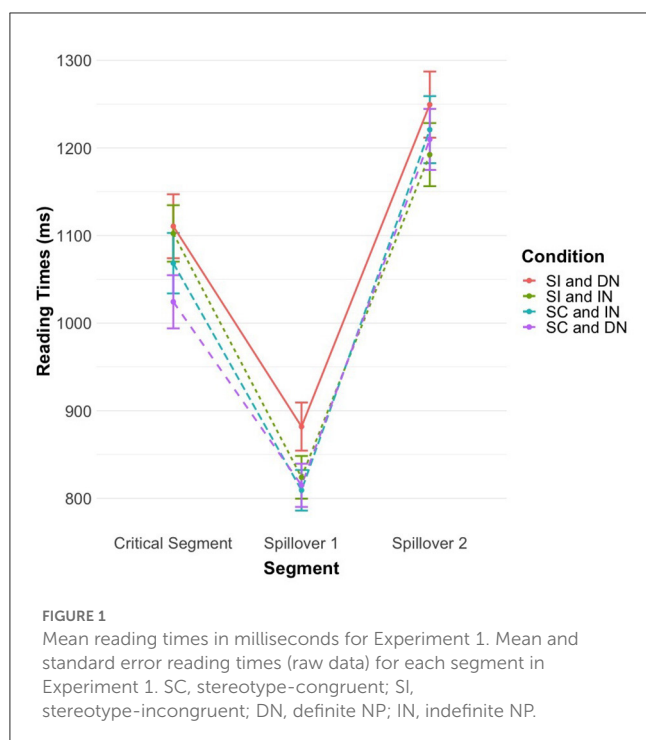
When looking at the effect of definite and indefinite NPs only, we found no significant differences again between definite ($M = 1,067.1$ ms, $SD = 589.7$) and indefinite NPs ($M = 1,085.5$ ms, $SD = 587.8$), $t(1, 107) = 0.552, p = 0.581$ (see Table 4).

No interaction effect between stereotype information and the article preceding the NP were observed, $t(347.7) = 0.109, p = 0.913$ (see Table 4).

2.2.4. Reading times for the two spillover segments

The two segments following the critical segment were analyzed to assess whether a processing difficulty emerged after reading a particular segment (Liversedge et al., 1998).

The analysis revealed a main effect of stereotype information on reading times for the first spillover segment, $t(82.12) = -2.4, p = 0.019$. Reading times of stereotype-incongruent information ($M = 852.61$ ms, $SD = 456.48$) were significantly longer than reading times of stereotype-congruent information ($M = 812.06$ ms, $SD = 420.55$). These results support Hypothesis 1, namely that information violating gender stereotypes is costlier to process than stereotype-congruent information.



A main effect of the article was also observed on the first spillover, with longer reading times after definite NPs ($M = 848.11$ ms, $SD = 458.79$) than after indefinite NPs ($M = 816.66$ ms, $SD = 418.55$), $t(1,076) = -2.18$, $p = 0.029$. No interaction effect between stereotype and NP was observed, $t(286.5) = 1.21$, $p = 0.229$ (see Table 4). Let us note that these results contradict the hypothesis of Relevance theory, namely that definite articles should be read more quickly than indefinite articles in plausible contexts. This issue is raised in the Section 2.3.

Contrast analyses were nonetheless conducted to assess whether within stereotype-incongruent conditions, longer processing times were observed with definite NPs as opposed to indefinite NPs (Hypothesis 2). These analyses revealed that reading times were significantly longer after reading stereotype-incongruent information introduced by a definite article ($M = 881.89$ ms, $SD = 480.11$) than when introduced by an indefinite article ($M = 824$ ms, $SD = 430.98$), $t(1,076) = 2.18$, $p = 0.029$ (see Figure 1). Moreover, reading an incongruent stereotype introduced by a definite article was significantly costlier than in any other condition (see Table 3).

The analysis of the second spillover segment revealed no effect of the stereotype information [$t(62.45) = -0.80$, $p = 0.429$], no effect of the NP [$t(178.65) = -1.33$, $p = 0.186$], and no interaction effect between stereotype and NP [$t(1,048) = 0.79$, $p = 0.428$] (see Table 4). These results suggest that the difficulty of processing emerged right after reading the critical segment and stopped immediately after the first spillover, namely, once the verb phrase was reached (see Figure 1).

2.3. Discussion

Experiment 1 replicated previous findings on the impact of gender stereotypes on processing times, using self-paced reading

tasks. Unlike previous studies which analyzed reading times of complete sentences (e.g., Carreiras et al., 1996; Cacciari et al., 1997; Reynolds et al., 2006; Dickinson, 2011; Lassonde, 2015) or acceptability judgements (e.g., Garnham et al., 2002; Sato et al., 2013; Gabriel et al., 2017), this study presented sentence segments of 1–3 words, allowing a moment-by-moment analysis of processing difficulty. The analysis revealed that the processing difficulty of stereotype-incongruent role nouns was delayed to the first spillover segment. This is in line with eye-tracking studies showing that reading times are significantly slowed down upon and/or after reading a pronoun that led to a mismatch between a role noun and its anaphoric pronoun (e.g., reading “electrician” followed by “she,” Reali et al., 2015, see also Kennison and Trofe, 2003; Duffy and Keir, 2004; Irmen, 2007; Esaulova et al., 2014). Experiment 1 thus replicates previous findings with French speaking Swiss participants: Information violating gender stereotypes is costlier to process than stereotype-congruent information.

