

University of Groningen

Processes underpinning gender and number disagreement in Dutch

Popov, Srdan; Bastiaanse, Roelien

Published in:
Journal of Neurolinguistics

DOI:
[10.1016/j.jneuroling.2018.01.001](https://doi.org/10.1016/j.jneuroling.2018.01.001)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Popov, S., & Bastiaanse, R. (2018). Processes underpinning gender and number disagreement in Dutch: An ERP study. *Journal of Neurolinguistics*, 46, 109-121. <https://doi.org/10.1016/j.jneuroling.2018.01.001>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Processes underpinning gender and number disagreement in Dutch: An ERP study

Srđan Popov^{a,b,*}, Roelien Bastiaanse^b

^a International Doctorate for Experimental Approaches to Language and Brain (IDEALAB), Universities of Groningen (NL), Newcastle (UK), Potsdam (DE), Trento (IT), Macquarie University (AU), Oude Kijk in 't Jatstraat 26, 9712 EK Groningen, The Netherlands

^b Center for Language and Cognition Groningen (CLCG), University of Groningen, Oude Kijk in 't Jatstraat 26, 9712 EK Groningen, The Netherlands

ARTICLE INFO

Keywords:

Gender
Number
Disagreement
P600
LAN

ABSTRACT

In the current experiment, participants read word-by-word sentences containing gender (adjective-noun) and number (article-noun) disagreement in Dutch while EEG was recorded. Number and gender disagreement were expected to elicit different responses due to several reasons. Firstly, gender is a lexical feature whose value (e.g., masculine or feminine) is stored in the lexicon, whereas number value is assigned depending on conceptual knowledge (numerosity). Also, Dutch marks number but not gender on the noun. Finally, due to the morphological nature of number, number disagreement provides more repair options than gender disagreement, thereby increasing the processing load. Both gender and number disagreement elicited a P600, but no LAN. The P600 effect was larger for number than gender disagreement in the late P600 stage. Since the observed effect was in the late P600 stage, we suggest that the most salient difference between the two types of disagreement lies in the increased repair complexity for number disagreement compared to gender disagreement.

1. Introduction

1.1. Number and gender as nominal features

Both gender and number are grammatical properties of nouns. In linguistic theory, they are often grouped together with person and labelled ‘phi-features’ (Adger & Harbour, 2008). All phi-features are always further specified with an appropriate value for their category, such as singular and plural for number. Quite often, this feature value also needs to be marked on an element different than the noun. For example, the article in Dutch has to have identical feature values as the noun it modifies. Therefore, if the noun is plural, the article has the plural form. However, the numerosity of the noun is obvious from the noun itself and feature reduplication onto the article is redundant in this case, as it does not bring any new information (Corbett, 2003). The process of establishing a relationship between two elements by transferring feature values from one to another is called agreement (e.g., Pesetsky & Torrego, 2007). The focus of this study is to investigate how number and gender features are decoded and processed in article/adjective-noun disagreement.

Agreement is, thus, a syntactic process whereby a relationship is established between elements at the syntactic level. But before agreement happens, the phi-features on the noun have to be valued. The way number and gender features receive and instantiate their value is where these features diverge most. In Dutch, gender is pervasively a lexical feature (e.g., Van Berkum, 1996). In terms of

* Corresponding author. Oude Kijk in 't Jatstraat 26, 9712 EK Groningen, The Netherlands.
E-mail address: s.popov@rug.nl (S. Popov).

<https://doi.org/10.1016/j.jneuroling.2018.01.001>

Received 11 July 2017; Received in revised form 21 December 2017; Accepted 8 January 2018

0911-6044/ © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

production theories, the gender value is stored as a separate node at the lemma level (e.g., [Levelt, Roelofs, & Meyer, 1999](#)). Therefore, the gender value is part of the noun's lexical syntax, and it is invariable. In some languages (e.g., Spanish and Italian), the gender feature is also overtly marked on the noun as a gender morpheme. Overt gender marking in Dutch is possible only when the noun contains a derivational suffix that is always associated with a specific gender value (e.g., nouns with the diminutive suffix *-(t)je* are always neuter). Number, however, is valued differently. Firstly, the speaker needs to assess the numerosity of the chosen concept: that is, whether there is only one or more than one entity. Once the value is determined, it is realized as a number morpheme on the noun (e.g., [Roelofs, 1997](#)). In case of Dutch, only plural nouns are morphologically marked (*-en, -s*).

The current study investigates the difference in the processing of the number and the gender feature embedded in agreement context in Dutch. This was achieved by carrying out an online event-related potentials (ERP) experiment based on a 'violation paradigm' (e.g., [Osterhout, McLaughlin, Kim, Greenwald, & Inoue, 2004](#)) comparing instances of grammatical gender and number agreement to gender and number disagreement/mismatch (determiner-noun). The goal of the study was two-fold. Firstly, we tried to find out whether the parser is sensitive to the inherent differences between gender and number features (i.e., lexical vs. conceptual; presence/absence of morphological marking). Secondly, since the experiment is based on a violation paradigm, we strived to show that structural repair processes are different for gender and number disagreement. More precisely, the number feature allows for more variation in its parameter settings at the level of a single noun (e.g., singular, *-inflection*; plural, *+inflection*), thereby increasing the complexity of repair processes which play an important role in ERP measures.

1.2. Previous ERP research on agreement

Previous studies on ERPs and sentence processing have mostly reported the presence of three language-related components: N400, left anterior negativity (LAN), and P600 ([Coulson, King, & Kutas, 1998](#); [Friederici, 1995](#); [Hagoort, Brown, & Groothusen, 1993](#); [Kutas & Hillyard, 1980](#); [Osterhout & Holcomb, 1992](#)). These components are usually elicited through violation paradigms consisting of two sets of identical sentences, differing minimally at a single point. The first set comprises meaningful grammatical sentences (baseline) against which identical sentences containing a grammatical or semantic violation are compared. In case of a semantic violation, the expected response is the N400. This component is construed as a marker of semantic and discourse integration difficulties ([Friederici, 2002](#); [Kutas & Federmeier, 2011](#); [Kutas & Hillyard, 1980](#)).

The LAN and P600 are usually described as markers of syntactic processing ([Friederici, 2002](#)), which makes them expected components in agreement studies. Indeed, a large number of studies on agreement reported a biphasic response to agreement mismatch in the form of the LAN followed by the P600 ([Barber & Carreiras, 2005](#); [Barber, Salillas, & Carreiras, 2004](#); [Gunter, Friederici, & Schriefers, 2000](#); [Molinaro, Vespignani, & Job, 2008](#)). As the name says, the left anterior negativity is a negative deflection peaking between 300 and 500 ms. It is usually left lateralized with anterior distribution (but see [Osterhout et al., 2004](#)). As to its function, it is characterized as being sensitive to morphosyntactic errors (e.g., [Friederici, 2002](#); [Molinaro, Barber, & Carreiras, 2011](#)). The LAN is usually followed by a positive deflection peaking at 600 ms (P600). Some authors make a distinction between the early and late P600 (e.g., [Hagoort & Brown, 2000](#)). The early P600 lasts from 500 ms to 700 ms post-stimulus onset and has a broad distribution, whereas the subsequent late P600 is the strongest in the parietal regions. In addition to different topography, the two stages are stipulated to be somewhat functionally different. The early stage reflects integration difficulty, which is followed by reanalysis and repair in the late stage.

