

available at www.sciencedirect.comwww.elsevier.com/locate/brainres**BRAIN
RESEARCH****Research Report****A deeper reanalysis of a superficial feature: An ERP study on agreement violations**Nicola Molinaro^{a,b,*}, Francesco Vespignani^a, Remo Job^a^aDiSCoF, University of Trento, Italy^bDepartment of Psychology, University of La Laguna, Campus de Guajara, 38205, San Cristobal de La Laguna, Tenerife, Spain

ARTICLE INFO

Article history:

Accepted 18 June 2008

Available online 26 June 2008

Keywords:

Language

Sentence processing

Syntactic agreement

Phonotactic

Lexicon

ERP

LAN

P600

ABSTRACT

A morphosyntactic agreement violation during reading elicits a well-documented biphasic ERP pattern (LAN+P600). The cognitive variables that affect both the amplitude of the two components and the topography of the anterior negativity are still debated. We studied the ERP correlates of the violation of a specific agreement feature based on the phonology of the critical word. This was compared with the violation of a lexical feature, namely grammatical gender. These two features are different both in the level of representation involved in the agreement computation and in terms of their role in establishing structural relations with possible following constituents. The ERP pattern elicited by the two agreement violations showed interesting dissociations. The LAN was distributed ventrally for both types of violation, but showed a central extension for the gender violation. The P600 showed an amplitude modulation: this component was larger for phonotactic violations in its late time window (700–900 ms). The former result is indicative of a difference in the brain structures recruited for the processing of violations at different levels of representation. The P600 effect is interpreted assuming a hierarchical relation among features that forces a deeper reanalysis of the violation involving a word form property. Finally the two features elicit distinct end-of-sentence wrap-up effects, consistent with the different roles they play in the processing of the whole sentence.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Every language restricts the number of possible sequences of sounds that can be produced, and these restrictions are expressed by the phonotactic constraints, which play a role mainly in the processing of single words (Goldsmith, 1990). However, in many languages, syntactic agreement interacts with some rules represented at the phonological level. This is the case of allomorphs, i.e. sentential particles sharing the same morphosyntactic features (as gender or number), whose superficial form is determined by the phonological context in

which they are expressed. Phonotactic constraints are mainly evident in the selection of determiners within noun phrases (NPs). An example is the indefinite article in English that has two allomorphic forms selected on the basis of the phoneme with which the following word begins: 'a' is used before words starting with a consonant while 'an' is used before words starting with a vowel.

A similar phenomenon in Italian concerns the masculine definite article allomorphy: the masculine article has two forms, 'il' vs. 'lo'. The choice of a specific form depends on the phonological characteristics of the word that follows it in the

* Corresponding author.

E-mail address: nicolaml@ull.es (N. Molinaro).

NP. The determiner 'lo' is selected when the following word's onset is a 'y', a consonant cluster of the form 's+consonant' or 'gn' or an affricate, like 'z'. In all other cases, the determiner 'il' is selected.¹

Phonotactic rules are crucial in language processing both to correctly produce a phonologically appropriate syllabic sequence and to segment spoken language (McQueen, 1998; Tabossi et al., 1995; Miozzo and Caramazza, 1999). As far as language comprehension is concerned, phonotactic constraints are context-sensitive, i.e. they depend crucially on the form of two adjacent words; they do not imply lexical, morphosyntactic information to be exploited, as grammatical gender, for example, does; moreover, their effects should be local, unlike gender and number agreement, which may span the entire sentence.

A qualitative different feature is grammatical gender: it is an arbitrary characteristic of many words (determiners, nouns, adjectives, pronouns) that is implemented also for inanimate referents that do not have a specific biological gender. Gender arbitrariness implies that different words with closely related referents (like synonyms) could have different grammatical gender. Moreover, the same concept can be expressed through masculine or feminine words across different languages. Gender features can be associated to morphological properties of a word: in Italian, for example, the '-a' suffix is mostly associated with feminine gender and '-o' with masculine. However, there are many exceptions: (i) irregular gender words, i.e. feminine gender words ending in '-o' and masculine gender words ending in '-a', and (ii) opaque gender words, i.e. words ending in '-e' that could be either masculine or feminine. Gender is an invariable property of nouns and, unlike phonology, it needs a complete lexical processing to be recognized (Corbett, 1991); moreover, this feature can be redundantly expressed across many words in the same sentence, like determiners, adjectives and pronouns.

In this study we investigated the processing of sentences containing phonotactic and gender violations while recording electrophysiological brain activity. Through the use of Event Related Potentials (ERPs) recorded on the scalp, time-locked to the processing of a particular event, it is possible to monitor the time course of the on-line processing of a particular event. ERPs, compared to other neurophysiologic techniques, have a high temporal resolution, but they can also vary on other dimensions, as the topography over the scalp or the amplitude of a component.

A large number of studies have focused on the processing of gender (and number) features, while the influence of phonotactic constraints has not been adequately studied. Phonological features, in fact, manifest their influence in many languages at the syntactic level, and their investigation could shed light on the cognitive processes involved in the computation of agreement.

In the last 15 years a considerable number of studies focused on the electrophysiological correlates of morphosyntactic agreement violation processing. The pattern of results showed that number and gender agreement violations elicit similar electrophysiological patterns, i.e. an early left-frontal

negativity around 300 ms, frequently classified as a LAN, followed by a P600.

Concerning the phonotactic agreement, the fact that the mismatching feature is not strictly syntactic in nature does not imply that the violation has no impact on syntactic processing: any difficulty in building an internal coherent representation of a phrase (the NP in our case) may recruit similar cognitive resources, this view is consistent with the minimalist frames (Chomsky, 1995) that does not assume separate and hierarchical processing of the different features that control agreement, but rather proposes the parallel use of a bundle of features in constructing structural relations between words. A neurobiological model of sentence comprehension, inspired by minimalist (Chomsky, 1995) and lexicalist (Jackendoff, 2002) theoretical frames, is the *unification model* (Hagoort, 2005; see also Vosse and Kempen, 2000). In this model, lexical items are retrieved sequentially with their associated structural frames and enter into a unification workspace incrementally, where high-level structured representations, such as phrases and utterances, are built. This operation consists of connecting lexical frames using not only syntactic features but also semantic and phonological information concurrently and interactively. Hence, not only morphosyntactic features, but also phonotactic rules constrain unification. As far as lexical frames cannot be structured in higher-level representations the syntactic structure of the sentence cannot be achieved and it is possible that a similar syntactic conflict emerges, independently from the nature of the feature that triggers the conflict.

LAN. Gender agreement violations elicited a LAN in some studies, i.e. an ERP effect also reported for number agreement violations (Spanish: Barber et al., 2004; Barber and Carreiras, 2005; Demestre et al., 1999; German: Gunter et al., 2000). Other authors (Deutsch and Bentin, 2001²) however reported a N400, an increased negativity around 400 ms, typically related to semantic/pragmatic processing difficulties (Kutas and Federmeier, 2000). Finally, Hagoort and Brown (1999) reported a not-lateralized anterior negativity in Dutch, when the target noun was in final position of the sentence, but no negative effects when the critical NP was in initial position (see also Hagoort, 2003). A possible explanation for these heterogeneous data could be that gender values could refer either to the biological gender of an animate referent, or to the grammatical gender of an inanimate referent, since in many languages nouns are always marked for gender. While the former value is considered a semantic feature, the latter is considered purely syntactic. Lamers et al. (2006) and Schmitt et al. (2002), for example, reported N400 effects for pronouns that did not match the biological gender of the antecedent. The only two studies that directly compared grammatical and semantic gender violations found contrasting data: in Spanish, Barber et al. (2004) reported LAN for both violations while in Hebrew Deutsch and Bentin (2001) found N400 in both cases.

While we do not focus on number agreement in this paper, it may be interesting to recall that in the only one study

¹ For a recent account of the *il/lo* selection in Italian see Russi (2006).

² Deutsch and Bentin (2001) reported also an eLAN with an onset at 50ms. However, since this effect has not been replicated we will not discuss it in the present paper.

(Barber and Carreiras, 2005) that directly compared the ERPs elicited by number and gender agreement violations the authors found the same LAN onset, around 300 ms for both violations. Even if it was not reported, the topography of the effects seem different (see Fig. 5 page 145): the gender violation elicited a left-lateralized negativity with a more central distribution compared to the number violation. This is in line with the previous literature: while the LAN elicited by number violations is consistently reported in the left-frontal area of the scalp, it is hard to dissociate between N400 and LAN effects for gender violations.

The different LAN topography in Barber and Carreiras (2005) experiment could indicate that number and gender features are processed in the same time window by different neural networks. The timing of this effects is consistent with studies on repetition priming and ERPs, which report a modulation of a positive component peaking around 300 ms in response to lexical information priming (Holcomb and Grainger, 2006). Assuming that word features cascade from the word processing level to the syntactic parser as soon as they have been identified, number and gender, both lexical features, could then manifest their influence on agreement computation after 300 ms, i.e. in the LAN time window. The study of Holcomb and Grainger (2006) (see also Grainger et al., 2006) also showed phonological information being processed before lexical information as indexed by an effect on a negative component around 250 ms, labelled N250. This latency difference between effects (N250 vs. P325) is consistent with models of visual word perception in which the phonological stage of processing precedes the lexical stage of processing (Coltheart et al., 2001; Perry et al., 2007). Assuming a cascading approach, we would then expect earlier effects for a phonotactic compared to a gender violation. Previous studies showed that ERPs could dissociate between qualitatively distinct syntactic violations that elicited LANs with a different onset. In her model, Friederici (2002) discusses the latency variability of the LAN as a function of the timing at which different syntactic word properties become available to the syntactic parser. One assumption behind her hypothesis is that the lexical access system communicates in a cascade fashion with the sentence processor: syntactic processing begins from word category information and, then, proceeds with morphosyntactic cues and semantic features. This proposal is based on the empirical observation that phrase structure violations (wrong grammatical word class) elicit an eLAN (early LAN) around 150 ms and morphosyntactic violations, such as agreement mismatches, a LAN around 300 ms.