Turning now to the effect of NPs, Experiment 1 revealed that definite NPs led to longer reading times than indefinite NPs on the spillover region. As noted above, these results contradict the assumption of Relevance theory, according to which definite articles should be less costly to process than indefinite articles within a plausible context. However, it should be noted that the observed longer processing time of definite articles was mainly driven by the processing of stereotype-incongruent information which generated a significant slowdown. Indeed, as revealed by contrast analyses, stereotype-incongruent information introduced by a definite NP (e.g., the surgeon_{female}) were significantly costlier than all other conditions. Within stereotype-congruent conditions (e.g., surgeon_{male}) definite articles were slightly (6 ms) costlier to process than indefinite articles. This is in line with Singh et al. (2016, p. 617), who also observed a slight slowdown with definite articles, as opposed to indefinite ones. It is likely that this experimental setup makes the processing of the definite articles costly, due to a difficulty to identify the referent in the previous context sentence. As we pointed out above (Section 1.2), the stimuli required a bridging inference, which is not necessary when the noun is preceded by an indefinite article, as it merely introduces a new referent.

Together these findings show that the processing of definite NPs, which requires the identification of a salient referent, is significantly affected by stereotypical representations. Normally, in plausible contexts, definite NPs should require little processing efforts (Sperber and Wilson, 1986, p. 706). However, the present experimental design suggests that the processing of definite descriptions interacts with social cognitive modules, generating a significant slowdown, despite a plausible context.

In sum, Hypothesis 1 was confirmed: Stereotypes are predictive of linguistic processing, where information incongruent with gender stereotypes is significantly costlier to process than stereotype-congruent information. Furthermore, Hypothesis 2 was also confirmed: Definite NPs were significantly costlier than indefinite NPs within the incongruent-stereotype condition. Regarding Hypothesis 2, it should be stressed that previous experiments on context plausibility (Singh et al., 2016; Müller and Mari, 2021) were not able to show a significant difference between definite and indefinite articles within implausible condition.

TABLE 3 Contrast analyses for the first spillover segment of Experiment 1.

Conditions	M (SD)	SC and IN	SC and DN	SI and IN
		t-test		
SC and IN	809 (406)			
SC and DN	815 (435)	$t(85.9) = 0.43, p = 0.672$		
SI and IN	824 (431)	$t(66.8) = 0.77, p = 0.447$	$t(79.2) = 0.38, p = 0.703$	
SI and DN	882 (480)	$t(66.7) = 2.72, p = 0.008$	$t(82.1) = 2.40, p = 0.019$	$t(1,076) = 2.18, p = 0.029$

SC, stereotype-congruent; SI, stereotype-incongruent; DN, definite NP; IN, indefinite NP.

3. Experiment 2

Experiment 2 assessed (a) whether information violating expectations about secondary social categories (nationality stereotypes) are costly to process as is information violating gender stereotypes (Hypothesis 3), and (b) whether definite articles are more quickly read than indefinite articles in redundant contexts (Hypothesis 4). Exploratory analyses were conducted to evaluate whether stereotype-incongruent information is easier to process when introduced by an indefinite article (single occurrence reading) than by a definite one (generalized representation of the social category).

3.1. Methods

3.1.1. Participants

For Experiment 2, 49 French speaking participants were recruited from a university in Switzerland. As in Experiment 1, only native French speakers were selected to participate in the experiment. The total sample size was set before data collection and based on a sample size estimation as conducted for Experiment 1. Using the website of Westfall et al. (2014; <https://jakewestfall.shinyapps.io/crossedpower/>), we set the values for “counterbalanced designs” as in the two previous experiments, namely with the “standard case” values of VPCs, a power of 0.80, a medium effect size of $d = 0.50$ and a number of stimuli of 20. The sample size estimation revealed that 48.8 participants were required. No additional participant was recruited once the pre-set sample size of 49 participants was reached. Similar to Experiment 1, we controlled that participants provided a minimum of 65% accuracy rate for comprehension questions. All participants responded with more than 65% accuracy. The final sample size resulted in 49 participants (28 women and 21 men; with an age mean of 23.06 years old, $SD = 3.53$).

3.1.2. Materials

The stimuli were constructed in a similar way to Experiment 1. They consisted in two sentences written in French, with the first sentence introducing the context, and the following sentence matching or violating a nationality-related stereotypes. The target sentence introduced a social category, i.e., inhabitants of a country. The NP introducing the social category, was either a definite NP (working as a presupposition, favoring a generalized and taxonomic representation of the social category) or an indefinite

NP (working as an assertion, favoring a single occurrence reading of the stereotype). As in Experiment 1, each stimulus varied across four conditions: (1) *stereotype-congruent with definite NP*, (2) *stereotype-congruent with indefinite NP*, (3) *stereotype-incongruent with definite NP*, and (4) *stereotype-incongruent with indefinite NP* (see Table 5).

Nationality-related stereotypes were based on folk stereotypes found in everyday speech (e.g., in movies, jokes, hearsay, comics, etc.) in the region of French speaking Switzerland. A list of 90 nationality stereotypes were pre-tested on another sample of 36 subjects (50% self-identified as women) from the same population as the final sample of Experiment 2. The pre-test was run on Qualtrics (Provo, UT) and asked participants to indicate on a 5-point Likert scale their opinion about diverse statements⁹. Response options included “agree,” “somewhat agree,” “neither agree nor disagree,” “somewhat disagree,” “disagree” (coded as 1 for “agree” and 5 for “disagree”). Statements that obtained the smallest scores ($M = 2.35, SD = 0.33$) were selected as nationality stereotypes and statements that obtained the highest scores ($M = 3.87, SD = 0.29$) were used as nationality counter-stereotype in the present study. In total, we used 20 stimuli, half matching nationality stereotype and half violating nationality stereotypes¹⁰. An additional set of 24 filler sentences was used to veil the purpose of the experiment. The complete list of stimuli and fillers is available at <https://osf.io/b8h5q/>.

3.1.3. Procedure

The procedure was the same as the one described in Experiment 1.

3.2. Results

3.2.1. Data analysis

As in Experiment 1, the effects of stereotype-congruent vs. incongruent information and the definite vs. indefinite article were measured by reading times. Four segments were considered

⁹ For example, *À quel point êtes-vous d'accord avec la proposition suivante: "les Japonais sont de grands séducteurs"* [To what extent do you agree with the following statement “the Japanese are great seducers”].

