



# Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity\*

ALICE FOUCART  
*University of Edinburgh*  
 CHERYL FRENCK-MESTRE  
*CNRS, Université de Provence*

(Received: May 19, 2009; final revision received: January 28, 2010; accepted: February 1, 2010)

*This study examines the effect of proficiency and similarity between the first and the second language on grammatical gender processing in L2. In three experiments, we manipulated gender agreement violations within the determiner phrase (DP), between the determiner and the noun (Experiment 1), the postposed adjective and the noun (Experiment 2) and the preposed adjective and the noun (Experiment 3). We compared the performance of German advanced learners of French to that of French native controls. The results showed a similar P600 effect for native and non-native speakers for agreement violations when agreement rules were similar in L1 and L2 (Experiment 1, depending on proficiency), whereas no effect was found for L2 learners when agreement rules varied across languages. These results suggest that syntactic processing in L2 is affected by the similarity of syntactic rules in L1 and L2.*

Keywords: second language acquisition, grammatical gender, ERPs

Grammatical gender (hereafter, gender), in contrast to natural or semantic gender, does not rely on semantic properties (e.g., sex, colour, shape, etc.) to categorize nouns into classes (Corbett, 1991). Gender is a classification system for nouns and is responsible for the syntactic cohesion of elements in a phrase through agreement. For instance, German, Spanish and French nouns assign gender to determiners, adjectives and participles which modify them and to pronouns which co-refer with them. The aim of the present article was to investigate how grammatical gender is processed in French as a second language (L2) and as compared to native speakers. While all native speakers acquire gender (Carroll, 1989; Clark, 1985; Müller, 1990; Perez-Pereira, 1991), how they do so and how gender may affect online processing is a matter of debate (Barber & Carreiras, 2005; Bolte & Conine, 2004; Deutsch & Bentin, 2001; Gunter, Friederici & Schriefers, 2000; Mirković, MacDonald & Seidenberg, 2005; Spinelli, Meunier & Seigneuric, 2006). This question becomes even more complex when one considers non-native speakers who start learning their L2 French late in life (i.e., after adolescence) (Bruhn de Garavito & White, 2002; Carroll, 1989; Hawkins & Chan, 1997; Herschensohn, 2007). Learning an L2 involves the acquisition of both competence and performance in

this language. For instance, learning gender in French involves acquiring both the knowledge of a word's gender (i.e., gender assignment) and how gender is expressed syntactically (i.e., gender agreement). The L2 learner must then develop the capacity to systematically produce and process this knowledge. In this article, we use the term PROCESS to refer to the realization of gender agreement between the noun and other elements related to it. We examine gender processing online via the recording of event related cortical potentials (ERPs) for late L2 learners in comparison to native processing and how performance is affected by similarity of the L1 and L2 on the other.

Several studies have used ERPs to examine the online processing of gender agreement in native speakers. ERP methodology is appropriate to the study of gender processing as violations can elicit either a lexico-semantic effect (N400) or syntactic effects (P600, LAN) and indeed the question has been raised as to whether gender is represented semantically or syntactically. The general finding of these studies is that gender agreement violations between two elements in sentence context provoke a P600 effect, reflecting difficulty in syntactic processing, which is sometimes preceded by a LAN (left anterior negativity) effect.

All of the studies to date that have examined grammatical gender processing in native speakers revealed a P600 effect in response to gender agreement violations (with the exception of violations occurring

\* This study was supported by a grant from the French Ministry of Foreign affairs.

Address for correspondence:

Alice Foucart, University of Edinburgh, 21, Hill Place, Edinburgh EH8 9DP, Scotland  
 alfoucart@googlemail.com

outside sentential context, in word pairs, which triggered an N400 effect). This effect was obtained regardless of the elements involved (e.g., article–noun, adjective–noun, reflexive–antecedent) or syntactic structure (within or outside the DP). The P600 effect was sometimes preceded by a LAN effect (Barber & Carreiras, 2005; Deutsch & Bentin, 2001; Gunter et al., 2000), but not consistently (Foucart & Frenck-Mestre, 2007; Frenck-Mestre, 2005; Hagoort & Brown, 1999). The LAN and P600 effects have been associated with syntactic processing, the LAN supposedly reflecting syntactic violation detection, and the P600 effect syntactic reanalysis or repair (Friederici, 2002; but see Osterhout, McLaughlin, Kim & Inoue, 2004). Thus, the consensus that emerges from these studies is that in L1, gender is represented syntactically, and that the online processing of grammatical gender is not a conceptual and/or semantic, but a syntactically driven process. The question raised here is whether L1 and L2 gender processing are similar in this regard.

Event related potentials have provided information that complements findings obtained with other methodologies regarding L2 processing. Notably, the capacity to distinguish semantic from syntactic processing has allowed researchers to show that lexico-semantic processing in L1 and L2 is very similar (Ardal, Donald, Meuter, Muldrew & Luce, 1990; Hahne, 2001), in contrast to syntactic processing, which has been claimed to differ in L1 and L2 at least as revealed by the processing of violations. On the one hand, a P600 effect similar to that found in native speakers has been revealed in L2 learners for syntactic anomalies (Hahne, 2001; Hahne, Mueller & Clahsen, 2006; Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008; Weber-Fox & Neville, 1996), suggesting that native-like syntactic processing can be achieved in L2. On the other hand, this effect was not found for all L2 learners (Hahne, 2001; Hahne & Friederici, 2001; Sabourin, 2003). Early negativities had not been reported in L2 (Hahne, 2001) until very recently (Hahne et al., 2006; Rossi, Gugler, Friederici & Hahne, 2006). Note, nonetheless, that these negativities have not been consistently reported in monolinguals and their interpretation is still in question (Müller & Hagoort, 2006). Moreover, recent ERP studies have put forward the hypothesis that differences between native and L2 syntactic processing are attributable to proficiency (Rossi et al., 2006) as well as to L1/L2 similarity (Osterhout, McLaughlin, Pitkänen, Frenck-Mestre & Molinaro, 2006; Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008; Tokowicz & MacWhinney, 2005).

Two recent ERP studies have revealed that native-like syntactic processing can be observed in the L2 provided a high level of proficiency as well as sufficient regularity of the grammatical rule in question. Rossi et al. (2006) presented sentences containing word category violations, morphosyntactic agreement violations or both types of

violations to high- and low-proficiency German and Italian L2 learners of Italian and German, respectively. The results for high-proficiency learners were similar to those found for native speakers (albeit with some differences in amplitude): an ELAN (early left anterior negativity) and a P600 effect for word category violations and a LAN and a P600 effect for morphosyntactic violations. In contrast, low-proficiency learners did not show any LAN effect and displayed a delayed P600 effect. The authors concluded that late L2 learners who achieve high proficiency can process language similarly to native speakers given sufficient exposure to the L2, and they suggested that Friederici's (2002) three-phase model could be applied to L2 language processing. In like manner, Hahne et al. (2006) concluded that native-like processing of inflectional morphology is attainable in the L2, with the caveat that it is constrained by the regularity (and thus learnability) of the system.

Other recent ERP studies have underlined the role of the similarity of the L1 regarding both the learning rate of grammatical features in the L2 (Osterhout et al., 2006; Tokowicz & MacWhinney, 2005) and the capacity to process knowledge online (Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008). Of particular interest for present purposes are three of these studies, which all examined the processing of grammatical gender. Sabourin and Haverkort (2003) studied German–Dutch learners; Sabourin and Stowe (2008) compared Romance and German learners of Dutch; and Tokowicz and MacWhinney (2005) examined English–Spanish learners.

In their first study, Sabourin and Haverkort (2003) investigated gender processing in L2 using determiner–noun and adjective–noun agreement violations. Participants were Dutch native speakers and German L1–Dutch L2 proficient learners. Their results showed a delayed P600 effect in sentence contexts for L2 German as compared to native Dutch speakers when reading a critical noun that did not match in gender with the immediately preceding definite determiner. Gender agreement errors between Dutch adjectives and nouns elicited a P600-like effect in the group of native speakers, but not in the group of proficient L2 learners, where basically no effect of gender agreement was found. It is noteworthy that gender agreement errors between adjectives and nouns also posed far greater problems for the German–Dutch learners offline, in a grammaticality judgment task. The authors concluded that L2 processing can reach native-like levels when syntactic constructions are similar in both languages but not when they differ in L1 and L2. It is important to note, however, that Sabourin and Haverkort (2003) compared overt marking of agreement on the definite determiner (where gender was neutralized on the pronominal adjective) to abstract gender for indefinite determiners where gender is only marked on the adjective (to illustrate, compare *Het kleine kind en de kleine tafel*

“the small child and the small table” to *Een klein kind en een kleine tafel* “a small child and a small table”). Thus different agreement rules were involved. These rules are similar in Dutch and German (*Das kleine Kind und der kleine Tisch* “the small child and the small table” vs. *Ein kleines Kind und ein kleiner Tisch* “a small child and a small table”). However, despite the similarity across languages, German–Dutch learners did not seem to be able to process agreement quickly enough to produce online effects for structures involving indefinite determiners.

In a second study, Sabourin and Stowe (2008) compared the performance of Romance and German learners for the definite determiner condition in Dutch (e.g. *Het<sub>neu</sub>\*De<sub>com</sub> kleine kind<sub>com</sub>* “the small child”). They observed a P600 effect for the German group and an uncharted frontal negativity in the group of Romance learners. The authors concluded that automatic gender processing in L2 not only depends on the presence of a grammatical gender system in the L1 but also requires overlapping of lexical gender. Unlike the commonality of agreement rules involved in Sabourin and Haverkort (2003), here the rules were either the same or differed across the L1 and L2. For German participants, agreement rules within the DP were the same in their L1 and L2, whereas for Romance speakers this was not the case, as not only the determiner but adjective as well must agree in gender with the noun in Romance languages (e.g., *la petite table* vs. *le petit enfant* “the small table” vs. “the small child”). Hence, differences in performance between the German and the Romance group may also have been due to cross-linguistic differences in agreement rules. Furthermore, conclusions from the results should be drawn with caution because of the variability between the groups of learners regarding *N* size (*N* = 8 Romance vs. 14 German learners).

Tokowicz and MacWhinney (2005) used ERPs to examine English–Spanish learners’ sensitivity to gender agreement violations between the determiner and the noun in visually presented Spanish sentences. Agreement violations provoked a P600 effect, showing that learners were indeed sensitive to gender agreement in their L2. In accordance with the ‘competition model’, these authors suggested that features that are not present in L1 (here, grammatical gender for native English speakers) should in fact be acquired faster than those that are in conflict (or ‘competition’) with L2 parameters (e.g., number agreement in English vs. Spanish). This hypothesis was corroborated by an absence of effect for nominal number agreement (largely absent in English) for these learners, but a small P600 effect to gender concord violations

The difference in results regarding the sensitivity to gender agreement in L2 in Sabourin and colleagues (2003, 2008) and Tokowicz and MacWhinney (2005) may be due to the regularity of gender marking in the languages involved in the studies. In Spanish, masculine

and feminine forms of the determiner are distinct, both in the singular and plural (*el/la, los/las*), whereas in Dutch the determiner is common for all genders in the plural. The consistency of gender marking in Spanish may have facilitated the online detection of gender agreement violations in this language compared to Dutch. Furthermore, gender agreement in Dutch has been shown to be learned late by native speakers and problematic for L2 learners, even early ones (Blom, Polisenska & Unsworth, 2008). Nevertheless, the results of Sabourin’s studies show a seemingly greater effect of the L1 on L2 processing than does that by Tokowicz and MacWhinney. It is important to underline the difference in proficiency across studies; participants were instructed beginners at university (Tokowicz & MacWhinney, 2005) versus three-year residents of the Netherlands whose proficiency was not independently tested (Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008).

### The present study

In the present series of experiments we aimed to answer two questions regarding gender processing; first, whether gender is processed in a similar manner by L2 learners and native L1 controls; and second, the extent to which such processing is constrained by similarities across the learners’ L1 and L2 regarding grammatical gender. To investigate these questions we compared German native speakers who were advanced learners of French to native French speakers. In three experiments, we manipulated gender agreement violations within the DP, between the determiner and the noun (Experiment 1), the postposed adjective and the noun (Experiment 2) and the preposed adjective and the noun (Experiment 3).

