Journal of Memory and Language 66 (2012) 226–248

Contents lists available at SciVerse ScienceDirect

Journal of Memory and Language

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

Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking ☆

Alice Foucart [a,*], Cheryl Frenck-Mestre [b,c]

[a] University of Edinburgh, Department of Psychology, 7, George Square, Edinburgh EH8 9JZ, Scotland, UK
[b] CNRS, France
[c] Université de Provence, France

ARTICLE INFO

Article history:
Received 29 June 2010
revision received 19 July 2011
Available online 30 August 2011

Keywords:
Second language acquisition
Grammatical gender
ERPs
Eye-tracking

ABSTRACT

We report a series of ERP and eye-tracking experiments investigating, (a) whether English–French learners can process grammatical gender online, (b) whether cross-linguistic similarities influence this ability, and (c) whether the syntactic distance between elements affects agreement processing. To address these questions we visually presented sentences which were either grammatically correct or contained noun–adjective gender agreement violations. In response to violations between the noun and a post-posed adjective (the canonical structure in French), both groups revealed a P600 effect. In contrast, violations between the noun and a pre-posed adjective (a less frequent order) triggered a P600 in French speakers but an N400 in L2 learners (implying that learners have not yet fully acquired native-like processing for pre-posed adjectives). Violations between the noun and the predicative adjective showed different effects for the native (P600) and non-native (no effect) groups with ERPs, but a similar pattern with eye-tracking. Overall, these results suggest that late L2 learners can acquire and process new features.

Introduction

Le table or *la table*? Grammatical gender often seems arbitrary to any learner whose native language (L1) does not possess this feature. Indeed, while semantic gender relies on extra-linguistic properties (e.g., sex, colour, shape, etc.) to categorise nouns into classes, grammatical gender is not as straightforward. French, like many other languages, has a grammatical gender system in which the intrinsic gender of the noun governs agreement between the elements of the sentence it co-refers with (e.g., determiner, adjectives, pronouns). Native speakers systematically acquire grammatical knowledge of gender assignment and of agreement rules at a relatively young age (Carroll, 1989; Clark, 1985; Müller, 1990; Perez-Pereira, 1991). This knowledge affects both lexical access (for recent reviews see Bolte & Connine, 2004; Spinelli, Meunier, & Seigneuric, 2006) and syntactic processing (Barber & Carreiras, 2005; Deutsch & Bentin, 2001; Foucart & Frenck-Mestre, 2011; Gunter, Friederici, & Schriefers, 2000; Osterhout & Mobley, 1995). In contrast, English relies predominantly on semantic gender and agreement is limited to pronouns and their animate antecedents (e.g., *he* and *she*). Hence, one may ask whether it is possible for English native speakers to acquire grammatical gender when learning French (here, and throughout the paper, we use the term 'acquire' to refer to the fact that the gender of a noun not only has to be learned but also processed online in a native-like manner). In other words, are they able to learn the gender of a noun and to process agreement online? Beyond this general level, one need also determine the factors that may affect online agreement. Is agreement processing affected by the similarity/difference of surface word order in first (L1) and second (L2; i.e., pre-posed and post-posed adjectives) and by

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

* Corresponding author.
E-mail address: alfoucart@gmail.com (A. Foucart).

0749-596X/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2011.07.007

syntactic structure (i.e., attributive vs. predicative adjective agreement)? These questions are the central issue of the present paper. Gender processing in L2 has already been the topic of numerous linguistic and psycholinguistic studies, particularly with respect to the influence of the first language and age of acquisition (Carroll, 1989; Foucart & Frenck-Mestre, 2011; Frenck-Mestre, Foucart, Carrasco, & Herschensohn, 2009; Gillon-Dowens, Vergara, Barber, & Carreiras, 2010; Hawkins & Franceschina, 2004; Keating, 2009; Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008; White, Valenzuela, Kozlowska-Macregor, & Leung, 2004); however, the questions addressed here are still in debate.