¹⁰ The plausibility was not pre-tested for Experiment 2, because the inhabitants introduced in the target sentence corresponded to those of the country presented in the context sentence (e.g., going to Japan and seeing Japanese is highly plausible).

TABLE 4 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)	6.874	0.049	(6.78, 6.97)	142.94	89.82	<0.001	Subjects intercept	0.09	0.30
Stereotype-congruent	-0.052	0.031	(-0.11, 0.11)	-1.64	115.5	0.103	Subjects slope (stereotype*NP)	0.01	0.09
Indefinite NP	-0.017	0.030	(-0.04, 0.08)	0.55	1,107	0.581	Stimuli intercept	0.005	0.07
Interaction	-0.005	0.045	(-0.08, 0.09)	0.11	347.7	0.913	Stimuli slope (stereotype)	0.001	0.04
First spillover									
(Intercept)	6.662	0.049	(6.56, 6.76)	137.03	76.99	<0.001	Subjects intercept	0.11	0.33
Stereotype-congruent	-0.065	0.027	(-0.12, -0.01)	-2.40	82.12	0.019	Subjects slope (stereotype)	0.004	0.06
Indefinite NP	-0.055	0.025	(-0.10, -0.005)	-2.18	1,076	0.029	Subjects slope (stereotype*NP)	0.003	0.05
Interaction	0.044	0.036	(-0.03, 0.12)	1.21	286.5	0.229	Stimuli intercept	0.003	0.06
							Stimuli slope (stereotype)	0.001	0.03
Second spillover									
(Intercept)	7.008	0.062	(6.89, 7.13)	113.52	62.41	<0.001	Subjects intercept	0.09	0.31
Stereotype-congruent	-0.026	0.033	(-0.09, 0.04)	-0.80	62.45	0.429	Subjects slope (stereotype)	0.01	0.11
Indefinite NP	-0.039	0.029	(-0.09, 0.02)	-1.33	178.56	0.186	Subjects slope (NP)	0.007	0.09
Interaction	0.030	0.038	(-0.04, 0.10)	0.79	1,048	0.428	Stimuli intercept	0.04	0.20
							Stimuli slope (stereotype)	0.004	0.06

DF, degrees of freedom; SE, standard error; CI, confidence interval; Var., Variance; SD, standard deviation. Values in bold are significant at $p < 0.05$ (calculated using Satterthwaites approximations). The selected mixed effects models are presented in Table 2.

TABLE 5 Example of a stimulus of Experiment 2 in the four experimental conditions.

Condition	Context sentence	Target sentence
Stereotype-congruent and definite NP	Mathilde est allée en Italie le week-end dernier.	Les Italiens ont joué aux grands séducteurs durant tout le séjour.
	Mathilde went to Italy last weekend.	The Italians played the great seducers during the whole stay.
Stereotype-congruent and indefinite NP	Mathilde est allée en Italie le week-end dernier.	Un Italien a joué au grand séducteur durant tout le séjour.
	Mathilde went to Italy last weekend.	An Italian played the great seducer during the whole stay.
Stereotype-incongruent and definite NP	Mathilde est allée au Japon le week-end dernier.	Les Japonais ont joué aux grands séducteurs durant tout le séjour.
	Mathilde went to Japan last weekend.	The Japanese played the great seducers during the whole stay.
Stereotype-incongruent and indefinite NP	Mathilde est allée au Japon le week-end dernier.	Un Japonais a joué au grand séducteur durant tout le séjour.
	Mathilde went to Japan last weekend.	A Japanese played the great seducer during the whole stay.

Vertical bars (|) mark presentation boundaries (i.e., sentences segments).

for the analysis: (a) the one containing the definite/indefinite NP that introduced the social category (i.e., inhabitants of a country), (b) the spillover segment to assess potential persistence of processing difficulty, (c) the segment presenting stereotype-congruent/incongruent information, and (d) its spillover segment [see example (7); vertical bars separate the sentence segments]:

(7) Mathilde | est allée | au Japon | le week-end dernier. | Les Japonais_{critical segment 1} | ont joué_{spillover 1} | aux grands séducteurs_{critical segment 2} | durant tout le séjour_{spillover 2}.
 [Mathilde | went to | Japan | last weekend. | The Japanese_{critical segment 1} | played_{spillover 1} | the great seducers_{critical segment 2} | during the whole stay_{spillover 2}.]

Similar to Experiment 1, reading times below 100 ms and above 4,000 ms have been excluded from the final dataset. This data exclusion resulted in the suppression of 1.1% of data and a final dataset of 969 datapoints (the dataset is available at <https://osf.io/b8h5q/>). The data were logarithmically transformed to meet the assumptions of mixed effects model analyses and data analysis was conducted on Rstudio (R Core Team, 2019, version 3.6.0), using the lme4 package (Bates et al., 2015b).

3.2.2. Model selection

We followed the same procedure as in Experiment 1 to specify the model (i.e., model selection based on the experimental design). The first segment under investigation in the present experiment did not mix the types of NPs and stereotype information. As illustrated in example (7), the first segment varies only in terms of the article used, namely definite or indefinite. The information violating/confirming stereotypes is only introduced in the seventh segment (critical segment 2). As a consequence, reading times on the first critical segment and the first spillover were assessed with the following maximal mixed effect model: `model1 <- lmer(log reading times ~ NP + (NP | subjects) + (NP | stimuli))`. On the other hand, reading times of the second critical segment and its spillover could be affected by both the type of NP and stereotype information. Therefore, reading times of those remaining segments were analyzed with the following maximal mixed effect model: `model2 <- lmer(log reading times ~ stereotype * NP + (stereotype*NP | subjects) + (stereotype*NP | stimuli))`.

TABLE 6 Resulting parsimonious models for reading times on the three analyzed segments of Experiment 2.