P600. The P600 has been classically assumed to reflect syntactic processing in sentence context (Osterhout et al., 2004). Following the discovery of a P600 after specific semantic/thematic violations (for reviews, Kolk and Chwilla, 2007; Kuperberg, 2007) a debate arose on the question whether these semantic effects on the P600 could still be interpreted syntactically (as suggested by Kim and Osterhout, 2005) or as the correlate of a thematic reanalysis (Kuperberg et al., 2003). According to Kim and Osterhout (2005) only syntactic incongruences trigger P600: this effect was recorded at the verb in sentences like ‘*The meal was devouring...*’ in which there is a strong bias (driven by the semantic relation between *meal* and

devour) for the passive form of the sentence. In other words, there is a strong expectation for the form ‘*devoured*’, and the -ing form is perceived as syntactically ill-formed. A more general hypothesis has been advanced by Kolk and Chwilla (2007): their *conflict monitoring model* (see also Vissers et al., 2006) proposes that this ‘semantic P600’ is a reprocessing stage triggered by a conflict between incompatible linguistic representations. Reprocessing would consist in a monitoring process of the cognitive representation of the sentence in order to detect the source of the incongruence and settle the sentence meaning. Crucially, the P600 is not triggered by syntactic incongruences only, but by any kind of conflict within a sentence (orthographic: Vissers et al., 2006; lexical: Federmeier et al., 2007; semantic: Kim and Osterhout, 2005; Kuperberg et al., 2003; Kolk et al., 2003).

It should be noted that the only P600 with the classical posterior distribution with onset at 500 ms is the ‘semantic P600’ that could still be interpreted as reflecting a syntactic difficulty (Kim and Osterhout, 2005), while the ‘orthographic P600’ reported by Vissers et al. (2006) has an earlier onset, around 400 ms, more consistent with a P300 interpretation (specifically, the P3b). Finally, the ‘lexical P600’ reported by Federmeier et al. (2007) has a right-anterior distribution. These components are different compared to the posterior P600 with onset around 500 ms usually triggered by syntactic violations. The picture that emerges from this literature is that the P600 is not a single component, but rather a family of late positive components, that could modulate in latency, amplitude and topography depending on the type of conflict within the sentence. Since we consider that both the phonotactic and the gender violations have an impact at the syntactic level, they should both elicit a posterior P600 with onset at about 500 ms.

Barber and Carreiras (2005) proposed an interesting interpretation of the P600 within the agreement violation literature. Based on the finding of a larger P600 effect in its late time window (between 700 and 900 ms) for gender compared to number agreement violations in Spanish, the authors recall a model proposed by Faussart et al. (1999). Agreement computation would be pursued in three distinct stages (Fig. 1): (i) an early stage of lexical access where the correct lexical entry is selected; (ii) a second stage of lexical recognition where the relevant features and the semantic properties are processed; (iii) a final stage where the item is evaluated with respect to the preceding context and consequently integrated: if the integration process fails, a reanalysis is triggered.³

Faussart and colleagues refer to the fact that while gender is an inherent feature of the target noun, number is not. This means that gender is generally a fixed property of the stem, since only one form (in Italian, feminine or masculine) is assigned to each noun, and it is stored with the corresponding lexical entry (Corbett, 1991). Number, on the opposite, is a non-inherent lexical entry: it signals the quantity of the referent, and being variable, independently combines with the selected lexical stem (Ritter, 1988).

As a consequence, gender is processed when the lexical entry is selected, i.e. in stage (i), while number is processed in

³ Faussart et al. (1999) model was inspired by a hypothesis on lexical retrieval proposed by Bradley and Forster (1987).

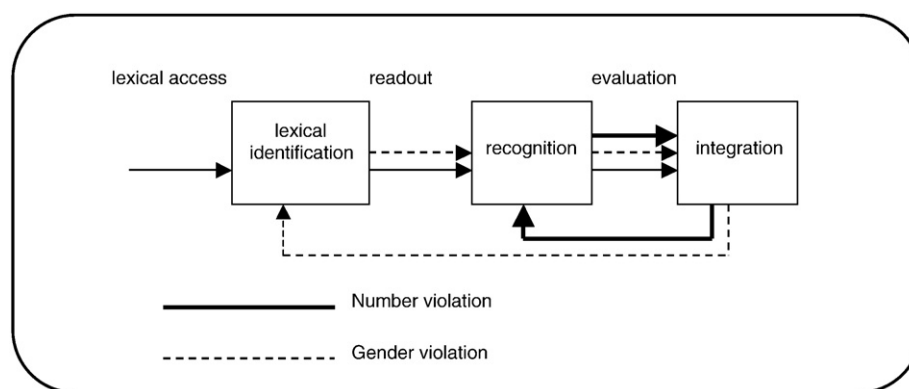


Fig. 1 – Model proposed by Faussart et al. (1999).

stage (ii). In terms of reanalysis a number agreement violation causes a rechecking of the outcome of stage (ii), i.e. complete lexical recognition; a gender agreement violation compels the language processor to reprocess the lexical item to verify its correspondence with the lexical entry, i.e. the system has to come back to stage (i), i.e. selection of the lexical entry.

Barber and Carreiras (2005) based their analysis on this model proposing that the distinction between regression to stage (ii) for reanalysis of number violations and regression to stage (i) for gender violations is reflected in the late amplitude of the P600: the more backward steps are needed to reanalyze a disagreement the bigger is the P600 amplitude. P600 thus would represent a diagnosis stage of the source of a linguistic incongruence (P600 onset: 500–700 ms) followed by the actual reanalysis of the incongruence to create a plausible interpretation of the message (P600 offset: 700–900 ms).

1.1. The present study

In the present paper we studied the computation of agreement between a noun and its determiner. We focus on the interaction between different stages of processing of the target noun and their syntactic relations at the sentence level. The critical noun phrases were inserted in sentences and we manipulated the determiner preceding the target noun as showed in Table 1. We compare the phonotactic violation, derived from the manipulation of the allomorphs of the Italian masculine determiner ('il' vs. 'lo'), with a gender agreement violation ('la'; examples in Table 1).

The selection of the nouns requiring 'lo' was performed on a group of 84 Italian masculine nouns with low written lexical frequency and with more than 4 letters; values were obtained using the ColFis written lexical frequency dictionary (Laudanna et al., 1995). This stimuli selection was aimed at postponing the lexical access of the target word as much as possible in order to possibly delay the onset of the LAN for the gender violation and make it better distinguishable from the onset of the phonotactic ERP effect. The comparison between the phonotactic and the gender violation could then dissociate possible effects of different stages of the target word processing during the computation of agreement. We choose the visual presentation of the stimuli, instead of the auditory one: in the auditory modality in fact the information is presented sequentially with a large variability between the onset of the word (at which the phonotactic mismatch could be detected) and the point at which the critical gender information could be extracted from the item.

Gender and phonotactic violations could modulate ERP effects in a rather different way. For what concerns the LAN we expect differences both in topography and in latency. Number and gender violations showed slightly different topographies in the study of Barber and Carreiras (2005), suggesting distinct neural systems for the processing of the two features. Since gender and phonological features are qualitatively different properties of the target noun, we expect different LAN topographies elicited by the two violations. Given the lexical status of gender features, gender violations could elicit left-anterior negativities with a more central

Table 1 – Examples of sentences used in the experiment

Control	La vecchina con The old woman with Le olive farcite con The olives stuffed with	lo the (+LO, +M) il the (+IL, +M)	scialle shawl (+M) peperone red pepper (+M)	cammina lentamente per la salita. walks slowly on the uphill. sono ottime. are very good.
Phonotactic violation	*La vecchina con *The old woman with *Le olive farcite con *The olives stuffed with	il the (+IL, +M) lo the (+LO, +M)	scialle shawl (+M) peperone red pepper (+M)	cammina lentamente per la salita. walks slowly on the uphill. sono ottime. are very good.
Gender violation	*La vecchina con *The old woman with *Le olive farcite con *The olives stuffed with	la the (+F) la the (+F)	scialle shawl (+M) peperone red pepper	cammina lentamente per la salita. walks slowly on the uphill. sono ottime. are very good.

distribution, similar to the N400 (as showed by Deutsch and Bentin, 2001). The N400 showed in fact to be sensitive to the lexical properties of the target word (Kutas and Federmeier, 2000). On the other side, since the phonotactic violation does not need a lexical analysis to be exploited, it shall trigger a left-anterior negativity not evident on the more central areas of the scalp. This would indicate that different brain networks are involved in the overprocessing caused by the two violations.

Assuming a cascading approach for the transmission of information between the word processor and the agreement computation, we also expect a modulation of the LAN onset based on the type of feature that is violated. Friederici (2002) showed that word category violations elicit an earlier LAN compared to morphosyntactic agreement violations. The target word features (phonology vs. gender) could as well cascade to the agreement computation process serially. We then expect the LAN elicited by the phonotactic violation with an earlier onset compared to gender violation. An alternative hypothesis is based on the Haagort's (2005) approach, according to which the lexical items are stored in memory with their relative syntactic frames. Their unification is pursued considering all the possible constraints at the same time; in other words, during agreement checking, all the word features that constrain agreement are evaluated in parallel. In this frame, phonotactic and gender mismatches should elicit the earlier ERP component (the LAN) with similar latency.

After the first attempt to construct a structured representation of the noun phrase, reflected in the LAN, the system should perform reanalysis for detecting the source of the error, indexed by a posterior P600. Assuming the two-stage reprocessing model of the P600 (Barber and Carreiras, 2005; Carreiras et al., 2004) phonotactic violations should elicit larger amplitude for phonotactic violations compared to gender violations in the P600 late time window. Since phonological processing of a word precedes the selection of the right lexical entry, the P600 triggered by phonotactic violations should be larger compared to the positive shift elicited by gender or number violations. In fact, the reanalysis of phonotactic mismatches requires a regression to a pre-lexical processing stage.

An alternative hypothesis is that a phonotactic violation does not need a demanding reanalysis, since this constraint has no consequences outside the determiner-noun relation, i.e. in the subsequent analysis of the sentence. Gender in Italian is more crucial on this dimension since a following pronoun or adjective may be marked for gender, helping the binding with the correct referent (De Vincenzi, 1999). Since the phonotactic violation does not require a revision of the deep sentence structure, one would expect a P600 with lower amplitude for this type of violation compared to a gender mismatch.⁴

Finally, since gender agreement violations have shown to elicit also consistent negative shifts after the last word of a sentence (Haagort and Brown, 1999), we monitor also this

end-of-sentence ERP correlate. This effect seems to be related to the consequences of the gender violation for the overall integration of the sentential information into one coherent message. More generally, it would reflect a processing mechanism that checks the well formedness of the sentence to ensure that all arguments in the construction have been represented correctly (wrap-up process: Just and Carpenter, 1980). As previously discussed, gender and phonotactic constraints play different roles at the syntactic level: while the phonotactic violation exerts its influence only within the minimal context of the NP, gender values can be expressed also on constituents outside the phrase. This is true not only for verbs and adjectives, but also for pronouns that can be expressed outside of the sentence, like tag pronouns for example. It becomes evident how the correct evaluation of the referents values within a sentence is crucial at the level of the discourse. If this is true, we expect an increased negativity for the sentences containing a gender violation after the last word of the sentence that should be practically absent for the sentences containing the phonotactic violation.