Grammatical gender agreement in French

French has a two-gender system (i.e., masculine and feminine) in which nouns and all elements within the determiner phrase (DP) must agree in gender and number. Bare nominals are generally not permissible in standard French (Deprez, 2005), such that nouns are systematically associated with a determiner. For singular definite determiners, gender is overtly realised in both the orthography and phonology (*le*$_m$ *livre*$_m$, the book/*la*$_f$ *chaise*$_f$, the chair) whereas it is neutralised in the plural (*les livres*/*les chaises*). Any adjectives within the DP must also agree in number and gender with the noun they modify, although the overt realisation of gender alternation on the adjective will vary depending upon its type (variable, invariable or suppletive). In French, due to noun-movement, the canonical order of adjectives is post-nominal, although French allows for both pre-nominal and post-nominal orders and there is a small class that is in fact predominantly pre-nominal. Linguistic theories suggest that noun movement is due to the strength value (weak or strong) of abstract grammatical features (see Bernstein, 1991; Carstens, 2000; Cinque, 1994; Laenzlinger, 2005; among others). According to these theories, when a noun is selected from the lexicon, it enters into a relationship with other categories, which are number and determiner, in order to create the syntactic structure of the DP. In Romance languages, it is proposed that the noun raises overtly higher than the adjective, from noun to number (due to the strong features in number), which gives the noun–adjective order, in contrast to the adjective–noun order in Germanic languages (due to weak features) in which attributive adjectives are systematically pre-nominal.

In French and in Romance languages in general, predicative adjectives (i.e., adjectives that are separated from the noun they agree with by, at least, the verb) also show agreement with the subject noun phrase, both in gender and in number. This is not the case for Germanic languages such as German, Dutch or English, for which predicative adjectives show no agreement with the subject noun phrase. Indeed, English lacks systematic grammatical gender agreement outside of the pronominal system. Given the cross-linguistic differences between French and English regarding both adjectival agreement and surface word order of attributive adjectives, one can ask whether English speakers can fully acquire the syntactic reflex of gender agreement when they learn French late in life. This question has been addressed in linguistic and psycholinguistic studies and contrasting theories have been proposed as reported in the next section.

Theories of gender processing in L2

Many linguistic studies have examined gender processing in an L2 (Anderson, 2008; Ayoun, 2007; Bruhn de Garavito & White, 2002; Carroll, 1989; Franceschina, 2001, 2005; Gess & Herschensohn, 2001; Hawkins, 1998, cited in Hawkins, 2001; Hawkins & Franceschina, 2004; White et al., 2004). Their findings diverge in relation to the hypothesized possibility for late L2 learners to acquire new features in the L2 if they are not present in their L1. These studies used the linguistic framework regarding gender agreement first proposed by Chomsky (1995) and later developed by Carstens (2000). According to this theory, agreement is due to a feature checking mechanism between the feature of the noun and other corresponding features in the structure of the DP. In other words, gender is an interpretable feature found in the noun: it conveys information required for semantic interpretation. In contrast, determiners and adjectives have uninterpretable gender features. These uninterpretable features are deleted by the checking mechanism that matches noun features (e.g., [+masc]) with the corresponding features of the determiner (head/head) and/or adjective (specifier/head).

From the results of several observational studies various authors have concluded that after a critical period (roughly adolescence), learners are no longer able to acquire abstract grammatical features not available in their L1 (Carroll, 1989; Hawkins, 1998, 2001; Hawkins & Chan, 1997; Hawkins & Franceschina, 2004). According to the failed functional features hypothesis (FFFH), proposed by Hawkins and collaborators, new interpretable features can be acquired in L2, whereas new uninterpretable features cannot. Thus the FFFH predicts that agreement checking should not be possible for learners whose native language does not possess these features (Hawkins, 1998; Hawkins & Franceschina, 2004). Consequently, if grammatical gender is not available in L1, then noun gender can be acquired, but gender agreement checking cannot. In essence, English–French learners can learn the gender of nouns, but do not go beyond a probabilistic selection of determiner based on noun phonology. In other words, they do not develop a syntactic reflex of gender agreement so agreement processing never truly becomes an integral part of their system and is easily affected by various factors (e.g., tiredness, speed of speech), which is not the case for L1 speakers and, assumedly, L2 speakers whose L1 possesses a gender system.

