The role of semantics and repair processes in article-noun gender disagreement in Italian

Popov, Srdan; Miceli, Gabriele; Curcic-Blake, Branislava; Bastiaanse, Roelien

Published in:
Brain and Language

DOI:
10.1016/j.bandl.2020.104787

Brain and Language 206 (2020) 104787

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

The role of semantics and repair processes in article-noun gender disagreement in Italian: An ERP study

Srđan Popov[a,b,*], Gabriele Miceli[c,d], Branislava Ćurčić-Blake[e], Roelien Bastiaanse[b,f]

[a] International Doctorate for Experimental Approaches to Language and Brain (IDEALAB), Universities of Groningen (NL), Newcastle (UK), Potsdam (DE), Trento (IT), Macquarie University (AU)
[b] Center for Language and Cognition Groningen, University of Groningen, The Netherlands
[c] Center for Mind/Brain Sciences (CIMeC), University of Trento, Rovereto, Italy
[d] Centro Interdisciplinare 'Beniamino Segre', Accademia dei Lincei, Rome, Italy
[e] Department of Neuroscience, University of Groningen, University Medical Center Groningen, The Netherlands
[f] Center for Language and Brain, National Research University Higher School of Economics, Russian Federation

ARTICLE INFO

Keywords:
Gender
Syntactic
Semantic
Agreement
P600
LAN

ABSTRACT

In this sentence reading study, we used event-related potentials (ERPs) to investigate the processing mechanism of article-noun gender disagreement in two kinds of nouns in Italian. The first are nouns with syntactic gender (*il treno*$_M$ 'train'; *la sedia*$_F$ 'chair') for which the processing and repair of gender disagreement entails only one repair option, namely for the article (morphosyntactic repair). The second kind are nouns with semantic gender (*il bambino*$_M$ 'boy', *la bambina*$_F$ 'girl'). Here, there are two options for processing and repairing gender mismatch: repairing the article (morphosyntactic repair) or repairing the noun (both morphosyntactic and semantic repair). Both classes of nouns elicited the LAN, indicating that gender disagreement is always registered at the morphosyntactic level. In addition, the P600 was elicited in both conditions, but was larger for semantic gender, reflecting a more complex repair for this class of nouns.

1. Introduction

1.1. Gender cues

Gender processing consists of decoding a number of available linguistic cues while simultaneously integrating them (see Vigliocco & Franck, 1999; Vigliocco & Hartsuiker, 2005). During sentence comprehension, gender can be decoded from two sources: it is either deduced from an agreeing element (probe) or from the noun (or pronoun) itself (goal). The distinction is crucial as only the goal is inherently (lexically) marked for gender. The value of the goal's gender feature gets copied onto the probe in the process called agreement (Bošković, 2011; Pesetsky & Torrego, 2007). Such an agreement relationship is usually morphologically marked as a gender suffix on the probe. However, the lexical gender feature on the goal can be represented at multiple levels as: syntactic information (lemma), morphological information (gender morpheme) or even semantic information (Vigliocco & Franck, 1999). The current study investigates how differences in gender encoding on the noun affect real-time gender processing by means of event-related potentials (ERPs).

Unlike number and person, *syntactic* (also called *grammatical*) gender does not bring any additional semantic or pragmatic information into context.[1] In turn, this means that its value (e.g., feminine or masculine) cannot be deduced from semantics. One reliable strategy to determine the gender value is to look at the gender inflection on the probe. Morphological marking is the only gender cue available on an element which is not lexically marked for gender.

Whereas probes have to be morphologically marked for gender, nouns may or may not be inflected for gender. For example, the overwhelming majority of Italian nouns ending in *-a* and *-o* are feminine and masculine, respectively, thus being gender-transparent. Still, almost 1/3 of all Italian nouns end in *-e* (Cacciari, 2011; D'Achille & Thornton, 2006, as cited in Caffarra, Siyanova-Chanturia, Pesciarelli, Vespignani, & Cacciari, 2015) and are, thus, gender-opaque (as an illustration, the

* Corresponding author at: Center for Language and Cognition Groningen, University of Groningen, PO Box 716 9700 AS, Groningen, the Netherlands.
E-mail address: s.popov@rug.nl (S. Popov).

[1] Some research suggests that (syntactic) gender may still be inherently linked to semantics, in the sense that speakers of a gendered language may assign prototypical gender features to the noun marked for that gender: for example, if *sun* is masculine and *moon* feminine, the speaker may attribute the quality of being powerful to the sun and that of being gentle to the moon (Boroditsky, Schmidt, and Phillips, 2003; see also Clarke et al., 1984; Clarke et al., 1981; Foundalis, 2002; Konishi, 1993; 1994). However, since this is still a contentious issue outside the scope of the current study, it will not be addressed any further.

https://doi.org/10.1016/j.bandl.2020.104787
Received 12 July 2019; Received in revised form 23 January 2020; Accepted 27 February 2020

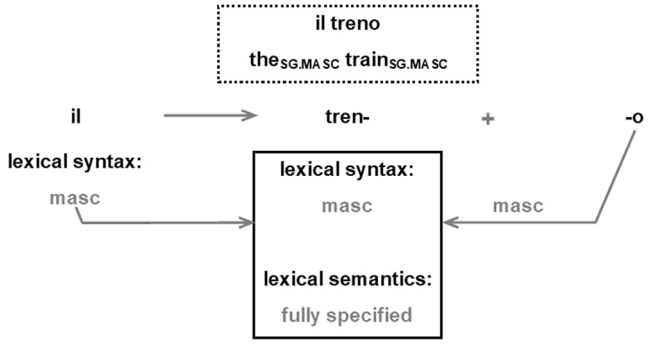

Fig. 1. A scheme of the assumed mechanism of visual syntactic gender processing.

noun *carne:* 'meat' is feminine, while *pesce:* 'fish' is masculine). Going one step further with morphological decomposition, it is often possible to interpret gender from a lexical suffix used for word derivation. In a number of languages, there is a one-to-one correspondence between a derivational suffix and gender (Hickey, 1999). For example, all Italian nouns ending in *–zione* are feminine, despite the final *-e* being an opaque gender marker.

Even though morphology can be a salient cue, it is not always enough for gender retrieval, as shown above for Italian. Regardless of any overt cues, gender information is first and foremost an intrinsic lexical-syntactic feature stored as a lexical representation at the lemma level (Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; Vigliocco, Antonini, & Garrett, 1997; but see Miozzo & Caramazza, 1997 for an alternative account).

Based on the already presented gender cues, gender decoding (in word-by-word reading) can be represented using the following scheme (Fig. 1). In the scheme, the masculine noun *treno:* 'train' is preceded by the definite masculine article *il*. The reader first accesses the masculine gender feature from the probe (article) and retains it for further computation. After that, the goal (noun) becomes available. Since the noun *treno* is a transparent noun, two possibly simultaneous processes are taking place.[2] Firstly, the parser is able to access the lemma of the noun (in which case, reading the stem itself is already sufficient), through which both lexical semantics and lexical syntax of the noun are available. At the same time, the gender feature from the gender marker (*-o*) is processed as masculine, thereby reaffirming the noun's gender feature value. In addition, while gender is being accessed from the lemma and the gender morpheme, it is also being checked against the gender value retrieved from the probe. Since all gender features are congruent (masculine), gender retrieval has been successful and the computation can continue unhindered.

Finally, there is one more gender cue, available only with a limited number of nouns. Nouns denoting people, professions, and some animals have real world referents that have biological sex (e.g., *man* and *woman*). This extra-linguistic feature is often preserved in the gender system of a language, meaning that if the referent is female in the real world, the noun denoting it is marked for feminine gender (Corbett, 1991). This correspondence between biological sex and lexical gender represents a very reliable cue, with rarely any exceptions. Such gender is usually called *semantic* (or *biological*) gender.

[2] We assume these two processes to be simultaneous in the sense that the visual information is probably available at the same time for both the stem and the gender marker. We do not make any claims regarding whether morphological decomposition and lemma access happen simultaneously or in any other order, as it is not immediately pertinent to the issue at hand. More information on the role of the form and lemma, as well as their interaction in gender retrieval, can be found in the dual-route model (comprehension) by Gollan and Frost (2001), and in the two-layer model (production) by Cubelli, Lotto, Paolieri, Girelli, and Job (2005).

However, semantic cues are sometimes difficult to disentangle from morphology. In Italian, nouns marked for semantic gender can be classified into two groups based on their morphological composition: opaque and transparent. Opaque nouns (e.g., *madre – padre* 'mother – father') end in the gender opaque morpheme *-e*. Moreover, the biological masculine-feminine pair usually consists of two different stems (e.g., *madr-e – padr-e*). In contrast, the gender of transparent nouns with semantic gender is determined by the gender morpheme (e.g., *bambin-o – bambin-a* 'boy – girl'), and the stem is shared between the pair members (e.g., *bambin-*). Since the current study focuses on the latter type of nouns, we propose a mechanism of (visual) gender processing for nouns with semantic gender with a shared stem (henceforth nouns with semantic gender).