Based on the previous study, the most reliable agreement processing marker seems to be the P600. It is almost unanimously reported, which is not the case with the LAN (e.g., [Aleman Bañón, Fiorentino, & Gabriele, 2012](#); [Nevins, Dillon, Malhotra, & Phillips, 2007](#)). As an illustration, several studies on Italian and Spanish gender and/or number agreement reported the LAN followed by the P600 (e.g., [Caffarra & Barber, 2015](#); [Caffarra, Siyanova-Chanturia, Pesciarelli, Vespignani, & Cacciari, 2015](#); [Dowens, Vergara, Barber, & Carreiras, 2010](#); [Molinaro et al., 2008](#); [O'Rourke & Van Petten, 2011](#)). However, studies on determiner-noun agreement in Dutch failed to report either LAN or the N400 ([Loerts, Stowe, & Schmid, 2013](#); [Meulman, Stowe, Sprenger, Bresser, & Schmid, 2014](#)). [Hagoort and Brown \(1999\)](#) proposed that the LAN can only be elicited by a morphologically overt violation, such as number agreement. This explains the absence of the LAN in gender violations in Dutch in which gender is a lexical feature. However, it still does not explain the lack of the LAN effect in number violations. Moreover, this explanation goes against the findings from Italian ([Caffarra et al., 2015](#)) and Spanish ([Caffarra, Barber, Molinaro, & Carreiras, 2017](#)) in which LAN was obtained for both transparent and opaque nouns, demonstrating that overt morphology may not be crucial for eliciting the LAN.

Similarly, two studies on Spanish determiner-noun gender agreement found conflicting results regarding the LAN. [Barber and Carreiras \(2005\)](#) reported the LAN followed by the P600, whereas [Wicha, Moreno, and Kutas \(2004\)](#) only found the P600. The volatility of the LAN in terms of its seemingly random distribution across studies has not been explained yet, even though several accounts have been offered.

In addition to the functional explanation by [Hagoort and Brown \(1999\)](#), [Molinaro, Barber, Caffarra, and Carreiras \(2014\)](#) indicated that methodological and technical factors could play a role regarding the LAN, such as the choice of the reference electrode. [Osterhout \(1997\)](#), [Tanner \(2015\)](#), and [Tanner and Van Hell \(2014\)](#) suggested that the presence of the LAN might be due to individual variations among participants (see also [Pakulak & Neville, 2010](#)), as well as to the averaging nature in obtaining ERP components.

1.3. Gender and number agreement in Spanish

Both number and gender are nominal phi-features. As such, they are often believed to affect processing mechanisms in the same way, to the extent that some studies lumped sentences with number and gender violations together as one syntactic condition (e.g.,

Hagoort, 2003; Martín-Loeches, Nigbur, Casado, Hohfeld, & Sommer, 2006). Indeed, agreement studies investigating gender (e.g., Gunter et al., 2000; Molinaro et al., 2008) and number (e.g., Münte, Szentkuti, Wieringa, Matzke, & Johannes, 1997) separately usually come to the same results and conclusions: both gender and number violations elicit the P600, which is sometimes preceded by the LAN. The processing mechanism is, hence, understood to be identical: the morphosyntactic violation is identified by the LAN already 300 ms post-stimulus onset, after which the violation is repaired, as indicated by the P600. However, in this way it is impossible to compare the effect size (e.g., amplitude size of the P600) between number and gender, which is as important as the presence/absence of a component or its distribution.

Barber and Carreiras (2005) tested Spanish determiner-noun and noun-adjective number and gender agreement in a single study. Gender in Spanish is a lexical feature, being part of the noun's lemma. However, unlike in Dutch, a large proportion of Spanish nouns are gender transparent (Teschner & Russell, 1984). Most nouns end in either *-o* or *-a* indicating that they are masculine or feminine, respectively. If the noun is used in plural, a suffix *-s* is added onto the gender suffix. The study showed that both gender and number disagreement elicit the LAN and the P600. Crucially, number and gender disagreement differed in the late P600 stage, in which the effect was larger for gender. The authors suggested that repair processes in gender are costlier due to the lexical nature of gender (Faussart, Jakubowicz, & Costes, 1999; Ritter, 1991, 1993). However, Aleman Bañón et al. (2012) looked into processing number and gender agreement in Spanish and failed to find any difference. Their rationale was that the parser processes both features in a similar fashion regardless of their inherent differences (Nevins et al., 2007).

An important aspect of gender in Spanish is that it is often transparent, with the word-final vowel indicating gender. This fact is of great importance for reading studies and processing accounts based on reading. In an ERP reading paradigm on sentence processing, words are presented one-by-one on the screen with an average duration of 300–350 ms. This is enough just for one fixation (Rayner & Clifton, 2009), that is, the eyes will fixate the word as a whole and the visual system will perceive it as one unit. An integral part of that unit is the gender morpheme, as well as the number inflection. In other words, as soon as the system perceives the word, it has the word's gender and number information at its disposal. Therefore gender, just like number, can be available from the suffix in addition to retrieving its value from the lemma (see Caffarra & Barber, 2015; Caffarra et al., 2015). Of course, this is only true in case of gender-transparent languages.

1.4. Current study

The current experiment was conducted in Dutch, as it allows for comparing a morphologically realized feature (number), and a feature lacking inflectional morphology (gender). Barring several exceptions (e.g., diminutives), nouns in Dutch are mainly gender-opaque, whereas plural nouns are always morphologically marked. As mentioned earlier, the gender feature is invariable, meaning that a noun, such as *het boek*_{N(outer)} ‘the book’ cannot become **de boek*_{C(common)} under any circumstances. Number, however, is derived from higher order semantics and our knowledge about the numerosity of the object in question. If the object is singular, it is not morphologically marked (*het boek*), and if there is more than one object, its form becomes *de boeken*.

Based on behavioural data, Lukatela, Kostić, Todorović, Carello, and Turvey (1987) proposed that the parser behaves in a binary way regarding syntactic violations. More precisely, the parser is only sensitive to the presence or absence of a violation, without any more detailed decomposition of the violation source. This is applicable to agreement mismatches, in which the parser detects the violation disregarding whether it is number, gender or case violation. In line with Lukatela et al.'s proposal are ERP results by Nevins et al. (2007), who found identical effects for both gender and number disagreement.

Based on the idea of parser's binarity, as well as on previous electrophysiological results, we expect that number and gender disagreement elicit the same syntactically-related components, that is, the LAN and P600. Functionally, the first effect should be the LAN as a marker of morphosyntactic incongruence. However, due to the lack of the LAN in previous studies on gender and number processing in Dutch (Loerts et al., 2013; Meulman et al., 2014), it is possible that the effect will be absent.

The first hypothesis is concerned with the nature of the nominal feature: gender is a lexical and invariable feature with a pre-determined value, whereas number is semantically-derived and can be either singular or plural. If such a distinction plays a role while processing agreement mismatch, it would be detectable either at the first syntactic parse (LAN) or at the level of structural integration (early P600).

Closely related to the previous hypothesis is the issue of how the feature value is realized. As already mentioned, gender feature is lexical and available directly from the lemma. Conversely, number is morphologically realized as a plural suffix *-s* or *-en*. We stipulate that the parser may be able to access the inflectional feature (number) faster, as it is perceptually more salient. Unlike gender, which has to be retrieved from the lemma, the number value is available off the suffix the same moment the whole word is fixated without accessing the lemma. Thus, higher perceptual saliency of number may be reflected in an earlier component onset (LAN/P600), as the number feature is probably accessed earlier than the gender feature.