Segment analyzed	Final parsimonious model
Critical segment 1	<code>lmer(log critical segment 1 ~ NP + (NP subjects) + (0 + NP stimuli))</code>
Spillover 1	<code>lmer(log spillover1 ~ NP + (NP subjects) + (1 stimuli))</code>
Critical segment 2	<code>lmer(log critical segment 2 ~ stereotype * NP + (1 subjects) + (stereotype + NP stimuli))</code>
Spillover 2	<code>lmer(log spillover 2 ~ stereotype * NP + (NP subjects) + (NP stimuli))</code>

Parsimonious models were selected after a random effect Principal Component Analysis, estimation of goodness of fit with likelihood ratio test, AIC, and BIC criteria (Bates et al., 2015a,b; Matuschek et al., 2017). Details of model selection are available at <https://osf.io/b8h5q/>.

The maximal mixed effect model for the four analyzed segments reached convergence. For the two critical segments and the two spillover segments, the built-in optimization procedures “nlminbwrap” and “bobyqa” of the lme4 package (Bates et al., 2015b) were used, respectively. The maximal mixed effect models for the four segments resulted however in a singular fit, indicating that the models were overparametrized. We thus conducted a random effect Principal Component Analysis, using the rePCA function of the lme4 package (Bates et al., 2015b). Goodness of fit was estimated with the likelihood ratio test (LRT) and AIC/BIC criteria (Bates et al., 2015a; Matuschek et al., 2017). The resulting models for reading times of the four segments are displayed in Table 6. The details of model selection and comparison are available at <https://osf.io/b8h5q/>.

3.2.3. Reading times for the first critical and spillover segments

The analysis revealed that there was no main effect of definite or indefinite NPs on reading times of the critical segment, $t(44.23) = 1.59, p = 0.118$. The next segment was also analyzed to assess if a processing difficulty emerged after reading the definite/indefinite NPs. The analyses revealed a significant difference between reading

times, $t(51.24) = 2.49$, $p = 0.016$, with longer reading times following an indefinite NP ($M = 884.69$ ms, $SD = 493.92$) compared to a definite NP ($M = 825.08$ ms, $SD = 439.37$). This result confirms Hypothesis 4, where definite NPs were expected to be read more quickly than indefinite ones (see Table 7). These results are compatible with Relevance theory, which argues that definite descriptions allow to spare cognitive efforts (compared with indefinite descriptions).

3.2.4. Reading times for the second critical and spillover segments

The analysis revealed that there was a main effect of stereotype information on reading times of the second critical segment. Reading times of stereotype-incongruent information ($M = 1,309.26$ ms, $SD = 676.07$) were significantly longer than reading times of stereotype-congruent information ($M = 1,199.43$ ms, $SD = 610.02$), $t(51.6) = -3.05$, $p = 0.004$. A main effect of stereotype was also observed for the spillover segment, with longer reading times for the stereotype-incongruent condition ($M = 1,066.05$ ms, $SD = 538.86$) compared to the stereotype-congruent condition ($M = 993.24$ ms, $SD = 523.52$), $t(831.7) = -3.04$, $p = 0.002$ (see Figure 2). These results support Hypothesis 3, according to which stereotype-incongruent information elicit longer reading times than stereotype-congruent information. Importantly, these results provide evidence for the persistence of stereotype effects with secondary social categories.

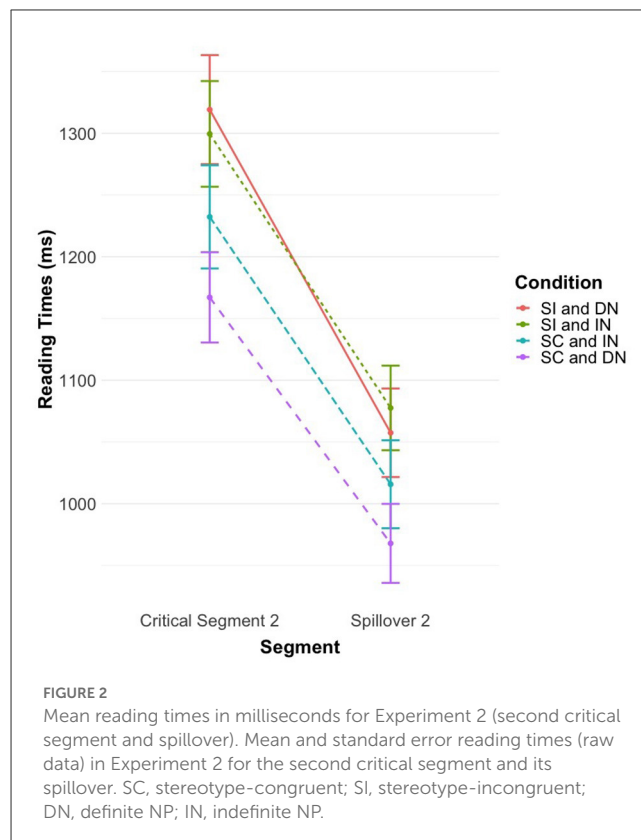
No effect of definite/indefinite NPs and no interaction effect were observed for both the second critical segment and its spillover (see Table 7).

Contrast analyses were nonetheless conducted on both segments to explore the possibility that stereotype-incongruent information might be easier to process when introduced by an indefinite article as opposed to a definite article. These analyses revealed that reading times in the *stereotype-incongruent with indefinite NP* condition were not significantly faster than in the *stereotype-incongruent with definite NP* condition, $t(64.7) = 0.27$, $p = 0.786$ (critical segment 2) and $t(49.7) = 0.43$, $p = 0.666$ (second spillover). In other words, the processing of stereotype-incongruent information does not appear to be affected by linguistic markers of definiteness (see Figure 2).