In contrast, the full transfer full access model (Schwartz & Sprouse, 1996; White, 1989, 2003) suggests that during initial stages of L2 acquisition the representation of grammatical features is based on the features available in the L1; however, the model assumes that throughout the lifespan learners have 'full access' to underlying universal grammar and that new features required by the L2 can be acquired, regardless of the age of acquisition. Thus, in principle, learners are able to acquire a native-like mental

representation of L2 abstract features. Hence, uninterpretable features can be acquired, feature-checking can be realised and agreement can be computed automatically in line with the findings of numerous linguistic studies (Bruhn de Garavito & White, 2002; Hopp, 2007; Schwartz & Sprouse, 1996; Taner, 2008; White, 1989; White et al., 2004).

At a more general level, competing accounts of L2 syntactic processing indeed exist. Recently, Clahsen and Felser (2006) proposed the Shallow Structure Hypothesis (SSH), which claims that L1 and L2 processing differ in that the syntactic analysis engaged in by late L2 speakers during language comprehension is not as in depth as that performed by native speakers. The authors suggested that L2 speakers rely on other available information such as lexico-semantic cues. Note, that such a hypothesis is not restricted to L2 processing; in the case of syntactically ambiguous structures native speakers may, as well, rely more on a surface interpretation rather than a full parse of the sentence (Christianson, Hollingworth, Halliwell, & Ferreira, 2001). In contrast, Hopp's (2007) Fundamental Identity hypothesis argues that grammatical representation and processing are similar in native and non-native speakers and that if differences are found, they are not due to a critical period but to factors related to L2 acquisition such as L1 transfer or performance factors.

Of interest for the present paper, the SSH proposes that native-like processing can be attained by L2 learners only in local contexts, i.e., elements of a phrase that are juxtaposed. The experiments reported in this paper addressed this issue by investigating gender agreement both in a local context, between the noun and attributive adjective within the DP (Experiments 1 and 2) and across a syntactic boundary, between the noun and the predicative adjective (Experiments 3 and 4) using online methodologies, namely ERPs and eye-tracking.

Online evidence of gender processing

The consensus from L1 online studies on gender agreement is that it is a syntactic process as opposed to being semantically driven. This conclusion stems from the results of numerous monolingual studies that have investigated gender processing using ERPs, all of which have revealed a P600 effect in response to gender agreement violations in sentence context. This effect has been reported in various languages, including Dutch (Hagoort & Brown, 1999), French (Foucart & Frenck-Mestre, 2011; Frenck-Mestre, 2005; Frenck-Mestre, Osterhout, McLaughlin, & Foucart, 2008; Frenck-Mestre et al., 2009), German (Gunter et al., 2000), Hebrew (Deutsch & Bentin, 2001) and Spanish (Barber & Carreiras, 2005; Gillon-Dowens et al., 2010). The P600 effect was obtained regardless of the elements involved (e.g., article–noun, adjective–noun, reflexive–antecedent) or the position of violations (within the DP or across a syntactic boundary). It was sometimes preceded by a LAN effect (Barber & Carreiras, 2005; Deutsch & Bentin, 2001; Gunter et al., 2000), but not consistently (Foucart & Frenck-Mestre, 2011; Frenck-Mestre, 2005; Frenck-Mestre et al., 2008, 2009; Hagoort & Brown, 1999; for a recent discussion of the LAN, see McLaughlin et al., 2010).