2. Results

2.1. Acceptability judgments

Participants responded at the acceptability judgments with an overall accuracy of 94%, with individual participants ranging from 87% to 98%, showing a very good accuracy in the detection of the ungrammaticalities. Phonotactic violations and gender violations were judged correctly respectively in the 94.5% and in the 95.2% of the cases, showing no difference in the accuracy of the acceptability judgments ($t(19) < 1$). One participant was not included in the following analysis given an excessive number of artifacts.

2.2. ERPs time-locked to the target noun

Grand average at the critical word for the Gender Violation compared with Control Condition is plotted in Fig. 2, while the comparison between the Phonotactic Violation and the Control is plotted in Fig. 3. For both Phonotactic and Gender violation, it was evident a frontal negativity starting at around 300 ms.

In Fig. 4 we reported the difference between waveforms (agreement minus disagreement) comparing Phonotactic and Gender disagreement at the midline electrodes. As it was evident from the differences between waveforms at the midline (Fig. 4) the topographical distribution of the negativity elicited by Gender violation showed its effect also on the frontal and central electrodes of the scalp (see maps in Fig. 5). Both violations elicited a consistent positive shift starting around 500 ms on the posterior area of the scalp, identifiable as a P600. Interestingly, the amplitude of the P600 elicited by the Phonotactic violation was larger in a late time window, after 700 ms, compared to the positivity elicited by the Gender violation (see P600 amplitude in Fig. 4).

We quantified ERPs as the mean amplitude values between 350 and 450 ms for the LAN evaluation and in two

⁴ For what concerns the P600 we do not expect any latency modulation, since this effect has been consistently reported after 500 ms for syntactic violations.

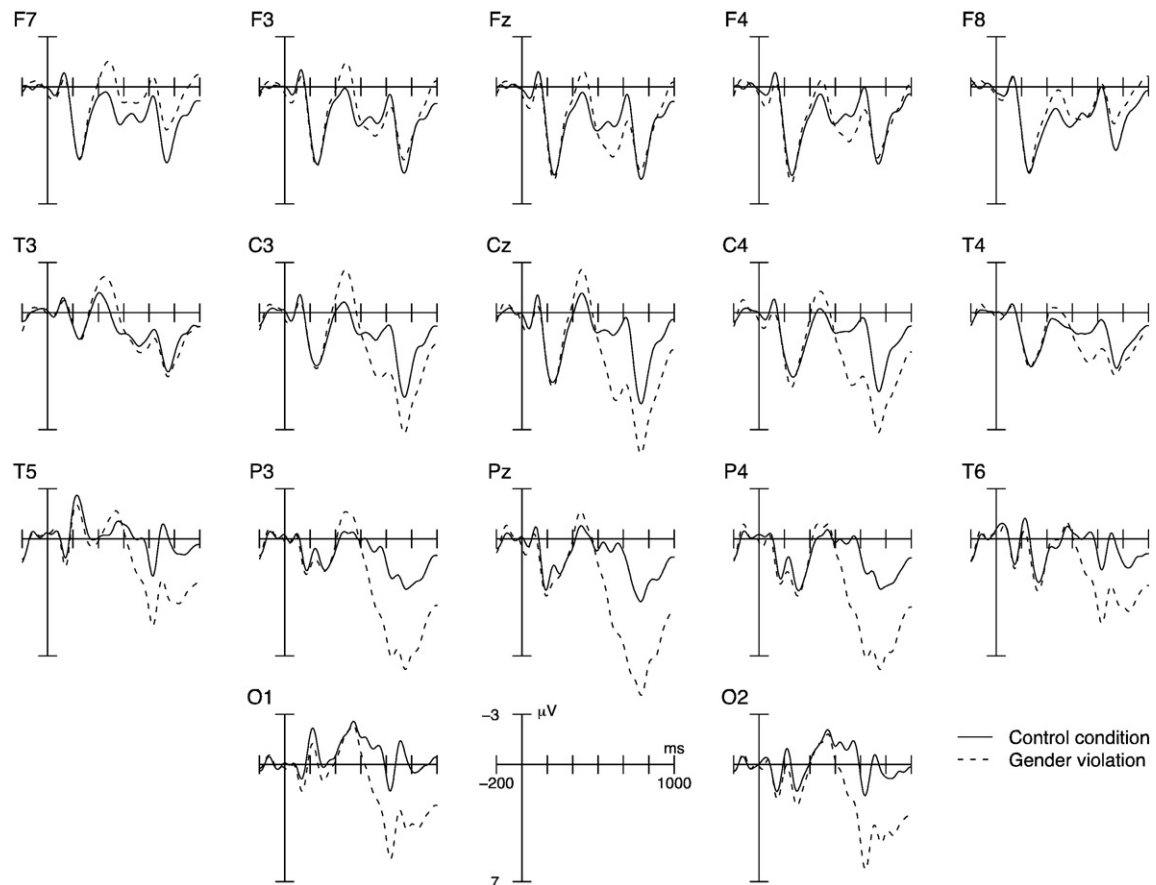


Fig. 2 – Grand-averages for the Control Condition (solid line) compared to the Gender Violation (dashed line). Vertical lines represent the onset of the critical noun presentation.

time-windows for the P600 analysis: between 500 and 700 ms and 700 and 900 ms after the critical noun processing. We performed two separate ANOVAs: one on the midline electrodes and one for the lateralized electrodes grouped in six clusters, three on the frontal, central and parietal areas of the left side of the scalp and three on the corresponding right areas (Fig. 9). The former analysis considered two factors: sentence type (the three conditions) and longitude (frontal, central and parietal sites); the latter analysis was conducted on the six groups considering also the hemisphere factor (left, right). For further details on the analysis see section 4.4.

2.2.1. Mean amplitude between 350 and 450 ms (LAN)

In the time window associated with LAN the six-clusters group of electrodes showed a main effect of sentence type (six-clusters group: $F(2,38)=7.066$, $p<0.01$) and an interaction of sentence type with hemisphere (six-clusters group: $F(2,38)=4.117$, $p<0.05$). While in the midline electrodes there was no significant effect, a lateralized negativity was elicited by both violations. In Table 2 we report the one-way ANOVAs for each cluster of electrodes comparing each experimental condition with the Control condition. These analyses showed that the topographical effect elicited by the gender violation was significant also in the frontal and central clusters on the right hemisphere, compared to the effect elicited by the phonotactic violation.

Beside the fact that a direct comparison of the two ungrammaticalities did only lead to marginal effects (Table 2, last column), the different topographies of the LAN were evident on the maps reported in Fig. 5. The difference was qualitatively maximal on Cz and typical analysis schemes (for example: Kim and Osterhout, 2005; Roehm et al., 2007) consider separate statistics on the midline and lateralized electrodes, as in our case; this is clearly a good solution to assess laterality of effects but it may not be optimal in distinguishing effects that differentially involve lateral and central scalp locations. A post-hoc analysis was thus performed on the middle coronal electrodes to distinguish the topographies of the LANs elicited by the two experimental violations. We run an ad-hoc ANOVA (Greenhouse–Geisser corrected) on the factors electrode (five levels: T3, C3, Cz, C4, T4) and sentence type (two levels: Gender and Phonotactic violation). In order to compensate for the absolute difference of the to-be-compared effects we applied the McCarthy and Wood (1985) procedure, scaling the voltages in each condition by the square root of the sum of the square voltages over all electrode locations. This analysis showed a significant interaction between electrode and sentence type: $F(4,76)=5.179$, $p<0.05$. This result confirms that the negativities elicited by the two violations had a different distribution on the scalp, since the Gender Violation caused a negativity over central and right coronal electrodes.

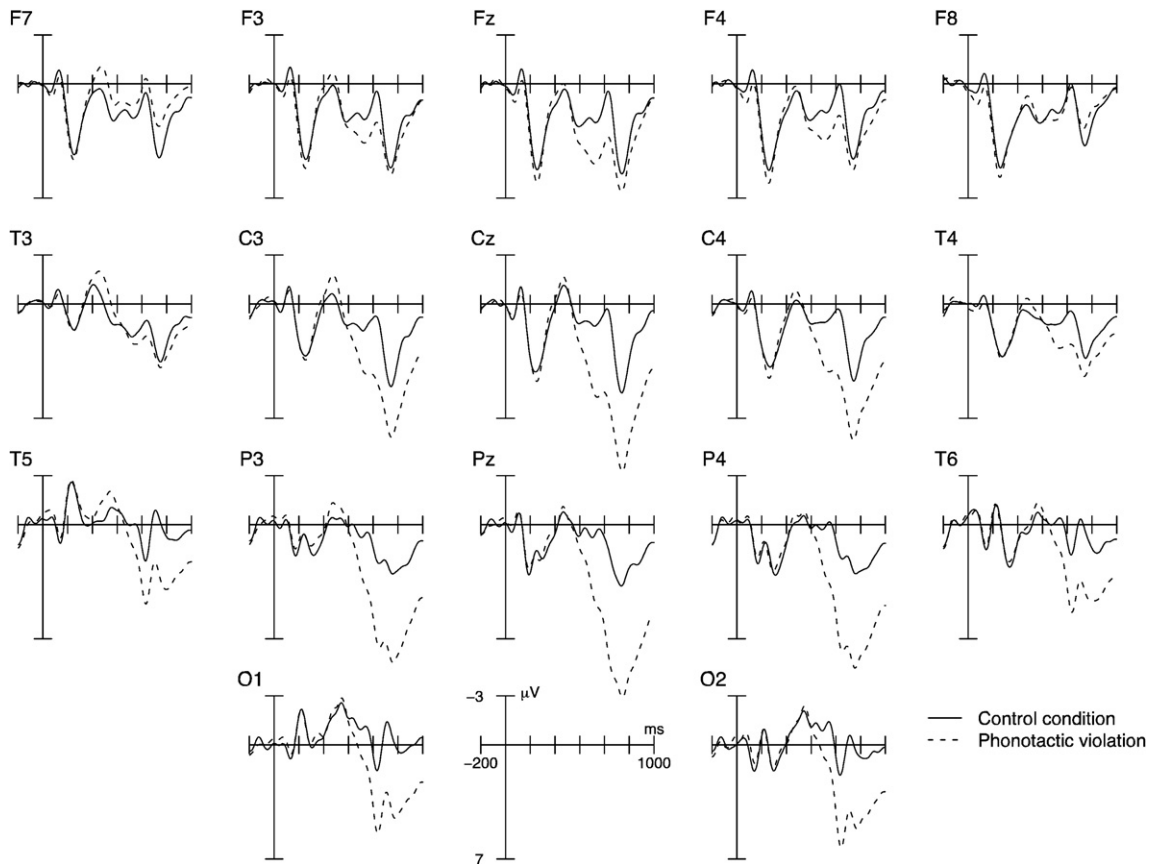


Fig. 3 – Grand-averages for the Control Condition (solid line) compared to the Phonotactic Violation (dashed line). Vertical lines represent the onset of the critical noun presentation.