The choice of population was made according to the grammatical properties of the native language in relation to the second language. We selected German L2 learners because German and French both have grammatical gender systems, but the systems differ in respect to the number of genders, adjective position relative to the noun and agreement of elements within the DP. French has a two-gender system (i.e., masculine and feminine) and determiner selection depends on gender and number only; in contrast, German has a three-gender system (i.e., masculine, feminine and neuter) and determiner selection depends on gender, number and case (i.e., nominative, accusative, dative and genitive). Regarding adjective position, two word orders exist in French (preposed and postposed) and follow the same agreement rules, whereas only one is found in German (preposed). In the present study, we manipulated gender agreement between the noun and the DEFINITE ARTICLE and the noun and the adjective (postposed and preposed) in the NOMINATIVE CASE; hence, it is important to note the difference of agreement rules for this particular type of DP (i.e., definite article, nominative case) across languages. In both

languages, the definite article is distinct for masculine and feminine in the singular (French: *le<sub>masc</sub> ballon* “the ball” / *la<sub>fem</sub> fleur* “the flower”; German: *der<sub>masc</sub> Tisch* “the table” / *die<sub>fem</sub> Tür* “the door”) but not in the plural (French: *les ballons* / *les fleurs*; German: *die Tische* / *die Türen*). For adjective agreement, however, the two languages differ. In French, attributive adjectives must agree in gender with the noun, independent of position (whether preposed: *le petit ballon* “the small ball” / *la petite fleur* “the small flower”; or postposed: *le ballon vert* “the green ball” / *la fleur verte* “the green flower”) and number (*les petits ballons* / *les petites fleurs*).<sup>1</sup> In German, attributive adjectives are invariable for singular nominative case (*der kleine Tisch* “the small table” / *die kleine Tür* “the small door”) and moreover gender is neutralized on both the determiner and adjective in the plural (*die kleine Tische* / *die kleine Türen*). Thus, the interesting question for the present study is whether L2 learners process gender independently in their L2 or whether they apply the system of their L1 to their L2. This interference could be lexical and/or ruled-based. In other words, German speakers may assign the gender of German nouns to French nouns even if gender is not shared across languages. For example, they could produce *\*le table* instead of *la table* since the noun “table” is feminine in French but masculine in German (French: *la<sub>fem</sub> table<sub>fem</sub>*; German: *der<sub>masc</sub> Tisch<sub>masc</sub>* “the table”). On the other hand, even if the correct gender of a noun is learned in the L2, learning the system of agreement may prove difficult. Hence, if L2 French (L1 German) speakers do apply their L1 gender system, gender processing in French will be hampered in the instance the systems diverge, and the effects that emerge in the German group (if any) should differ from those of native speakers. The three experiments reported below addressed these questions.

Experiment 1 examined gender agreement between the definite article and the noun in the singular so that gender is marked on the determiner. In Experiments 2 and 3 we used the plural so that gender was marked only on the adjective. Such a manipulation allows direct comparison with Sabourin and Haverkort’s (2003) study

<sup>1</sup> In French, gender agreement on the adjective can be phonologically realized (*le<sub>masc</sub> petit<sub>masc</sub> [pəti] livre<sub>masc</sub>* “the small book”; *la<sub>fem</sub> petite<sub>fem</sub> [pəti] table<sub>fem</sub>* “the small table”), phonologically silent (*le<sub>masc</sub> livre<sub>masc</sub> bleu<sub>masc</sub> [blø]* “the blue book”; *la<sub>fem</sub> table<sub>fem</sub> bleu<sub>fem</sub> [blø]* “the blue table”) or unmarked (*le<sub>masc</sub> livre<sub>masc</sub> rouge<sub>masc</sub> [ʁuʒ]* “the red book”; *la<sub>fem</sub> table<sub>fem</sub> rouge<sub>fem</sub>* “the red table”). In the present study, all the experimental adjectives were phonologically realized. Regarding the plural, it is marked by the addition of a final “s” at the end of the adjective (*les<sub>masc</sub> petits<sub>masc</sub> [pəti] livres<sub>masc</sub>* “the small books”; *les<sub>fem</sub> petites<sub>fem</sub> [pəti] tables<sub>fem</sub>* “the small tables”). The “s” is phonologically silent unless followed by a vowel (*les petits [pətiʒ] avions* “the small planes”). Some adjectives are modified in the plural but they are grammatical exceptions (e.g., *le livre médical* [medikal] “the medical book”; *les livres médicaux* [mediko] “the medical books”).

in which gender was also marked only on the adjective. Moreover, had we not performed this manipulation, the violation would have occurred between the determiner and the adjective (*le<sub>masc</sub> petit<sub>masc</sub>/\*petite<sub>fem</sub> livre<sub>masc</sub>* “the small book”), rather than only between the adjective and the noun (*les petits<sub>masc</sub>/\*petites<sub>fem</sub> livres<sub>masc</sub>*). The present study investigated gender processing in L2 when agreement rules are similar in L1 and L2 (Experiment 1) and when rules differ across languages (Experiments 2 and 3), using the same type of DP across experiments.

## Experiment 1

In Experiment 1 we manipulated gender agreement between the determiner and noun in sentence contexts with French native speakers and German–French L2 learners. Based on the literature, we expected gender agreement errors to elicit a P600 effect in our group of native French speakers (Foucart & Frenck-Mestre, 2006; Frenck-Mestre, 2005). It is also possible that a LAN could precede the P600 effect, in line with the results of some studies (Barber & Carreiras, 2005; Deutsch & Bentin, 2001; Gunter et al., 2000; but see Hagoort & Brown, 1999; Osterhout, Bersick & McKinnon, 1997). In the L2 learner group we also expected to find a P600 effect elicited by these gender agreement violations, given the results reported by Sabourin and Haverkort (2003) as well as Sabourin & Stowe (2008) for German–Dutch bilinguals. Whether the effect in the bilingual group would be delayed and/or of different magnitude compared to native French speakers is an open question. It is worth noting that German and French are farther removed from each other than are German and Dutch, hence it is possible that our bilingual participants could demonstrate less-efficient second language processing than those studied by Sabourin and colleagues (2003, 2008). Sabourin and Stowe (2008) also examined Romance learners of Dutch and found a different pattern of ERP effects for these learners compared to native speakers and German learners. However, these results have to be considered cautiously because the study involved a small number of participants from various Romance languages. Last, the concomitant presence of a LAN effect in the bilingual group vs. the absence of such an effect could provide useful information depending upon whether the effect was found in the native participant group (Hahne, 2001; Hahne & Friederici, 2001; but see Foucart & Frenck-Mestre, 2006; Frenck-Mestre, 2005; Osterhout et al., 2004; Osterhout et al., 2006).

## Method

### Participants

Sixteen French native speakers and 16 German–French learners received 20 euros for their participation. The

mean age of all participants was 22.3 years, ranging from 20 to 27. They had normal or corrected-to-normal vision. All were students at the University of Provence. The L2 learners were Erasmus students; they had all studied French at school (mean 9.3 years) and passed the required exam to attend courses in a French university (Diplôme d'Etudes de Langue Française, DELF; individual results not available). After the experiment, they were asked to complete an offline test which consisted of circling the correct gender-marked article of the words presented during the experiment (mean errors: 4.8%; *SD* 3.6). They were also asked to self-rate their level of French on a scale from 1 to 6 (1 = very poor; 6 = excellent) for different aspects of language comprehension and production (written comprehension, 4.7; oral comprehension, 4.5; written production, 4; oral production, 4). Hence, based on both number of years of study and successful completion of the DELF (level 2 or better), participants' proficiency level was considered to be advanced.

### Materials

Ninety-six concrete French nouns, 48 masculine and 48 feminine, of low to medium frequency (frequency mean per million: 35.2, *Lexique 2*; New, Pallier, Brysbaert & Ferrand, 2004) and in length from 3 to 8 letters (mean 5.8) served as target words. These stimuli were selected such that: (1) all stimuli had a single unambiguous translation in German; (2) no homographs or cognates across languages were included; (3) all stimuli had a consonant as initial phoneme in French; (4) half (48) of the target words shared the same gender in French and German (e.g., French, *la montre*<sub>fem</sub>; German, *die Uhr*<sub>fem</sub> "the watch") and half (48) had opposite gender across languages (e.g., French, *la clef*<sub>fem</sub>; German, *der Schlüssel*<sub>masc</sub> "the key"). For words that had opposite gender, half (24) were masculine in French (feminine in German) and half (24) were feminine in French (masculine in German). For words that shared the same gender across languages, half (24) were masculine and half (24) were feminine. Printed frequency and length were matched across masculine and feminine gender and across same versus opposite gender in the two languages.

The target nouns were embedded in short sentence contexts. All sentences were simple declarative structures and followed the same pattern: adverbial phrase, definite determiner, critical noun, copula, complement. All critical nouns were seen in sentence pairs which were identical with the exception of the gender of the definite determiner preceding the critical noun, which either agreed in gender with the subsequent target noun or violated gender agreement. Two lists were created such that only one member of a sentence pair was seen per list. In each list, there were 24 sentences in each of 4 conditions defined by the factors Language Coherency of the critical noun (same vs. opposite gender in French and German)

Table 1. Example of all the conditions presented in Experiment 1.

| Condition                        | Example   |
|----------------------------------|---|
| Opposite gender across languages |   |
| <b>Correct</b>                   | <i>Hier la</i> <sub>fem</sub> <i>chaise</i> <sub>fem</sub> <i>était dans le salon.</i>  |
| <b>Incorrect</b>                 | <i>Hier *le</i> <sub>masc</sub> <i>chaise</i> <sub>fem</sub> <i>était dans le salon.</i><br>"Yesterday the <sub>fem/masc</sub> * <i>chair</i> <sub>fem</sub> was in the living room." |
| Same gender across languages     |   |
| <b>Correct</b>                   | <i>Hier le</i> <sub>masc</sub> <i>balai</i> <sub>masc</sub> <i>était dans la cave.</i>  |
| <b>Incorrect</b>                 | <i>Hier *la</i> <sub>fem</sub> <i>balai</i> <sub>masc</sub> <i>était dans la cave.</i><br>"Yesterday the <sub>masc/fem</sub> * <i>broom</i> <sub>masc</sub> was in the cellar."       |

and Agreement between the determiner and critical noun (gender agreement vs. violation). In addition, 60 syntactically correct filler sentences of varying structures were presented. The sentences were presented in a fixed-random order, with the restriction that the same condition was not immediately repeated. Two fixed-random orders were created per list. Each list began with four warm-up sentences. Each participant saw only one list. Example sentences illustrating the four conditions are presented in Table 1.

### Procedure

Participants were seated comfortably in a dimly lit, sound attenuated, electrically shielded room during recording. They were requested not to move any part of their body or to make eye-movements outside the rest periods, and to blink only when instructed to do so by a visual prompt on a CRT screen placed 60 cm in front of them. The prompt followed each sentence. Three short rest periods were provided at regular intervals during the experiment. Sentences were presented visually at a rate of 650 ms per word (500 ms presentation followed by 150 ms blank screen) in the centre of the screen. Following the last word of each sentence, a "yes/no" prompt was presented and participants were requested to respond whether they considered the sentence SEMANTICALLY acceptable or not (independently of its grammaticality). Responses to the questions were recorded. The entire session (including placement of the electrode cap) lasted roughly two hours.

### EEG recording

EEG activity was recorded continuously from 13 scalp locations, using tin electrodes attached to an elastic cap (Electrocap International). Scalp sites included standard International 10–20 locations (Jasper, 1958) over frontal, central, temporal, posterior temporal, parietal and occipital areas (F7, F8, Fz, Cz, Pz, T5, T6, O1, O2) of the

left and right hemispheres. In addition, electrodes were placed centrally between homologous central (FC5, FC6) and parietal sites (CP5, CP6). Horizontal eye-movements were monitored by means of an electrode placed at the outer canthus of the left eye while blinks and vertical eye-movements were monitored via an electrode beneath the right eye. All electrodes were referenced to the left mastoid. An electrode was placed over the right mastoid to ascertain whether any effects of experimental variables were visible on the mastoid recordings (none were found). The EEG was amplified with a bandpass of 0.1–40 Hz (3dB cut-off) by means of an SAI Bioamp 32 channel Model and was digitized online at 200 Hz. EEGs were later filtered below 15 Hz. The electrode impedance threshold value was set to 3 k $\Omega$  for scalp electrodes and 10 k $\Omega$  for face electrodes. Epochs began 100 ms prior to stimulus onset and continued 1100 ms thereafter. Average ERPs were formed offline for trials free of muscular and/or ocular artefacts and amplifier blocking (artefact rejection was performed by a computerized routine).