Only a handful of studies have investigated the online processing of gender agreement in an L2. The two main methods used to address this question were the recording of ERPs and of eye-movements while L2 learners processed grammatical gender in sentential context. The picture that arises from these studies is yet unclear. Whereas all have shown that L2 learners can process grammatical gender online to some extent, the degree of "nativeness" of performance varied across studies. To account for this variability, some authors have argued strongly that L1 grammatical features are crucial (Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008) whereas others have underlined the role of L2 experience (Foucart & Frenck-Mestre, 2011; Gillon-Dowens et al., 2010; Keating, 2009) and possible competition between L1 and L2 systems (Foucart & Frenck-Mestre, 2011; Tokowicz & MacWhinney, 2005). These results and conclusions of these studies are summarized below.

In two ERP experiments using the same materials but different L2 populations, Sabourin and colleagues (Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008) compared the ERP responses of Dutch native speakers and L2 learners to gender agreement manipulations in visually presented sentences. They manipulated determiner–noun and adjective–noun agreement within the DP. From the pattern of ERP responses elicited by gender agreement violations, Sabourin and Haverkort (2003) concluded that L2 learners can attain native-like processing for grammatical gender when syntactic structures are similar in L1 and L2 but not when they differ across languages. This conclusion was extended by Sabourin and Stowe (2008) who suggested that automatic gender processing in L2 depends not only on the presence of a grammatical gender system in the L1 but also requires overlapping of lexical gender. The claim for processing differences as a function of the overlap of L1 and L2 grammatical features should, however, be considered with caution because of the variability between the L2 groups (both regarding *N* size and learners' native language, which comprised several languages for a given class in some cases).

In another ERP study, Tokowicz and MacWhinney (2005) examined beginning English–Spanish learners' sensitivity to grammatical violations as a function of the presence of L2 structures in the native language. The results showed a small but reliable P600 effect in response to violations that occurred in structures that were similar in L1 and L2 (i.e., tense marking) and structures that were unique to the L2 (i.e., gender agreement), but not in structures that were different across languages (i.e., nominal number agreement). Of particular interest for the present purposes is the finding that L2 learners showed online sensitivity to gender agreement violations, despite the lack of grammatical gender in the learners' L1 (English). The authors suggested, contrary to the claims forwarded by Sabourin and colleagues, that features that are not present in L1 – here, grammatical gender for native English speakers – cannot only be acquired but 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).

The conclusion that L2 syntactic processing in general and gender agreement in particular is affected by the

similarity of syntactic rules across L1 and L2 was recently supported by Foucart and Frenck-Mestre (2011) in an ERP study. They compared the sensitivity of French native speakers and German–French learners to three different gender agreement manipulations within the determiner phrase: between the determiner and the noun, a post-posed attributive adjective and the noun, and a pre-posed attributive adjective and the noun. French native speakers revealed a classic P600 effect in response to gender agreement violations independently of where the violation occurred (i.e., on the determiner or adjective). In contrast, L2 learners only showed sensitivity to violations between the determiner and the noun. The authors accounted for the divergence of results between the two groups with cross-linguistic differences in agreement rules (in this case, gender agreement in German is 'neutralised' in the plural whereas it is overtly realised in French on the adjective, with some exceptions). These results implied that L2 learners transferred the agreement rules from their L1 to their L2 and thus did not compute agreement in plural NPs in their L2, French. This argument supports both the idea of cross-linguistic competition, forwarded by Tokowicz and MacWhinney (2005), and of transfer, in the framework of the full-transfer full-access model (Schwartz & Sprouse, 1996).