2.2.2. Mean amplitudes between 500 and 700 ms and between 700 and 900 ms (early and late stage of the P600)

In the 500–700 ms time window the P600 elicited by both violations resulted in a main effect of sentence type (midline group: $F(2,38)=12.946$, $p<0.001$; six-clusters group: $F(2,38)=6.042$, $p<0.01$) and an interaction between sentence type and longitude (midline group: $F(4,76)=10.765$, $p<0.001$; six-clusters group: $F(4,76)=16.665$, $p<0.001$).

In the following time window (700–900 ms) similar results were evident from the analyses: a main effect of sentence type (midline group: $F(2,38)=20.057$, $p<0.001$; six-clusters group: $F(2,38)=11.411$, $p<0.001$) and its interaction with longitude (midline group: $F(4,76)=36.509$, $p<0.001$; six-clusters group: $F(4,76)=44.686$, $p<0.001$).

Post-hoc comparisons were performed on the midline electrodes in order to clarify differences in the amplitude of the P600 evident in the grand-average. Results reported in Table 3 showed a significant difference between the two violations in the late time window (700–900 ms) confirming the larger amplitude of the P600 for the Phonotactic violation, compared to the Gender violation, evident in the grand-average.

2.3. ERPs time-locked to the end of the sentence

Grand-averages recorded at the end of the sentence for each condition at the midline electrodes are showed in Fig. 6. The

waveforms associated to the three conditions diverged 200 ms after the onset of the last word of the sentence: the Gender violation showed a negative trend compared to both the Control condition and the Phonotactic Violation; also the Phonotactic violation showed a negative trend that was less negative compared to the Gender Violation. This negative shift was evident all over the scalp but it was more pronounced over the centro-parietal areas for the gender mismatch (see maps in Fig. 7).

2.3.1. End of sentence effects as the mean amplitude between 200 and 400 ms

In order to statistically evaluate the negativities evident in the grand-average we quantified the effects as the average amplitude of the waveforms associated with each condition in a time window between 200 and 400 ms. In both the midline and the six-clusters analysis a main effect of sentence type emerged (midline group: $F(2,38)=11.174$, $p<0.0001$; six-clusters group: $F(2,38)=9.052$, $p<0.001$). The analyses performed on the midline group showed a marginal interaction between sentence type and longitude (midline group: $F(4,76)=2.793$, $p<0.1$), thus suggesting that the negative effect was mainly evident over the centro-parietal electrodes. In order to statistically dissociate between the effects of the three conditions we run one-way ANOVAs on each electrode of the midline comparing the three critical conditions two by two. Results are reported in Table 4.

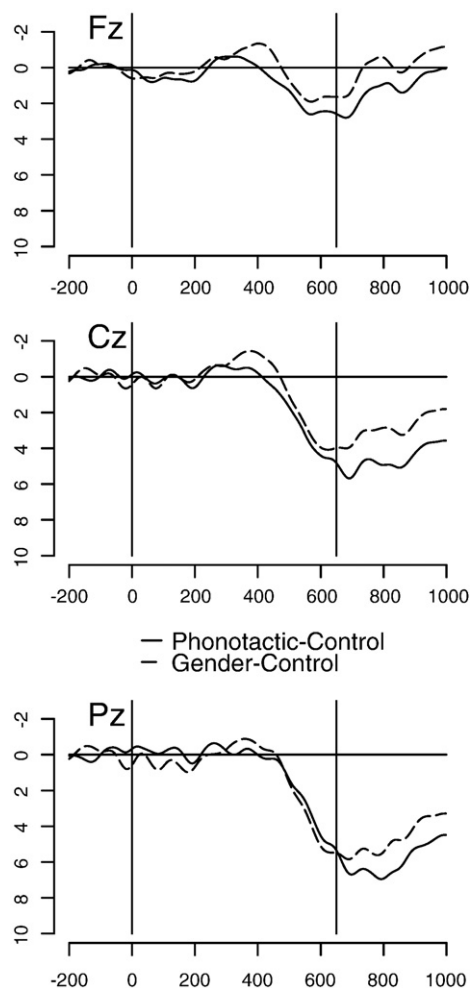


Fig. 4 – Difference waves (disagreement less agreement) comparing Phonotactic (solid line) and Gender Violation (dashed line) at the midline electrodes.

These data confirm that the Gender Violation elicited a sustained negativity, with a centro-parietal distribution, compared to both the Phonotactic Violation and the Control Condition. The Phonotactic Violation and the Control condition statistically differed only in the frontal electrodes, suggesting less difficulty in the final integration of this condition compared to the Gender Violation.

2.4. Summary of the results

Statistical analyses showed that phonotactic mismatches elicited the ERP pattern previously reported in literature for morphosyntactic mismatches, i.e. a LAN followed by the P600. Thus, also the phonotactic violation can be correlated with the electrophysiological pattern previously reported for other morphosyntactic violations.

Moreover, the LAN did not modulate in latency, since both violations elicited an effect with an onset around 300 ms. There was a significant difference in the topographical distribution of the two components: while the Phonotactic violation LAN did not show any effect over the central electrodes, the Gender violation was evident also over Cz. With respect to the P600 component, Gender and Phonotactic violations diverged in a late time window, after 700 ms, with the Phonotactic violation eliciting a larger positivity. This P600 pattern time-locked to the critical noun was reversed at the end of the sentence, with the Gender Violation eliciting an increased negativity compared to the Phonotactic. In the overall interpretation of the utterance, gender incongruence seemed to cause more wrap-up difficulties compared to the phonological incongruence.

3. Discussion

The aim of the present paper was to investigate how the electrophysiological components related to agreement

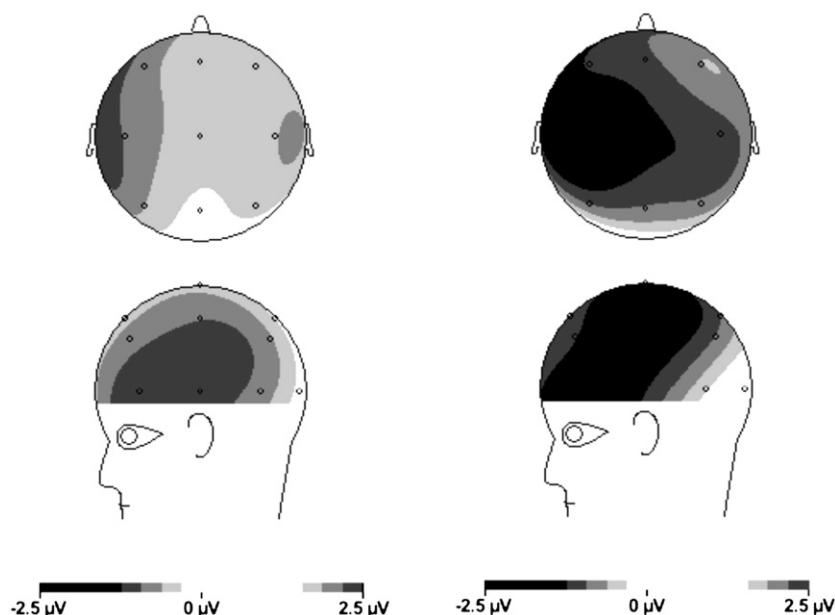


Fig. 5 – Scalp distribution of the LAN effects, mean voltage difference between ungrammatical and grammatical versions between 350 and 450 ms, respectively elicited by the Phonotactic Violation (left image) and the Gender Violation (right image).

Table 2 – F values and significance levels resulting from post hoc comparisons between the average amplitude values, in the 350–450 ms time-windows, of the experimental conditions for each lateralized group

df=1,19	Phonotactic violation vs. Agreement		Gender violation vs. Agreement		Phonotactic violation vs. Gender violation	
	Left	Right	Left	Right	Left	Right
Frontal	5.666*	-a	21.472***	6.258*	3.995#	3.345#
Central	12.685**	-a	27.645***	4.328*	-a	-a
Parietal	5.852*	-a	8.000*	-a	-a	-a

-aF<1, #p<0.1, *p<0.05, **p<0.01, ***p<0.001.

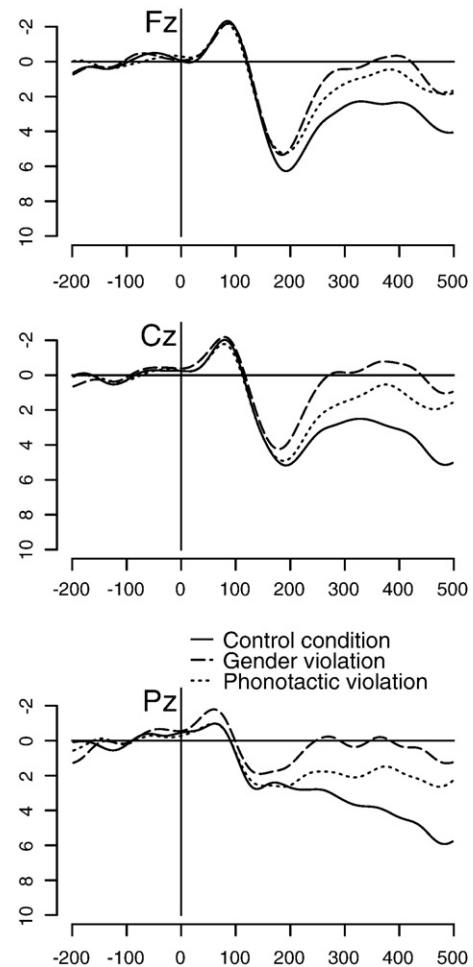
computation interact with different levels of processing (i.e. phonological and lexical), during sentence comprehension. We found that the violation of these constraints elicited the ERP pattern previously described in correlation with agreement violations: a left-lateralized frontal negativity followed by the P600. These ERP components were differently affected by the type of violation (i.e. phonotactic vs. gender agreement violation). We also found end of sentence effects, consistently with previous studies that compared sentences with syntactic violations to well-formed sentences.