Finally, we hypothesize that structural repair processes are of different complexity in gender and number disagreement. The difference is expected to be recorded in the late stage of the P600 in which repair and reanalysis processes are stipulated to take place (Friederici, 2002; Hagoort & Brown, 2000). Barber and Carreiras (2005) suggest that gender disagreement is costlier to repair than number disagreement due to the lexical nature of gender as opposed to morphologically marked number. However, in the current study, we assume that number disagreement is more complex to repair as it offers more repair options, which does not seem to be the case for the (sentence) stimuli used by Barber and Carreiras. For example, in gender disagreement the neuter noun *boek* is preceded by an adjective marked for common gender **een grote*_C *boek*_N ‘a big book’. The parser repairs the incongruity by correcting the gender inflection on the adjective (**een grote*_C *boek* > *een groot*_N *boek*_N). In number disagreement, in addition to repairing the singular article into the plural article (equivalent to repairing the gender inflection on the adjective)

(**het*_{SG} *boeken*_{PL} > *de*_{PL} *boeken*_{PL}), the parser can also repair the inflectional morpheme on the noun (**het*_{SG} *boeken*_{PL} > *het*_{SG} *boek*_{SG}). In other words, the parser can apply only one operation on gender disagreement, but two operations on number disagreement. As a consequence of the increased processing load due to multiple repair processes in number, we expect that the P600 effect will be larger for number than gender disagreement.

2. Method

2.1. Participants

Thirty participants were tested for this experiment (8 male; mean age 22.1; age range 19–33). Out of the total number, 4 participants were excluded prior to analysis due to low scores on the grammaticality judgment task (accuracy below 80%). Furthermore, 2 participants were excluded due to an excessive number of artifacts, leaving 24 participants whose results are reported in this study. All participants were right-handed, which was asserted with a Dutch version of the Edinburgh Handedness Inventory (Oldfield, 1971). The participants declared no reading impairment, or any history of psychiatric or neurological illness. They all had normal or corrected-to-normal vision. All the participants signed a consent form prior to the experiment, and received a € 20 voucher for their participation. The experiment was approved by the local ethics committee (Research Ethics Committee (CETO), Faculty of Arts, University of Groningen).

2.2. Acceptability ratings for the materials

An acceptability judgment survey was carried out in order to verify the grammaticality/ungrammaticality of experimental stimuli and fillers. Both grammatical and ungrammatical sentences were included in the survey. The question asked was: 'Is this sentence grammatically correct?' The offered answers were 'yes' and 'no'. The instructions included an example. There were a total of 480 sentences divided over 4 lists. The survey was set up on an online platform SurveyMonkey (www.surveymonkey.com). In total, 32 native speakers of Dutch took the survey (9 male; mean age 33.18; age range 21–62). Each participant filled only one survey. In order to include a stimulus into the experimental set, it had to have an approval rate of at least 80%, that is, 80% of the participants had to correctly judge the sentence as grammatical or ungrammatical. Out of 480 sentences, 16 sentences did not reach the inclusion threshold. These sentences were modified and assessed by another 3 speakers of Dutch who unanimously approved the modified sentences.

2.3. Materials

The materials used in the experiment consisted of 320 experimental sentences and 160 fillers. The experimental sentences were created on a basis of 40 unique nouns, half of which were monosyllabic and the other half trisyllabic. The nouns were controlled for noun-verb homophony, phonological alternations, and animacy. In addition, all nouns had to use the suffix *-en* exclusively for the plural form, which in turn could not be homophonous to an infinitival verb form (e.g., *boek* 'book' > *boeken* 'books', but also 'to book'). Nouns could not display any graphemic/phonological alternations between the singular and the plural form (e.g., voicing *huis* 'house' > *huizen* 'houses' or irregular plurals *stad* 'city' > *steden* 'cities'), and the vowel length had to be maintained (e.g., no nouns with short-long vowel alternation, such as *pad* 'path' > *paden* 'paths'). Finally, all nouns were inanimate.

Each noun yielded 4 discrete sentences, with each sentence used once as grammatical and once as ungrammatical. If the full repertoire of nouns had been used, a participant would have been exposed to the same noun 8 times. In order to reduce this, as well as to make the experiment shorter, stimuli were divided over 2 lists. Each participant was exposed to only one list. Each target noun appeared 4 times in a list, always in a different sentence. Items were counterbalanced between lists in such a way that if a grammatical sentence was in the first list, its ungrammatical counterpart was in the second. Consequently, each participant read 160 experimental and 80 filler sentences.

All experimental sentences were divided into two conditions: gender (80 sentences per list) and number (80 sentences per list). In Dutch, the common gender article *de* is homophonous with the plural article *de*, which is used for both common and neuter nouns in plural. This could lead to a possible ambiguity in the gender violation condition, such as **de*_C *mes*_N 'the knife', in which the violation becomes apparent only at (the end of) the noun. Since *de* is always used as the definite plural article, the parser may expect to encounter *de* *messen* 'the knives'. Once it becomes obvious that there is no plural suffix *-en*, the parser registers a violation. Without any additional context, this violation is ambiguous between number and gender violation. More precisely, since *mes* is a neuter (*het*) noun, **de* *mes* may be recognized as a gender violation provided that *de* is reanalyzed as a singular common article. However, if the parser classifies *de* as a plural article, the violation is perceived as a number violation. We also do not rule out a possibility that the parser may perceive **de* *mes* as a double violation. In order to overcome this obstacle, gender violations were created between the indefinite article *een* (used only in singular for both genders) and a gender marked adjective followed by a noun; for example: **een* *mooie*_C *dorp*_N 'a beautiful village'. The indefinite article indicates to the parser that it should expect a singular noun. The following adjective (*mooie*) is marked with *-e*, which according to the inflectional rules indicates the noun has to be of common (*de*) gender. By combining the two pieces of information, the parser is ready to encounter a singular *de* noun. However, in the violated sentences the following noun is always a singular *het* noun. Therefore, an unambiguous gender violation is created between the indefinite article-adjective complex and the target noun.

There was an additional reason for including only nouns of the *het* type. As already mentioned, Dutch gender is a lexical feature, that is, it has to be retrieved from the lemma. The only exception is nouns derived with a gender specific suffix, such as the diminutive suffix. Almost all Dutch nouns can have a diminutive form (derived through the suffix *-tje* and its allomorphs), the usage of which is also relatively high (Shetter, 1959). All diminutives are of *het* type (e.g., *de tafel* ‘the table’, *het tafeltje* ‘the little table’). Consequently, in a violated condition, such as *een *rood_N tomaat_C* ‘a red tomato’ the parser recognizes the violation only at the end of the word. Again, the indefinite article means the noun should be singular, whereas the adjective form indicates the noun is of the *het* type. However, since almost any Dutch noun can be used as a diminutive, it is possible that the parser’s strategy is to expect a diminutive noun. In order to avoid this possible strategy, we decided not to use *de* nouns as experimental items.

All gender sentences were created in two structural ‘molds’. In the first mold (1), the sentence started with an expletive subject (e.g., *er* ‘it/there’) or a general place adverbial (e.g., *hier* ‘here’). The subject was followed by a verb, either lexical or auxiliary, after which the indefinite article *een* ‘a/one’ was presented. The second sentence type (2) started with a personal pronoun (e.g., *hij/zij* ‘he/she’) followed by a verb, which was followed by the indefinite article. In both sentence types, the indefinite article was followed by an adjective. In non-violated sentences, the adjective was always inflected with the suffix *-e*. In violated sentences, the adjective had a zero marking used with *het* nouns. The target noun was placed after the adjective. Also, the target noun was never at the end of a sentence; it was always followed by a prepositional phrase, adverbial, or a lexical verb.