#### Data analysis

The ERP data were quantified by calculating the mean voltage amplitudes and peak latencies (relative to a 100 ms pre-stimulus baseline), for four time windows: 50–150, 150–300, 300–500 and 500–800 ms post presentation of the critical noun. These windows were selected based on prior studies of visual processing of linguistic stimuli, and roughly correspond to the temporal windows associated with the N1, P2 and/or ELAN, N400 and/or LAN, and P600 components that are frequently observed in these studies. The main component of interest, based on prior ERP studies of grammatical gender, was the P600 (defined as the mean positive amplitude 500–800 ms post stimulus), although analyses were performed on the previous time windows to ascertain whether any earlier differences occurred between sentence conditions. Prior to analyses, trials with artefact were rejected (French: 4.2% and 3.2%, German: 8.4% and 6.8% for correct and incorrect conditions, respectively and no significant difference emerged between groups,  $F < 1$ ). Analyses were conducted on the data acquired at midline and lateral sites. At both, repeated measures ANOVAs were performed with two levels of Agreement (agreement vs. violation), two levels of Language Coherency (same vs. opposite gender of critical noun in French and German), and either three levels of Electrode at midline (frontal, central and posterior) or four levels of electrode at lateral sites for each hemisphere (F7/8, FC5/6, CP5/6, T5/6). The Greenhouse-Geisser (1959) correction was applied to all repeated measures with more than one degree of freedom. All significant differences involving more than two conditions were confirmed by post-hoc comparisons.

## Results

### Behavioural data

The total number of positive responses at sentence end was calculated for all sentences. Even though participants were asked to assess the semantic acceptability of sentences independently of their grammaticality, their responses were apparently affected by syntactic violations. For the French participants, there was an effect of Agreement, whereby participants rejected sentences that contained a gender agreement error more frequently than sentences that did not contain such errors (35% vs. 15% rejection rate, respectively;  $F(1,15) = 6.66, p < .02$ ). German–French learners also responded negatively more often to sentences that contained a gender agreement error than to sentences that contained no error (30% vs. 23% negative responses;  $F(1,14) = 4.32, p < .05$ ).

### ERP data

No significant differences emerged as a function of experimental factors prior to the 500–800 ms window after the target word. Beginning at roughly 500 ms post presentation of the critical noun, a widely distributed positive deflection, which peaked at approximately 700 ms, was observed for nouns that did not agree in gender with the immediately previous definite determiner as compared to those that did. This P600 effect for Agreement was confirmed at midline ( $F(1,30) = 22.92, p < .001$ ) and tended to be modified by an interaction with Group and Electrode ( $F(2,60) = 2.39, p = .10$ ). At lateral sites, the effect of Agreement ( $F(1,30) = 14.35, p < .001$ ) was modified by an interaction with Hemisphere and Language Coherency ( $F(1,30) = 10.61, p < .003$ ) as well as by the higher-order interaction involving Group, Hemisphere and Language Coherency ( $F(1,30) = 5.07, p < .03$ ). Independent analyses were subsequently performed on the two participant groups.

*French native speakers:* The P600 effect elicited by Agreement violations was significant at midline ( $F(1,15) = 8.85, p < .009$ ) and varied as a function of electrode ( $F(2,30) = 3.97, p < .03$ ), being slightly larger over posterior than anterior sites, as can be seen in Figures 1 and 2. At lateral sites, the P600 effect for Agreement was marginally significant ( $F(1,15) = 3.50, p = .08$ ) and tended to vary as a function of Electrode ( $F(3,45) = 2.76, p < .05$ ); differences were significant at mid-lateral sites (FC5, CP5, FC6, CP6) but not at anterior or posterior lateral sites (F7, F8, T5, T6). These effects are depicted in Figures 1 and 2 for nouns sharing the same gender across French and German and for nouns with opposite gender, respectively. No differences were found as a function of the factor Language Coherency, nor did this factor interact with Agreement.

*German–French L2 learners:* The P600 effect elicited by Agreement violations was highly significant at both

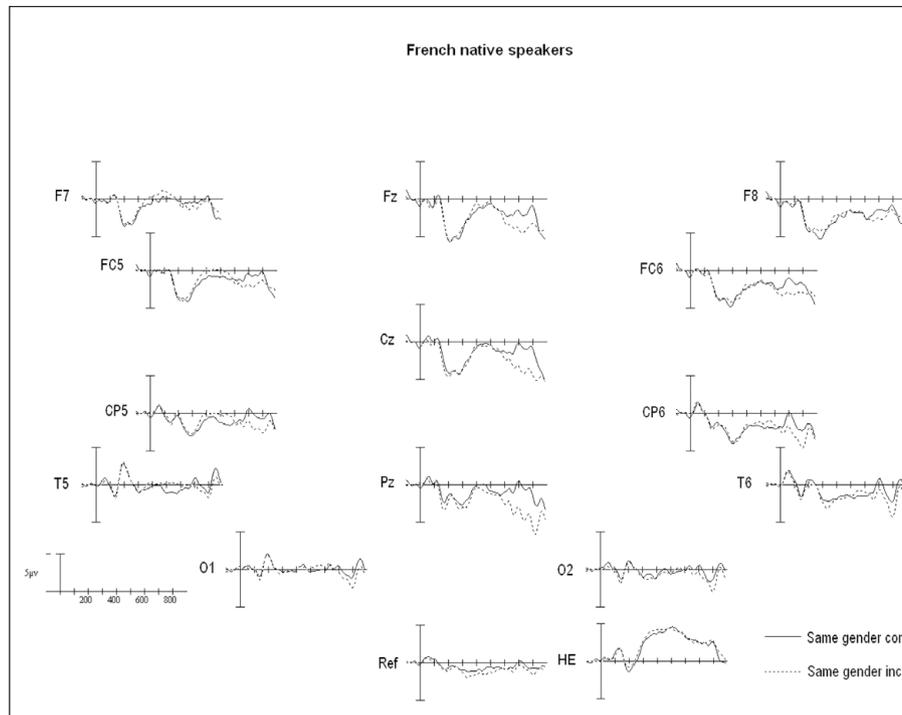


Figure 1. P600 effect for French native speakers in the case of gender violations between determiners and nouns of the same gender across languages.

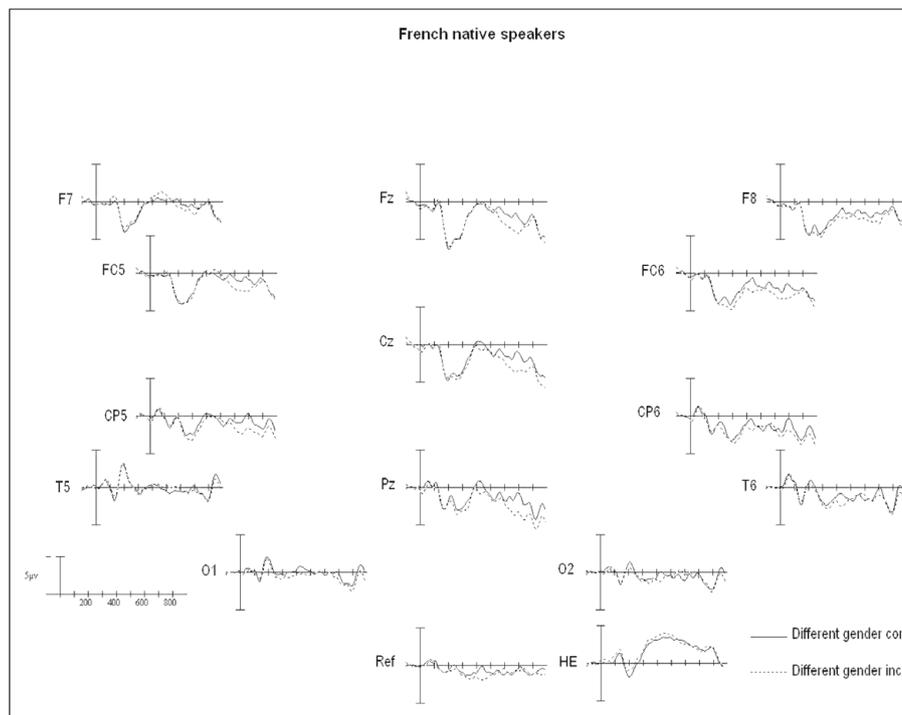


Figure 2. P600 effect for French native speakers in the case of gender violations between determiners and nouns of different gender across languages.

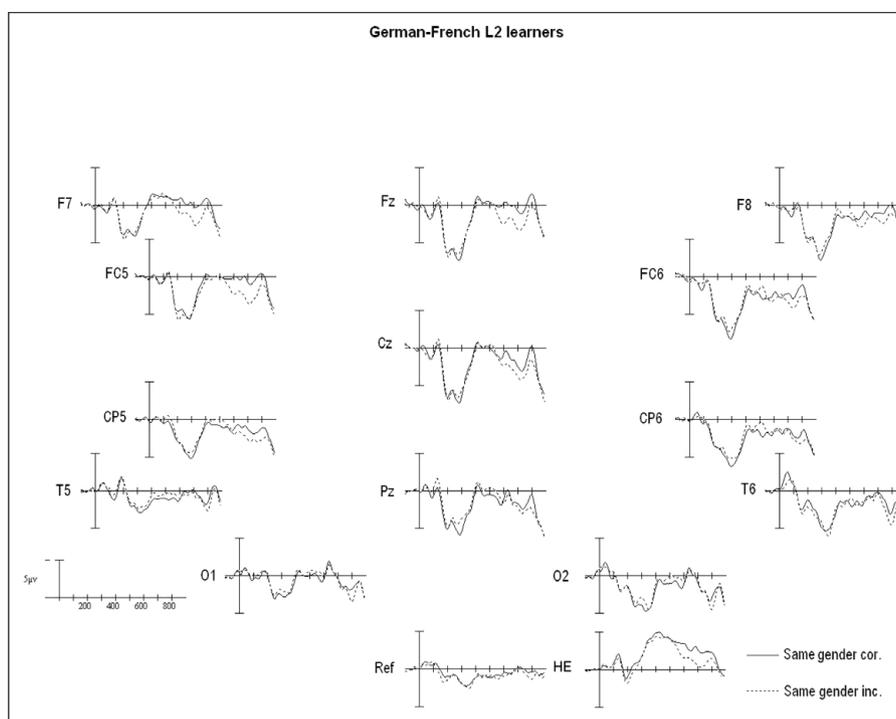


Figure 3. P600 effect for all German–French bilinguals in the case of gender violations between determiners and nouns of the same gender across languages.

midline ( $F(1,15) = 11.99, p < .003$ ) and lateral sites ( $F(1,15) = 16.65, p < .001$ ), as can be seen in Figures 3 and 4, for nouns sharing the same gender across French and German and for nouns with opposite gender, respectively. No other main effects were observed. At lateral sites, a significant interaction obtained between Agreement, Language Coherency and Hemisphere ( $F(1,15) = 7.53, p < .02$ ); post-hoc comparisons (Newman-Keuls) revealed that for nouns that had the same gender across German and French, agreement errors elicited a significant P600 effect larger on the right ( $p < .001$ ) than on the left hemisphere ( $p < .08$ ), while for nouns that had opposite gender across languages, agreement errors provoked a significant P600 effect larger on the left ( $p < .04$ ) than on the right hemisphere ( $p > .25$ ). No ready explanation of this hemispheric difference is available. No other interactions were observed.