We discuss these effects separately in the following paragraphs.

3.1. The LAN

In the Introduction we discussed a possible modulation in latency of the LAN in response to the manipulation of the type of constraint that was violated. Some models on sentence processing (Friederici, 2002; MacDonald et al., 1994) propose that the target word information could cascade to the sentence processor serially. In particular, since we used nouns with phonologically opaque endings for gender, which were also low frequency and long, we predicted that the lexical computation needed to recognise the gender of the target noun would have required more time than identifying the phonological properties of the same item. We then expected an earlier LAN onset for the phonotactic violation.

However, the pattern of data showed exactly the same timing in the ERP pattern elicited by the phonotactic and gender

**Fig. 6 – Grand-averages for the Control Condition (solid line), the Phonotactic Violation (dotted line) and the Gender Violation (dashed line) time-locked to the last word of the sentence.**

violations, namely a left-lateralized negativity with an onset around 300 ms that we recognized as a LAN. A similar component has been reported for number agreement violations in many languages (Barber and Carreiras, 2005; Coulson et al., 1998; De Vincenzi et al., 2003; Kutas and Hillyard, 1983; Osterhout and Mobley, 1995; Rossi et al., 2005), violations of gender agreement (Barber and Carreiras, 2005; Gunter et al., 2000), tense (Morris and Holcomb, 2005), person (Hinojosa et al., 2003) and other ungrammatical verb inflections (Friederici et al., 2003; Linares et al., 2006; Penke et al., 1997; Rodriguez-Fornells et al., 2001; Vos et al., 2001). However, these works were not specifically aimed to detect latency differences within agreement violations.⁵ On the other side large latency variations derive from word category violations, such as eLAN (Friederici, 2002).

⁵ Rizzi (personal communication) did find a latency shift of the peak of the LAN comparing number mismatches with tense mismatches, with the LAN for number mismatches 20 ms before the tense violation. To our knowledge this is the only study that reported latency shifts for this component and suggests that further studies have to be run in order to better understand the LAN phenomenon.

Table 3 – F values and significance levels resulting from post hoc comparisons between the average amplitude values, in the 500–700 ms and 700–900 ms time-windows, of the experimental conditions for each electrode of the midline (Fz, Cz, Pz)

df=1,19	Midline	Phonotactic violation vs. Agreement	Gender violation vs. Agreement	Phonotactic violation vs. Gender violation
500–700 ms	Frontal	18.316***	3.697#	-a
	Central	21.114***	10.292**	-a
	Parietal	23.866***	23.588***	-a
700–900 ms	Frontal	8.333**	-a	5.393*
	Central	42.988***	10.593**	6.741*
	Parietal	65.779***	41.816***	4.159*

-aF<1, #p<0.1, *p<0.05, **p<0.01, ***p<0.001.

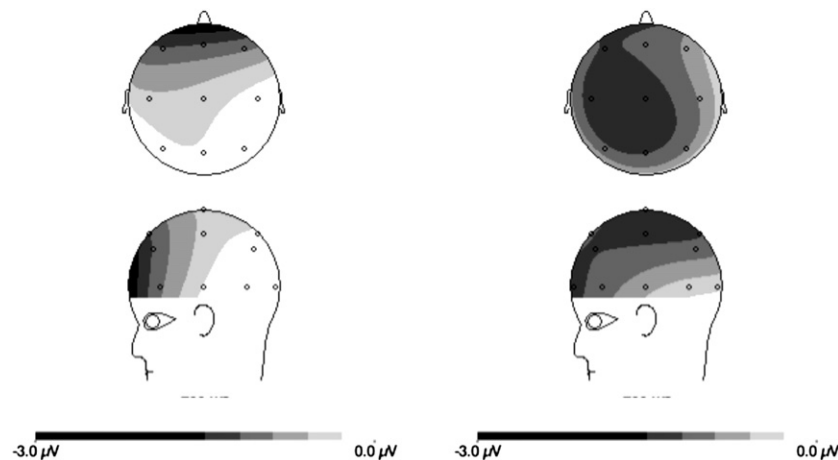


Fig. 7 – Scalp distribution of the end of sentence effects, mean voltage difference between ungrammatical and grammatical versions between 200 and 400 ms, respectively elicited by the Phonotactic Violation (left image) and the Gender Violation (right image).

The lack of time differences, then, does not corroborate the assumption that distinct word features cascade from lexical access to the parser during agreement computation. The only information that elicits an earlier LAN onset is the word category violation (Friederici, 2002), probably because the word category is processed before any other word property during sentence processing and not only because it is earlier available. Our data are consistent with the *unification model* (Haagort, 2005). A similar LAN onset for different agreement violations suggests that the parser does not check for agreement on the target word within the phrase until its lexical processing has been completed. The only fact that an incorrect phonology is perceived after a wrong determiner does not help the system to diagnose which was the possible error source (either incorrect parsing or speaker error) and solution; only after full lexical access of the target noun the system tries to integrate the feature involved in agreement computation. On the opposite, separately checking for agreement of different features (number, gender and allomorphic form whenever this information is needed) would imply multiplication of agreement processing. Moreover, the aim of the parser is not detecting possible grammatical errors but getting the message meaning and interpreting it once all the complete lexical representation of the target noun is available. As a consequence, the onset of the LAN does not modulate probably because this interval represents the time needed to process all the target word features before integrating it with the previous context. In this way LAN onset can converge with other findings (Dambacher et al., 2006; Yang et al., 2007) that assume 300 ms as a typical upper bound for lexical access.

The main dissociation we found between the phonotactic and the gender violation in the early time window concerned the LAN topography: while the effect elicited by the phonotactic mismatch has a left ventral distribution, the effect elicited by gender violation is not clearly distinguishable between a LAN and a N400 (see Fig. 5). Gender violations instead have been associated both with N400 (Deutsch and Bentin, 2001) and with not-lateralized frontal negativities (Hagoort and Brown,

1999). Many authors have also reported a LAN (Barber et al., 2004; Barber and Carreiras, 2005; Gunter et al., 2000).

Barber et al. (2004) suggested that the topography of the LAN could differ on the basis of the type of gender: they compared the effects related to gender agreement violation where gender has a strict grammatical role and effects related to semantic (biological) gender. Their hypothesis was that the former violation elicits a LAN while the latter elicits a N400. For example, the study of Deutsch and Bentin (2001) on subject–verb gender agreement violations reports N400 modulations based on the animacy of the previously processed subject. A similar effect was reported in German by Lamers et al. (2006) and Schmitt et al. (2002), when the gender of a pronoun did not match the biological gender of the antecedent. Barber et al. (2004) manipulated the animacy of the constituent where the violation is detected. This study is relevant for the present experiment, since the stimuli selected for our experiment included both inanimate nouns for which gender information has only a syntactic valence (18 per condition, like ‘sciale’ (+M), shawl) and nouns for which this feature has a biological meaning, the corresponding referent being animate (12 per condition, like ‘scia-tore’ (+M)/‘scia-trice’ (+F), skier).⁶ as a consequence, the effect we report for the gender violation could be a mix of LAN and N400. Nevertheless, since Barber et al. (2004) reported the same electrophysiological pattern when agreement violations of two types of gender (semantic and grammatical) were compared in the same experiment, we exclude this biological confound. The biological factor in gender agreement processing does not affect the (topographically central) distribution of the LAN.

This gender effect in Italian is similar to the ones reported in Spanish (Barber and Carreiras, 2005; Barber et al., 2004; Demestre et al., 1999) and in Hebrew (Deutsch and Bentin, 2001). As in our experiment, in the study of Barber and

⁶ This choice was pursued in order to extend the number of nouns requiring the ‘lo’ article in our material, and thus obtaining a consistent set of epochs to isolate the electrophysiological components related to the phonotactic manipulation.

Carreiras (2005) there seems to be a topographical dissociation: the reported maps show a qualitatively more ventral distribution of the LAN triggered by a number mismatch compared to the gender violation, that triggers a negativity even in more central areas of the scalp. In the present study the topography of the LAN elicited by gender violations has a scalp distribution that is more central but not so different to be easily detected with separate ANOVAs for midline scalp locations and more lateral ones. As we have shown, the use of statistics specifically suited for detecting differences between lateral and central distributions of the effects allows the discrimination of these topographical differences clearly visible on the topographical maps.

Overall, our data suggest that the computation of agreement between the target noun and the preceding context proceeds in parallel for the different features, involving distinct neural substrates depending on the violated feature. The negativity over the left anterior area of the scalp, evident also on the central areas, elicited by the gender violation, could indicate the interaction of the target word lexical features at the syntactic level. Lexical effects have been in fact reported to affect the N400 component during sentence processing (for a review Kutas et al., 2006). When the inherent lexical properties of the target word are not involved at the syntactic level of processing, the agreement violation does not trigger any negativity over the central area of the scalp, but elicits a ventrally distributed LAN. This explanation is consistent with the distribution of the LAN elicited by number agreement violations: since number is more independent than gender from the lexical properties of the target noun (it is in fact a non-inherent feature), it triggers a LAN with no central distribution. The phonotactic violation LAN shows a left ventral distribution either: in this case the computation of agreement is disrupted and the syntactic integration cannot proceed; also this violation however is independent from the lexical status of the target word, and it does not elicit any central negativity.⁷

3.2. The P600

The second component that is elicited by agreement violations is the P600. This ERP effect is not specific for agreement violations, but it could be informative on the type of reanalysis that is triggered by a linguistic incongruence. In the present study we recorded a larger posterior P600 effect for the phonotactic violation compared to the one elicited by the gender violation in the time window between 700 and 900 ms. Since there is neither topographical nor latency modulations of the P600 across conditions, the same neural population is recruited in this later stage of processing for both agreement mismatches. As discussed in the Introduction, a recent debate arose on the functional meaning of the P600, which has been reported even for orthographic (Visser et al., 2006), lexical (Federmeier et al., 2007) and semantic conflicts (Kim and

Osterhout, 2005; Kolk et al., 2003; Kuperberg et al., 2003). Excluding the former two cases (orthographic and lexical conflicts) in which respectively the latency and the topography of the positive shift are not consistent with the classical posteriorly distributed effect, the latter studies (semantic conflicts) could still be interpreted as a syntactic conflict (as suggested by Kim and Osterhout, 2005).