- (1) Er lag een mooi dorp vlakbij de grote stad.
there lay a beautiful_N village_N near the big city
‘A beautiful village was close to the big city.’

*Er lag een mooie dorp vlakbij de grote stad.
there lay a beautiful_C village_N near the big city.
- (2) Zij verliet een prachtig dorp met pijn in haar hart.
she left a beautiful_N village_N with pain in her heart
‘She left a beautiful village with an aching heart.’

*Zij verliet een prachtige dorp met pijn in haar hart.
she left a beautiful_C village_N with pain in her heart

All sentences in the number condition had the same structure (3, 4). They began with the plural article *de* in grammatical sentences and the singular neuter article *het* in violated sentences. The article was followed by an inflected adjective, which is the correct form for both singular *het* and plural *de* nouns, after which the target noun was presented. Like in the gender condition, the noun was always followed by a prepositional phrase, adverbial phrase, or a lexical verb.

- (3) De gezellige dorpen trekken veel toeristen in de zomer.
the_{PL} nice_{PL} villages_{PL} attract many tourists in the summer
‘The nice villages attract many tourists in the summer.’

*Het gezellige dorpen trekken veel toeristen in de zomer.
the_{SG} nice_{SG} villages_{PL} attract many tourists in the summer
- (4) De noordelijke dorpen hebben last van aardbevingen.
the_{PL} northern_{PL} villages_{PL} have trouble from earthquakes
‘The northern villages are troubled by earthquakes.’

*Het noordelijke dorpen hebben last van aardbevingen.
the_{SG} northern_{SG} villages_{PL} have trouble from earthquakes

We explained the reasons for having only *het* nouns as experimental items. A consequence of this choice is that it can facilitate a learning strategy in participants. In other words, when a participant encounters a sentence starting with *het*, he may learn after a few items that such sentences are always ungrammatical. Similarly, any instance of a zero-marked adjective indicates an ungrammatical sentence. To prevent this possibility, 160 filler item sentences with a pattern reversed to that of the experimental stimuli were included. Half of the filler items contained *het* nouns (5) and the other half *de* nouns (6). The *de* noun-group was used to counterbalance the gender condition, whereas the *het* noun-group counterbalanced the number condition. Consequently, it was impossible to judge a sentence as grammatical or ungrammatical based only on the article or the adjective. The participant had to pay attention to everything preceding the noun, as well as the noun itself, in order to correctly judge the sentence.¹

¹ Based on the stimulus opening (definite article for number, everything else for gender), participants might develop a strategy for identifying whether a sentence belonged to the gender or number condition. As a consequence, the processing of the critical noun region could be different compared to an experimental setting in which such a strategy would be impossible. However, we highlight again that it was impossible to guess whether the sentence was grammatical or ungrammatical. Therefore, even if participants employed the guessing strategy, it would have been applied to both grammatical and ungrammatical sentences in both conditions. Since our results are based on the difference between grammatical and ungrammatical sentences, being able to guess which condition a sentence belonged to should not interfere with our results.

- (5) Er ligt een rotte tomaat in de koelkast.
 there lie a rotten_C tomato_C in the fridge
 ‘There is a rotten tomato in the fridge.’
 *Er ligt een rot tomaat in de koelkast.
 there lie a rotten_N tomato_C in the fridge
- (6) Het oude paspoort is niet meer geldig.
 the_{N,SG} old_{N,SG} passport_{N,SG} is not more valid
 ‘The old passport is not valid anymore.’
 *De oude paspoort is niet meer geldig.
 the_{C/PL} old_{C/PL} passport_{N,SG} is not more valid

2.4. Procedure

Participants were seated in front of the screen at a distance of 70–80 cm. The experiment was presented in E-Prime (Psychology Software Tools, Inc). The passive task was to read the sentences presented word-by-word on the screen. The active task was to reply to a randomly assigned grammaticality judgment question. On average, a question appeared once for every five sentences. The purpose of the active task was to keep participants focused. The experiment opened with written instructions that were repeated by the experimenter. There was a brief practice session (4 sentences) in order to ensure that participants had understood the instructions and were able to follow the stimulus presentation. After the last practice item, participants had a chance to ask for clarifications or more detailed instructions. Once they were ready, they could proceed to the experimental part by pressing any keyboard button. Each trial opened with a fixation cross (500 ms) and a break (200 ms), after which the first word was presented (400 ms). The stimulus onset asynchrony (the time between the onsets of two subsequent words) was 600 ms. The last word in the sentence was presented with a full stop. Sentences were shown on a black background with white letters. The letter font was Arial and letter size was 24 pt. After the last word was presented, the screen remained blank for 500 ms. In case there was a grammaticality judgment question, a question mark appeared after the 500 ms break. The question remained on the screen for 3 s, during which participants were supposed to press either ‘p’ or ‘q’ (counterbalanced across participants), depending on whether the previous sentence was grammatical or not.

The experiment lasted approximately 30 min. Stimuli were divided into 4 blocks, each containing 40 experimental and 20 filler items. There were 12 grammaticality judgment questions per block, 8 for experimental and 4 for filler items. The presentation order within a block was random, as determined by the software. The participants were advised to take a short break after each block.

2.5. EEG data acquisition and processing

Continuous EEG data were recorded using the ASA-Lab system (ANT Neuro Inc, Enschede, The Netherlands) from 64 Ag/AgCl scalp electrodes fitted in an elastic cap (WaveGuard). Electrodes were positioned according to the extended 10–20 system. Eye movements were recorded using one bipolar channel for horizontal movements (HEOG; the electrodes were placed at the outer canthus of the eyes) and one for vertical movements (VEOG; placed above and below the left eye). Electrode impedances did not exceed 10 k Ω , and were kept at 5 k Ω or below in the large majority of cases. Data were sampled at 512 Hz with the common average reference.

Data were pre-processed with Brain Vision Analyzer 2.04 (Brain Products, GmbH, Munich, Germany). The first step was to down-sample the data to 256 Hz in order to speed up the analysis; this was followed by re-referencing the offline data to the average of the left and right mastoid. Afterwards, a band-pass filter was applied (0.1–40 Hz) after which an automatic ocular correction was performed. The continuous data were segmented into 1700 ms long epochs, starting 200 ms before the trigger marker (target noun onset). The automatic artifact rejection ($\pm 100 \mu\text{V}$ threshold, minimal activity $0.1 \mu\text{V}$) was performed in the interval of -100 ms to 1000 ms for each epoch. Approximately 4% of all trials were excluded, with no difference in the rejection rate between the four conditions (*gender grammatical*: 3.95%, *gender ungrammatical*: 4.38%, *number grammatical*: 4.98%, *number ungrammatical*: 3.55%; $F(3, 69) = 1.04, p > .1$). Electrodes with a high artifact contamination rate ($> 20\%$) were interpolated (1 electrode in 5 participants). Finally, the baseline correction was applied starting -100 ms until 0 ms after which data were averaged per subject and per condition. If a participant had fewer than 70% averaged trials in one or more conditions, his data were excluded from the analysis. This resulted in excluding the data of 2 participants.