In the outlined scheme (Fig. 2), the parser first accesses the gender feature from the masculine article (*il*) and retains it for checking against the goal. Once the goal becomes available, the parser has immediate access to the entire word-form. Again, we assume that two simultaneous processes take place. Unlike nouns such as *tren-o* whose stem contains fully specified gender and semantic information, the stem of *bambin-o* is underspecified both for lexical syntax (gender value) and semantics (the stem denotes a child without specifying its biological sex). Therefore, the parser has to revert to decoding gender information off the gender morpheme (*-o*), which it uses to specify both the syntactic and semantic information of the lemma. In this case, the parser has to completely rely on the noun's gender morpheme for complete syntactic and semantic specification. However, in case the noun is preceded by a probe (e.g., gender marked article), the gender of the goal is assumed to match that of the probe, meaning that gender specification comes both from the probe and is confirmed against the gender information from the noun's morpheme. The problem arises when the gender information from the probe and from the gender morpheme do not match. In such instances, there is competition for specifying the gender value of the lemma (both in terms of lexical syntax and lexical semantics) from either the probe or the gender morpheme. This issue will be addressed in more detail at a later point.

Finally, the last gender cue, which is closely related to semantic gender, is gender based on stereotypes. In addition to nouns whose referents' denotation contains the immutable biological gender property (e.g., *boy* and *girl*), there is a group of nouns for which the biological gender is assumed based on stereotypes. Such nouns mainly denote professions, for example, there is a tendency to assume that the referent of the noun *judge* is male and is referred to using the masculine pronoun, whereas a *babysitter* is expected to be feminine and, thus, referred to using a feminine pronoun by default. However, unlike with nouns whose gender information is part of the referent's denotation, the gender of nouns based on stereotypes can be altered with additional contextual information (e.g., *We have just found a babysitter, she is coming later tonight.* and *We have just found a babysitter, his name is John. He is coming later tonight.*). Previous studies have shown that participants indeed assign gender values based on stereotypes, and that this happens regardless of whether it is in short distance agreement, such as article-noun agreement (Molinaro, Su, & Carreiras, 2016), or long distance agreement, such as noun-reflexive pronoun agreement (Canal, Garnham, & Oakhill, 2015; Osterhout, Bersick, & McLaughlin, 1997).

The current study revolves around how gender is processed in the case of nouns denoting semantic gender (which is not based on stereotypes) by varying the gender morpheme (e.g., *bambin-o* – 'boy', *bambin-a* – 'girl'). As explained above, visual gender processing in these nouns is a complex process which relies on interpreting gender cues coming from different sources. Most importantly, gender assignment in these nouns does not only affect the noun's lexical syntax (gender feature value), but also determines the noun's lexical semantics. Since this issue is at the interface of morphosyntactic and semantic processing, ERPs seem to be the best method to tackle it.

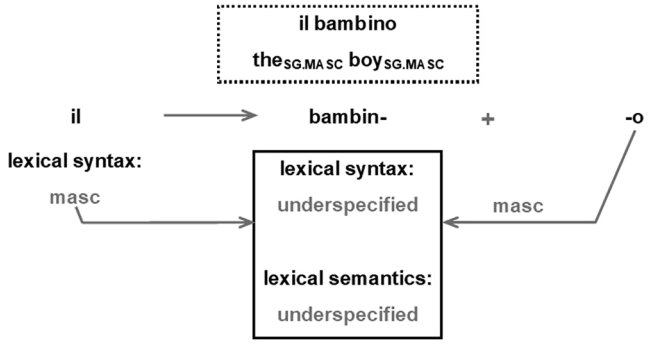

Fig. 2. A scheme of the assumed mechanism of visual semantic gender processing.

1.2. ERP gender agreement studies

Like a vast majority of ERP studies in the field of agreement processing (e.g., Friederici, Steinhauer, & Frisch, 1999; Hagoort, Brown, & Groothusen, 1993; Kutas & Hillyard, 1980; Osterhout & Nicol, 1999), the current one also employs a widely used violation paradigm, whereby instances of grammatical and ungrammatical usage of gender are compared. Such a comparison yields well-documented ERP components, namely N400 (Deutsch & Bentin, 2001; Molinaro et al., 2016; Schmitt, Lamers, & Münte, 2002), left anterior negativity (LAN; Barber & Carreiras, 2005; Barber, Salillas, & Carreiras, 2004; Caffarra et al., 2015; Gunter, Friederici, & Schriefers, 2000; Molinaro, Vespignani, & Job, 2008), and the P600 (Barber & Carreiras, 2005; Barber et al., 2004; Caffarra et al., 2015; Deutsch & Bentin, 2001; Gunter et al., 2000; Molinaro et al., 2008; Popov & Bastiaanse, 2018; Schmitt et al., 2002; Wicha, Moreno, & Kutas, 2004), which are believed to reflect underlying language-related processes. In addition to investigating the mechanism underlying gender (dis)agreement processing, the study also focuses on the repair processes believed to be reflected in the P600 component (e.g., Friederici, 1995, 2002; Friederici & Jacobsen, 1999; Friederici & Meyer, 2004; Kaan, Harris, Gibson, & Holcomb, 2000; Molinaro, Barber, & Carreiras, 2011; Popov & Bastiaanse, 2018). The study investigates in depth whether nouns with semantic gender are repaired in a different way than nouns with syntactic gender.

In order to process gender, the parser has to apply a number of operations. To reiterate, the source of the gender information is the noun. Gender is part of the lexical syntactic information (lemma) of that noun. In case of disagreement, such as gender mismatch, most studies report a biphasic pattern of the LAN and P600 (e.g., Barber & Carreiras, 2005; Barber et al., 2004; Gunter et al., 2000; Molinaro et al., 2008). This is the expected outcome, considering that both the LAN and P600 are usually elicited in response to syntactic manipulations (e.g., Caffarra, Mendoza, & Davidson, 2019; Kaan et al., 2000; Molinaro et al., 2011; for an alternative view of P600, see Brouwer, Fitz, & Hoeks, 2012; Sassenhagen, Schlesewsky, & Bornkessel-Schlesewsky, 2014). Since gender processing consists of either morphological or lexical syntax decoding, it is safe to label gender processing as a syntactic process (Hagoort & Brown, 1999).

The presence of the LAN is understood to indicate anomaly detection during morphosyntactic processing (e.g., Friederici, 2002; Molinaro, Barber, Caffarra, & Carreiras, 2014). That is, the LAN arises as an automatic response to an agreement mismatch, typically 300–500 ms post-stimulus onset. It is followed by a centro-parietal positive shift (P600), from 500 ms on. The P600 is assumed to represent integration difficulty in the form of repair and reanalysis (Friederici, 1995, 2002; Friederici & Jacobsen, 1999; Friederici & Meyer, 2004; Kaan, Harris, Gibson, & Holcomb, 2000; Molinaro, Barber, & Carreiras, 2011). This is the stage at which the parser tries to repair the syntactic incongruity and integrate it with the rest of the discourse at the structural level. Unlike the LAN and P600, the semantically-related N400 (centro-posterior distribution, peaking 300–500 ms post-stimulus onset) has rarely been reported in disagreement studies.

Since the current study mostly focuses on the LAN and P600, we would like to address several issues regarding the status and interpretation of these components. When it comes to the LAN, the problem is its unreliable elicitation in (dis)agreement studies. Unlike the P600, which is almost unanimously reported, the LAN seems to be somewhat more volatile (for a review, see Molinaro et al., 2011). For example, in an almost identical paradigm measuring determiner-noun gender agreement in Spanish, Barber and Carreiras (2005) reported the LAN followed by the P600, whereas Wicha et al. (2004) reported only the P600. In addition, studies on gender agreement in Italian (e.g., Caffarra et al., 2015; Molinaro et al., 2008) almost always confirm the presence of the LAN, whereas the effect seems to be absent in Dutch (Hagoort & Brown, 1999). The underlying cause may be language specific, that is, it may depend on the way a language encodes the gender feature (Friederici & Weissenborn, 2007). The large majority of Italian nouns end in a gender marking suffix, whereas Dutch nouns are gender opaque, unless they contain a gender specific derivational suffix. Hagoort and Brown (1999) suggested that the LAN is sensitive only to phonologically overt inflectional morphology, which explains the difference between Dutch and Italian. However, such an explanation cannot account for both the absence of the LAN in Spanish as reported by Wicha et al. (2004) and its presence as reported by Barber and Carreiras (2005) or the presence of the LAN in gender disagreement with both regular (transparent) and irregular (opaque) nouns (Caffarra & Barber, 2015; Caffarra et al., 2015; for word pair reading, see Caffarra, Janssen, & Barber, 2014).

Furthermore, the status of the LAN as a component in its own right has also been questioned. Several studies have suggested that the LAN may actually be an artefact, obtained through averaging multiple trials across multiple participants who respond to morphosyntactic violations either with a negative or a positive effect in the LAN time window (Osterhout, McLaughlin, Kim, Greenwald, & Inoue, 2004; Tanner & Van Hell, 2014). However, other studies have disputed this claim and provided evidence that the LAN is an independent component which reflects the cost of morphosyntactic processing (Caffarra et al., 2019; Molinaro, Barber, Caffarra, & Carreiras, 2014).

The P600, as already mentioned, is rather consistently reported in agreement studies, and it is understood as reflecting (structural) repair and reanalysis (Molinaro et al., 2011). However, the P600 component has also been shown to be more domain general, and not only related to repair and reanalysis as a result of syntactic manipulation. Rather, the P600 has been reported in studies on irony processing (e.g., Regel, Gunter, & Friederici, 2011; Weissman & Tanner, 2018) or in studies on the 'semantic illusion effect' (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kim & Osterhout, 2005; Brouwer et al., 2012) in which semantically anomalous sentences elicited the P600 instead of the expected N400. Since the focus of this study is on agreement processing, we will from this point on focus only on the role of the P600 in repair and reanalysis.