3.3. Discussion

Experiment 2 revealed that stereotype-incongruent information about nationalities is longer to process than stereotype-congruent information. This finding provides further evidence that information confirming stereotypical expectations is easily processed, whereas information violating stereotypical expectations about secondary social categories is difficult to process (Hypothesis 3). Interestingly, the effects of information confirming/violating nationality stereotype already appeared on the critical segment and persisted in the spillover segment.

Regarding the effects of definite/indefinite NPs, Experiment 2 tested whether indefinite descriptions were costlier to process than definite descriptions, as proposed by Relevance theory. The present



study confirmed Hypothesis 4, showing that definite NPs led to faster reading times than indefinite ones. This finding is all the more interesting in light of Experiment 1, where the fast reading of definite descriptions was disrupted because of information violating gender stereotypes. Finally, we explored whether indefinite articles (i.e., the representation of a single occurrence within a kind) could facilitate the processing of stereotype-incongruent information. These analyses revealed that indefinite articles could not make stereotype-incongruent information easier to process than when subjected to a generalization (i.e., a plural definite). Thus, in this experimental setup, the processing effects of stereotypes appear to be stronger than linguistic markers.

4. General discussion

The present study investigated the effects of social modules of cognition on the relevance-guided comprehension heuristic across two experiments, in order to shed light on the relevance comprehension heuristic. Both experiments assessed the extent to which stereotypes impact the processing of specific linguistic information. Experiment 1 aimed to replicate previous findings on the effects of gender stereotypes on reading cross-linguistically. Experiment 2 sought to investigate the effects of secondary social categories, i.e., nationalities.

The results of Experiment 1 showed that information violating gender stereotypes is longer to process than stereotype-congruent information (Hypothesis 1). This finding goes in line with previous studies that investigated, in different languages, the effect of gender

TABLE 7 Statistical results of the selected parsimonious models for Experiment 2.

	Fixed effects						Random effects		
	Estimate	SE	CI (95%)	t-value	DF	p-value		Var.	SD
Critical segment 1									
(Intercept)	6.783	0.044	(6.69, 6.87)	152.78	52.98	<0.001	Subjects intercept	0.08	0.29
Indefinite NP	0.050	0.031	(−0.01, 0.11)	1.59	44.23	0.118	Subjects slope (NP)	0.02	0.13
							Stimuli (NP)	0.002	0.05
Spillover 1									
(Intercept)	6.600	0.051	(6.50, 6.70)	129.10	55.97	<0.001	Subjects intercept	0.11	0.33
Indefinite NP	0.060	0.024	(0.01, 0.11)	2.49	51.24	0.016	Subjects slope (NP)	0.006	0.07
							Stimuli intercept	0.002	0.05
Critical segment 2									
(Intercept)	7.063	0.059	(6.95, 7.18)	119.07	71.81	<0.001	Subjects intercept	0.10	0.32
Stereotype-congruent	−0.103	0.034	(−0.17, −0.03)	−3.05	51.61	0.004	Stimuli intercept	0.02	0.14
Indefinite NP	−0.008	0.031	(−0.07, 0.05)	−0.27	69.63	0.787	Stimuli slope (stereotype)	0.005	0.07
Interaction	0.053	0.042	(−0.03, 0.14)	1.26	864.35	0.207	Stimuli slope (NP)	0.002	0.04
Spillover 2									
(Intercept)	6.864	0.051	(6.76, 6.97)	133.79	70.98	<0.001	Subjects intercept	0.09	0.31
Stereotype-congruent	−0.085	0.028	(−0.14, −0.03)	−3.04	831.65	0.002	Subjects slope (NP)	0.008	0.09
Indefinite NP	0.015	0.033	(−0.05, 0.08)	0.45	67.86	0.654	Stimuli intercept	0.006	0.08
Interaction	0.026	0.039	(−0.05, 0.10)	0.648	831.09	0.517	Stimuli slope (NP)	0.09	0.31

DF, degrees of freedom; SE, standard error; CI, confidence interval; Var., Variance; SD, standard deviation. Values in bold are significant at $p < 0.05$ (calculated using Satterthwaites approximations). The selected mixed effects models are presented in [Table 6](#).

stereotypes on reading and anaphora resolution (cf. Section 1.1). Our findings with French speaking Swiss participants support the cross-linguistic evidence that gender is rapidly encoded during reading (Garnham et al., 2002; Gygax et al., 2021) and affects processing depending on whether the information communicated matches one's stereotypical expectations.

Furthermore, Experiment 1 revealed that stereotype-incongruent information makes the processing of presuppositional contents (definite articles) significantly costlier than assertions (indefinite descriptions) (Hypothesis 2). This is because the identification of a salient referent (required for definite NPs) is inconsistent with the encoding of stereotype-incongruent information. Importantly, these findings offer promising opportunities for the study of the relevance comprehension heuristic. While previous studies failed to reach conclusive results when using a broad category of plausibility (Singh et al., 2016; Müller and Mari, 2021), the present study revealed that the processing of definite descriptions is affected by stereotype information, generating a significant slowdown in narrowly defined plausible contexts. We suggest that the “plausibility of contexts”, used in these previous studies, conflated different variables, such as surprise effects, comprehension problems as well as typicality effects.

In Experiment 2, we further assessed the effect of stereotypes about secondary social categories on processing, and revealed that information violating nationality stereotypes was costly to process (Hypothesis 3). This is consistent with Lassonde's (2015) study, showing that sentences containing stereotype-incongruent information about diverse social categories (e.g., nuns, rockstars) are costly to process. Together, these findings have some interesting implications, suggesting that any kind of stereotype-incongruent information would be difficult to process because more cognitive effort is required to access the information from associative memory. Unfortunately, this possibility has rarely been addressed, as most studies to date have focused only on gender stereotypes. In an endeavor to determine whether the effects of stereotypes on processing are consistent, future studies should investigate, with varying methodologies (e.g., response times, self-paced reading, eye-tracking, or event-related brain potentials) and across cultures, how and whether information processing is similarly affected by stereotype about various social categories.