2.6. Analysis

For the analysis, we used averaged participant values (in μV) per condition, level of grammaticality, and regions of interest (ROI). Regions of interest (Fig. 1) were created by averaging the values of 5–6 adjacent electrodes (50 in total), which resulted in 9 ROIs: 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). For the statistical analysis, 4 time windows were created based on the literature and visual data inspection: 1) 300–450 ms corresponding to the LAN/N400; 2) 450–600 ms onset of the P600; 3) 600–800 ms early P600; 4) 800–1000 ms late P600. The LAN time window is identical to the one used by Barber and Carreiras (2005) in a similar agreement study. The P600 window was mainly based on the visual inspection, and the

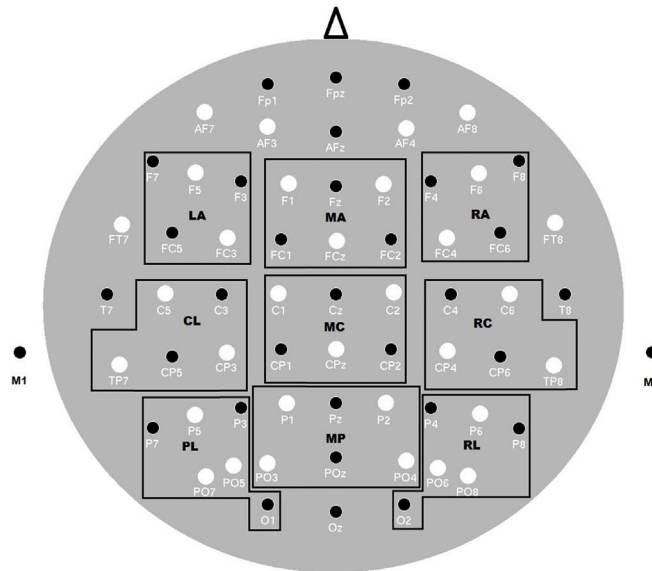


Fig. 1. Electrode positions and the 9 regions of interest used in the analysis: left anterior (LA), left central (LC), left posterior (LP), midline anterior (MA), midline central (MC), midline posterior (MC), right anterior (RA), right central (RC), right posterior (RP).

450–600 ms time window was used for the detection of the onset of the P600 effect. The early P600 (600–800 ms) and late P600 (800–1000 ms) windows are identical to the time windows in an agreement study by Nevins et al. (2007), in which the early and late P600 were reported.

The following within subject factors were included in a repeated measure ANOVA: 1) condition (2 levels: gender and number); 2) grammaticality (2 levels: grammatical and ungrammatical); 3) hemisphere (2 levels: left and right); 4) anteriority (3 levels: anterior, central, and posterior). The global analysis for each time window was performed by two separate ANOVAs. The first ANOVA analyzed only lateral the regions and it included all 4 factors. A second omnibus ANOVA was run on the midline regions only, excluding the factor hemisphere. In case the assumption of sphericity was violated, the Greenhouse and Geisser (1959) correction was applied. The significance level was set to $p < .05$. Follow-up tests were performed only for interactions that were at least marginally significant ($p < .1$) and that included factor grammaticality. Bonferroni adjustment was used for multiple comparisons. Finally, incorrectly judged trials were not included in the analysis.

3. Results

3.1. Accuracy results

The cut-off for including a participant in the data analysis was set at 80%, meaning that each participant had to correctly respond to 51 out of 60 questions. Of the 30 participants, 4 were excluded due to a low score on the grammaticality judgment task. The remaining 26 participants had an accuracy rate of 94% (average number of errors: 3.6, SD 2.6). Four participants performed at ceiling. Since the only purpose of the grammaticality judgment question was to ensure the participants' alertness throughout the entire experiment, these data were not further analyzed.

3.2. ERP results

A visual inspection of the waveforms indicated a centro-parietal positive effect from approximately 500 ms post-stimulus onset. The effect was caused by ungrammatical sentences in both conditions. The positivity seemed to be of somewhat smaller amplitude in the gender condition (Fig. 2) compared to the number condition (Fig. 3). Contrary to expectations, the positivity was not preceded by a left-lateralized negative effect in either condition.

The first time window (300–450 ms) did not yield any significant effects or interactions. This holds true for both the lateral and midline analysis. Since the LAN was expected in this time window, we performed a hypothesis-driven ANOVA in the left anterior region only. Still, the factor grammaticality did not reach significance ($F(1, 23) = 0.278, p > .1$).

In the following time window (450–600 ms), the omnibus ANOVA on the lateral regions revealed a main effect of grammaticality ($F(1, 23) = 4.313, p < .05$), with ungrammatical sentences showing a more positive waveform.

The midline results mirrored the lateral results. The global midline ANOVA produced a main effect of grammaticality ($F(1, 23) = 8.448, p < .01$), with ungrammatical sentences eliciting a more positive waveform than grammatical sentences.

The positive effect continued into the 600–800 ms time window, with ungrammatical sentences in both conditions yielding a