1.3. Gender cues in ERP agreement studies

Due to some variability in ERP findings on gender processing (e.g., presence or absence of the LAN), researchers started exploring how different gender cues affect gender processing. The most obvious example is the contrast between gender-transparent and gender-opaque nouns. As the former are always marked with a gender suffix, gender retrieval from the lemma is practically redundant. However, in the case of nouns with opaque gender, the only way to access the gender feature is through the lemma. Caffarra et al. (2015) compared the processing of gender transparent and gender opaque nouns in Italian. As mentioned, nouns ending in *-o* and *-a* are masculine and feminine, respectively, with very few exceptions. This means that simple morphological decomposition is enough for successful gender retrieval. Still, there are a

few Italian nouns whose gender does not correspond to the suffix (e.g., *la mano*$_F$ 'the hand'). For such gender opaque nouns the only way to retrieve gender is from the lemma. The study showed no difference in processing transparent and opaque gender in a gender mismatch condition, with each condition eliciting identical LAN and P600 effects. However, when only grammatical instances were compared, a difference in the waveform between transparent and opaque gender was observed. Consequently, it is possible that the LAN and P600 are attuned to certain processes, such as morphosyntactic violation detection and repair, without reflecting the fine-grained difference between morphological and lemma gender mismatch.

One of the questions addressed in the current study is what happens when a semantic component interacts with syntactic processes. Measuring effects in semantic gender processing should provide an answer. As mentioned, nouns with semantic gender have real-world referents whose biological sex corresponds to their language gender. Therefore, it should be theoretically possible to draw on the semantic information in order to infer the noun's gender. In a behavioural production study, Vigliocco and Franck (1999) and Vigliocco and Zilli (1999) demonstrated that nouns with semantic gender cause both non-brain-damaged participants and people with aphasia to produce fewer errors. Therefore, the authors proposed a dual-route model for nouns with semantic gender, according to which the gender feature can be retrieved both from the lemma and semantic information. For all other nouns, the only route is through the lemma. Even though this model is based on oral production, the idea of an additional semantic route has been applied in a number of ERP comprehension studies on semantic gender (Barber et al., 2004; Deutsch & Bentin, 2001; Hammer, Jansma, Lamers, & Münte, 2005; Osterhout et al., 1997). Since the N400 is usually related to semantic processing and is taken to indicate difficulty integrating information in a wider semantic and discourse context (Kutas & Federmeier, 2011), the expectation in these studies was that the N400 would reflect the activation of semantic information in semantic gender access.

The results of these studies on semantic gender are not clear-cut. Deutsch and Bentin (2001) studied subject-verb agreement in Hebrew, a language in which the subject and verb agree in gender. In their design, in which the subject was either animate or inanimate, the verb followed the subject and was the target word. They reported the N400 only for disagreement with animate stimuli, whereas gender violation with inanimate nouns elicited the P600. The presence of the N400 was interpreted as a hallmark of semantic processing, as it was present only with animate nouns. In another subject-predicate agreement processing study, Barber et al. (2004) manipulated the gender type of the subject in Spanish. Based on the results from Deutsch and Bentin (2001) study, Barber et al. (2004) expected to observe the LAN and P600 for nouns with syntactic gender, and the N400 for nouns with semantic gender. Interestingly, both conditions elicited identical effects: the LAN and P600. However, it is worth pointing out that the LAN effect was statistically weakly supported in the left lateral regions, as the overall lateral analysis did not yield an interaction between agreement and electrode site. The presence of the LAN was established by running post-hoc tests, despite the lack of an interaction, for each lateral electrode, which revealed an agreement effect in the left anterior and left central electrodes. This finding, taken together with the statistically significant agreement effect in the midline anterior sites, was interpreted as the LAN effect. Therefore, according to the authors, gender processing in general is a syntactic process. The N400 is obtained for outright semantic violations: something that semantic gender violation is not. The different results, as compared to Deutsch and Bentin (2001) study, were explained by a typological distance between Hebrew and Spanish.

Similarly, two studies on pronoun reference processing reported contradictory results regarding the role of semantic gender. Schmitt et al. (2002) manipulated the pronoun's antecedent in German. The antecedent was an animate referent marked for semantic gender (e.g., der Bub$_M$ 'the boy') in the first condition, and a diminutive version of the nouns in the second. In German, diminutives are derived through suffixation, during which all nouns are assigned neuter gender regardless of the noun's semantic information (e.g., das$_N$ Bübchen$_N$ 'the little boy'). In the violated non-diminutive condition, both the N400 and P600 were reported, whereas only the P600 was attested in the violated diminutive condition. The authors concluded that establishing pronoun reference is a syntactic process eliciting the P600 in both conditions. In addition, if the antecedent noun is marked for semantic gender, semantic processes are also activated, which is reflected in the presence of the N400. These findings were challenged by a different pronoun study in German in which no effect of semantic gender was detected, and both antecedents with grammatical and semantic gender elicited the P600 (Hammer et al., 2005). There was no N400 associated with the semantic gender condition.

Finally, ERP studies have also addressed the role of stereotypical gender (e.g., a *nurse* is stereotypically referred to using feminine gender) in gender (dis)agreement processing. Molinaro et al. (2016) investigated the processing of stereotypical gender in article-noun gender (dis)agreement in Spanish. They showed that agreement mismatch based on stereotypes (e.g., *las*$_F$ *mineras*$_F$ the (women) miners) elicits the N400 on the noun. However, the syntactic disagreement condition (*las*$_F$ *mineros*$_M$ the$_F$ miners$_M$) in this study also elicited the N400, possibly due to the paradigm setup. Whereas the experiment provides evidence that stereotypical gender (i.e., semantic knowledge of the referent's gender) can override syntax, the study is somewhat weakened by the presence of the N400 in the syntactic disagreement condition which was expected to elicit the P600. In addition, Canal et al. (2015) and Osterhout et al. (1997) looked into the processing of reflexive pronouns with nouns marked for stereotypical gender (e.g., an engineer… himself/herself). Both studies elicited the P600, with Canal et al. (2015) also reporting the Nref. Their interpretation is that the violation of stereotypical gender by a gender mismatching reflexive is processed as an agreement violation. Even though it is plausible to see such violations as violating the noun's semantics, it is important to note that the effect is measured on the pronoun, whereas the establishing of the gender stereotype from our world knowledge takes place at the level of the noun. Furthermore, co-reference establishing between the noun and the reflexive requires a formal feature evaluation, which is a syntactic process that either excludes any semantics from the noun or overrides it (see Osterhout et al., 1997; Van Berkum, Koornneef, Otten, & Nieuwland, 2007). This is another reason why it is important to measure the gender (dis)agreement effect at the source of the gender information (i.e., the noun) if we are interested in whether semantics plays a role in agreement.

1.4. P600 as a marker of repair and reanalysis

As already mentioned, most ERP agreement studies employ the violation paradigm to study the underlying agreement mechanism or the retrieval of a lexical syntactic feature such as gender. One of the consequences of such a paradigm is that the parser will try to repair the syntactically incongruent structure. According to previous studies (Friederici, 1995, 2002; Friederici & Jacobsen, 1999; Friederici & Meyer, 2004; Kaan, Harris, Gibson, & Holcomb, 2000; Molinaro, Barber, & Carreiras, 2011; Popov & Bastiaanse, 2018), the repair (and reanalysis) processes take place in the late syntactic stage (P600). Moreover, Hagoort and Brown (2000) suggest that the P600 consists of two distinct stages, an early stage in which integration processes take place (usually 500–700 ms post-stimulus onset), and a late stage in which repair and reanalysis take place (700–900 ms post-stimulus onset).

In line with this distinction, Gunter et al. (2000) claim that semantic processes can modulate late syntactic processing. This is compatible with Friederici (2002) model of auditory sentence comprehension in which syntactic and semantic processes are held separate until 500 ms,

after which they interact. This means that semantic processing can be reflected in the P600 component, in addition to the N400, especially when there is an interaction with syntax.

Popov and Bastiaanse (2018) showed that the complexity of repair can modulate the amplitude of the P600, with more complex repair processes eliciting a P600 of greater amplitude. In their study, instances of gender and number disagreement were compared, with number disagreement exhibiting a more complex repair operation (two repair options as opposed to a single option for gender disagreement). The amplitude of the P600 was larger for number disagreement than for gender disagreement. Since an increase in effect's amplitude indicates a larger processing load for that condition (Otten & Rugg, 2005), the result was taken to indicate that the repair process was more complex in number than in gender. Furthermore, the difference was found in the late time window (800–1000 ms), which is in line with Hagoort and Brown's (2000) claim.

Still, it is worth noting that not all studies on gender and number agreement found the same pattern. Barber and Carreiras (2005) reported a larger P600 effect for gender disagreement than number disagreement in Spanish article-noun pairs. The difference between the two studies may stem from a different way the gender feature is encoded in Dutch (mostly opaque) and Spanish (moastly transparent) or from methodological differences (e.g., different time windows for the statistical analysis).

1.5. Repair mechanism for article-noun disagreement in nouns with syntactic and semantic gender

In the current study, we employed the violation paradigm in the form of article-noun gender disagreement to two groups of nouns in Italian: nouns with syntactic and nouns with semantic gender. As we explained earlier in the introduction, accessing the gender feature in these nouns works in a different way in the visual modality. In nouns with syntactic gender, the gender feature is available both from the noun's gender morpheme and the stem (through the lemma). Also, the value of the gender feature (masculine or feminine) does not affect the noun's lexical semantics, at least not to the extent it does in nouns with semantic gender (but see Boroditsky, Schmidt, and Phillips, 2003; Clarke, Losoff, McCracken, & Rood, 1984; Clarke, Losoff, McCracken, & Still, 1981; Foundalis, 2002; Konishi, 1993; 1994). The situation with nouns with semantic gender is more complex, as their stem is underspecified both syntactically and semantically. Consequently, the gender feature must be retrieved from the suffix, which will evaluate the noun's gender feature as masculine and feminine, and specify the animate referent as either male or female. Due to these differences, we propose that the repair mechanism in these two groups of nouns works differently, which will be reflected in electrophysiological responses.