Experiment 2 also revealed that definite descriptions are less costly than indefinite descriptions when the context is redundant. These results align with Relevance theory, which argues that definite descriptions allow to spare cognitive efforts (compared with indefinite descriptions). Moreover, our exploratory analysis showed that when a single occurrence was encoded linguistically (e.g., “A Japanese played the great seducer”, as opposed to “The Japanese played the great seducers”), it did not facilitate the processing of incongruent-stereotype information. This suggests that the processing of incongruent-stereotype information cannot be modulated by linguistic markers¹¹.

Before concluding, let us note that the present study's findings bear some important considerations. Given the current context and issues, it seems particularly important to study how stereotypical information is processed. For instance, although much effort and attention has been paid to gender equality in the 21st century (e.g., the increasing use of inclusive language, the promotion of STEM professions among girls and women, or strikes for women's right), we still observed that some conceptions of gender roles remain unchanged. Moreover, current crises (namely the COVID-19 pandemic, the war in Ukraine, or the climate crisis) led people to rapidly form negative stereotypes about inhabitants of certain countries. By documenting how people process stereotypes during reading, the present study showed that information confirming a stereotype is easily processed and thus, might not be questioned or noticed, while still being significant in the relevance comprehension heuristic. Furthermore, this could play a role in the maintenance of stereotypical expectations and the emergence of prejudices.

We also stress two important limitations to the current study. First, it does not allow to make direct comparisons between the two tested stereotypes. Indeed, gender stereotype in/congruence occurred within the NP (e.g., “une chirurgienne”; “a surgeon_{female}”) whereas nationality stereotype in/congruence stood in the relation between the NP and the predicate (e.g., “A/The Japanese *played the great seducer/s*”). In a future study, it would be worth testing these two stereotypes, and others, in a comparable way. Furthermore, our study focused only on narrow linguistic phenomena (definite vs. indefinite descriptions). While this may be an asset experimentally (limiting other variables weighing in the processing speed), further studies are needed to see if stereotypes also constrain the processing of other linguistic markers, such as other presupposition triggers.

Overall, the present study's findings suggest that stereotypes bring significant constraints on the processing of linguistic information. These elements are of interest for Relevance theory, insofar as they confirm that the comprehension heuristic is constrained by information which goes beyond propositional cues, such as the listener's knowledge about social categories. Across both experiments, stereotype-incongruent information was less salient than stereotype-congruent ones, making it cognitively more costly to process. Moreover, these findings suggest a possible hierarchy between social and linguistic information in the derivation of meaning: Indeed, while stereotype-incongruent information slows down the processing of definite descriptions (which are normally processed quickly), we did not observe that the processing of stereotype-incongruent information was facilitated when preceded by an indefinite article (single occurrence reading).

Data availability statement

The original contributions presented in the study are available at <https://osf.io/b8h5q/>, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Commission d'Éthique de la Recherche (CER),

¹¹ Let us note that an anonymous reviewer drew our attention to the potential problems of confounding variables in the experimental setup. For this reason, further experiments with a setup that better isolates the variables should be conducted to test this hypothesis.

Université de Neuchâtel. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Data analysis, reporting, and data interpretation was conducted by MAM. All authors equally developed the study concept and contributed to the study design. Data collection was performed by all authors. The first draft of the manuscript was written by all authors, and all authors reviewed and revised it together. All authors approved the final version of the manuscript for submission.

Funding

This research was funded by the University of Neuchâtel.