The results for the German–French bilingual participants revealed apparently similar P600 effects for gender agreement errors independent of whether the French target noun shared the same gender in German or not. Nonetheless, inspection of the ERP data showed a larger P600 effect elicited by agreement errors, in particular at midline, for nouns with same gender across the native and second language of learners in comparison to nouns with opposite gender across languages (cf. Figures 3 and 4). In line with previous monolingual and bilingual studies (Inoue & Osterhout, 2005; Osterhout

et al., 2006; Osterhout & Mobley, 1995) we used two different indicators to determine whether our group of L2 learners included subgroups. On the one hand, the size of the P600 effect elicited by gender agreement violations was calculated at midline per participant, for nouns that differed in gender across French and German. Two subgroups of eight participants could be established based on the median split. On the other hand, sensitivity to gender agreement errors was determined based on behavioural responses. In the group that was above the median regarding the P600 effect, the majority of participants (7 out of 8) also showed a greater tendency to respond negatively to sentences that contained a gender agreement error than to those that did not (34% vs. 20% negative responses, respectively). In the group that was below the median, no such trend was apparent (27% vs. 27%, for the two sentence types, respectively).

Based on the above two criterion, ERP data for the L2 learners were reanalyzed at midline as a function of Group, Agreement, Language Coherency and Electrode. The reanalysis revealed a main effect of Agreement ( $F(1,14) = 22.08, p < .001$ ), which was modified by Group ( $F(1,14) = 13.66, p < .002$ ) as well as by the higher-order interaction between Agreement, Group and Language Coherency ( $F(1,14) = 5.617, p < .03$ ). The independent analysis of data for the two subgroups revealed the following. In the subgroup that was behaviourally sensitive to gender agreement errors and

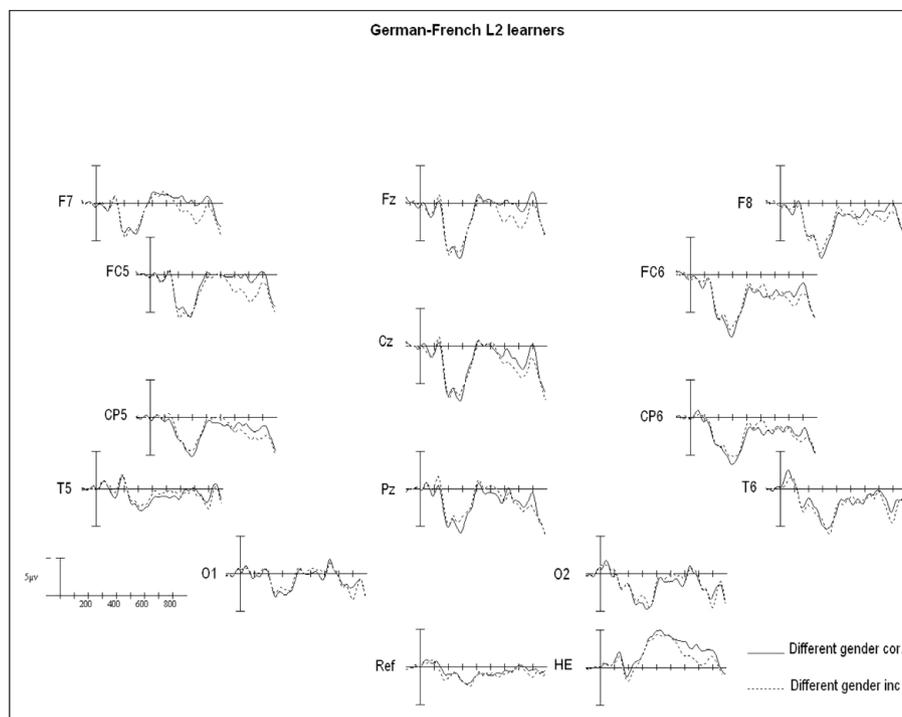


Figure 4. P600 effect for all German–French bilinguals in the case of gender violations between determiners and nouns of different gender across languages.

showed a larger P600 effect, there was a main effect of Agreement ( $F(1,7) = 32.55, p < .001$ ) which did not interact with Language Coherency ( $F < 1$ ). In this group, gender agreement errors on the critical noun elicited a large P600 effect, and this was true both when the noun had opposite gender in German and French and when it had the same gender across languages (Figure 5). The other subgroup of learners did not show a main effect of Agreement but revealed an interaction between Agreement and Language Coherency ( $F(1,7) = 6.01, p < .04$ ). Post-hoc analyses (Scheffé) showed that a P600 effect was elicited by Agreement violations for nouns that had the same gender across languages but not for nouns that had opposite gender in French and German (Figure 6).

In addition, we directly compared the effect found in native speakers and in the L2 group that showed a P600 effect independently of gender consistency across languages. At midline there was a main effect of Agreement ( $F(1,22) = 28.5, p < .001$ ) which was modified by the interaction with Group  $\times$  Electrode due to a larger effect at anterior sites for the L2 group than for native speakers.

### Discussion

The results showed that for both French native speakers and German L1–French L2 learners a P600 effect

was elicited by gender agreement errors between a target noun and a preceding determiner in sentence contexts. Overall, we found that the gender agreement effect was independent of the overlap of lexical gender across languages, which is to be expected in the group of French native speakers, but is somewhat surprising in the group of German–French bilinguals. However, behavioural responses provided during the ERP experiment showed that some of our learners were more sensitive to the determiner–noun gender mismatch than others. Furthermore, the inspection of individual ERP data revealed that there were in fact two sub-populations among our German–French participants. Whereas all the learners showed sensitivity to gender agreement for nouns that had the same gender in French and German, only a subset of these participants showed sensitivity to gender violations for nouns that had opposite gender across their native and the later-learned language.

Sabourin and Haverkort (2003) and Sabourin and Stowe (2008) put forward the hypothesis that although second language learners are able to learn new information and incorporate it at a lexical level, they may not attain syntactic competence. That is, although a learner may learn the grammatical gender of a noun in his/her non-native language, be able to correctly classify it offline and moreover show sensitivity to gender agreement online in a task that taps into lexical processing, the same sensitivity will not necessarily be apparent during

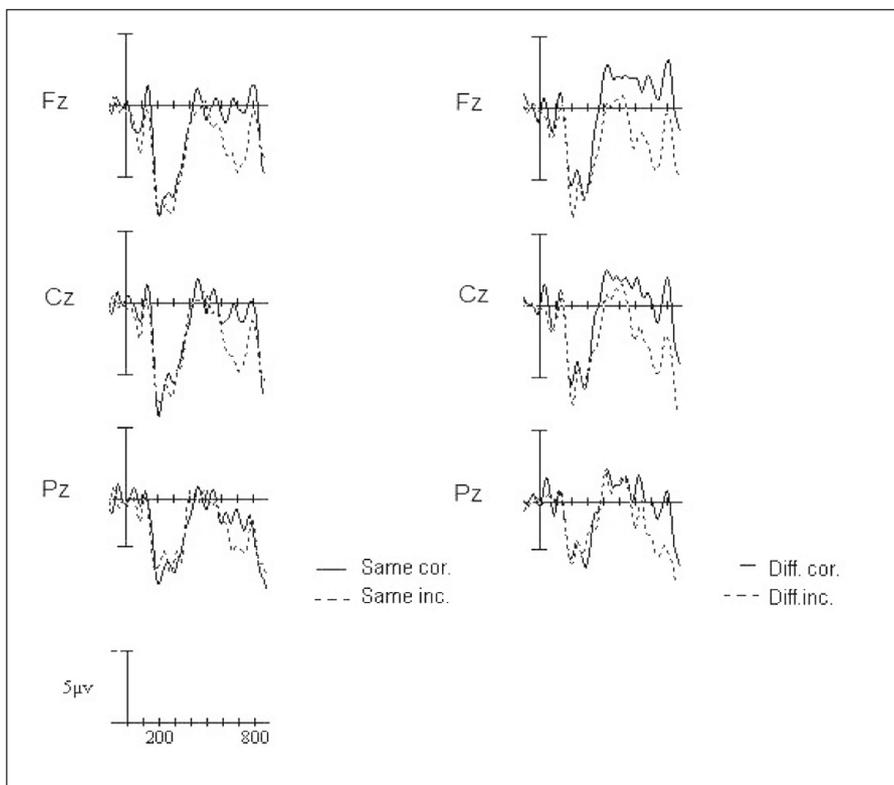


Figure 5. P600 effect for German–French bilinguals (subgroup A) in the case of gender violations between determiners and nouns of both the same and different gender across languages.

online syntactic processing (see Hopp, 2007, for further supporting evidence). This claim was made on the basis of ERP data from advanced German–Dutch learners, who showed a significant P600 effect to gender agreement errors between determiners and nouns in short Dutch sentences but failed to show a significant effect to gender agreement errors between nouns and adjectives. Our results do not support Sabourin and colleagues' proposal as they clearly show that one can find ERP effects in the L2 that are highly similar in amplitude and latency to that of native speakers. Note that, as outlined by Herschensohn (2006, 2007), several authors have shown differences in difficulty with adjective and determiner agreement as well as a dissociation between the acquisition of number and gender marking. As Herschensohn points out, the differences in rate of acquisition both within a domain (e.g., gender agreement for different elements within the DP, number vs. gender agreement within the DP, etc.) and across domains (e.g., nominal vs. verbal) clearly demonstrate that there is no simple answer to the question of the critical period hypothesis and acquisition of parameterized functional features. Moreover, as our own results demonstrate, even within a seemingly homogeneous population there can be different learning rates for the same grammatical phenomenon (see also Dewaele & Véronique, 2001; Osterhout et al., 2006).

In the next experiments, we manipulated gender violations between nouns and adjectives to investigate whether our German–French learners were as sensitive to agreement violations between these elements as they were between determiners and nouns.

## Experiment 2

Linguistic studies of L2 acquisition and processing have shown that, as in L1 development, agreement between the article and the noun seems to be more accurate and more rapidly acquired than agreement between the adjective and the noun (Bartning, 2000; Bruhn de Garavito & White, 2002; Dewaele & Véronique, 2001; Grandfeldt, 2000). However, the acquisition process of the noun phrase seems to be different in L1 and L2, in part because of a transfer from the L2 learner's native language system (Grandfeldt, 2000; Parodi, Schwartz & Clahsen, 1997). To investigate whether our L2 learners were differentially sensitive to adjective–noun than to determiner–noun agreement online, we conducted a further experiment where gender agreement was manipulated between the noun and the postposed adjective. In French, the canonical order of adjectives is post-nominal due to verb raising, whereas in German it is pre-nominal (Bernstein, 1991; Laenzlinger, 2005). In line with previous studies examining gender

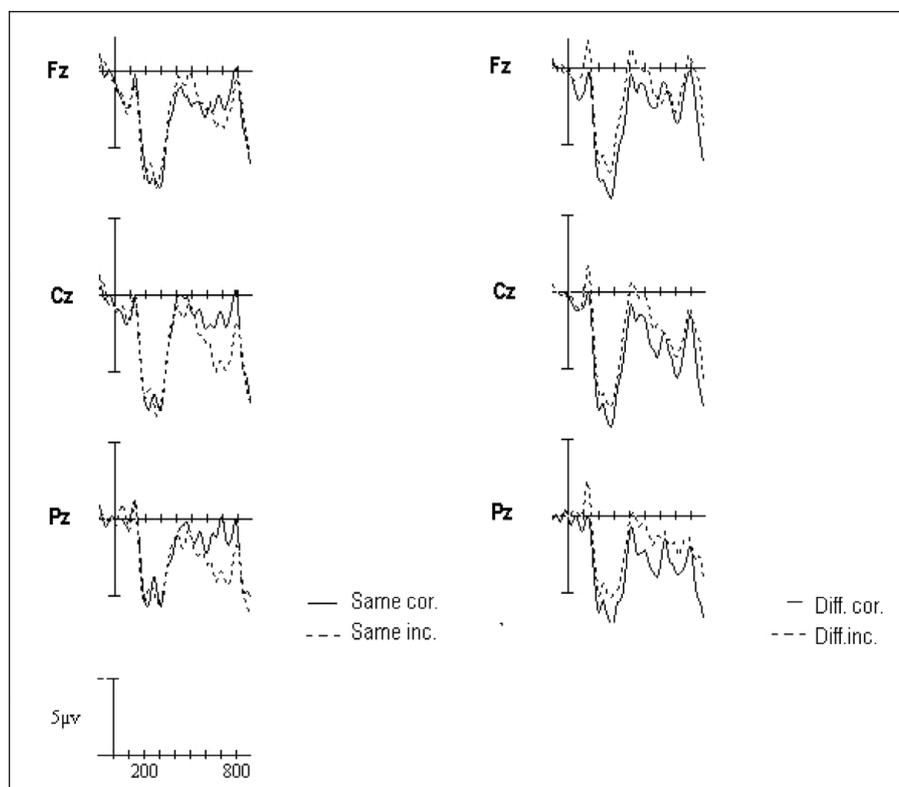


Figure 6. P600 effect for German–French bilinguals (subgroup B) in the case of gender violations between determiners and nouns of both the same and different gender across languages.

agreement violations within the NP, we expected a P600 effect for French native speakers, either preceded or not by an early negativity (Gunter et al., 2000; Hagoort & Brown, 1999). For German–French L2 learners, in view of the results of Experiment 1 as well as those of previous studies of L2 gender processing (Sabourin & Stowe, 2008; Tokowicz & MacWhinney, 2005) we expected to find a P600 effect elicited by gender agreement violations. However, if these learners experience greater difficulty processing adjective–noun than determiner–noun agreement online, as suggested in linguistic studies and by recent ERP studies (Sabourin & Haverkort, 2003), it is possible that no effect will be elicited by gender violations between the noun and the adjective.