In case of an article-noun string with a syntactic gender mismatch, we expect the parser to have only one straightforward option for repairing the violation (Fig. 3). If the incongruous feminine article *la* precedes the masculine noun *tren-o* ('train'), both the stem (lemma) and the suffix establish that the noun is masculine. Therefore, the repair process can go only one way, which is repairing the offending feminine article *la* into the correct masculine article *il*.

The repair process for nouns with semantic gender is more complex as there are more repair options. In the Fig. 4, the article used is feminine *la* followed by a masculine noun *bambin-o* ('boy'). The stem *bambin-* is underspecified for gender and semantics, as it needs the gender value to specify the referent's biological sex (the semantics of the stem is that of a child, and the gender morpheme determines whether it is male or female). In case of a mismatch between the article and the gender morpheme, the parser has two options. It may take the

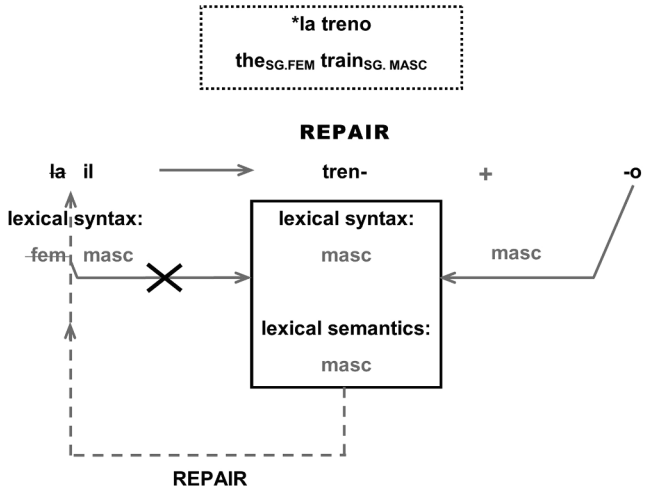

Fig. 3. Repair mechanism for nouns with syntactic gender.

gender value of the noun's suffix to be correct (*-o* masculine), and thus repair the feminine article (*la*) into the masculine article (*il*). Or, it may assume that the article's gender value is correct and that the noun suffix needs to be repaired (from *bambino-o* 'boy' to *bambin-a* 'girl'). Both options are equally plausible, as no context at this point induces a preferred interpretation.

It follows from the repair mechanisms given above that nouns with syntactic gender have a single repair possibility, whereas for nouns with semantic gender there is a double repair option, making them more complex to repair. In addition to the sheer number of options, the underlying repair mechanism operates at different levels for the two groups of nouns. Nouns with syntactic gender are repaired at the morphosyntactic level, as it only involves gender retrieval from the lemma and morphology. However, for nouns with semantic gender, both the same morphosyntactic process and a lexical semantic process are involved. Namely, any gender mismatch in these nouns creates a semantic incongruity, as the specification of the referent's sex (lexical semantics) depends on the gender value (lexical syntax). Therefore, when a mismatching gender feature from the article is checked against the noun, a mismatch is created both at the syntactic and semantic level.

As a consequence, along with the repair of the gender feature (lexical syntax), nouns with semantic gender also undergo the repair or reanalysis of lexical semantics. When the masculine noun *bambin-o* is repaired into *bambin-a*, its semantic specification for biological sex is also repaired (from boy to girl) in addition to its morphosyntactic gender.

1.6. Experimental predictions

The current study focuses on article-noun gender disagreement processing in Italian with a particular emphasis on the role of repair. The novelty of this study is that we compared two groups of nouns (with semantic and syntactic gender), which enabled us to isolate and compare the role of morphology, lexical syntax, and lexical semantics in gender disagreement processing.

In line with previous studies on agreement (e.g., Barber et al., 2004; Barber & Carreiras, 2005; Gunter et al., 2000; Molinaro et al., 2008), our first prediction is that article-noun gender disagreement will show its morphosyntactic nature, which should be reflected in the elicitation of the P600 component and possibly the LAN. Our reservation about the

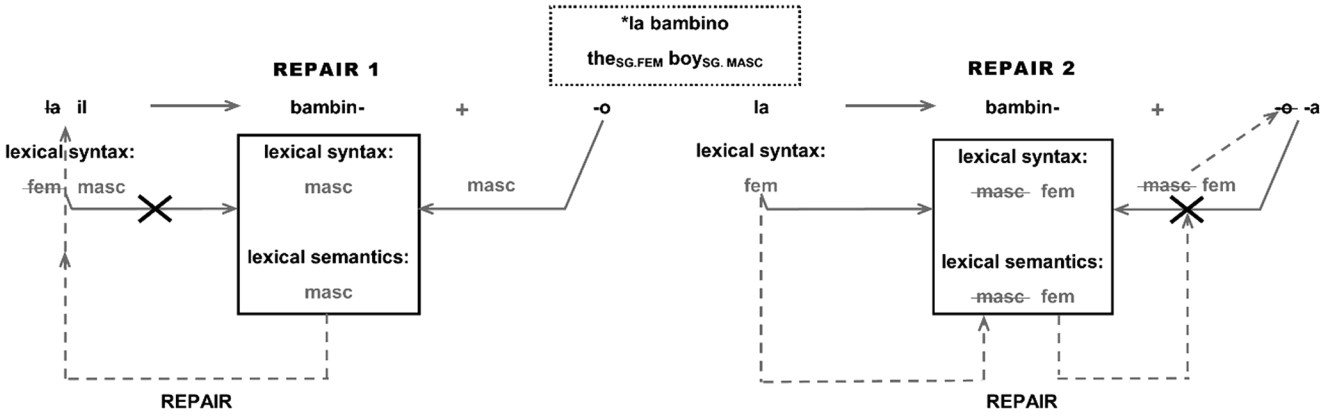

Fig. 4. Repair mechanism for nouns with semantic gender.

LAN has two main reasons. Firstly, the LAN has been inconsistently elicited in previous studies on (dis)agreement, for yet unclear reasons (see Caffarra et al., 2019; Molinaro et al., 2014; Tanner & Van Hell, 2014). The second reason is related to the function of the LAN. The LAN is believed to reflect an automatic morphosyntactic violation detection (Friederici, 2002; Molinaro et al., 2014). In our study, if the LAN were elicited, it should occur for nouns with syntactic gender, which contain a straightforward morphosyntactic violation. Nouns with semantic gender, however, contain both a morphosyntactic and a semantic violation. A gender mismatch between the article and the noun also violates the noun's semantics, as the referent's biological sex is determined through morphological means in visual gender processing. According to previous studies on semantic gender (Barber et al., 2004; Deutsch & Bentin, 2001; Hammer et al., 2005; Molinaro et al., 2016; Osterhout et al., 1997), if this violation is truly legible at the semantic level, it can be expected to elicit the N400. This is another novelty of this study; no other sentence processing ERP study has measured the effect of semantic gender at the noun as the source of gender information. Once the gender feature is transmitted onto the probe, its origin is most often irrelevant for any further computation. Therefore, if the aim is to find out whether semantics is activated and engaged during gender processing, the effect should be measured at the noun.

However, it is equally plausible that nouns with semantic gender also elicit the LAN. According to Friederici (2002), syntax and semantics are seen as separate in their respective stages (represented by the LAN and N400), and they can only start interacting at a later point (P600). Therefore, all things being equal, gender disagreement causes a violation that should be legible only at the syntactic level. An open question, though, is whether such a violation can still be legible to the semantic system in case of semantic gender.

Finally, we expect to be able to measure two effects in the P600 stage. Firstly, as mentioned above, an interface between syntax and semantic is only allowed at a late processing stage (P600). If this is the case, a semantic incongruity in nouns with semantic gender should not be reflected in the N400 but in the P600, and in particular at the early P600 stage, at which syntax and semantics interact. This interaction leads to detect a violation at the semantic level, and consequently its repair/reanalysis. Therefore, we expect the early P600 effect to be larger in amplitude for nouns with semantic gender than for nouns with syntactic gender. The second expectation regarding the P600 is that the late stage, seen as reflecting repair and reanalysis processes, will also be larger for nouns with semantic gender due to more complex repair.

2. Method

2.1. Participants

For this experiment, 25 native speakers of Italian (17 male, mean age 23.4, range 19–33) were recruited. They all reported themselves as right-handed, which was confirmed with the Edinburgh Handedness Questionnaire (Oldfield, 1971). All participants had at least 13 years of education, corresponding to finished high school. No participant reported any history of neurological or psychiatric disease. They all had normal or corrected-to-normal vision.

2.2. Materials

The experimental materials consisted of 160 sentences, half of which contained nouns marked for semantic gender (semantic condition), and the other half nouns marked with syntactic gender (syntactic condition). The frequency of the two noun types was matched (semantic gender average LogF: 2.08, syntactic gender average LogF: 2.32; $t(38) = 1.33$, $p > 0.1$). Each noun was used four times as: (1) singular grammatical; (2) singular ungrammatical; (3) plural grammatical; (4) plural ungrammatical. Half of the nouns were masculine, and half feminine.