Across this family of effects, in the present study we report a posteriorly distributed P600 effect with an onset around 500 ms. In our opinion it represents a two-stage reanalysis triggered by a syntactic conflict, since both the phonotactic and the gender violation represent a mismatch at a syntactic level. Barber and Carreiras (2005; see also Carreiras et al., 2004) first suggested that this positive shift could be decomposed into two distinct stages: an early stage of diagnosis of the incongruence within a sentence (P600 onset: 500–700 ms) followed by the final reprocessing of the stimulus (P600 offset: 700–900 ms). The P600 onset has a more wide distribution over the scalp (both on anterior and posterior areas) and may be related to the reactivation of the target word to operate a fine-grained analysis of the incongruence and detecting the source of the error. This stage of reactivation (i.e. tracing back the error probably checking also the contextual information) should then be sensitive to working memory limitations. This seems confirmed in some studies by Vos (Vos et al., 2001; Vos and Friederici, 2003) in which the onset of the P600 was sensitive to the reading span levels of the participants. In the present study we did not find any difference in the onset of the P600 between the two violations (gender vs. phonotactic violation: Table 3), probably because they both reactivated the same critical context, i.e. the determiner preceding the noun.

The P600 offset has a more posterior distribution and its amplitude depends on the stage of processing of the critical word at which the system has to come back to perform reanalysis. In other words, the processor may be uncertain about whether or not the relation between the target word and the preceding sentence fragment was processed exactly at each level of analysis; the processor is then compelled to go back to the stage of processing where the incongruence arises in order to reprocess that particular information. As proposed by Barber and Carreiras (2005; based on Faussart et al., 1999), the more number of stages the system has to come back, the larger the amplitude of the P600 offset is.

Table 4 – F values and significance levels resulting from post hoc comparisons between the average amplitude values after the onset of the last word of each sentence, in the 200–400 ms time-windows, of the experimental conditions for each electrode of the midline (Fz, Cz, Pz)

df=1,19	Midline	Phonotactic violation vs. Agreement	Gender violation vs. Agreement	Phonotactic violation vs. Gender violation
200–400 ms	Frontal	6.423*	9.509**	^a
	Central	4.087#	19.876***	9.603**
	Parietal	4.169#	25.114***	8.081*
^a F<1, #p<0.1, *p<0.05, **p<0.01, ***p<0.001.				

⁷ A caveat for the present hypothesis concerns the visual presentation of the stimuli: it is possible that the phonotactic violation is detected later because the participants process foveally the whole word form (i.e. with all feature information present).

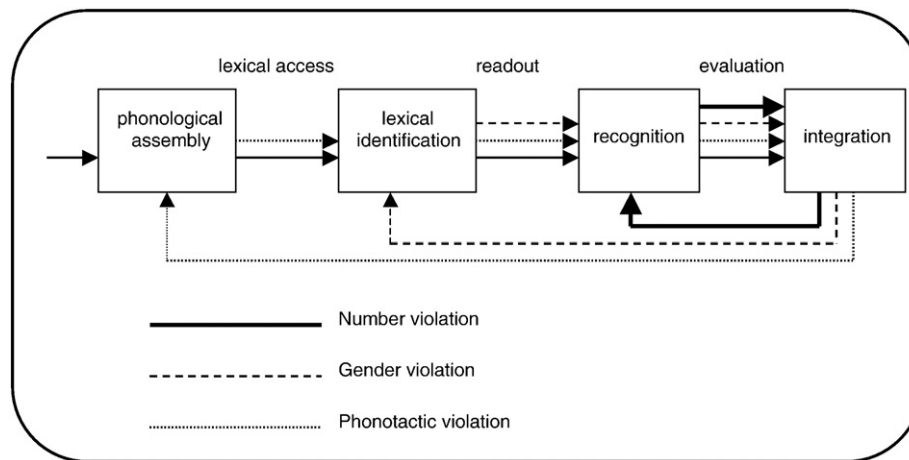


Fig. 8 – Extension of Faussart et al. (1999) model, proposed in the present paper.

Our data nicely fit with their P600 model of reanalysis. Moreover, many models on single word processing show that the phonological processing precedes the selection of the correct lexical entry (Coltheart et al., 2001; Holcomb and Grainger, 2006; Perry et al., 2007). The reanalysis of a phonotactic violation, where the system recognizes an illegal syllabic sequence, thus requires the system to go back to a stage of phonological processing, before the stage (i) in Faussart et al. (1999) model. This is confirmed by the fact that the phonotactic violation elicits a P600 with larger amplitude compared to the gender violation (Fig. 4) in its late time window (Table 3).

In Fig. 8 we present an extension of the model proposed by Faussart et al. (1999), integrated with the data reported in the present study.

The hypothesis that the late P600 amplitude reflects the regression needed for reanalysis is consistent with a number of ERP studies that compared different types of violation. For instance, the P600s elicited by syntactic violations have a different topography compared to the ones elicited by unpreferred continuations, like garden-path sentences (Friederici et al., 2002; Kaan et al., 2000; Kaan and Swaab, 2003; Kuperberg et al., 2003). While syntactic violations elicit P600 with a posterior distribution, syntactic ambiguities elicit a positivity with the same latency but a frontal topographical distribution. Critically, a substantial difference is identifiable also in the late amplitude of the components: violations elicit larger late P600 effects compared to unpreferred continuations. Even these findings can fit with our model of the P600: a morphosyntactic violation requires the system to go back to the lexical analysis of the critical item (stages (i) or (ii) in Faussart et al. model) and triggers a bigger P600 compared to unpreferred structures that compels the system to rerun only the integration stage (stage (iii)), since there is no morpho-syntactic violation.

Another example comes from a study by Munte et al. (1998) that compared the ERPs elicited by semantic violations, morphosyntactic violations and orthographic violations (where the critical word was substituted by a pseudo-homophone, i.e. an orthographically different word sharing the same phonological properties) in German. They found P600 not only for morpho-

syntactic violations but also for orthographic violations. Interestingly, Munte et al. (1998) reported a larger P600 elicited by pseudo-homophones, compared to the morphosyntactic violation of the morphosyntactic case. This could mirror the fact that the pseudo-homophone compels the system to come back to the stage of orthographic processing of the critical word, while the violation of case triggers a reprocessing of the lexical stage of processing. A following study run in Dutch by Vissers et al. (2006) showed that the P600 for pseudo-homophones arises only in a high-cloze probability context. These authors interpret their P600 within the conflict monitoring frame (Kolk and Chwilla, 2007). On the one hand the pseudo-homophone is clearly unacceptable, on the other hand it is highly expected not only semantically, but also phonologically, since the phonological form of the pseudo-homophone confirms the semantic expectation. It thus seems that this conflict compels the system to reprocess the word at an orthographic level of processing, eliciting a larger P600 with respect to a syntactic violation.

3.3. End-of-sentence effects

The processing difficulties triggered by a syntactic incongruence well correlate with an electrophysiological component triggered by the last word of the sentence. If the sentence is anomalous (both syntactically and semantically) an enhanced negativity is triggered at the end of the string. This electrophysiological correlate has been consistently correlated with wrap-up processes (Osterhout and Holcomb, 1995), i.e. a reinterpretation of all the arguments of the sentence in order to check if they have been represented correctly for obtaining the whole meaning of the utterance.

In this study we report a larger end of sentence negativity elicited by the violation of gender agreement compared to the phonotactic violation. Interesting to note, the pattern of the end of sentence effect for the two violations presents the opposite trend compared to the P600 amplitude time-locked to the critical noun: while the P600 was larger for the phonotactic violation, the end of sentence negativity was larger for the gender violation. As discussed above, we consider the P600 amplitude to reflect the backward steps, in the target noun processing, the system has to do in order to perform

reanalysis. Even if the causes of the wrap-up effects have never been fully identified (cf. [Frazier, 1999](#)), they are considered to include all the processes of semantic interpretation of the sentence in a broad sense, such as establishing its true-value properties, establishing the referents of free pronouns, establishing the speech act of the sentences (i.e., if it is a question or an assertion).

In the present experiment, the larger negativity for the gender violation could be caused by the fact that gender values are often expressed referentially across sentences, also through the use of tag pronouns. As a consequence, selecting the gender value associated to each sentential argument is critical. This is not the case of the phonotactic mismatch, where all the values (number and gender) associated to the determiner-noun pair are congruent. This difference could explain the enhanced negativity of the sentences containing gender violations compared to the effect associated to phonotactic violations that is only marginally significant ([Table 4](#)).

3.4. Conclusions

In the present paper we developed a fine-grained analysis of the electrophysiological effects elicited by qualitatively different types of agreement violations: we compared the influence of phonotactic constraints and lexical constraints during agreement processing. Through this study we bring evidence that agreement computation starts after the lexical processing of the target word is completed, evaluating interactively the different feature that constrain agreement. Agreement checking shows however a qualitatively different processing for distinct features, since distinct neural networks reacts to different types of violations. The following reprocessing procedure tries to isolate the source of the error, in order to get a coherent interpretation of the message. The present study thus shows that both the topography of the early negativity and the amplitude of the late positivity could be informative for investigating syntactic computations.

4. Experimental procedures

4.1. Participants

Twenty-one undergraduate students of the University of Padova took part at the experiment (12 females, age-range: 19–28 years; mean: 22.3 years). All of them were native Italian speakers. All participants were right-handed as assessed by the Edinburgh Inventory for handedness ([Oldfield, 1971](#)). They reported having no neurological disorder and having normal or corrected-to-normal vision. They were assigned course credits for their participation.

4.2. Material

Ninety Italian masculine nouns were selected: half of them required the ‘il’ determiner and half the ‘lo’ determiner. We then presented a questionnaire to 18 Italian speakers asking for the right masculine article for each item on a seven-point scale (with the two articles at the extremes). The critical items

were intermixed with the same number of nouns requiring the ‘il’ determiner. In general, there was less confidence in recognizing the right determiner for the nouns requiring ‘lo’. From this starting list we choose the 45 ‘lo’ nouns that were more consistently assigned with the right determiner. For each of these items, we selected a corresponding ‘il’ noun with similar frequency and number of letters.