### Method

#### Participants

Fourteen French native speakers and 14 German–French learners received 20 euros for their participation. The mean age of all participants was 22 years, ranging from 19 to 28. They had normal or corrected-to-normal vision. All were students at the University of Provence. The L2 learners were Erasmus students; they had all studied French at school (mean 8 years) and passed the required exam to attend courses in a French university (DELF,

individual results not available). After the experiment, they were asked to complete an offline test which consisted of circling the correct gender-marked article of the words presented during the experiment (mean errors 4.2%, *SD* 3.7). They were also asked to self-rate their level of French on a scale from 1 to 6 (1 = very poor; 6 = excellent) for different aspect of language (written comprehension, 4.8; oral comprehension, 4.8; written production, 3.8; oral production, 4.2). Hence, the proficiency level of participants was considered to be advanced, based on both number of years of study and successful completion of the DELF (level 2 or better).

#### Materials

Ninety-Six grammatical/ungrammatical sentence pairs were created. Grammaticality was determined by gender agreement between the noun and postposed adjective. The same 96 nouns used in Experiment 1 were used in the present experimental sentences. In addition to the 96 critical nouns, a set of 40 critical adjectives were selected (mean frequency per million: 42.5; mean length 5.9 letters, range between 4 and 8, *Lexique 2*; New et al., 2004). The critical adjectives were both orthographically and phonologically modified when inflected for the feminine (e.g., *vert*<sub>masc</sub> [vɛR] vs. *verte*<sub>fem</sub> [vɛRt] “green”). These 40 adjectives were paired with the 96 nouns, with each

Table 2. Example of all the conditions presented in Experiment 2.

| Condition                        | Examples  |
|----------------------------------|---|
| Opposite gender across languages |   |
| <b>Correct</b>                   | <i>En été, les chaises<sub>fem</sub> blanches<sub>fem</sub> sont dans le jardin.</i>  |
| <b>Incorrect</b>                 | <i>En été, les chaises<sub>fem</sub> blancs<sub>masc</sub> sont dans le jardin.*</i><br>“In summer, the <u>white<sub>fem/masc</sub></u> chairs <sub>fem</sub> are in the garden.” |
| Same gender across languages     |   |
| <b>Correct</b>                   | <i>Cette année, les montres<sub>fem</sub> blanches<sub>fem</sub> sont très tendance.</i>  |
| <b>Incorrect</b>                 | <i>Cette année, les montres<sub>fem</sub> blancs<sub>masc</sub> sont très tendance.*</i><br>“This year, <u>white<sub>fem/masc</sub></u> watches <sub>fem</sub> are very trendy.”  |

adjective presented between 1 and 6 times as determined by semantic fit between the adjective and noun. Each sentence comprised (an adverb), a plural definite article, noun, critical adjective, copula and a final complement (see Table 2). The plural form of the article was used so that no gender information was provided by this element (in French, the plural form of the definite determiner is identical for masculine and feminine words; e.g., *le<sub>masc</sub> livre<sub>masc</sub>* “the book”, *la<sub>fem</sub> table<sub>fem</sub>* “the table” → *les livres, les tables*). An additional 96 sentences were added as fillers. All fillers were syntactically correct. Half were semantically anomalous due to restrictions of the noun in relation to the adjective (e.g., *les vélos féroces* “the ferocious bikes”), while the other half were both semantically and syntactically correct. For these latter, invariable adjectives, which do not mark gender (*rouge<sub>fem/masc</sub>* “red”), were paired with a set of filler nouns that shared the same gender in French and German. The structure of filler sentences was similar to experimental sentences. For all sentences (experimental stimuli and fillers), the semantic acceptability was verified by French native speakers in an offline task prior to the main experiment. Two lists were created such that each noun–adjective pair was seen in both conditions (gender agreement vs. disagreement) but in only one condition for a given participant. In each, there were 24 sentences per condition, defined by Agreement (gender agreement between the noun and the critical adjective) and Language Coherency (nouns of same vs. opposite gender in French and German). The sentences were presented in a fixed-random order, with the restriction that no list begin with an ungrammatical sentence and no more than two ungrammatical sentences follow each other, there being

six fixed-random orders per list. Each list began with four training sentences. Each participant saw only one list.

### Procedure

Sentences were presented visually, at a rate of 650 ms (500 ms presentation followed by 150 ms blank screen) in a single block of 196 sentences. Following each sentence, a “yes/no” prompt was presented, and participants were requested to judge whether the sentence was correct or not (syntactically and/or semantically). Half of the participants made positive responses with the left hand; the other half used their right hand. Responses to the questions were recorded. Participants were seated comfortably in a dimly lit, sound attenuated, electrically shielded room during recording. They were requested not to move any part of their body or to make any eye-movements outside the rest periods. A short break was provided in the middle of the experiment.

### EEG Recording

This was identical to Experiment 1 with the exception that EEG activity was recorded continuously from 21 scalp locations. Scalp sites included standard International 10–20 locations (Jasper, 1958) over frontal, temporal, central, posterior temporal, parietal and occipital areas of the left and right hemispheres (FP1/2, F7/8, F3/4, C3/4, T5/6, P3/4, O1/2), as well as over midline (Fz, Cz, Pz). In addition, electrodes were placed centrally between homologous anterior and central sites (Fc5/6), central and parietal sites (Cp5/6). Average ERPs were formed offline from trials free of muscular and/or ocular behaviour and amplifier blocking (behavioural rejection was performed by a computerized routine and led to less than 6% of rejections per stimulus category overall). Averaging was performed without regard to behavioural responses.

### Data analysis

The ERP data were quantified post presentation of the critical adjective, for the same time windows as in Experiment 1. Prior to analyses, trials with artefact were rejected (French: 4.3% and 4.8%, German: 0.8% and 1.25% for correct and incorrect conditions, respectively and no significant difference emerged between groups,  $F < 1$ ). At midline, a three-way repeated measures ANOVA was performed, with two levels of Agreement (gender agreement vs. violation), two levels of Language Coherency (same vs. opposite gender in French and German) and three levels of Electrode (Fz, Cz and Pz). At lateral sites, four-way ANOVAs were performed with repeated measures on Agreement, Language Coherency, Hemisphere and Electrode, with three levels of electrode at anterior lateral sites (F7/8, F3/4, Fc5/6) and four levels of electrode at lateral, centro-parietal sites (C3/4, Cp5/6, P3/4 and T5/6). The factor Noun Gender (masculine vs. feminine words) was not included in the analyses as grand

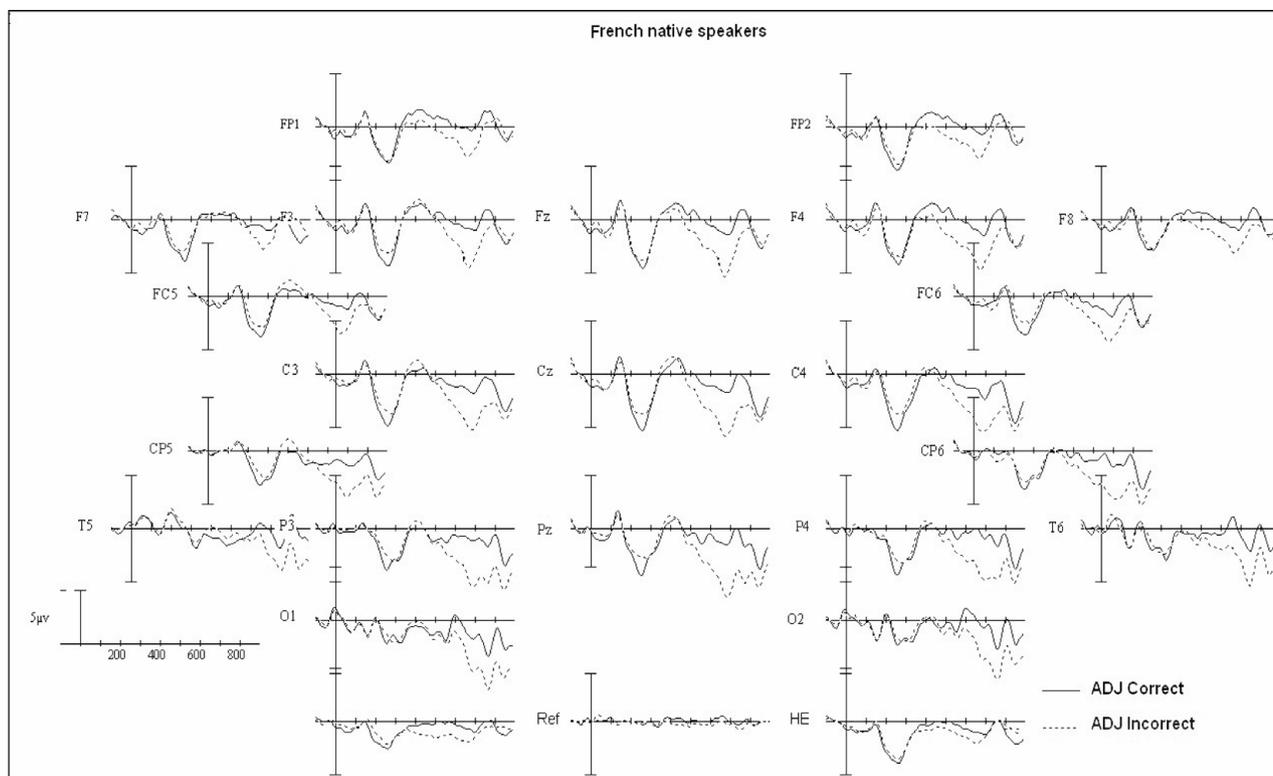


Figure 7. French native speakers: P600 effect when the noun and the postposed adjective disagree in gender.

averages revealed no differences for this factor ( $F < 1$ ). The Greenhouse-Geisser (1959) correction was applied to repeated measures with greater than one degree of freedom. All significant differences involving more than two conditions were confirmed by post-hoc comparisons.

### Results

The grand means revealed differences in the waveforms for adjectives that agreed in gender with the preceding noun as compared to those that did not. French native speakers showed a positive deflection in the waveform for sentences containing gender agreement violations, between 500–800 ms after the onset of the critical adjective, corresponding to a P600 effect (Figure 7). German–French learners showed a negative deflection in the waveforms in the 100–180 ms time window, corresponding to the N100 component, for sentences containing gender-agreement violations (Figure 8). These differences were confirmed in ANOVAs performed on the mean voltages obtained for each sentence condition as a function of time window and electrode sites.

In the N100 (80–180 ms) time window, no main effects were found, but there was a tendency at midline for the Group  $\times$  Agreement interaction ( $F(1,26) = 3.21, p = .08$ ), which reached significance at anterior ( $F(1,26) = 6.49, p < .02$ ) and centro-parietal ( $F(1,26) = 4.7, p < .04$ ) sites. Post-hoc comparisons revealed a negativity for French

L2 speakers but not for native speakers. No main effects or interactions were found in either the P200 (160–280 ms) or N400 (300–500 ms) time windows. In the P600 (500–800 ms) window, an effect of Agreement emerged at midline ( $F(1,26) = 4.97, p < .04$ ) and at centro-parietal sites ( $F(1,26) = 7.42, p < .01$ ), while a tendency was found at anterior lateral sites ( $F(1,26) = 3.7, p = .06$ ). The effect was modified by a significant interaction with Group at midline ( $F(2,26) = 4.23, p < .05$ ), anterior lateral ( $F(1,26) = 8.11, p < .008$ ) and centro-parietal ( $F(1,26) = 10.24, p < .004$ ) sites. Post-hoc comparisons (Scheffé) revealed a P600 effect for French native speakers but not for French L2 speakers. Given the interactions with Group, subsequent ANOVAs were performed on the data for each group independently.