As a consequence of a balanced design (each noun used as singular and plural, grammatical and ungrammatical), the participant would have been exposed to the same noun 4 times. In order to minimise repetition, experimental items were split between 2 lists, so that every participant saw each noun only twice, once as singular and once as plural. Also, the noun was used once in a grammatical and once in an ungrammatical context.

Therefore, each subject was exposed to one list, that is, each subject read 80 experimental and 80 filler sentences. Experimental items were divided into 4 conditions of 20 sentences each: (1) semantic gender grammatical; (2) semantic gender ungrammatical; (3) syntactic gender grammatical; (4) syntactic gender ungrammatical. All nouns were gender-transparent, meaning that all masculine nouns ended in either *-o* (singular) or *-i* (plural) and feminine nouns ended in either in *-a* (singular) or *-e* (plural). Each sentence started with a temporal adverb (e.g., often, yesterday, rarely), followed by a definite article, which was followed by a noun. The goal noun phrase was always in the subject position. Depending on the gender and number of the noun, the article took one of 4 forms: *il* masculine singular; *i* masculine plural; *la* feminine singular; *le* feminine plural.[3] The noun phrase was followed by the verb, and by either an object or an adverb. Half of the experimental sentences contained a morphosyntactic violation in the form of gender disagreement between the article and the noun. For example, a masculine singular noun, such as *nonno* 'grandfather', was preceded by the feminine singular article *la*, resulting in an ungrammatical sentence. Example sentences for each condition are given in (1) and (2).

[3] An equal number of singular and plural nouns was used in each condition. A post-hoc comparison between singular and plural nouns in each condition did not result in any differences.

Experimental stimuli.

(1) Semantic gender					
Spesso	il	nonno	fumava	la	pipa.
often	$the_{SG.M}$	$grandfather_{SG.M}$	smoked	the	pipe
'The grandfather often smoked his pipe.'					
*Spesso	la	nonno	fumava	la	pipa.
often	$the_{SG.F}$	$grandfather_{SG.M}$	smoked	the	pipe
''The grandfather often smoked his pipe.'					
(2) Syntactic gender					
Generalmente	il	treno	parte	in	ritardo.
usually	$the_{SG.M}$	$train_{SG.M}$	leaves	in	late
'The train usually leaves late.'					
*Generalmente	la	treno	parte	in	ritardo.
usually	$the_{SG.F}$	$train_{SG.M}$	leaves	in	late
'The train usually leaves late.'					

There were 80 fillers per list (3). Each filler item was presented once as grammatical and once as ungrammatical. The ungrammatical sentence lacked subject-verb agreement, having the predicate in infinitive.

(3) Filler items							
Le	scale	lassú	sono	troppo	ripide	per	me.
the	stairs	up there	are	too	steep	for	me
'The stairs up there are too steep for me.'							
*Le	scale	lassú	essere	troppo	ripide	per	me.
the	stairs	up there	be	too	steep	for	me
*'The stairs up there be too steep for me.'							

2.3. Procedure

Participants were seated at a comfortable distance from the screen (app. 80 cm) while the EEG was being recorded. Their task was to read the sentences carefully and answer the grammaticality judgment question after every sentence. The presentation software used was E-Prime 2.0 (Psychology Software Tools, Inc.). Before the experiment started, the experimenter explained the task in detail and provided a few examples. Instructions were also presented on the screen, followed by 5 practice sentences. Additional clarification was given, if requested, before the experiment started.

Sentences were presented word-by-word, written in white on a black background (Arial, 24pt). Each stimulus started with a fixation cross (500 ms), followed by blank screen for 300 ms. Each word was presented for 300 ms and was followed by a 300 ms break (black screen). The last word always appeared with a full stop. After the last word, the screen remained blank for 1 s, after which a question mark appeared. The question mark was a cue for the participant to press the appropriate key ("p" or "q") depending on whether the sentence was grammatical or not. The assignment of the keys to Yes and No responses was counterbalanced across participants. There were 4 blocks, each containing 40 sentences, with a break after each block. Recording took 20 min on average.

2.4. EEG Recording and Data processing

The continuous electroencephalogram was recorded from 64 Ag/AgCl scalp electrodes (WaveGuard) using the ASA-Lab system (ANT Neuro Inc, Enschede, The Netherlands). Additional bipolar electrodes were used to record horizontal (HEOG; at the outer canthus of each eye) and vertical (VEOG; above and below the left eye) eye movements. Impedances were kept below 10 kΩ. Data were acquired at 512 Hz sampling rate with the common average reference.

The offline processing was carried out in Brain Vision Analyzer 2.0.4 software (Brain Products, GmbH, Munich, Germany). Data were down-sampled to 256 Hz before being re-referenced to the average of the mastoids. Offline filtering was performed using a band-pass filter (0.1–40 Hz), followed by automatic eye-blink correction. Data were segmented into epochs starting 200 ms before the onset of the critical word (the noun) and lasting until 1500 ms post-word onset. Only correct trials were segmented. The (automatic) artifact rejection (+/-100 μV threshold) was performed only on the section of each epoch (−200–1000 ms) used in the statistical analysis. In total, 6.4% of all trials were excluded from the analysis (semantic gender grammatical: 8.2%, semantic gender ungrammatical: 6.8%, syntactic gender grammatical: 5.4%, syntactic gender ungrammatical: 5.2%). There was no difference in the number of retained epochs between conditions ($F(3, 72) = 1.37$, $p < .1$). Data were corrected relative to the 200 ms prestimulus baseline.

Finally, data were averaged per subject and per condition. No participants were excluded due to excessive noise or artefacts, as they were all above the threshold of 60% of averaged trials in all conditions.

2.5. Analysis

Averaged values (in μV) were extracted per participant, per condition, and per region of interest. Scalp electrodes were divided into 9 regions of interest, each containing either 5 or 6 electrodes (Fig. 2.1): left anterior (F7, F5, F3, FC3, FC5), midline anterior (F1, Fz, F2, FC1, FCz, FC2), right anterior (F4, F6, F8, FC4, FC6), left central (TP7, C5, C3, CP5, CP3), midline central (C1, Cz, C2, CP1, CPz, CP2), right central (C4, C6, CP4, CP6, TP8), left posterior (P7, P5, P3, PO7, PO5, O1), midline posterior (P1, Pz, P2, PO3, POz, PO4), and right posterior (P4, P6, P8, PO6, PO8, O2). The analysis was carried out in three independent time windows, from 300 to 450 ms (roughly corresponding to the LAN and N400), followed by the 500–700 ms window (early P600), and ending with the 700–900 ms window (late P600).

For the statistical analysis, which was carried out using IBM SPSS (Version 23), repeated measures ANOVAs were used with the following within subject factors: condition (2 levels: syntactic and semantic gender), grammaticality (2 levels: grammatical and ungrammatical), hemisphere (2 levels: left and right hemisphere), and anteriority (3 levels: anterior, central, and posterior). The significance level was set to $p < .05$. For each time window, 2 global repeated measures ANOVAs were performed; first for the lateral regions (all factors included), and then for the midline regions (factor hemisphere excluded). Post-hoc paired *t*-tests were carried out with those interactions that turned out at least marginally significant ($p < .1$), and that included the factor grammaticality. Pair-wise comparisons were corrected using the Bonferroni correction. If the assumption of sphericity was violated, the Greenhouse and Geisser (1959) correction was applied. Lastly, only correctly judged trials were included in the analysis.

3. Results

3.1. Accuracy Data

Overall accuracy in the grammaticality judgment task was 97.1%. Out of 25 participants, 6 achieved 100% accuracy; the number of errors in the rest ranged from 1 to 6. No further analysis was performed due to the very low number of errors and their equal distribution across conditions.

3.2. ERP results

Visual inspection showed a left-lateralised anterior negativity and posterior-central positivity (semantic gender: Fig. 5; syntactic gender: Fig. 6). The negativity had the same distribution in both conditions, encompassing the left anterior region but also spreading into the left central region. In accordance with most previous reading studies on agreement, negativity started approximately 300 ms post-stimulus onset, and lasted until 450 ms. It was immediately followed by a positive effect lasting from 500 ms until almost 1200 ms in both conditions. Visually, the positive effect appeared to be of greater amplitude and with a somewhat wider distribution in the semantic gender condition. There seemed to be no difference in amplitude, time-window or