We selected unambiguous masculine nouns, avoiding nouns like ‘docente’ (teacher) that is bi-gendered. The starting grapheme for the ‘lo’ nouns was always the ‘s-’, while the ‘il’ nouns were more heterogeneous: 5 with ‘b-’, 10 ‘c-’, 3 ‘d-’, 4 ‘f-’, 1 ‘g-’, 2 ‘l-’, 3 ‘m-’, 4 ‘p-’, 1 ‘q-’, 10 ‘s-’ and 1 ‘t-’. The ending grapheme in most of the ninety nouns was ‘-e’, which could be the suffix both for masculine and for feminine nouns. Given the low number of opaque and low frequent nouns requiring ‘lo’ we selected 12 transparent items ending with ‘-tore’, a morphological suffix referring to a masculine referent in Italian.

Most of the nouns were concrete with high imaginability, but the list comprised some concrete nouns with low imaginability (like ‘il paese’, the town) and some abstract items (like ‘lo stile’, the style). Neither the number of abstract nouns in the two lexical categories (9/45 for ‘il’ and 14/45 for ‘lo’, Chi-sq (1)=0.93) nor the number of low-imaginability nouns (19/45 for ‘il’ and 19/45 for ‘lo’) was statistically different in the ‘il/lo’ lists.

The ninety nouns were then divided in three different groups, each of them containing 15 ‘il’ nouns and 15 ‘lo’ nouns. The logarithm of the written frequency of the nouns was balanced (‘il’ average LogF: 1.93; ‘lo’ average LogF: 1.86; $t(88)=0.21$). Each group contained 12 animate and 18 inanimate nouns.

In order to control whether the material was balanced on a number of dimensions, we performed separate ANOVAs with Article (2 levels) and Group (3 levels) as factors on the following parameters: number of letters (mean across conditions: 8.64) and logarithm of the frequency (mean cross conditions: 1.89). We did not find any significant effect.

For each noun we composed a sentence containing the noun and the corresponding determiner. The NP was not in initial or final position of the sentence. Sentences contained from 8 to 13 words (mean 10.8, s.d. 1.3). The syntactic structure of the sentences was variable. We checked that the critical NPs were not inserted in high-cloze contexts, in order to avoid possible semantic anticipatory effects (as described by [DeLong et al., 2005](#)). The fragment of the sentence preceding the critical NP was not predictive of the NP itself, allowing for any determiner (‘il’, ‘lo’ and ‘la’) to be plausible.

We composed three different lists based on the groups above described. In each list the stimuli were divided in the three following conditions: Control Condition, Phonotactic Violation, Gender Violation (see [Table 1](#)).

In addition to the experimental material, the list contained 130 filler sentences (correct sentences, garden-path sentences, semantic violations and number agreement violations), for a total of 220 sentences: half of them were grammatical and half ungrammatical.

4.3. Procedure

Participants were randomly assigned to one stimulus list. Each participant was seated in front of a computer monitor and was

instructed to read as normally as possible and to try to understand the sentences. Each trial consisted of the following events: a fixation cross appeared in the centre of the screen for 400 ms, after which a stimulus sentence was presented word by word. Each word appeared in the centre of the screen for 300 ms, followed by a blank-screen interval of 300 ms. Sentence-ending words appeared followed by a full stop. Words were presented in white fonts (font: Courier New; size: 18) on a dark grey background (RGB values: 60, 60, 60).

A 1000 ms blank-screen interval followed each sentence, after which a prompt appeared asking participants to decide if the preceding sentence was a normal sentence. Participants were instructed to answer 'SI' (yes) if the sentence was semantically coherent and grammatically well formed and 'NO' (no) otherwise. Participants responded by pressing one of two buttons, which were counter-balanced (left and right) across participants. Participants were instructed not to blink or move their eyes during sentence presentation. To familiarize them with the procedure, a practice block was presented before the experimental trials with twenty items (half with syntactic errors), none of which had the structure used in the experimental manipulation. On average the experiment took about 1 h and 40 min per participant, including electrode montage and debriefing.

4.4. Data acquisition and analysis

This experiment was run using a BrainAmp acquisition and analysis system. EEG was continuously recorded from 21 electrodes placed on the scalp with the aid of an elastic cap at standard positions (10–20 system). Recordings were obtained from left and right pre-frontal (Fp1, Fp2), frontal (F3, F4), inferior frontal (F7, F8), temporal (T3, T4), central (C3, C4), parietal (P3, P4), posterior temporal (T5, T6), and occipital (O1, O2) locations, and from three midline locations (Fz, Cz and Pz). Additional external electrodes of the same material were placed on mastoids A1, A2 and around eyes Ve1, Ve1, He1, He2. Monopolar differential recording was referenced to the left mastoid. Impedance was kept below 5 k Ω for mastoid and scalp electrodes, and below 10 k Ω for EOG electrodes. Data were acquired at a sampling rate of 250 Hz.

Offline EEG amplitudes were filtered with a 30 Hz low-pass filter and segmented in epochs of (-200, 1000) ms, with respect to the target word onset (the critical noun). Epochs with an absolute difference between two values larger than 80 μ V were rejected in order to exclude from the following analyses trials characterized by eye blinks or muscle artefacts. On average, the following percentage of epochs was rejected: control condition: 7%; Phonotactic violation: 8%; Gender violation: 8%. Number of rejected trials did not statistically differ across conditions. One subject was excluded from the analyses since the number of epochs per condition was too low.

Epochs were then corrected for baseline activity for each single channel as the mean voltage in the (-200, 0) ms range and averaged with respect to the different experimental conditions. For each of the described conditions the obtained single subject waveforms were (i) used to calculate the mean voltage on given time-windows used for the statistical analysis and (ii) averaged between subjects in order to plot grand-average waveforms for the qualitative analyses.

We quantified ERPs through the amplitude of the components of interest as the mean voltage of single subjects' waveforms within a window of activity. The following windows were employed: 350–450 ms (LAN/N400), 500–700 ms and 700–900 ms for the P600. Repeated measures analyses of variance (ANOVAs) were performed on the above dependent measures. The Greenhouse–Geisser (Greenhouse and Geisser, 1959) correction was applied to all repeated measures with greater than one degree of freedom in the numerator. In such cases, the corrected *p*-value is reported.

Data acquired at midline electrodes and data acquired at the lateral sites were treated separately to allow for quantitative analysis of hemispheric differences. On the data from midline sites, two-way ANOVAs were performed, with repeated measures on three levels of sentence type (Phonotactic Violation, Gender Violation and Control) and three levels of longitude (Frontal, Central and Parietal).

Data acquired over the lateral sites were averaged in different spatially homogeneous groups in order to reduce global variance and number of levels for following evaluation with repeated measure ANOVAs. This grouping, involved in the analyses of the lateral electrodes was defined in the following six clusters, three per hemisphere in the longitudinal dimension (Fig. 9): left Frontal (F7, F3), Right Frontal (F8, F4), Left Central (T3, C3), Right Central (T4, C4), Left Posterior (T5, P3) and Right Posterior (T6, P4). The six clusters grouping was analyzed in terms of two spatial factors that could interact with the sentence type factor: longitude (three levels: Frontal, Central, Parietal) and hemisphere (two levels: left, right).

The midline analysis was primarily used to analyze the P600 effect, which is typically not lateralized and maximal at the centro-parietal sites along the midline, while the six clusters grouping was primarily used to analyze LAN activity and differentiate lateralized activity between conditions. Effects related with longitude factor or hemisphere factor were

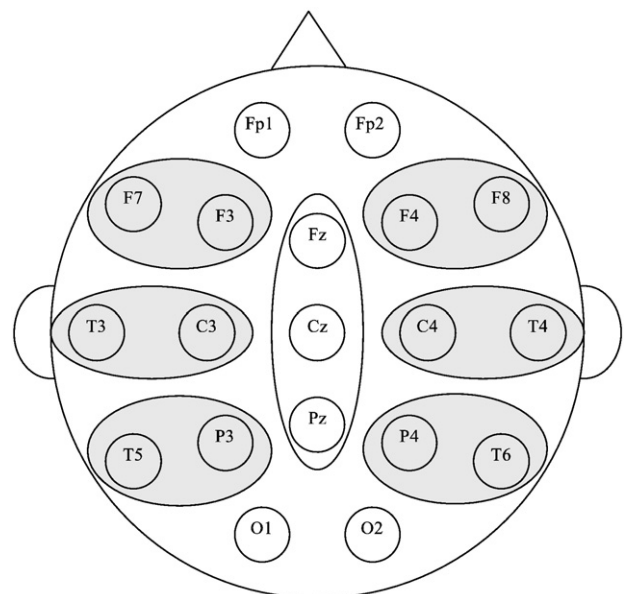


Fig. 9 – Electrodes used in the ERP experiment and relative groups used for the analyses.

considered when they interacted with the experimental manipulations. In order to better discriminate the topographical distribution of the LAN effects, when significant interactions emerged from the overall ANOVA, we performed one-way ANOVAs on each cluster of electrodes comparing two experimental conditions. A similar analysis scheme was conducted for the P600 evaluation: however, since the P600 is usually not lateralized, we performed one-way ANOVAs between two experimental conditions on each midline electrode.

Acknowledgments

This research was conducted at the Cognition and Language Lab of the Università degli Studi di Padova, Padova (Italy) and was supported by a PRIN 2005 grant. NM is partially founded by the “Juan de la Cierva” program and the grant SEJ2006-09238 from the Spanish Ministry of Education and Science. Preliminary results of the present study were presented at the ‘Architecture of Language’ meeting held in Pisa (Italy), June 8–10, 2006.