*French native speakers:* No significant differences emerged as a function of experimental factors prior to the 500–800 ms window after the target word. In the 500–800 ms window, a significant effect of Agreement was found at midline ( $F(1,13) = 6.23, p < .03$ ) and at anterior ( $F(1,13) = 7.29, p < .02$ ) and centro-parietal ( $F(1,13) = 12.66, p < .004$ ) sites. Adjectives that disagreed in gender with the previous noun provoked a P600 effect (see Figure 7). As might be expected for our monolingual controls, no effect of Language Coherency was found ( $F < 1$ ), nor did this factor interact with Agreement ( $F < 1$ ).

*German–French L2 learners:* The results for the German–French learners (Figure 8) revealed only a

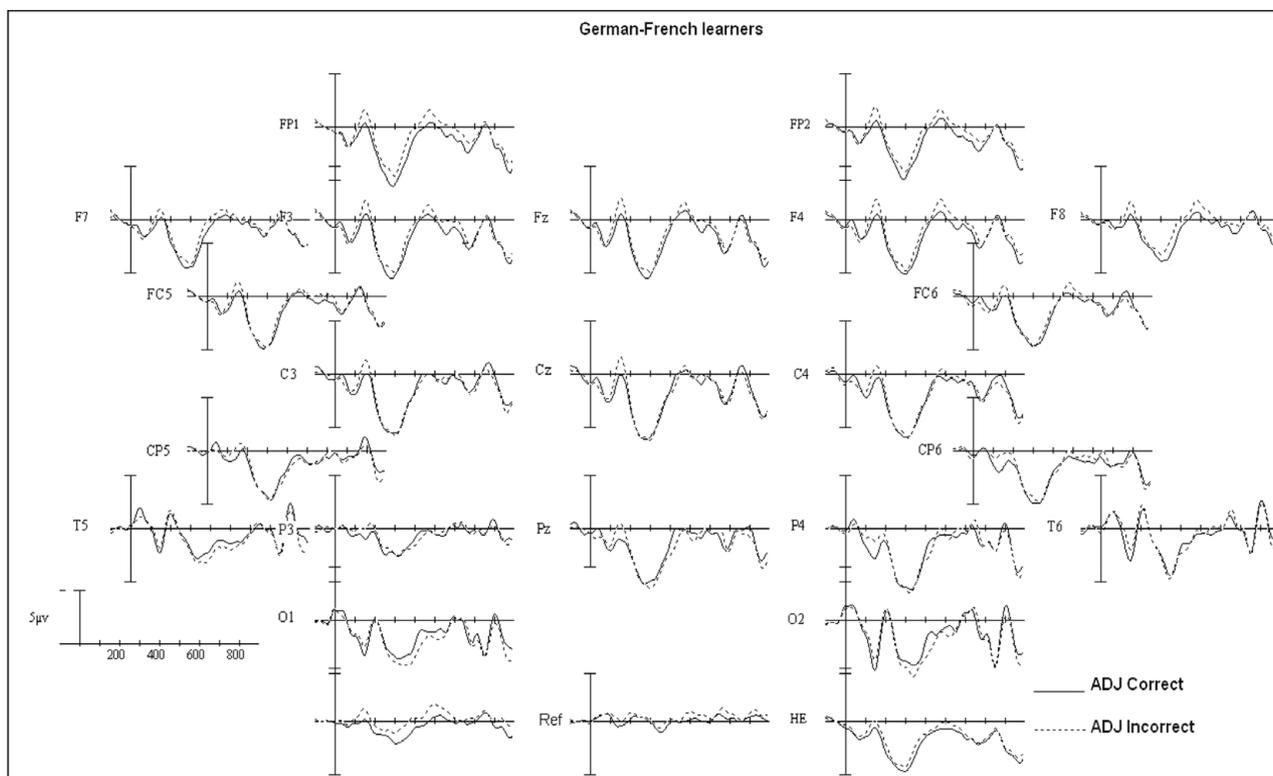


Figure 8. German–French learners: ERP traces for gender agreement errors between the noun and the postposed adjective and for correct sentences.

significant effect of Agreement in the N100 time window (midline ( $F(1,13) = 5.31, p < .04$ ), anterior lateral sites ( $F(1,13) = 9.16, p < .009$ ) and centro-parietal ( $F(1,13) = 5.52, p < .04$ ) sites). The broad scalp distribution of this effect is not consistent with an ELAN, and no ready explanation of its significance is available. No other main effects or interactions were significant, in any other time window.

### Discussion

In the present experiment we manipulated gender agreement between the noun and the postposed adjective to investigate whether L2 learners show online sensitivity to adjective–noun agreement, or if, as suggested by Sabourin and Haverkort (2003), this is not computed fast enough to be detected via ERPs, in contrast to determiner–noun agreement.

Before looking at German–French late bilinguals' performance, it is essential to examine native speakers' results. As predicted, native speakers displayed a P600 effect in response to gender agreement violations between a noun and a postposed adjective. This suggests that grammatical gender is processed at the syntactic level, in

line with the results of Experiment 1 and those of previous studies (Hagoort & Brown, 1999; Osterhout & Mobley, 1995).

German L2 learners showed only an early negativity to gender agreement errors. The significance of this early effect is not immediately clear. Several recent studies of native processing have reported very early ERP response to syntactic manipulations at times in the absence of any later positivity (Hasting & Kotz, 2008; Hasting, Kotz & Friederici, 2007; Pulvermüller & Shtyrov, 2003; Pulvermüller, Shtyrov, Hasting & Carlyon, 2008). However, these effects have most often been found for verbal agreement manipulations (but see Malaia, Wilbur & Weber-Fox, 2009), which may indeed be more automatic than the kind of gender agreement examined here, between the noun and adjective, and, moreover, they have generally been found in rather impoverished contexts, consisting of only a pronoun and inflected verb and in a mismatch paradigm (but see Hasting & Kotz, 2008; Malaia et al., 2009). As such, it is perhaps less surprising to find such a rapid response to violations under those conditions than here. Whether the early negativity to agreement violations we found in the group of L2 is in fact indicative of sensitivity to syntactic agreement is questionable. First, no such response was found in our

control group of native French speakers, and second, no subsequent evidence of processing difficulty was found in the ERP trace for the L2 speakers, whereas native controls demonstrated a clear P600 effect to anomalies. Further work is necessary to elucidate this effect.

The absence of a P600 effect for this group can be accounted for by two potential explanations. First, it is important to remember that, as mentioned earlier, in German, there is no gender distinctions in the plural of determiners, adjectives and pronouns in the nominative case (e.g., *die kleinen Tische<sub>masc</sub>*, *die kleinen Türen<sub>fem</sub>*, *die kleinen Autos<sub>neuter</sub>* “the small tables, doors, cars”). Thus, the first explanation could be that German learners apply rules from their L1 to their L2. This would explain why German learners showed a P600 effect similar to native speakers for gender agreement violations between a definite article and a singular noun in French (Experiment 1), but do not show such an effect here, when agreement involved a plural noun and adjective. We refer to this hypothesis hereafter as ‘the common plural agreement’ hypothesis.

The second explanation is related to adjective–noun agreement. As mentioned before, adjective–noun agreement seems to be less accurate and later acquired than determiner–noun agreement (Bartning, 2000; Bruhn de Garavito & White, 2002; Dewaele & Véronique, 2001; Grandfeldt, 2000). The absence of a P600 effect for gender agreement between the noun and the postposed adjective for our L2 learners seems to confirm this hypothesis. However, we have to point out again that the position of the adjective in Romance and Germanic languages differs. Indeed, the fact that the word order noun + adjective does not exist in German may hinder these learners in the syntactic process of gender agreement when the adjective is in a post-nominal position. This hypothesis is in line with that proposed by Sabourin and Stowe (2008) regarding the influence of the L1 on the L2. In their study, Romance speakers revealed a pattern similar to that of native Dutch speakers when structures were identical in both languages, but failed to do so when structures differed. This hypothesis is referred to as ‘the surface structure’ hypothesis hereafter. Note that while the results of a previous study (White, Valenzuela, Kozłowska-Macregor & Leung, 2004) suggested that agreement can in fact be observed in new syntactic structures, these data were obtained in an offline task. To test these two hypotheses we ran a third experiment which involved gender agreement with preposed adjectives which exist in our L2 learners’ native language, German.

### Experiment 3

In Experiment 2, we suggested that the absence of a P600 effect for gender mismatch for German speakers

processing French noun–adjective gender agreement could be due to one of two factors. The first is the fact that gender is neutralized in the plural, i.e., there is no gender distinction for determiners or adjectives in the plural in German (i.e., the ‘common plural agreement’ hypothesis; see discussion of Experiment 2); the second is the fact that adjective–noun agreement is less accurate and acquired later than determiner–noun agreement in L2 (see discussion of Experiment 2). However, we also pointed out that while postposed adjectives are predominant in French, they do not exist in German. To eliminate the hypothesis that the absence of a native-like effect was due to a ‘surface structure’ difference, we conducted a third experiment in which gender agreement was manipulated between the preposed adjective and the noun. The aim of this comparison was to determine whether computing gender agreement is costlier when occurring within a surface structure that differs in the native and second language, i.e., postposed adjectives (Experiment 2), than within a structure that is identical in L1 and L2, i.e., preposed adjectives (Experiment 3). In sum, if German speakers show a P600 effect to gender agreement violations for preposed adjectives, then the ‘surface structure’ hypothesis would be supported, and would account for the absence of native-like sensitivity to gender-agreement violation on the postposed adjective in Experiment 2. On the other hand, if no P600 effect is revealed, we can conclude that the difference of surface structure between the two languages does not account for the difficulty in gender processing in L2. For French native speakers, we expected a P600 effect in line with the results obtained in Experiments 1 and 2.

### Method

#### Participants

Fourteen French native speakers and 14 German–French learners received 20 euros for their participation.<sup>2</sup> The mean age of all participants was 21.6 years, ranging from 19 to 28. They all had normal or corrected-to-normal vision. L2 learners were Erasmus students at the University of Provence. They had all studied French at school (mean 8 years) prior to their arrival in France. They had all passed the exam that allows foreign students to attend courses in a French university (DELF; individual results not available). Participants had to complete an offline test which consisted in circling the gender of

<sup>2</sup> Five German speakers had taken part in the Experiment 2 to reduce the variability due to participants’ knowledge of French. However, as some prior participants had gone back to their home country, they had to be replaced. Experiment 3 was conducted five months after Experiment 2, and analyses revealed no significant difference between the results for the group who had taken part in the previous experiment and the new group.

Table 3. Example of all the conditions presented in Experiment 3.

| Condition                        | Examples  |
|----------------------------------|---|
| Opposite gender across languages |   |
| <b>Correct</b>                   | <i>Les petites<sub>fem</sub> chaises<sub>fem</sub> sont dans le salon.</i>  |
| <b>Incorrect</b>                 | <i>Les petits<sub>masc</sub>* chaises<sub>fem</sub> sont dans le salon*.</i><br>“The small <sub>fem/masc</sub> * chairs <sub>fem</sub> are in the living room.” |
| Same gender across languages     |   |
| <b>Correct</b>                   | <i>Les petits<sub>masc</sub> balais<sub>masc</sub> sont dans la cave.</i>   |
| <b>Incorrect</b>                 | <i>Les petites<sub>fem</sub>* balais<sub>masc</sub> sont dans la cave.*</i><br>“The small <sub>masc/fem</sub> * brooms <sub>masc</sub> are in the cellar.”      |

the stimuli (mean: 5.6% errors, *SD* 3.5), as well as on the self-assessment of their level in French (written comprehension, 4.6; oral comprehension, 4.4; written production, 4; oral production, 3.9). Thus, according to both number of years of study and successful completion of the DELF (level 2 or better), the proficiency level of participants was considered to be advanced.