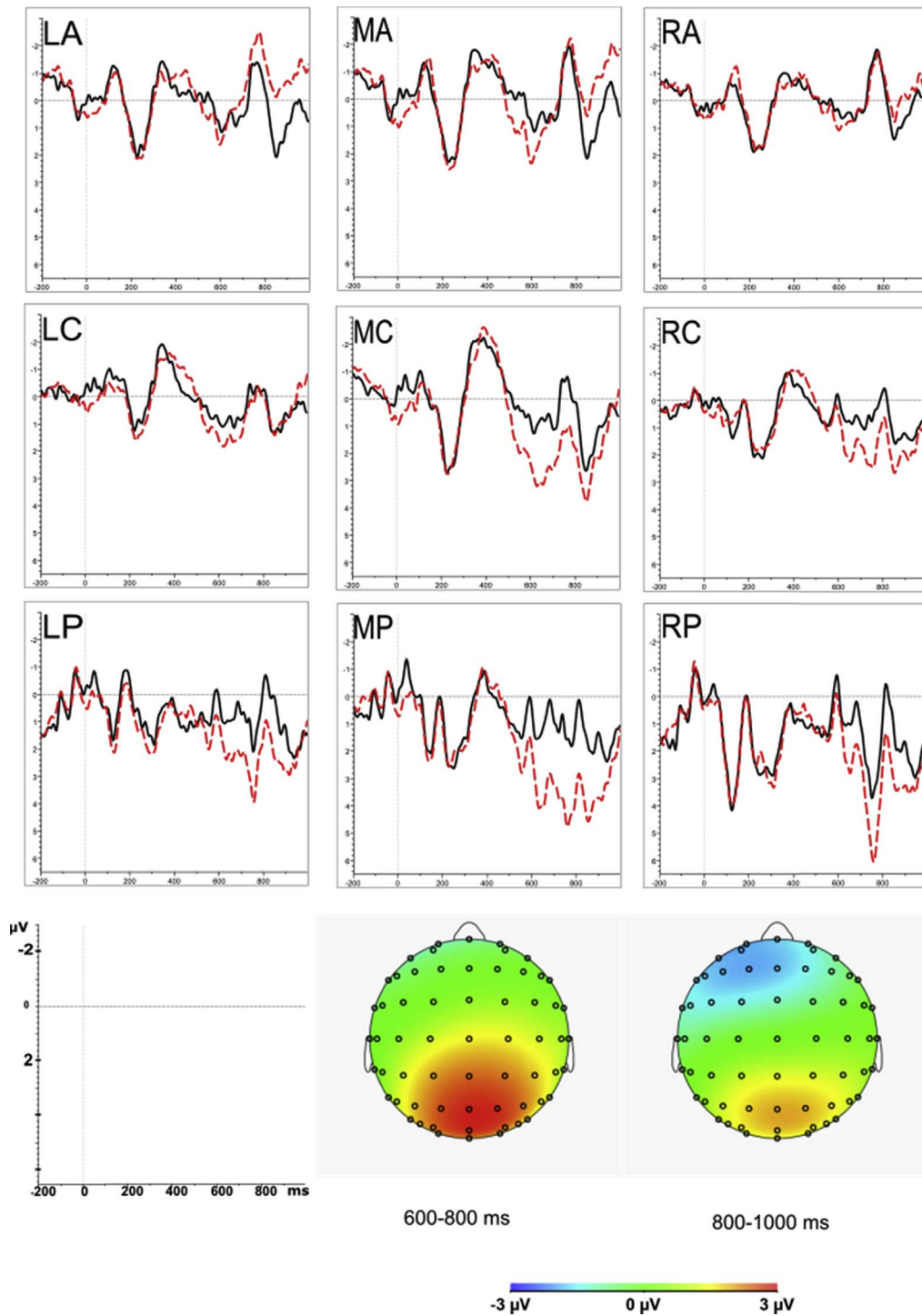


Fig. 2. Grand average ERPs for the gender condition across all 9 ROIs: full black line represents correct sentences and dashed red line represents violated sentences. The topographic maps represent a difference between ungrammatical and grammatical sentences. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

more positive waveform (grammaticality: $F(1, 23) = 10.45, p < .01$). In addition, the effect of grammaticality interacted with anteriority ($F(2, 46) = 14.22, p = .001$), and it also entered into a marginal three-way interaction with anteriority and hemisphere ($F(2, 46) = 3.283, p < .1$). Follow up t -tests for each lateral region of interest showed that ungrammatical sentences elicited a more positive waveform than grammatical sentences in both posterior regions (left: $t(23) = -4.315, p < .01$; right: $t(23) = -4.355, p < .01$); and also in the right central region ($t(23) = -3.68, p = .01$).

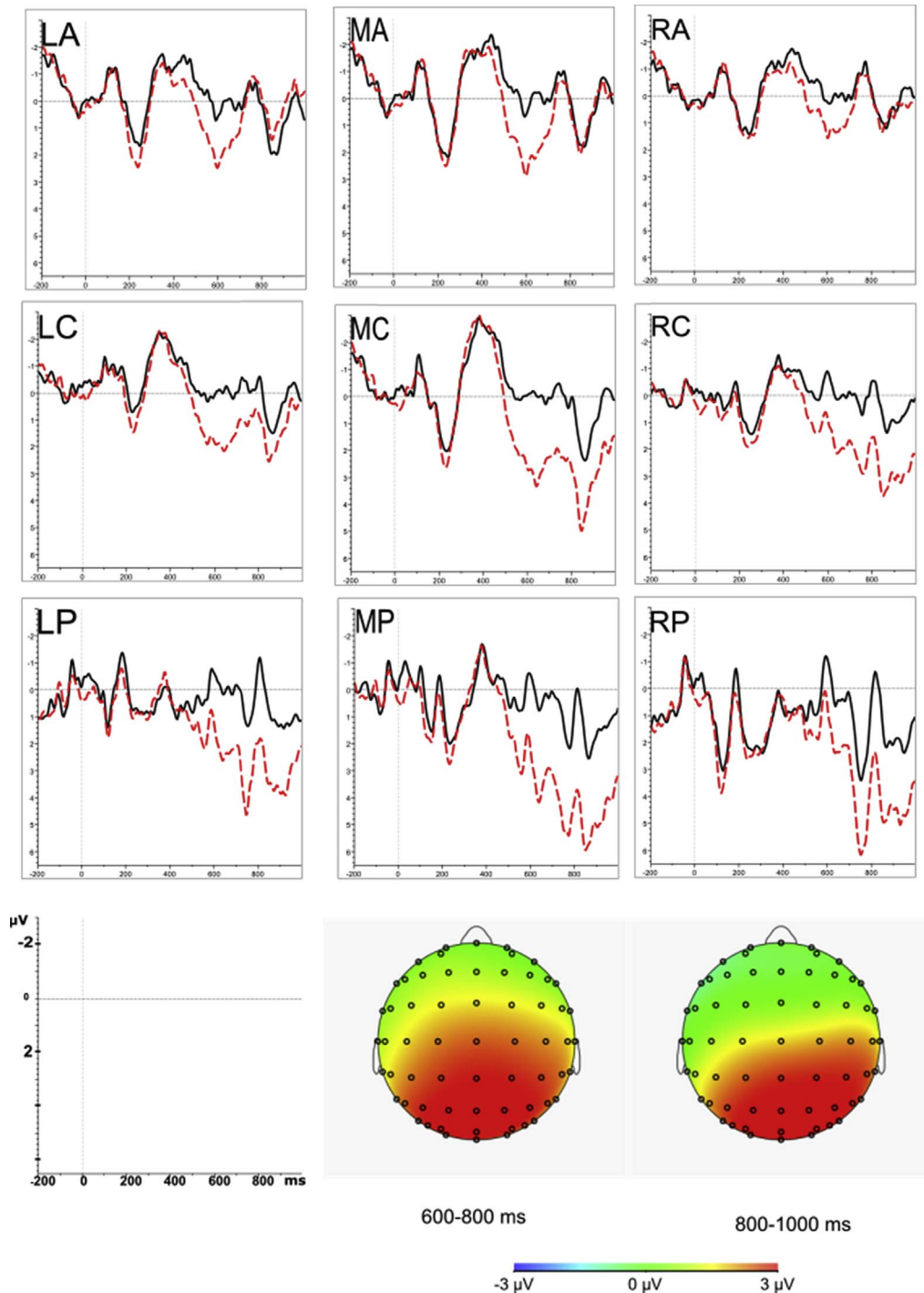


Fig. 3. Grand average ERPs for the number condition across all 9 ROIs: full black line represents correct sentences and dashed red line represents violated sentences. The topographic maps represent a difference between ungrammatical and grammatical sentence. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

The midline analysis revealed a main effect of grammaticality ($F(1, 23) = 13.519, p = .001$), which also interacted with anteriority ($F(2, 46) = 22.123, p < .001$). The positive effect like the one from the midline analysis was significant in the posterior ($t(23) = -5.058, p < .001$) and central region ($t(23) = -3.701, p < .01$).

Finally, the positive deflection caused by ungrammatical sentences in both conditions persisted in the latest time window of 800–1000 ms, but only in an interaction with anteriority ($F(2, 46) = 43.49, p < .001$) and hemisphere ($F(23) = 16.602, p < .001$)

or both ($F(2, 46) = 5.581, p < .05$). The main effect of grammaticality was only marginally significant ($F(1, 23) = 3.572, p < .1$). A set of 6 follow-up t-tests was performed for each region of interest in order to pinpoint the exact distribution of the positive effect. Identically to the previous time window, the positive effect was the strongest in the posterior regions (left: $t(23) = -3.757, p < .01$; right: $t(23) = -4.698, p < .001$). Importantly, a close-to-significant interaction between condition and grammaticality was obtained in the overall ANOVA ($F(1, 23) = 3.674, p < .1$). The follow-up tests showed that the effect of grammaticality was significant in the number condition ($t(23) = -5.675, p < .001$), while being absent from the gender condition ($t(23) = -0.146, p > .1$). The follow-up also showed that ungrammatical number sentences elicited a significantly more positive waveform than ungrammatical gender sentences ($t(23) = -6.577, p < .001$), while there was no difference between grammatical number and grammatical gender sentences ($t(23) = 0.471, p > .1$).