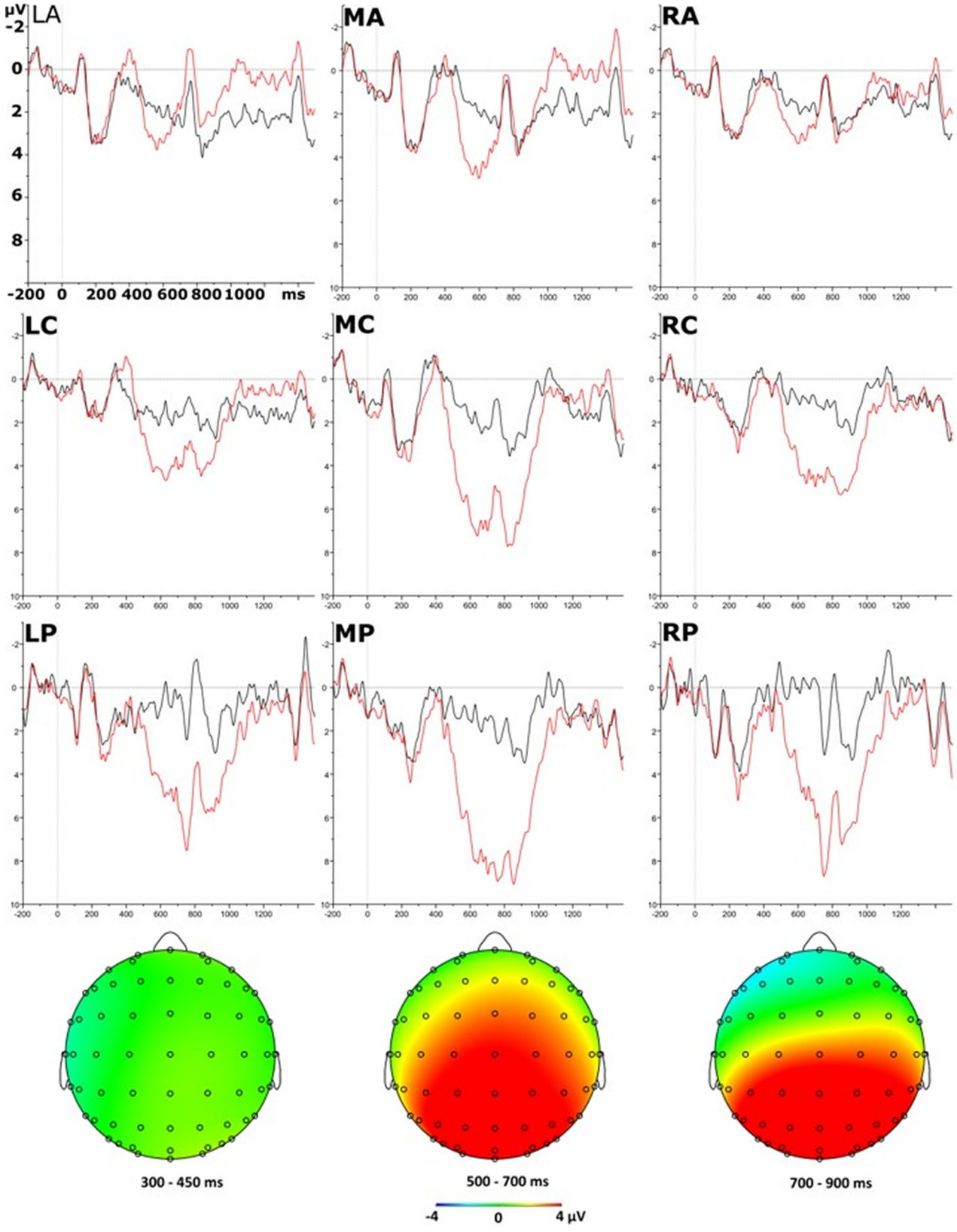

Fig. 5. Grand average ERPs for the semantic gender condition across all 9 ROIs: the black line represents correct sentences and the red line represents sentences with gender mismatch. The topographic maps represent a difference between ungrammatical and grammatical sentences.

distribution of the negative effect between the two conditions.

In the first time window (300–450 ms), the overall lateral ANOVA showed an interaction between grammaticality and hemisphere ($F(1, 24) = 14.74$, $p < .001$, $\eta_p^2 = 0.038$), as well as a three-way interaction between grammaticality, hemisphere, and anteriority ($F(2, 48) = 4.29$, $p < .05$, $\eta_p^2 = 0.15$). The follow-up tests of the three-way interaction revealed that ungrammatical sentences elicited a more negative waveform than grammatical sentences in the left anterior region

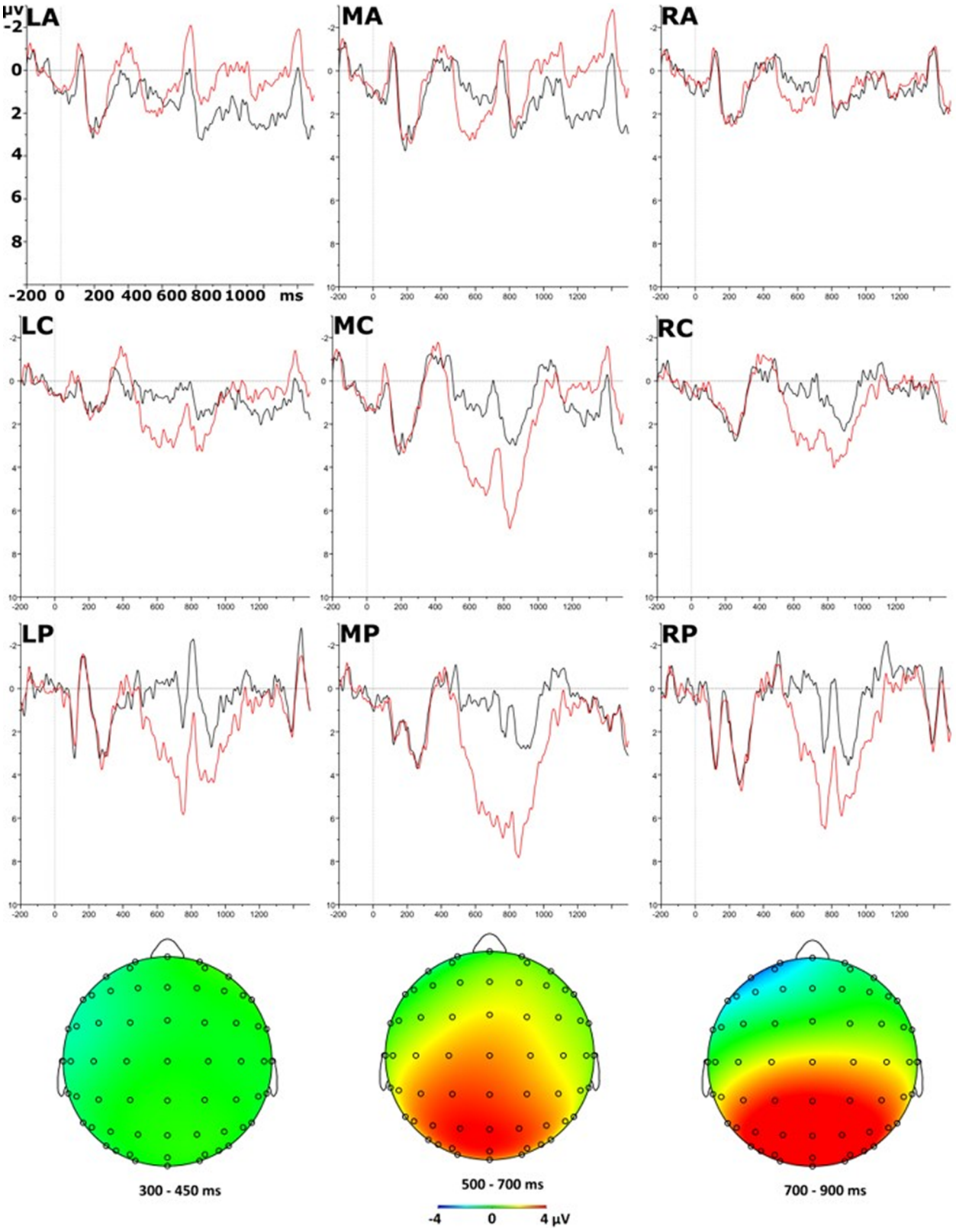

Fig. 6. Grand average ERPs for the syntactic gender condition across all 9 ROIs: the black line represents correct sentences and the red line represents sentences with gender mismatch. The topographic maps represent a difference between ungrammatical and grammatical sentences.

only ($t(24) = 2.37, p < .05$). No significant effect was observed in the midline analysis.

The main effect of grammaticality was significant in the lateral analysis of the 500–700 ms time window ($F(1, 24) = 20.59, p < .001, \eta_p^2 = 0.46$), with ungrammatical sentences eliciting a more positive waveform. The effect also interacted with anteriority ($F(2, 48) = 11.94, p < .01, \eta_p^2 = 0.33$), as well as anteriority, hemisphere, and condition in a four-way interaction ($F(2, 48) = 5.344, p < .05$). The follow-up

on the four-way interaction revealed that the positive effect was present or trending in the following regions: left central (semantic gender/SEM: $t(24) = -3.79$, $p < .01$, syntactic gender/SYN: $t(24) = -3.07$, $p < .05$); left posterior (SEM: $t(24) = -5.60$, $p < .001$, SYN: $t(24) = -3.92$, $p < .01$); right anterior (only SEM: $t(24) = -2.51$, $p < .1$); right central (SEM: $t(24) = -5.04$, $p < .001$, SYN: $t(24) = -2.394$, $p < .1$); right posterior (SEM: $t(24) = -6.08$, $p < .001$, SYN: $t(24) = -2.94$, $p < .05$). Also, no statistically significant difference between grammatical sentences in the two conditions was observed (all p's > 0.1). However, ungrammatical sentences did elicit a significantly more positive effect in the semantic gender condition than in the syntactic gender condition in the following regions: left central ($t(24) = 3.18$, $p < .05$); left posterior ($t(24) = 3.27$, $p < .05$); right central ($t(24) = 3.82$, $p < .01$); right posterior ($t(24) = 3.47$, $p < .01$).

The midline analysis showed an effect of grammaticality ($F(1, 24) = 31.69$, $p < .001$, $\eta_p^2 = 0.57$), which also interacted with anteriority ($F(2, 46) = 13.73$, $p < .001$, $\eta_p^2 = 0.36$). Unlike in the lateral analysis, ungrammatical sentences yielded a more positive waveform in all three regions across the anterior-posterior axis (anterior: $t(24) = -3.09$, $p < .01$; central: $t(24) = 5.57$, $p < .001$; posterior: $t(24) = -6.62$, $p < .001$).