### Materials

Ninety-six experimental sentence pairs were created, one member being grammatically well-formed and the other ill-formed. Grammaticality was determined by gender agreement between the preposed adjective and the noun. The same 96 nouns as in Experiments 1 and 2 were used, paired with 42 adjectives in short sentence contexts (mean frequency of adjectives per million: 130.6; mean length 5.4 letters, range between 3 and 8, *Lexique 2*; New et al., 2004). The adjectives were selected such that they could be placed pre-nominally. The semantic acceptability of sentences was checked by French native speakers prior to the main experiment. Each adjective was seen between one and six times. Twenty-four sentences were presented per condition, defined by Gender Agreement (gender agreement vs. mismatch between the preposed adjective and critical noun), Language Coherency (nouns of same vs. opposite gender in French and German) and Noun Gender (masculine vs. feminine nouns). Two lists were created such that all critical nouns were seen in both gender agreement conditions but in only one condition per list. The pattern of sentences was: adverb (or adverbial phrase), plural definite article, adjective, critical noun, copula and complement. In addition to the 96 experimental sentences, 96 syntactically correct filler sentences involving various syntactic structures were presented. The sentences were presented in a fixed-random order, and six fixed-random orders were created per list. Each participant saw only one list. The experiment proper was preceded by four warm-up sentences. An example of all experimental conditions is presented in Table 3.

### Procedure and EEG recording

These were identical to Experiment 2.

### Data analysis

This was identical to Experiment 2, with the exception of the P600 time window, which was reduced to a shorter period, i.e., from 500–700 ms, due to visual inspection which revealed both a shorter latency and smaller amplitude than obtained in Experiment 2 and in relation to many L1 studies (Friederici, Hahne & Saddy, 2002; Hahne & Friederici, 2001). Prior to analyses, trials with artefact were rejected (French: 1.4% and 1.12%, German: 0.8% and 0.5% for correct and incorrect conditions, respectively and no significant difference emerged between groups,  $F < 1$ ).

### Results

Gender agreement violations between the adjective and the following noun provoked differences in the waveforms that depended both on participant group and time window. French native speakers showed a positive deflection in the waveform for these agreement violations between 500–700 ms after the onset of the critical noun (Figure 9), whereas German–French learners did not show any differences, in any time window (Figure 10). ANOVAs were performed on these data.

No main effects or interactions were significant prior to the 500–700 ms time window. In this window, a significant effect of Agreement emerged at centro-parietal sites ( $F(1,26) = 4.48, p < .04$ ) and tended towards significance at midline ( $F(1,26) = 3.12, p = .09$ ). The effect was modified by an interaction with Group (midline ( $F(1,26) = 11.96, p < .002$ ) anterior lateral ( $F(1,26) = 7.47, p < .011$ ) and centro-parietal ( $F(1,26) = 9.52, p < .005$ ) sites). Post-hoc analyses confirmed the presence of a P600 effect to agreement violations for French native speakers but not for German learners. No other experimental factors were significant, nor did they interact with Agreement. Subsequent ANOVAs were performed on the data for each group independently.

*French native speakers:* Gender agreement violations between the adjective and the following noun provoked a P600 effect (Figure 9) as revealed by a significant effect of Agreement (midline ( $F(1,13) = 15.4, p < .002$ ), anterior lateral ( $F(1,13) = 6.41, p < .02$ ) and centro-parietal sites ( $F(1,13) = 12.48, p < .004$ )). No other effects reached significance, nor were any interactions observed.

*German–French L2 learners:* No differences were observed as a function of experimental factors or their interactions (all  $F_s < 1$ ; Figure 10).

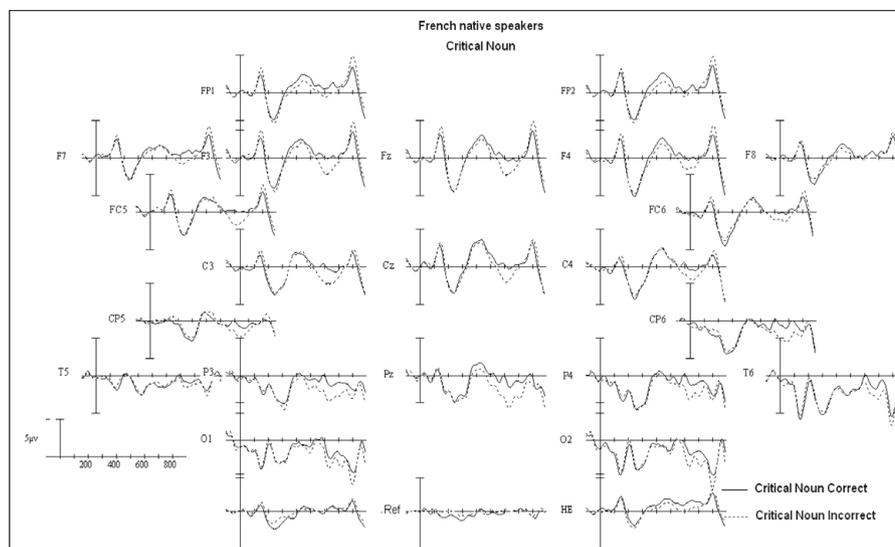


Figure 9. French native speakers' ERP response to gender agreement violations between the preposed adjective and the critical noun.

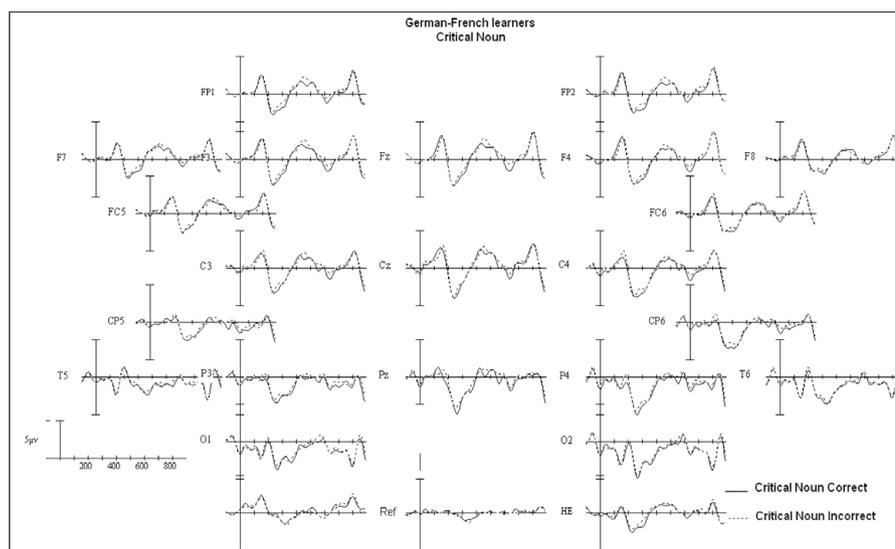


Figure 10. German–French learners' ERP response to gender agreement violations between the preposed adjective and the critical noun.

### Discussion

In the present experiment, we manipulated gender agreement between preposed adjectives and nouns to investigate whether our L2 learners would experience lesser difficulties computing gender agreement when occurring within a surface structure that was the same in the native and second language (i.e., preposed adjectives) compared to a surface structure that differed across the L1 and L2, (i.e., postposed adjectives), which we tested in Experiment 2.

In line with previous studies (Barber & Carreiras, 2005), our native speakers displayed a P600 effect in

response to gender agreement violations between the preposed adjective and the noun. These results are consistent with those obtained in Experiments 1 and 2 for gender agreement violations between a singular definite article and noun and between a postposed adjective and noun, respectively.

In contrast, German–French learners did not show any effect of grammatical gender agreement errors on critical nouns following pre-nominal adjectives. This pattern replicates the absence of P600 effect for gender agreement in Experiment 2 for these participants. Hence, we can argue that the absence of an effect of gender agreement is not due to a different surface structure in French

and German (i.e., the absence of postposed adjectives in German). However, from our results, we cannot conclude whether the absence of native-like online sensitivity to gender is due to the fact that adjective agreement is less accurate and later acquired than determiner agreement in L2, or whether our L2 learners realized a transfer from their L1 to their L2 as suggested by the ‘common plural agreement’ hypothesis. In the present case, the rule that was apparently transferred from the L1 and hindered gender agreement processing in L2 French is that of common agreement for all genders in the plural in German. It is likely that once they become more proficient, these learners will apply the rules of the French system procedurally and will show evidence of online processing agreement akin to what is found for native speakers, however, our current data do not allow us to test this assumption. Further research on German native speakers who have been exposed to French for several years is required to investigate this question.

### General discussion

In the present article we report three experiments investigating whether gender is processed in a similar way by native and non-native speakers, and whether processing in an L2 is influenced by the native language. In these experiments we used ERPs to compare French native speakers and German–French learners. In Experiment 1, participants read sentences containing gender agreement violations between the definite article and the noun. The recordings of participants’ brain activity (ERPs) showed sensitivity to these violations, as revealed by a P600 effect, for native speakers as well as for the L2 learners. However, the response in the L2 group was less consistent than that of native speakers. Indeed, whereas all L2 learners showed sensitivity to gender agreement for nouns that had the same gender in French and German, only a subset of these participants showed the same sensitivity for nouns that had opposite gender across their native and later-learned language.

In Experiment 2, we manipulated gender agreement between the noun and the postposed adjective to investigate whether agreement between the noun and the adjective was less accurate and later acquired than that between the determiner and the noun in L2, as suggested in linguistic studies (Bartning, 2000; Bruhn de Garavito & White, 2002; Dewaele & Véronique, 2001; Grandfeldt, 2000) as well as recent ERP studies of L2 processing (Sabourin & Haverkort, 2003). Results for French native speakers showed a P600 effect in response to syntactic violations; in contrast, German learners did not show any effect (except for an unclear early negativity). We argued that this absence of native-like sensitivity to gender agreement between the noun and the postposed

adjective could be due to two factors. On the one hand, we suggested that German–French learners transferred the plural agreement rule from their L1 to the L2, since in German the determiner is common for all genders in the plural (i.e., *die*). That is, although abstract gender is present, gender is not overtly marked in the plural. This seemingly hindered processing in L2 French for the L1 German participants, who apparently applied the same rule in French and failed to perform the required gender agreement in their L2, or at least not rapidly and/or systematically enough for an ERP response to become visible. On the other hand, we suggested that the absence of a P600 effect might be due to the fact that adjective agreement was indeed less accurate and later acquired than determiner agreement in L2. However, before we drew this conclusion it was necessary to check whether the difference of surface structure between the two languages was not responsible for the absence of sensitivity to gender agreement violations. Indeed, while postposed adjectives are both canonical and highly frequent in French, they do not exist in German. Hence, we conducted a final experiment (Experiment 3) in which gender agreement was manipulated between the preposed adjective and the noun; this word order is common to German and French. Again, French native speakers showed a P600 effect in response to gender agreement violations whereas L2 learners did not reveal any effect. Thus, we concluded that the absence of an effect in the L2 group was not mainly a question of different surface structure in L1 and L2. We admit, however, that the possibility remains that the difference across German and French for canonical adjective position within the NP may have caused a general difficulty in acquiring adjectival agreement in L2 French for these learners. Nonetheless, our results leave open two possibilities regarding the absence of an ERP response to noun–adjective gender agreement in our L1 German–L2 French learners. Either they had greater difficulty with adjectival agreement (independent of word order), in line with the results of various studies (Bartning, 2000; Bruhn de Garavito & White, 2002; Dewaele & Véronique, 2001; Grandfeldt, 2000; Sabourin & Haverkort, 2003) or the difference in German and French gender agreement for plural DPs could account for the pattern we obtained. This question is currently under investigation in our research laboratory.

Sabourin and colleagues (Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008) suggested that although L2 learners are able to learn new information and incorporate it at a lexical level, they may not attain syntactic competence (see Hopp, 2007, for further supporting evidence). They argued that constructions that are not grammatically similar in both languages are processed differently by non-native speakers. They reached this conclusion because in their study Romance language learners of Dutch did not reveal a P600 effect in the case

of gender agreement violations between the determiner and the noun in Dutch, whereas German–Dutch learners did. They accounted for this difference between their L2 groups by linguistic proximity, as German and Dutch are more closely related languages than are Dutch and the Romance languages. However, the results of our first experiment only partially corroborate this conclusion. Our results for German L1–French L2 learners showed a P600 effect for gender agreement violations between the determiner and the noun (Experiment 1) despite the linguistic distance between these two languages. Nonetheless, this effect was modified by the overlap of lexical gender in the L1 and L2 in our study. Half of our participants were sensitive to gender violations independently of gender coherency between French and German, whereas the other half showed an effect only when nouns shared the same gender across languages. The difference between our results and Sabourin's may stem from the proficiency of the participants or experimental design (they did not distinguish the native language of their Romance language speakers, and the number of participants was limited). It is possible that only the most proficient L2 learners will show online sensitivity to grammatical gender manipulations when their L1 and L2 do not have overlapping lexical gender and agreement rules. Further online investigations are required.