The midline analysis revealed a main effect of grammaticality ($F(1, 23) = 5.087, p < .05$), as well as an interaction between grammaticality and anteriority ($F(2, 46) = 48.22, p < .001$). Further testing showed that the positive effect in the midline was present in the posterior region only ($t(23) = -4.978, p < .001$). Most importantly, a significant interaction surfaced between condition and grammaticality ($F(1, 23) = 4.47, p < .05$). Upon further testing, it turned out that the positive effect was entirely driven by the difference between grammatical and ungrammatical sentences in the number condition ($t(23) = -2.808, p < .05$), while being absent from the gender condition ($t(23) = -0.607, p > .1$). Lastly, ungrammatical number sentences showed a trend towards eliciting a significantly more positive waveform than ungrammatical gender sentences ($t(23) = -2.498, p < .1$), while the comparison between grammatical sentences of both conditions did not yield a significant result ($t(23) = 0.522, p > .1$).²

3.3. Summary of ERP results

The statistical analysis confirmed a significant centro-parietal positivity elicited by ungrammatical sentences (P600). The P600 was not preceded by either a left lateralized negativity (LAN) or a central negativity (N400). In the P600 time windows, the distribution of the effect always included the two lateral posterior regions, as well as the right central region. In the midline, the distribution included the posterior and central region from 600 to 800 ms, retreating to the posterior region only in the last time window (800–1000 ms). Most importantly, the P600 effect was longer lasting in the number condition, spanning a large time window from 450 ms until 1000 ms. The effect in the gender condition was, however, statistically detectable from 450 ms until 800 ms.

4. Discussion

We tested processing of gender and number disagreement between the article, the adjective and the noun in Dutch. Results are in line with previous research in that a robust posterior positivity (P600) was elicited by ungrammatical sentences (Barber et al., 2004; Gunter et al., 2000; Molinaro et al., 2008). The P600 is interpreted as a stage in which repair and reanalysis take place (Friederici, 2002; Hagoort & Brown, 2000). It is often preceded by the LAN, which arises in response to morphosyntactic violations (Friederici, 2002; Molinaro et al., 2011, 2014). However, even though the current experiment contained a morphosyntactic violation, it failed to elicit the LAN.

4.1. Lack of biphasic response

A number of authors talk about a biphasic response to agreement violations, that is, the LAN followed by the P600 (e.g., Barber & Carreiras, 2005; Molinaro et al., 2011). In the first stage (LAN), the parser automatically identifies the morphosyntactic violation, after which it tries to integrate and repair it (P600) during the late syntactic stage (Friederici, 2002). The presence of the P600 has been reported in almost all agreement studies, which is not the case with the LAN. As an illustration, Barber and Carreiras (2005) reported the LAN followed by the P600 in number and gender disagreement in Spanish, both for article-noun and noun-adjective pairs. However, Wicha et al. (2004) tested gender disagreement between article-noun in Spanish, which produced only the P600. Similarly, Aleman Bañón et al. (2012) tested adjective-noun and noun-adjective disagreement in Spanish, neither of which elicited the LAN.

There are at least three accounts that attempted to explain the inconsistency in obtaining the LAN. Hagoort and Brown (1999) suggested that the LAN is sensitive only to phonologically overt morphosyntactic violations. In terms of the present study, this means that the LAN should have been recorded in the number condition only. Since gender is a lexical feature in Dutch, there was no inconsistency between the article/adjective and the noun's inflectional morphology for gender, as the noun is not morphologically marked for gender. However, number violation is a typical example of an article disagreeing with the noun because of the noun's number morphology (e.g., *het boek* 'the book', **het boek-en* 'the books'). Therefore, our findings are incompatible with the proposed account.

Molinaro et al. (2014) proposed that methodological factors may influence whether or not the LAN can be elicited or not. For example, the reference choice may play a role, since studies using the left mastoid as the reference reported the LAN less frequently

² As pointed out by one of the reviewers, comparing ERP effects across conditions, in the manner in which it was done in this study, is possible only if the baseline (waveform elicited by grammatical sentences) does not differ between the conditions. In order to ascertain that there was no difference between grammatical gender and grammatical number sentences, we performed an additional analysis for each time window with grammatical sentences only. There was no statistical difference between grammatical sentences showing that a comparison between the number condition and gender condition is justifiable.

than studies using the average of the mastoids. This account is not supported by the current experiment either, since we used the average of the mastoids. As for other methodological factors, we used a fairly common number of stimuli per condition (80 per condition), as well as the often-used stimulus asynchrony (600 ms). Thus, failure to elicit the LAN cannot be accounted for by either the Hagoort and Brown (1999) or Molinaro et al.'s (2014) account. A third account, proposed by Osterhout (1997), focuses on individual differences and ERP response. According to this hypothesis, some people react to violation either in the form of a positive (P600) or negative deflection (N400). Once the individual data are averaged, they result in a P600, which is sometimes preceded by the LAN. The LAN is seen as a residual distribution of the negative deflection that some participants exhibit and that was cancelled out in other regions by the P600 (see also Tanner, 2015; Tanner & Van Hell, 2014). Even though our data may be compatible with the account, any claim in its favour or against it would require an in-depth analysis of individual data, which is outside the scope of this study. Regardless of the absence of the LAN, current results are in line with a large body of studies that failed to elicit the biphasic pattern, and only yielded the P600 (e.g., Aleman Bañón et al., 2012; Nevins et al., 2007; Wicha et al., 2004).

4.2. P600 in gender and number disagreement

Due to all the issues associated with the LAN, the main hypotheses of the current study relied on the P600. We predicted that both kinds of violation would elicit the P600, with possible differences in distribution and/or amplitude. The difference was, indeed, recorded in the last of the three time windows in which the P600 was expected.

We hypothesized that the parser is sensitive to three main differences between gender and number disagreement. The first difference is related to the way each feature is encoded. In Dutch, gender is mainly a lexical feature, whereas number is an inflectional feature. We proposed that the onset of the P600 may be modulated by the way the targeted feature is encoded. Barber and Carreiras (2005) suggested that accessing gender is costlier than accessing number due to the lexical (gender) – morphological (number) opposition. That is, retrieving a lexical feature is a more arduous process than decoding a feature from a suffix. However, our data did not show any difference in the onset of the P600. The effect was first detected in the 450–600 ms time window in both conditions simultaneously.

The P600 effect in the ensuing time window (600–800 ms) was also identical in both conditions. This time window corresponds roughly to the early stage of the P600, which is stipulated to represent integration processes (Hagoort & Brown, 1999). Since no difference was detected, we assume that both gender and number are integrated in the same manner. The parser seems to lean towards a binary behaviour (Lukatela et al., 1987), meaning that it is only sensitive to the presence/absence of an agreement violation. There is no evidence in this stage that the parser takes into account the fact that gender is a formal lexical feature, whereas number is based on semantics.