In the last time window (700–900 ms), ungrammatical sentences continued to elicit a more positive waveform, which was confirmed by a main effect of grammaticality ($F(1, 24) = 16.24$, $p < .001$, $\eta_p^2 = 0.4$). In addition, grammaticality interacted with the topographic factors of anteriority ($F(2, 48) = 103.88$, $p < .001$, $\eta_p^2 = 0.81$) and hemisphere ($F(1, 24) = 7.92$, $p < .01$, $\eta_p^2 = 0.25$), as well as with both of them ($F(2, 48) = 18.17$, $p < .001$, $\eta_p^2 = 0.43$). Finally, a non-significant trend was observed in the three-way interaction between condition, grammaticality, and anteriority ($F(2, 48) = 2.9$, $p < .1$, $\eta_p^2 = 0.11$). The follow-up tests on the latter interaction revealed that the positive effect was confined to the central ($t(24) = -4.92$, $p < .001$) and posterior region ($t(24) = -8.6$, $p < .001$) in the semantic gender condition. In the syntactic gender condition, however, the effect was reliably present in the posterior region only ($t(24) = -5.38$, $p < .001$), while falling short of significance in the central region ($t(24) = -2.7$, $p < .1$). As in the previous time window, there was no difference between grammatical sentences in the two conditions. Also, ungrammatical sentences again elicited a more positive effect in the semantic gender condition than in the syntactic gender condition (central region: $t(24) = 3.64$, $p < .01$; posterior region: $t(24) = 3.47$, $p < .01$).

In line with the lateral results, the midline analysis yielded an overall effect of grammaticality ($F(1, 24) = 22.19$, $p < .001$, $\eta_p^2 = 0.48$), as well as an interaction between grammaticality and anteriority ($F(2, 48) = 92.17$, $p < .001$, $\eta_p^2 = 0.79$). Further testing proved that the positivity was present in the central ($t(24) = -5.18$, $p < .001$) and posterior midline region ($t(24) = -8.19$, $p < .001$).

3.3. Summary of ERP results

In both conditions, ungrammatical sentences elicited a left-lateralised anterior negativity in the 300–450 ms time window. The effect corresponds to the left anterior negativity component (LAN). The LAN was followed by a broadly distributed positivity (P600) in the subsequent time window of 500–700 ms, which was also elicited by ungrammatical sentences. The positive effect was most reliably present in the left central region and posterior lateral regions, as well as in all midline regions. Its distribution was slightly broader in the semantic gender condition. Also, the positive effect elicited by ungrammatical sentences in the semantic gender condition was significantly larger (larger amplitude) in all lateral regions compared to the positivity elicited by ungrammatical sentences in the syntactic gender condition. The P600 effect persisted into the last time window (700–900 ms). Again, the effect had a broader distribution (and larger amplitude) in the semantic gender condition than in the syntactic gender condition. Finally, no difference was found between the waveforms elicited by grammatical sentences in the two conditions.

4. Discussion

We compared the processing of article-noun gender disagreement in Italian nouns with syntactic and Italian nouns with semantic gender. Our findings indicate that the same mechanism underlies gender (dis) agreement processing in both nouns with syntactic and nouns with semantic gender, as evidenced by the presence of the LAN and P600 in both conditions. As the experimental manipulation in nouns with syntactic gender is clearly morphosyntactic, and we obtained the same results for both syntactic and semantic gender (LAN and P600), we can conclude that gender (dis)agreement in semantic gender is also processed at the morphosyntactic level. The difference in the amplitude of the P600 between the two conditions indicates that integration and repair processes can be influenced both by syntax and by semantics.

4.1. LAN

The left-anterior negativity is understood to represent an automatic (morphosyntactic) violation detection (Friederici, 2002; Molinaro et al., 2014). In the current experiment, the LAN was expected for nouns with syntactic gender, as they contain a straightforward morphosyntactic violation in the form of gender disagreement between the article and the noun. The gender on the noun is marked both morphologically and lexically (lemma), and is in mismatch with the gender retrieved from the preceding article. The situation is more complicated for nouns with semantic gender. Since their lemma is underspecified for gender, the gender value has to be determined via the gender morpheme. In other words, the noun's gender value and lexical semantics are fully specified inflectionally, via either the probe's gender value or the noun's morpheme. When there is a mismatch between the two, in addition to a morphosyntactic violation, the noun's semantics remains underspecified. The reason for this is that in our experimental nouns with semantic gender the stem is shared between the male–female pair (e.g., the stem *bambin-* denotes a child and the appropriate suffix indicates whether it is a boy or a girl). Thus, if the probe indicates masculine gender and the noun's morpheme is feminine, there is a mismatch at the semantic level regarding whether the noun's semantics is that of a female or male member of the pair. For this reason, a plausible expectation was that violations in nouns with semantic gender would be perceived as semantic and elicit the N400 instead of the LAN.

The experimental manipulation in this study elicited the LAN in the LAN/N400 time window, peaking at approximately 350 ms, and confined to the left lateral and left central regions. Its spatial, as well as its temporal distribution, is in line with previous studies reporting the LAN (Friederici, 2002; Molinaro et al., 2014). More importantly, the effect was identical in both the syntactic and the semantic gender conditions. Since the LAN is assumed to represent an automatic violation detection process, it seems that the gender violation is recognised/processed in the same manner for both types of nouns. The parser retrieves the gender information in the same way for both nouns with syntactic and semantic gender.

The presence of the LAN, instead of the N400, in nouns with semantic gender is not completely unexpected. Firstly, nouns with semantic gender do contain a clear morphosyntactic violation. It is just that in addition to this violation, there is a mismatch in semantics, at the level of the referent's real-world biological sex. Since this information has to be first retrieved via morphological processes (probe's value and noun's gender morpheme), it is plausible that in this case the morphosyntactic violation is more prominent. Secondly, the noun's semantics is not violated in the prototypical way that would include semantic incongruity of the noun with the preceding context. Rather, the mismatch is with a functional unit (article/gender morpheme), which means the violation is at the interface of semantics and

morphosyntax. It is possible that such a violation is legible only at the morphosyntactic level. Consequently, the system does not recognise a semantic gender violation as an outright semantic violation, which was also suggested by Barber et al. (2004).

Moreover, semantics may be bypassed at this processing stage, perhaps because it does not yet interact with syntax (Friederici, 2002). Be this as it may, our stronger hypothesis was that the semantics would be affected by gender disagreement, but at the level of integration and repair rather than at the detection level. We expected that the effect would be obvious already at the early P600 stage, which is believed to reflect integration processes in which syntax and semantics are allowed to interact.

Finally, Caffarra et al. (2015) and Caffarra and Barber (2015) looked into the role of morphological gender transparency in article-number (dis)agreement in Italian and Spanish, respectively, by comparing gender transparent and gender opaque nouns. For the disagreement comparison, they reported both the LAN and P600. The LAN was not modulated by the type of gender cue, that is, the LAN effect was the same regardless of whether the gender feature was retrieved from the form (noun suffix) or the lemma. Our finding is in line with the findings reported by Caffarra and Barber (2015) and Caffarra et al. (2015), as we also manipulated gender retrieval from different gender cues (morphology, lexical syntax/lemma, and lexical semantics) without any difference in the LAN effect. This finding contributes further to the idea that the LAN may not be sensitive enough to discern different operations applied by the parser to retrieve the gender value (e.g., accessing the gender feature from morphological decomposition vs. deducing the gender feature from lexical semantics). The LAN seems to be only sensitive to the detection of a morphosyntactic anomaly, in this case gender disagreement, without reflecting the exact mechanism of the creation of that anomaly.

4.2. P600

The results show a positive deflection with mainly posterior-central distribution, corresponding to the P600 (e.g., Friederici, 1995, 2002; Friederici & Jacobsen, 1999; Friederici & Meyer, 2004; Kaan, Harris, Gibson, & Holcomb, 2000; Molinaro, Barber, & Carreiras, 2011). On visual inspection, the effect had a somewhat broader distribution in the earlier time window. The processing of (dis)agreement is generally understood to be a morphosyntactic process, which the P600 is sensitive to. In the current study, the P600 effect was of significantly greater amplitude in nouns with semantic gender than in nouns with syntactic gender. We interpret increased effect size in the early time window as indicative of syntax-semantics interaction in nouns with semantic gender. Since the semantic violation was caused by syntax, the N400 could not be elicited. Still, the violation was registered at a later stage once syntactic and semantic information was integrated.

The last expectation addressed the issue of the differences in the repair process between nouns with syntactic and semantic gender. The late stage of the P600 is assumed to reflect reanalysis and repair processes (Friederici, 1995, 2002; Friederici & Jacobsen, 1999; Friederici & Meyer, 2004; Kaan, Harris, Gibson, & Holcomb, 2000; Molinaro, Barber, & Carreiras, 2011; Popov & Bastiaanse, 2018), whose purpose is to enable the computation to continue unhindered, despite an illegal entry in the form of a syntactic violation (Molinaro et al., 2008). We proposed that the article-noun disagreement in nouns with syntactic gender contains a single repair option, whereas violations in nouns with semantic gender can be repaired in two ways. Therefore, we expected to find differences in the P600 effect between nouns with semantic and syntactic gender.