REFERENCES

- Barber, H., Carreiras, M., 2005. Grammatical gender and number agreement in Spanish: an ERP comparison. *J. Cogn. Neurosci.* 17, 137–153.
- Barber, H., Salillas, E., Carreiras, M., 2004. Gender or genders agreement? In: Carreiras, M., Clifton, Jr., C. (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERP, and Beyond*. Psychology Press.
- Bradley, D.C., Forster, K.I., 1987. A reader's view of listening. *Cogn.* 25, 103–134.
- Carreiras, M., Salillas, E., Barber, H., 2004. Event-related potentials elicited during parsing of ambiguous relative clauses in Spanish. *Cogn. Brain Res.* 20, 98–105.
- Chomsky, N., 1995. *The Minimalist Program*. MIT press.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., Ziegler, J., 2001. DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256.
- Corbett, G.G., 1991. *Gender*. Cambridge University Press, Cambridge, UK.
- Coulson, S., King, J., Kutas, M., 1998. Expect the unexpected: event-related brain response to morphosyntactic violations. *Lang. Cogn. Processes* 13, 21–58.
- Dambacher, M., Kliegl, R., Hofmann, M., Jacobs, A.M., 2006. Frequency and predictability effects on event-related potentials during reading. *Brain Res.* 1084, 89–103.
- DeLong, K.A., Urbach, T.P., Kutas, M., 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nat. Neurosci.* 8, 1117–1121.
- Demestre, J., Meltzer, S., Garcia-Albea, J.E., Vigil, A., 1999. Identifying the null subject: evidence from event-related potentials. *J. Psycholinguist. Res.* 28, 293–312.
- Deutsch, A., Bentin, S., 2001. Syntactic and semantic factors in processing gender agreement in Hebrew: evidence from ERPs and eye movements. *J. Mem. Lang.* 45, 200–224.
- De Vincenzi, M., 1999. Differences between the morphology of gender and number: evidences from establishing coreferences. *J. Psycholinguist. Res.* 28, 537–553.
- De Vincenzi, M., Job, R., Di Matteo, R., Angrilli, A., Penolazzi, B., Ciccarelli, L., Vespignani, F., 2003. Differences in the perception and time course of syntactic and semantic violations. *Brain Lang.* 85, 280–296.
- Faussart, C., Jakubowicz, C., Costes, M., 1999. Gender and number processing in spoken French and Spanish. *Riv. Linguist.* 11, 75–101.
- Federmeier, K.D., Wlotko, E.W., De Ochoa-Dewald, E., Kutas, M., 2007. Multiple effects of sentential constraint on word processing. *Brain Res.* 1146, 75–84.
- Frazier, L., 1999. *On Sentence Interpretation*. Kluwer Academic Publishers, Dordrecht, Holland.
- Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* 6, 78–84.
- Friederici, A.D., Hahne, A., Saddy, D., 2002. Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *J. Psycholinguist. Res.* 31, 45–63.
- Friederici, A.D., Ruschmeyer, S.A., Hahne, A., Fiebach, C.J., 2003. The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cereb. Cortex* 13, 170–177.
- Goldsmith, J.A., 1990. *Autosegmental and Metrical Phonology*. Basil Blackwell, p. Oxford, UK.
- Grainger, J., Kiyonaga, K., Holcomb, P.J., 2006. The time course of orthographic and phonological code activation. *Psychol. Sci.* 17, 1021–1026.
- Greenhouse, S., Geisser, S., 1959. On methods in the analysis of profile data. *Psychom.* 24, 95–112.
- Gunter, T.C., Friederici, A.D., Schriefers, H., 2000. Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *J. Cogn. Neurosci.* 12, 56–568.
- Hagoort, P., 2003. Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *J. Cogn. Neurosci.* 15, 883–899.
- Haagort, P., 2005. On Broca, brain and binding: a new framework. *Trends Cogn. Sci.* 9, 416–423.
- Hagoort, P., Brown, C.M., 1999. Gender electrified: ERP evidence on the syntactic nature of gender processing. *J. Psycholinguist. Res.* 28, 715–728.
- Hinojosa, J.A., Martin-Loeches, M., Casado, P., Munoz, F., Rubia, F.J., 2003. Similarities and differences between phrase structure and morphosyntactic violations in Spanish: an event-related potentials study. *Lang. Cogn. Processes* 18, 113–142.
- Holcomb, P.J., Grainger, J., 2006. On the time course of visual word recognition: an event-related potential investigation using masked repetition priming. *J. Cogn. Neurosci.* 18, 1631–1643.
- Jackendoff, R., 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press.
- Just, M.A., Carpenter, P., 1980. A theory of reading: from eye fixations to comprehension. *Psychol. Rev.* 85, 109–130.
- Kim, A., Osterhout, L., 2005. The independence of combinatory semantic processing: evidence from event-related potentials. *J. Mem. Lang.* 52, 205–225.
- Kaan, E., Swaab, T.Y., 2003. Repair, revision, and complexity in syntactic analysis: an electrophysiological differentiation. *J. Cogn. Neurosci.* 15, 98–110.
- Kaan, E., Harris, A., Gibson, E., Holcomb, P., 2000. The P600 as an index of syntactic integration difficulty. *Lang. Cogn. Processes* 15, 159–201.
- Kolk, H.H.J., Chwilla, D.J., 2007. Late positivities in unusual situations. *Brain Lang.* 100, 257–261.
- Kolk, H.H.J., Chwilla, D.J., Van Herten, M., Oor, P.J.W., 2003. Structure and limited capacity in verbal working memory: a study with event-related potentials. *Brain Lang.* 85, 1–36.
- Kuperberg, G.R., 2007. Neural mechanisms of language comprehension: challenges to syntax. *Brain Res.* 1146, 23–49.
- Kuperberg, G.R., Sitnikova, T., Caplan, D., Holcomb, P.J., 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cogn. Brain Res.* 17, 117–129.
- Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470.

- Kutas, M., Hillyard, S.A., 1983. Event-related brain potentials to grammatical errors and semantic anomalies. *Mem. Cogn.* 11, 539–550.
- Kutas, M., Van Petten, C., Kluender, R., 2006. Psycholinguistic electrified II: 1994–2005. In: Traxler, M., Gernsbacher, M.A. (Eds.), *Handbook of Psycholinguistics*, 2nd Edition. Elsevier, New York, pp. 659–724.
- Lamers, M.J.A., Jansma, B.M., Hammer, A., Munte, T.F., 2006. Neural correlates of semantic and syntactic processes in the comprehension of case marked pronouns: evidence from German and Dutch. *BMC Neurosci.* 7, 1–12.
- Laudanna, A., Thornton, A.M., Brown, G., Burani, C., Marconi, L., 1995. Un corpus dell'italiano scritto contemporaneo dalla parte del ricevente. In: Bolasco, S., Lebart, L., Salem, A. (Eds.), *III Giornate internazionali di Analisi Statistica dei Dati Testuali*, vol. I. Cisu, Roma, pp. 103–109.
- Linares, R.E., Rodriguez-Fornells, A., Clahsen, H., 2006. Stem allomorphy in the Spanish mental lexicon: Evidence from behavioral and ERP experiments. *Brain Lang.* 97, 110–120.
- MacDonald, M.C., Pearlmutter, N.J., Seidenberg, M.S., 1994. The lexical nature of syntactic ambiguity resolution. *Psychol. Rev.* 101, 676–703.
- McCarthy, C., Wood, C.C., 1985. Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalogr. Clin. Neurophysiol.* 62, 203–208.
- McQueen, J.M., 1998. Segmentation of continuous speech using phonotactics. *J. Mem. Lang.* 39, 21–46.
- Miozzo, M., Caramazza, A., 1999. The selection of lexical-syntactic features in noun phrase production: evidence from the picture-word interference paradigm. *J. Exp. Psychol.: Learn. Mem. Cogn.* 25, 907–922.
- Morris, J., Holcomb, P.J., 2005. Event-related potentials to violations of inflectional verb morphology in English. *Cogn. Brain Res.* 25, 963–981.
- Munte, T.F., Heinze, H.J., Matzke, M., Wieringa, M., Johannes, S., 1998. Brain potentials and syntactic violations revisited: no evidence for specificity of the syntactic positive shift. *Neuropsychologia* 36, 217–226.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Osterhout, L., Holcomb, P.J., 1995. Event-related brain potentials and language comprehension. In: Rugg, M.D., Coles, M.G.H. (Eds.), *Electrophysiology of Mind: Event-related Brain Potentials and Cognition*. Oxford University Press, Oxford.
- Osterhout, L., Mobley, L.A., 1995. Event-related brain potentials elicited by failure to agree. *J. Mem. Lang.* 34, 739–773.
- 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: Carreiras, M., Clifton, C. (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERP, and Beyond*. Psychology Press.
- Penke, M., Weyerts, H., Gross, M., Zander, E., Munte, T.F., Clahsen, A., 1997. How the brain processes complex words: an ERP study of German verb inflections. *Cogn. Brain Res.* 6, 37–52.
- Perry, C., Ziegler, J.C., Zorzi, M., 2007. Nested incremental modeling in the development of computational theories: the CDP+ model of reading aloud. *Psychol. Rev.* 114, 273–315.
- Ritter, E., 1988. A head-movement approach to construct state noun phrases. *Linguist.* 26, 909–929.
- Rodriguez-Fornells, A., Clahsen, H., Lleo, C., Zaake, W., Munte, T.F., 2001. Event related brain responses to morphological violations in Catalan. *Cogn. Brain Res.* 11, 47–58.
- Roehm, D., Bornkessel-Schlesewsky, I., Rosler, F., Schlesewsky, M., 2007. To predict or not to predict: influences of task and strategy on the processing of semantic relations. *J. Cogn. Neurosci.* 19, 1259–1274.
- Rossi, S., Gugler, M.F., Hahne, A., Friederici, A.D., 2005. When word category information encounters morphosyntax: an ERP study. *Neurosci. Lett.* 384, 228–233.
- Russi, C., 2006. A 'usage-based' analysis of the allomorphy of the Italian masculine definite article. *Stud. Lang.* 30, 575–598.
- Schmitt, B.M., Lamers, M., Munte, T.F., 2002. Electrophysiological estimates of biological and syntactic gender violation during pronoun processing. *Cogn. Brain Res.* 14, 333–346.
- Tabossi, P., Burani, C., Scott, D., 1995. Word identification in fluent speech. *J. Mem. Lang.* 34, 440–467.
- Visser, C.W.M., Chwilla, D.J., Kolk, H.H.K., 2006. Monitoring in language perception: The effect of misspellings of words in highly constrained sentences. *Brain Res.* 1106, 150–163.
- Vosse, T., Kempen, G.A.M., 2000. Syntactic structure assembly in human parsing: a computational model based on competitive inhibition and lexicalist grammar. *Cogn.* 75, 105–143.
- Vos, S.H., Friederici, A.D., 2003. Intersentential syntactic context effects on comprehension: the role of working memory. *Cogn. Brain Res.* 16, 111–122.
- Vos, S.H., Gunter, T.C., Kolk, H.H.J., Mulder, G., 2001. Working memory constraints on syntactic processing: an electrophysiological investigation. *Psychophysiology* 38, 41–63.
- Yang, C.L., Perfetti, C.A., Schmalhofer, F., 2007. Event-related potential indicators of text integration across sentence boundaries. *J. Exp. Psychol.: Learn. Mem. Cogn.* 33, 55–89.