4.3. P600 as a marker of repair

Our final prediction regarded repair processes which are believed to take place in the late P600 stage (Hagoort & Brown, 1999). Due to the inflectional nature of number, number disagreement should be a more complex process to repair, as it offers two repair options compared to a single repair option in gender. As an illustration, article/adjective-noun gender disagreement can trigger only one repair process in which the adjective has to be repaired (**een grote_C boek* > *een groot_N boek_N* ‘a large book’). This is equivalent to the first repair option in number in which the preceding singular article is repaired into the plural article (*het_{SG} boeken_{PL}* > *de_{PL} boeken_{PL}* ‘the books’). However, the parser has an additional repair option at its disposal, and that is repairing (deleting) the plural suffix on the noun (*het_{SG} boeken_{PL}* > *het_{SG} boek_{PL}* ‘the book’).³ The double repair option in number may be more demanding than a single repair in gender. The increase in processing demand was expected to be reflected as a difference in the late stage of the P600, either as higher amplitude or a broader distribution in the number condition. The former scenario turned out to be true in the current study. These results are not in line with what Barber and Carreiras (2005) reported for number and gender disagreement in Spanish. They found that the P600 in its late stage was larger for gender, which they attributed to the lexical nature of gender which makes it more difficult to process. However, such results were not replicated by Aleman Bañón et al. (2012) in a similar study on Spanish in which no difference between gender and number disagreement was reported. The only psycholinguistic explanation we can offer is that gender is most often realized as a transparent morpheme in Spanish (-o for masculine and -a for feminine) on top of which a plural number suffix can be added. It is, thus, plausible that repair processes are affected in a different way in Spanish and Dutch due to the morphological nature of gender. Alternatively, there was an important methodological difference between our study and Barber and Carreiras (2005). Namely, the P600 time windows used in the current study (600–800 ms and 800–1000 ms) are somewhat later than the ones in the study by Barber and Carreiras (500–700 ms and 700–900 ms), which may also have influenced the difference in the results.

³ Note that in our stimuli there was an intervening adjective between the article and the noun in the number condition. The role of the adjective was to make stimuli between the two conditions as similar as possible (gender: indefinite article, adjective, target noun; number: definite article, adjective, target noun). All adjectives preceded by a definite article in Dutch are marked with the suffix -e. Therefore, the repair process in number disagreement only needs to target the article, since the adjective is morphologically already compatible with both singular and plural nouns.

5. Conclusion

The current study delineates which ERP effect are caused by the inherent difference between number and gender, as opposed to the differences arising as a consequence of using the violation paradigm/disagreement. Our results are in line with the majority of agreement studies. The experimental manipulation (article/adjective-noun gender or number disagreement) elicited the P600. The LAN was not observed in the current study, which is in line both with a large number of studies (e.g., Hagoort & Brown, 1999; Wicha et al., 2004), but simultaneously at odds with an equally substantial body of research (e.g., Barber & Carreiras, 2005; Molinaro et al., 2011). The lack of the LAN seems to be part of a larger controversy revolving around this component that is well worth further research. The most notable result of this study was obtained in the very late processing stage of the P600. More precisely, number disagreement elicited a P600 which lasted well through the late P600 stage (longer than 800 ms post-stimulus onset), whereas the P600 in gender disagreement was short-lived (until 800 ms) and it never reached the late stage. We interpret the longevity of the P600 effect as a marker of increased processing load due to a more complex repair in number than in gender disagreement.

Funding

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.

Conflicts of interest

None.

References

- Adger, D., & Harbour, D. (2008). Why phi. In D. Harbour, D. Adger, & S. Bejar (Eds.), *Phi-theory: Phi features across interfaces and modules* (pp. 1–34). Oxford: Oxford University Press.
- Aleman Bañón, J., Fiorentino, R., & Gabriele, A. (2012). The processing of number and gender agreement in Spanish: An event-related potential investigation of the effects of structural distance. *Brain Research*, 1456, 49–49.
- 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). Brighton, UK: Psychology Press.
- Caffarra, S., & Barber, H. (2015). Does the ending matter? The role of gender-to-ending consistency in sentence reading. *Brain Research*, 1605, 83–92.
- Caffarra, S., Barber, H., Molinaro, N., & Carreiras, M. (2017). When the end matters: Influence of gender cues during agreement computation in bilinguals. *Language, Cognition and Neuroscience*, 1–17.
- 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.
- Corbett, G. G. (2003). Agreement: Canonical instances and the extent of the phenomenon. *Morphology: Selected papers from the third mediterranean morphology meeting* (pp. 109–128). September 20–22, 2001.
- Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language & Cognitive Processes*, 13, 21–58.
- Dowens, M. G., Vergara, M., Barber, H., & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience*, 22, 1870–1887.
- Faussart, C., Jakubowicz, C., & Costes, M. (1999). Gender and number processing in spoken French and Spanish. *Rivista di Linguistica*, 11, 75–101.
- 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, 259–281.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 78–84.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95–112.
- Gunter, T., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal autonomy and late interaction. *Journal of Cognitive Neuroscience*, 12, 556–568.
- Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, 15, 883–899.
- 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. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, 38, 1531–1549.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language & Cognitive Processes*, 8, 439–483.
- 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. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38.
- Loerts, H., Stowe, L. A., & Schmid, M. S. (2013). Predictability speeds up the re-analysis process: An ERP investigation of gender agreement and cloze probability. *Journal of Neurolinguistics*, 26, 561–580.
- Lukatela, G., Kostić, A., Todorović, D., Carello, C., & Turvey, M. T. (1987). Type and number of violations and the grammatical congruency effect in lexical decision. *Psychological Research*, 49, 37–43.
- Martin-Loeches, M., Nigbur, R., Casado, P., Hohlfeld, A., & Sommer, W. (2006). Semantics prevalence over syntax during sentence processing: A brain potential study of noun–adjective agreement in Spanish. *Brain Research*, 1093, 178–189.
- Meulman, N., Stowe, L. A., Sprenger, S. A., Bresser, M., & Schmid, M. S. (2014). An ERP study on L2 syntax processing: When do learners fail? *Frontiers in Psychology*, 5, 1072.
- Molinaro, N., Barber, H., 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., Vespignani, F., & Job, R. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research*, 1228, 161–176.
- Münte, T. F., Szentkúti, A., Wieringa, B. M., Matzke, M., & Johannes, S. (1997). Human brain potentials to reading syntactic errors in sentences of different complexity. *Neuroscience Letters*, 235, 105–108.
- Nevins, A., Dillon, B., Malhotra, S., & Phillips, C. (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research*, 1164, 81–94.
- O'Rourke, P. L., & Van Petten, C. (2011). Morphological agreement at a distance: Dissociation between early and late components of the event-related brain potential. *Brain Research*, 1392, 62–79.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language*, 59, 494–522.
- Osterhout, L., & Holcomb, P. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785–806.
- 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.). *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond* (pp. 271–308). Psychology Press.
- Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, 22, 2728–2744.
- 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). John Benjamins Publishing.
- Rayner, K., & Clifton, C. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology*, 80, 4–9.
- Ritter, E. (1991). Two functional categories in noun phrases. *Syntax and Semantics*, 25, 37–62.
- Ritter, E. (1993). Where is gender? *Linguistic Inquiry*, 24, 795–803.
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249–284.
- Shetter, W. Z. (1959). The Dutch diminutive. *The Journal of English and Germanic Philology*, 58, 75–90.
- Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: A commentary on “grammatical agreement processing in reading: ERP findings and future directions” by Molinaro et al., 2014. *Cortex*, 66, 149–155.
- Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, 56, 289–301.
- Teschner, R. V., & Russell, W. M. (1984). The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics*, 1, 115–132.
- Van Berkum, J. (1996). *The psycholinguistics of grammatical gender: Studies in language comprehension and production*. Doctoral dissertation. Nijmegen, Netherlands: Nijmegen University Press Max Planck institute for Psycholinguistics.
- 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.