References

- Banaji, M. R., and Hardin, C. D. (1996). Automatic stereotyping. *Psychol. Sci.* 7, 136–141. doi: 10.1111/j.1467-9280.1996.tb00346.x
- 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, 255–278. doi: 10.1016/j.jml.2012.11.001
- Bartholow, B. D., and Dickter, C. L. (2008). A response conflict account of the effects of stereotypes on racial categorization. *Soc. Cogn.* 26, 314–332. doi: 10.1521/soco.2008.26.3.314
- Bates, D., Kliegl, R., Vasishth, S., and Baayen, H. (2015a). Parsimonious mixed models. *arXiv [Preprint]*. arXiv: 1506.04967.
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015b). Fitting linear mixed-effects models using Lme4. *J. Statist. Softw.* 67, 1–51. doi: 10.18637/jss.v067.i01
- Beukeboom, C. J., and Burgers, C. (2019). How stereotypes are shared through language: a review and introduction of the social categories and stereotypes communication (SCSC) framework. *Rev. Commun. Res.* 7, 1–37. doi: 10.12840/issn.2255-4165.017
- 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, 389–411. doi: 10.1037/met0000159
- Brewer, M. B. (1988). "A dual process model of impression formation," in *Advances in Social Cognition*, eds R. S. Wyer, and T. K. Srull (Hillsdale: Lawrence Erlbaum Associates), 1–36.
- Cacciari, C., Carreiras, M., and Cionini, C. B. (1997). When words have two genders: anaphor resolution for Italian functionally ambiguous words. *J. Mem. Lang.* 37, 517–532. doi: 10.1006/jmla.1997.2528
- Carreiras, M., Garnham, A., Oakhill, J., and Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: evidence from English and Spanish. *Q. J. Exp. Psychol. Hum. Exp. Psychol.* 49, 639–663. doi: 10.1080/713755647
- Cave, T., and Wilson, D. (2018). *Reading Beyond the Code: Literature and Relevance Theory*. Oxford: Oxford University Press.
- Clark, H. H. (1975). "Bridging," in *Proceedings of the 1975 Workshop on Theoretical Issues in Natural Language Processing - TINLAP '75* (Morristown, NJ: Association for Computational Linguistics), 169–174. doi: 10.3115/980190.980237
- Clifton, C. (2013). Situational context affects definiteness preferences: accommodation of presuppositions. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 487–501. doi: 10.1037/a0028975
- Devine, P. G. (1989). Stereotypes and prejudice: their automatic and controlled components. *J. Pers. Soc. Psychol.* 56, 5–18. doi: 10.1037/0022-3514.56.1.5
- Dickinson, J. (2011). The impact of 'violating the heterosexual norm' on reading speed and accuracy. *Psychology* 2, 456–459. doi: 10.4236/psych.2011.25071
- Duffy, S. A., and Keir, J. A. (2004). Violating stereotypes: eye movements and comprehension processes when text conflicts with world knowledge. *Mem. Cognit.* 32, 551–559. doi: 10.3758/BF03195846
- Elbro, C., and Buch-Iversen, I. (2013). Activation of background knowledge for inference making: effects on reading comprehension. *Sci. Stud. Read.* 17, 435–452. doi: 10.1080/10888438.2013.774005
- Esaulova, Y., Reali, C., and Stockhausen, L., v. (2014). Influences of grammatical and stereotypical gender during reading: eye movements in pronominal and noun phrase anaphor resolution. *Lang. Cogn. Neurosci.* 29, 781–803. doi: 10.1080/01690965.2013.794295
- Fiske, S. T. (1998). "Stereotyping, prejudice, and discrimination," in *The Handbook of Social Psychology*, eds D. T. Gilbert, S. T. Fiske, and G. Lindzey (Boston, MA: McGraw-Hill), 357–411.
- Gabriel, U., Behne, D. M., and Gygax, P. M. (2017). Speech vs. reading comprehension: an explorative study of gender representations in Norwegian. *J. Cogn. Psychol.* 29, 795–808. doi: 10.1080/20445911.2017.1326923
- Garnham, A., Oakhill, J., and Reynolds, D. (2002). Are inferences from stereotyped role names to characters' gender made elaboratively? *Mem. Cogn.* 30, 439–446. doi: 10.3758/BF03194944
- Garrod, S., and Sanford, A. (1977). Interpreting anaphoric relations: the integration of semantic information while reading. *J. Verbal Learning Verbal Behav.* 16, 77–90. doi: 10.1016/S0022-5371(77)80009-1
- Gerrig, R. J., and McKoon, G. (1998). The readiness is all: the functionality of memory-based text processing. *Discour. Process.* 26, 67–86. doi: 10.1080/01638539809545039
- Gilbert, D. T., and Hixon, J. G. (1991). The trouble of thinking: activation and application of stereotypic beliefs. *J. Pers. Soc. Psychol.* 60, 509–517. doi: 10.1037/0022-3514.60.4.509
- Graesser, A. C., Singer, M., and Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychol. Rev.* 101, 371–395. doi: 10.1037/0033-295X.101.3.371
- Grice, H. P. (1957). Meaning. *Philos. Rev.* 377–388. doi: 10.2307/2182440
- Grice, H. P. (1975). "Logic and conversation," in *Syntax and Semantics: Speech Acts*, eds P. Cole and J. L. Morgan (New York, NY: Academic Press), 41–58.
- Gygax, P., Sato, S., Öttl, A., and Gabriel, U. (2021). The masculine form in grammatically gendered languages and its multiple interpretations: a challenge for our cognitive system. *Lang. Sci.* 83, 101328. doi: 10.1016/j.langsci.2020.101328
- Haviland, S. E., and Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *J. Verb. Learn. Verb. Behav.* 13, 512–521. doi: 10.1016/S0022-5371(74)80003-4
- Heim, I. (1982). *The semantics of definite and indefinite noun phrases* (Ph. D. Dissertation). University of Massachusetts, Amherst, MA, United States.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Irmen, L. (2007). What's in a (Role) name? Formal and conceptual aspects of comprehending personal nouns. *J. Psycholinguist. Res.* 36, 431–456. doi: 10.1007/s10936-007-9053-z
- Kendeou, P., McMaster, K. L., and Christ, T. J. (2016). Reading comprehension: core components and processes. *Policy Insights Behav. Brain Sci.* 3, 62–69. doi: 10.1177/2372732215624707
- Kennison, S. M., and Trofe, J. L. (2003). Comprehending pronouns: a role for word-specific gender stereotype information. *J. Psycholinguist. Res.* 32, 355–378. doi: 10.1023/A:1023599719948
- Klein, O., and Bernard, P. (2015). "Stereotypes in social psychology," in *International Encyclopedia of the Social & Behavioral Sciences*, ed J. D. Wright (Amsterdam: Elsevier), 446–452. doi: 10.1016/B978-0-08-097086-8.24097-X
- Krieglmeyer, R., and Sherman, J. W. (2012). Disentangling stereotype activation and stereotype application in the stereotype misperception task. *J. Pers. Soc. Psychol.* 103, 205–224. doi: 10.1037/a0028764
- Kutas, M., and Federmeie, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trend. Cogn. Sci.* 4, 463–470. doi: 10.1016/S1364-6613(00)01560-6