The results we obtained for adjectival agreement (Experiments 2 and 3) support Sabourin and colleagues' hypothesis that automatic, native-like processing of gender will only occur in L2 learners to the extent that their L1 provides a basis for the transfer of both lexical gender and rules of agreement. Indeed, even if our German–French speakers were able to assign gender to French nouns offline, as evidenced by the offline test, their online processing showed evidence of such only when agreement occurred between the noun and the determiner but not when it occurred between the noun and the adjective. Nevertheless, since it has been shown that, as in L1, adjective agreement in L2 is less accurate and later acquired than determiner agreement (Bartning, 2000; Bruhn de Garavito & White, 2002; Dewaele & Véronique, 2001; Grandfeldt, 2000), we predict that more advanced learners should grammaticalize this knowledge and show sensitivity to gender agreement between the adjective and the noun, as they do for agreement between the determiner and the noun, despite differences in the L1 and L2 grammatical system. Hence, we can argue from our results that high-proficiency L2 learners who receive enough exposure to their L2 can process gender in a similar way to native speakers; at present our results and those of others (Sabourin & Haverkoort, 2003) show this to be true for obligatory elements, i.e., determiner–noun (bare nouns are not permissible in French), but not for other agreeing elements in the DP. Recently, Gillon Dowens, Vergara, Barber and Carreiras

(2010) showed that late L2 learners (L1 English) who had been exposed to their second language (Spanish) for at least 12 years revealed similar results as native speakers in the case of syntactic violation between the determiner and the noun (i.e., early negativity and P600 effect), but not in case of violation between the noun and the predicate adjective (only P600 effect in L2 learners). These results are in line with previous studies suggesting that near-native syntactic processing can be attained, but that it depends on proficiency (Hahne & Friederici, 2001; Hahne et al., 2006; Rossi et al., 2006).

Our results are in line with previous monolingual and bilingual studies that have examined gender agreement within the determiner phrase in sentence contexts, showing a P600 effect to a violation of this agreement rule (Hagoort & Brown, 1999; Sabourin & Haverkoort, 2003). We did not find an earlier negativity, as has been found in some studies of native speakers for this type of processing (Barber & Carreiras, 2005; Gunter et al., 2000), in either the native group of French speakers or the German–French learner group. The main issue for present purposes is the similarity of the ERP response across the native and learner groups, i.e., the finding of a significant P600 response that was similar in distribution, size and latency across the two groups (Experiment 1). This result differs from that reported in several previous bilingual studies which looked at the ERP effects associated with processing various types of syntactic anomalies. Some have reported a delayed P600 effect in non-native speakers (Weber-Fox & Neville, 1996); others have reported a similar P600 effect in native and non-native participant groups, but an absence of a concomitant early negativity in the non-native group that was present in the native group (Hahne, 2001; Hahne & Friederici, 2001). Our results for determiner agreement (Experiment 1) differ in that we found a significant P600 effect only to gender agreement errors, independent of the native language status of our participants.

As a final note, Friederici and collaborators (Rossi et al., 2006) have recently suggested that the three-phase model proposed to represent universal syntactic processing in monolinguals (Friederici, 2002) could serve as a theoretical framework for bilinguals. This model proposes an initial stage of autonomous phrase structure construction (reflected by an ELAN), a second phase where morphosyntactic processing occurs (reflected by a LAN) and a third phase of reanalysis and repair (reflected by the P600). In line with previous studies (Hagoort & Brown, 1999; Sabourin & Haverkoort, 2003), our native speakers displayed a classic P600 in response to gender-agreement violations between two elements of the noun phrase, supporting the claim that gender is represented syntactically, and that the online processing of grammatical gender is not a conceptual and/or semantic, but a syntactically driven process. However, we did not find

evidence of a biphasic response (i.e., LAN + P600 effects). Hence, our results do not support Friederici's model and its adaptation to syntactic processing in bilinguals as recently proposed (Rossi et al., 2006). Furthermore, it has been suggested recently that linguistic and methodological factors influence ERP effects, and therefore that early ERP effects (such as the ELAN and LAN) may not reflect strict successive processing stages (Hasting & Kotz, 2008). It is important to note, however, that we do not disagree with the proposal that L1 models can be extended to account for L2 processing, as suggested by the conclusions we drew from the present study that a near-native level of syntactic processing can be reached by highly proficient L2 speakers who have received enough exposure to their L2. Further research is required to support this proposal.

## References

- Ardal, S., Donald, M. W., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language*, 39, 187–205.
- Barber, H., & Carreiras, M. (2005). Grammatical gender and number in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, 17, 137–153.
- Bartning, I. (2000). Gender agreement in L2 French: Pre-advanced vs. advanced learners. *Studia Linguistica*, 54, 225–237.
- Bernstein, J. (1991). DPs in French and Walloon: Evidence for parametric variation in nominal head movement. *Probus*, 3, 101–126.
- Blom, E., Polisenska, D., & Unsworth, S. (2008). The acquisition of grammatical gender in Dutch. *Second Language Research*, 24, 259–265.
- Bolte, J., & Conine, C. (2004). Grammatical gender in spoken word recognition in German. *Perception & Psychophysics*, 66 (6), 1018–1032.
- Bruhn de Garavito, J., & White, L. (2002). L2 acquisition of Spanish DPs: The status of grammatical features. In A. T. Pérez-Leroux & J. M. Liceras (eds.), *The acquisition of Spanish morphosyntax: The L1/L2 Connection*, pp. 151–176. Dordrecht: Kluwer.
- Carroll, S. (1989). Second language acquisition and the computational paradigm. *Language Learning*, 39, 535–594.
- Clark, Eve (1985). The acquisition of Romance, with special reference to French. In D. I. Slobin (ed.), *The crosslinguistic study of language acquisition 1*, pp. 687–782. Hillsdale, NJ: Erlbaum.
- Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.
- Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: Evidence from ERPs and eye movements. *Journal of Memory and Language*, 45, 200–224.
- Dewaele, J. M., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition*, 4, 275–297.
- Foucart, A., & French-Mestre, C. (2006). Processing of adjectives in French as first and second language: Evidence from ERPs. Presented at CUNY, New York, USA, 23–25 March.
- Foucart, A., & French-Mestre, C. (2007). Is adjective processing in L2 affected by L1? Evidence from ERPs and eye-tracking. Presented at Architecture and Mechanisms of Language Processing, Turku, Finland, 24–27 August.
- French-Mestre, C. (2005). Eye-movement as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research*, 21, 175–198.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 78–84.
- Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research*, 31 (1), 45–63.
- Gillon Dowens, M., Vergara, M., Barber, H., & Carreiras, M. (2010). Morpho-syntactic processing in late L2 learners. *Journal of Cognitive Neuroscience*, 22 (8), 1870–1887.
- Grandfelt, J. (2000). The acquisition of the determiner phrase in bilingual and second language French. *Bilingualism: Language and Cognition*, 3 (3), 263–280.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95–111.
- Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, 12, 556–568.
- Hagoort, P., & Brown, C. M. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research*, 28, 715–728.
- Hahne, A. (2001). What's different in second language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, 30, 251–266.
- Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, 4, 123–141.
- Hahne, A., Mueller, J., & Clahsen, H. (2006). Morphological processing in a second language: Behavioural and event-related potential evidence for storage and decomposition. *Journal of Cognitive Neuroscience*, 18, 121–134.
- Hasting, A., & Kotz, S. (2008). Speeding up syntax: On the relative timing and automaticity of local phrase structure and morphosyntactic processing as reflected in event-related brain potentials. *Journal of Cognitive Neuroscience*, 20 (7), 1207–1219.
- Hasting, A., Kotz, S., & Friederici, A. (2007). Setting the stage for automatic syntax processing: The mismatch negativity as an indicator of syntactic priming. *Journal of Cognitive Neuroscience*, 19 (3), 386–400.
- Hawkins, R., & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition: The "failed functional features hypothesis". *Second Language Research*, 13, 187–226.

- Herschensohn, J. (2006). Français langue seconde: From functional categories to functionalist variation. *Second Language Research*, 22, 95–113.
- Herschensohn, J. (2007). *Language development and age*. Cambridge: Cambridge University Press.
- Hopp, H. (2007). Ultimate attainment at the interfaces in second language acquisition: Grammar and processing. Unpublished doctoral thesis, University of Groningen.
- Inoue, K., & Osterhout, L. (2005). Honorifics in Japanese: An ERP study. Presented at the Annual CUNY Sentence Processing Conference, Tucson, March.
- Jasper, H. H. (1958). The ten twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 5, 196–214.
- Laenzlinger, C. (2005). French adjective ordering: Perspectives on DP-internal movement types. *Lingua*, 115, 645–689.
- Malaia, E., Wilbur, R. B., & Weber-Fox, C. (2009). ERP evidence for telicity effects on syntactic processing in garden-path sentences. *Brain and language*, 108, 145–158.
- Mirković, J., MacDonald, M. C., & Seidenberg, M. S. (2005). Where does gender come from? Evidence from a complex inflectional system. *Language and Cognitive Processes*, 20, 139–168.
- Müller, N. (1990). Gender and number agreement within DP. In J. M. Meisel (ed.), *Bilingual first language acquisition: French and German grammar development*, pp. 53–88. Amsterdam: Benjamins.
- Müller, O. & Hagoort, P. (2006). Access to lexical information in language comprehension: Semantics before syntax. *Journal of Cognitive Neuroscience*, 18, 84–96.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004) Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36 (3), 516–524.
- Osterhout, L., Bersick, M., & McKinnon, R. (1997). Brain potentials elicited by words: Word length and frequency predict the latency of an early negativity. *Biological Psychology*, 46, 143–168.
- Osterhout, L., McLaughlin, J., Kim, A., & 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, Jr. (eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond*, pp. 271–308. New York: Psychology Press.
- Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: A paradigm for exploring the neurocognition of second-language processing. *Language Learning*, 56, 199–230.
- Osterhout, L. & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739–773.
- Parodi, T., Schwartz, B. D., & Clahsen, H. (1997). On the L2 acquisition of morphosyntax of German nominals. *Essex Research Reports in Linguistics*, 15, 1–43.
- Perez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, 18, 571–590.
- Pulvermüller, F., & Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *NeuroImage*, 20, 159–172.
- Pulvermüller, F., Shtyrov, Y., Hasting, A. S., & Carlyon, R. P. (2008). Syntax as a reflex: Neurophysiological evidence for early automaticity of grammatical processing. *Brain and Language*, 104, 244–253.
- Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, 18, 2030–2048.
- Sabourin, L. (2003). Grammatical gender and second language processing: an ERP study. Unpublished doctoral dissertation, University of Groningen.
- Sabourin, L., & Haverkort, M. (2003). Neural substrates of representation and processing of a second language. In R. van Hout, A. Hulk, F. Kuiken and R. J. Towell (eds.), *The lexicon–syntax interface in second language acquisition*, pp. 175–195. Amsterdam: John Benjamins.
- Sabourin, L., & Stowe, L. A. (2008) Second language processing: When are L1 and L2 processed similarly? *Second Language Research*. 24 (3), 397–430.
- Spinelli, E., Meunier, F., & Seigneuric, A. (2006). Does gender information influence early phases of spoken word recognition? *The Mental Lexicon*, 1, 277–297.
- Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event related potential investigation. *Studies in Second Language Acquisition*, 2, 173–204.
- Weber-Fox, C. M., & Neville, H. J. (1996). Maturation constraints on functional specializations for language processing: ERP and behavioural evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231–256.
- White, L., Valenzuela, E., Kozłowska-Macregor, M., & Leung, Y. I. (2004). Gender and number agreement in non-native Spanish. *Applied Psycholinguistics*, 25, 105–133.