In the current study, the late P600 effect was larger for nouns with semantic gender than for nouns with syntactic gender. According to Otten and Rugg (2005), if an effect's distribution is the same for two conditions, the underlying cognitive processes are likely to be identical. By contrast, a larger amplitude in one condition signals a heavier processing load. This assumption aligns neatly with our hypothesis that nouns with semantic gender are more complex to repair. In case of a simple article-noun gender mismatch (e.g., *la*$_F$ *treno*$_M$ 'the train'), the parser can only repair the article (*la*$_F$ *treno*$_M$ > *il*$_M$ *treno*$_M$). If the disagreeing noun is marked for semantic gender as well as a gender transparent suffix (*-o* for masculine, *-a* for feminine), the parser has an extra repair option. In addition to repairing the article (*la*$_F$ *bambino*$_M$ > *il*$_M$ *bambino*$_M$ 'the boy'), it can also repair the noun (*la*$_F$ *bambino*$_M$ 'the boy' > *la*$_F$ *bambina*$_F$ 'the girl'). By doing this, the parser tackles two different linguistics levels: it alters the noun's gender feature stored at the lemma level (syntax), but also the referent's biological sex (semantics).

5. Conclusion

In the current study, we addressed several theoretically relevant issues. Firstly, we confirmed that semantic violations caused by syntax are not recognised as typical semantic violations, as demonstrated by the presence of the LAN instead of the N400. However, the interaction between the two processing levels is possible at the (early) stage of the P600. Secondly, we showed that an increase in repair complexity leads to an increase of the P600 amplitude in the late time window. In addition to adding to the existing knowledge of the processing stages, this study also addresses the question of semantic gender processing in (dis) agreement studies by measuring the effect on the source of gender information - the noun. We showed that there is indeed a semantic component in processing nouns with semantic gender. However, semantic processing is reflected by the P600 rather than the N400, as commonly expected in the literature.

6. Statement of significance

The current study focuses on the relationship between specific language operations during (dis)agreement processing and their reflection in ERP components. This is done for the first time by comparing gender violations on nouns with syntactic and nouns with semantic gender. The results show that violations at the syntax-semantic interface (in the semantic gender condition) are reflected in a larger P600 component than purely syntactic violations (in the syntactic gender condition). This suggests that violations at the syntactic-semantic interface are costlier to process and repair than solely syntactic violations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded through Erasmus Mundus Joint International Doctorate for Experimental Approaches to Language and Brain (IDEALAB) of the Universities of Groningen (NL), Newcastle (UK), Potsdam (DE), Trento (IT) and Macquarie University, Sydney (AU), under Framework Partnership Agreement 2012-0025 - specific grant agreement number 2013-1458/001-001-EMII EMJD by the European Commission to the first author.

Roelien Bastiaanse is partially supported within the framework of a subsidy by the Russian Academic Excellence Project '5-100'.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bandl.2020.104787.

References

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*, 137–153.

Barber, H., Salillas, E., & Carreiras, M. (2004). Gender or genders agreement? In M. Carreiras & C. Clifton (Eds.), On-line study of sentence comprehension; Eye-tracking, ERP and beyond (pp. 309–328). New York, Psychology Press.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews, 59*, 55–73.

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. Language in Mind: Advances in the Study of Language and Thought, 61-79.

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research, 1446*, 127–143.

Bošković, Ž. (2011). On unvalued uninterpretable features. In Proceedings of the North East Linguistics Society Annual Meeting (Vol. 39), 109–120.

Cacciari, C. (2011). Psicologia del linguaggio [Psychology of language]. Bologna, Italy: Il Mulino.

Caffarra, S., & Barber, H. A. (2015). Does the ending matter? The role of gender-to-ending consistency in sentence reading. *Brain Research, 1605*, 83–92.

Caffarra, S., Janssen, N., & Barber, H. A. (2014). Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia, 63*, 124–134.

Caffarra, S., Mendoza, M., & Davidson, D. (2019). Is the LAN effect in morphosyntactic processing an ERP artifact? *Brain and Language, 191*, 9–16.

Caffarra, S., Siyanova-Chanturia, A., Pesciarelli, F., Vespignani, F., & Cacciari, C. (2015). Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology, 52*, 1019–1030.

Canal, P., Garnham, A., & Oakhill, J. (2015). Beyond gender stereotypes in language comprehension: Self sex-role descriptions affect the brain's potentials associated with agreement processing. *Frontiers in Psychology, 6*, 1953.

Clarke, M. A., Losoff, A., McCracken, M. D., & Rood, D. S. (1984). Linguistic relativity and sex/gender studies: Epistemological and methodological considerations. *Language Learning, 34*, 47–64.

Clarke, M. A., Losoff, A., McCracken, M. D., & Still, J. (1981). Gender perception in Arabic and English. *Language Learning, 31*, 159–169.

Corbett, G. G. (1991). *Gender*. Cambridge: Cambridge University Press.

Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., & Job, R. (2005). Grammatical gender is selected in bare noun production: Evidence from the picture–word interference paradigm. *Journal of Memory and Language, 53*, 42–59.

D'Achille, P., & Thornton, A. M. (2006). I nomi femminili in –o [Feminine nouns ending in –o]. In E. Cresti (Ed.). *Prospettive nello studio del lessico italiano, Atti SILFI* (pp. 473–481). Firenze, Italy: FUP.

Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender-agreement in Hebrew: Evidence from ERPs and eye movements. *Journal of Memory and Language, 45*, 200–224.

Foundalis, H. E. (2002). Evolution of gender in Indo-European languages. *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 24). .

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language, 50* 259 259.

Friederici, A. D (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*, 78–84.

Friederici, A. D., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research, 28*, 467–484.

Friederici, A. D., & Meyer, M. (2004). The brain knows the difference: Two types of grammatical violations. *Brain Research, 1000*, 72–77.

Friederici, A. D., Steinhauer, K., & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory & Cognition, 27*, 438–453.

Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax–semantic interface. *Brain Research, 1146*, 50–58.

Gollan, T. H., & Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research, 30*, 627–651.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–112.

Gunter, T., Friederici, A., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal autonomy and late interaction. *Journal of Cognitive Neuroscience, 12*, 556–568.

Hagoort, P., & Brown, C. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research, 28*, 715–728.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8*, 439–483.

Hammer, A., Jansma, B. M., Lamers, M., & Münte, T. F. (2005). Pronominal reference in sentences about persons or things: An electrophysiological approach. *Journal of Cognitive Neuroscience, 17*, 227–239.

Hickey, R. (1999). The phonology of gender in Modern German. In Rissanen, Matti, & Unterbeck (Eds.), Gender in grammar and cognition (pp. 621–663). Berlin: Mouton-de Gruyter.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15*, 159–201.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*, 205–225.

Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research, 22*, 519–534.

Konishi, T. (1994). The connotations of gender: A semantic differential study of German and Spanish. *Word, 45*, 317–327.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647.

Kutas, M., & Hillyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.

Levelt, W. (1989). *Speaking: From intention to articulation.* Cambridge, Mass.: MIT Press.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–38.

Miozzo, M., & Caramazza, A. (1997). Retrieval of lexical–syntactic features in tip-of-the tongue states. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1410–1423.

Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2014). On the left anterior negativity (LAN): The case of morphosyntactic agreement. *Cortex, 66*, 156–159.

Molinaro, N., Barber, H., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex, 47*, 908–930.

Molinaro, N., Su, J. J., & Carreiras, M. (2016). Stereotypes override grammar: Social knowledge in sentence comprehension. *Brain and Language, 155*, 36–43.

Molinaro, N., Vespignani, F., & Job, R. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research, 1228*, 161–176.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In M. Carreiras & C. Clifton (Eds.), On-line study of sentence comprehension; Eye-tracking, ERP and beyond (pp. 271–308). New York, Psychology Press.

Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition, 25*, 273–285.

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes, 14*, 283–317.

Otten, L. J., & Rugg, M. D. (2005). Interpreting Event-Related Brain Potentials. In T. C. Handy (Ed.). *Event-related potentials: A methods handbook* (pp. 1–16). Cambridge, Mass: The MIT Press.

Pesetsky, D., & Torrego, E. (2007). The syntax of valuation and the interpretability of features. In S. Karimi, V. Samiian, & W. D. Wilkins (Eds.), Phrasal and clausal architecture: Syntactic derivation and interpretation (pp. 262–294). Amsterdam, John Benjamins Publishing.

Popov, S., & Bastiaanse, R. (2018). Processes underpinning gender and number disagreement in Dutch: An ERP study. *Journal of Neurolinguistics, 46*, 109–121.

Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience, 23*, 277–293.

Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language, 137*, 29–39.

Schmitt, B. M., Lamers, M., & Münte, T. F. (2002). Electrophysiological estimates of biological and syntactic gender violation during pronoun processing. *Cognitive Brain Research, 14*, 333–346.

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morpho-syntactic processing. *Neuropsychologia, 56*, 289–301.

Van Berkum, J. J., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research, 1146*, 158–171.

Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science, 8* 314 314.

Vigliocco, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language, 40* 455 455.

Vigliocco, G., & Hartsuiker, R. J. (2005). Maximal input and feedback in production and comprehension. In A. Cutler (Ed.). *Twenty-first century psycholinguistics: Four cornerstones* (pp. 209–228). NJ, Lawrence Erlbaum: Mahwah.

Vigliocco, G., & Zilli, T. (1999). Syntactic accuracy in sentence production: The case of gender disagreement in Italian language-impaired and unimpaired speakers. *Journal of Psycholinguistic Research, 28*, 623–648.

Weissman, B., & Tanner, D. (2018). A strong wink between verbal and emoji-based irony: How the brain processes ironic emojis during language comprehension. *PLOS one, 13*, e0201727.

Wicha, N. Y. Y., Moreno, E., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*, 1272–1288.