- Lassonde, K. A. (2015). Reducing the impact of stereotypical knowledge during reading. *Discourse Process.* 52, 149–171. doi: 10.1080/0163853X.2014.917221
- Leinbach, M. D., Hort, B. E., and Fagot, B. I. (1997). Bears are for boys: metaphorical associations in young children's gender stereotypes. *Cogn. Dev.* 12, 107–130. doi: 10.1016/S0885-2014(97)90032-0
- Lepore, L., and Brown, R. (1997). Category and stereotype activation: is prejudice inevitable? *J. Person. Soc. Psychol.* 72, 275–287. doi: 10.1037/0022-3514.72.2.275
- 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.
- 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
- Mercier, H., and Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behav. Brain Sci.* 34, 57–74. doi: 10.1017/S0140525X10000968
- Misersky, J., Gygax, P. M., Canal, P., Gabriel, U., Garnham, A., Braun, F., et al. (2014). Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behav. Res. Methods* 46, 841–871. doi: 10.3758/s13428-013-0409-z
- Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., and Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *J. Pers. Soc. Psychol.* 77, 167–184. doi: 10.1037/0022-3514.77.1.167
- Müller, F., and Rothermund, K. (2014). What does it take to activate stereotypes? Simple primes don't seem enough. *Soc. Psychol.* 45, 187–193. doi: 10.1027/1864-9335/a000183
- Müller, M.-L., and Mari, M. A. (2021). Definite descriptions in the light of the comprehension vs. acceptance distinction: comparing self-paced reading with eye-tracking measures. *Front. Commun.* 6, 1–17. doi: 10.3389/fcomm.2021.634362
- O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., and Halleran, J. G. (1998). Updating a situation model: a memory-based text processing view. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 1200–1210. doi: 10.1037/0278-7393.24.5.1200
- Perfetti, C., and Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Sci. Stud. Read.* 18, 22–37. doi: 10.1080/10888438.2013.827687
- Quadflieg, S., and Macrae, C. N. (2011). Stereotypes and stereotyping: what's the brain got to do with it? *Eur. Rev. Soc. Psychol.* 22, 215–273. doi: 10.1080/10463283.2011.627998
- Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., and Pascalis, O. (2002). Representation of the gender of human faces by infants: a preference for female. *Perception* 31, 1109–1121. doi: 10.1068/p3331
- Rapp, D. N., and van den Broek, P. (2005). Dynamic text comprehension: an integrative view of reading. *Curr. Dir. Psychol. Sci.* 14, 276–279. doi: 10.1111/j.0963-7214.2005.00380.x
- Real, C., Esaulova, Y., Öttl, A., and Stockhausen, L. V. (2015). Role descriptions induce gender mismatch effects in eye movements during reading. *Front. Psychol.* 6, 1–13. doi: 10.3389/fpsyg.2015.01607
- Rees, H. R., Ma, D. S., and Sherman, J. W. (2020). Examining the relationships among categorization, stereotype activation, and stereotype application. *Pers. Soc. Psychol. Bull.* 46, 499–513. doi: 10.1177/0146167219861431
- Reynolds, D., Garnham, A., and Oakhill, J. (2006). Evidence of immediate activation of gender information from a social role name. *Q. J. Exp. Psychol.* 59, 886–903. doi: 10.1080/02724980543000088
- Roberts, C. (2003). Uniqueness in definite noun phrases. *Linguist. Philos.* 26, 287–350. doi: 10.1023/A:1024157132393
- Robinson, H. M. (2005). *Unexpected (in) Definiteness: Plural Generic Expressions in Romance*. New Brunswick, NJ: The State University of New Jersey.
- Rubio-Fernández, P. (2013). Associative and inferential processes in pragmatic enrichment: the case of emergent properties. *Lang. Cogn. Process.* 28, 723–745. doi: 10.1080/01690965.2012.659264
- Sato, S., Gygax, P. M., and Gabriel, U. (2013). Gender inferences: grammatical features and their impact on the representation of gender in bilinguals. *Biling. Lang. Cogn.* 16, 792–807. doi: 10.1017/S1366728912000739
- Schwarz, F. (2009). *Two Types of Definites in Natural Language* (Ph. D. Dissertation). University of Massachusetts Amherst, Massachusetts, United States. Available online at: https://scholarworks.umass.edu/open_access_dissertations/122 (accessed February 20, 2023).
- Schwarz, F. (2019). "Definites, domain restriction, and discourse structure in online processing," in *Grammatical Approaches to Language Processing*, eds C. J. Clifton, K. Carlson, and J. D. Fodor (Berlin: Springer), 187–208. doi: 10.1007/978-3-030-01563-3_10
- Shutts, K., Banaji, M. R., and Spelke, E. S. (2009). Social categories guide young children's preferences for novel objects. *Dev. Sci.* 13, 599–610. doi: 10.1111/j.1467-7687.2009.00913.x
- Singh, R., Fedorenko, E., Mahowald, K., and Gibson, E. (2016). Accommodating presuppositions is inappropriate in implausible contexts. *Cogn. Sci.* 40, 607–634. doi: 10.1111/cogs.12260
- Sperber, D., Clément, F., Heintz, C., Mascaró, O., Mercier, H., Origg, G., et al. (2010). Epistemic vigilance. *Mind Lang.* 25, 359–393. doi: 10.1111/j.1468-0017.2010.01394.x
- Sperber, D., and Wilson, D. (1986). *Relevance: Communication and Cognition*. Vol. 142. Cambridge, MA: Harvard University Press.
- Sperber, D., and Wilson, D. (2015). Beyond speaker's meaning. *Croat. J. Philos.* 15, 117–149.
- Weinraub, M., Clemens, L. P., Sockloff, A., Ethridge, T., Gracely, E., and Myers, B. (1984). The development of sex role stereotypes in the third year: relationships to gender labeling, gender identity, sex-typing toy preference, and family characteristics. *Child Dev.* 55, 1493–1503. doi: 10.2307/1130019
- 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, 2020–2045. doi: 10.1037/xge000014
- Wharton, T., and Strey, C. (2019). "Slave of the passions: making emotions relevant," in *Relevance, Pragmatics and Interpretation, 1st Edn*, eds K. Scott, B. Clark, and R. Carston (Cambridge University Press), 253–266. doi: 10.1017/9781108290593.022
- White, K. R., Crites, S. L., Taylor, J. H., and Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. *Soc. Cogn. Affect. Neurosci.* 4, 191–198. doi: 10.1093/scan/nsp004
- Wilson, D. (2016). Reassessing the conceptual-procedural distinction. *Lingua* 175, 5–19. doi: 10.1016/j.lingua.2015.12.005
- Winter, B. (2019). *Statistics for Linguists: An Introduction Using R*. New York, NY: Routledge.
- Winter, B., and Wieling, M. (2016). How to analyze linguistic change using mixed models, growth curve analysis and generalized additive modeling. *J. Lang. Evol.* 1, 7–18. doi: 10.1093/jole/lzv003