Two recent studies using ERPs (Gillon-Dowens et al., 2010) and eye-tracking (Keating, 2009) investigated gender agreement processing between the noun and the predicative adjective in L2 Spanish. In both of these studies, sensitivity to agreement violations was compared for native Spanish speakers and English–Spanish late learners. While Gillon-Dowens et al. (2010) studied highly proficient learners who had been immersed for an extended period in their L2 (at least 12 years), Keating (2009) studied classroom learners, who were classified as beginner, intermediate and advanced English–Spanish late learners. In both studies, advanced L2 learners showed online sensitivity to grammatical gender agreement within the determiner phrase that was in many ways similar to that found for native speakers, suggesting that late L2 learners can reach native-like processing even for features that do not exist in the learner's L1. However, in line with the 'full access' theory (Schwartz & Sprouse, 1996), processing seems to be more rapidly acquired and more deeply processed for features that are shared across languages (similar conclusion were drawn in a more recent paper by Gillion-Dowens and colleagues with a group of Chinese–Spanish proficient learners for the same type of violations; Gillon Dowens, Guo, Guo, Barber, & Carreiras, 2011). Furthermore, the authors of both studies argued that processing complex structures might be more difficult for L2 speakers than for native speakers. In support of this, Gillion-Dowens et al. found that Spanish-L1 speakers and English-L1 late learners of Spanish showed similar patterns of ERPs for gender and number agreement violations between the determiner and noun in Spanish noun phrases (i.e., early negativity and P600 effect), but not for violations between the noun and the predicative adjective (i.e., early negativity and P600 effect for native speakers but only P600 effect for L2 speakers). Keating (2009) found a similar pattern (i.e., longer fixation times in the incorrect condition) for advanced L2 learners to Spanish native speakers when violations occurred within the determiner phrase but not when they occurred in non-local contexts. The authors suggested that L2 learners may be unable to keep the information required for gender processing in working memory and simultaneously process other elements in between the noun and the adjective (e.g., the verb).

In summary, the L2 online studies reported above have reached contrasting conclusions regarding late learners' capacity to acquire features that are not present in their L1. While some claim that after a critical age processing cannot reach native-like level unless the L1 and L2 systems overlap, others argue that age of acquisition may not be the crucial factor and that proficiency and exposure to L2 are equally important. In the present study we address three questions to further investigate L2 gender processing.

The present study

In the present study we investigated L2 learners' capacity to acquire new grammatical features (i.e., gender) and to process them online. We compared French native speakers to a group of English–French adult learners whose proficiency was high enough to permit them to attend courses in the French language, in a French university. We manipulated gender agreement between the noun and the adjective to examine L2 learners' sensitivity to agreement violations depending on both the position of the adjective in the sentence (pre-nominal vs. post-nominal attributive adjectives) and its syntactic role (attributive vs. predicative). Manipulating the position and syntactic role of the adjective allowed us to address different questions:

(a) Can late English–French learners acquire gender and process it online?

To investigate this question we manipulated gender agreement between the noun and the post-posed adjective, (Experiment 1) which is the canonical order for attributive adjectives in French and that which learners are most exposed to in the classroom as well as in everyday speech.

(b) Is gender agreement processing hindered when the surface order is similar in L1 and L2 but the language properties of the two languages differ?

As mentioned above, it has been suggested that processing within new structures is acquired more rapidly than in structures that are shared in L1 and L2 (Tokowicz & MacWhinney, 2005). English and French properties differ as French possesses gender but English does not. Thus, while adjectives can be positioned before the noun in both languages, they do not have to agree in gender in English, but they must in French. Hence, if L2 learners are able to acquire gender at all, it may be easier to process it in a new surface order (i.e., post-posed) than in a surface order that exists in both languages but do not share the same syntactic rules due to the language properties. In Experiment 2, we manipulated gender agreement violations

between the pre-posed adjective and the noun to investigate this question.

(c) Can L2 learners process gender when the two agreeing elements are separated by a syntactic boundary (i.e., predicative adjectives)?

According to the SSH (Clahsen & Felser, 2006), when L2 learners must process elements that intervene between the noun and the adjective, they are not able to retain the information required for gender processing in their working memory due to a lack of processing resources. In Experiments 3 and 4, we manipulated gender agreement violations between the noun and the predicative adjective but they were only separated by a copula to reduce the number of words between the noun and the adjective and to limit the complexity of the structure the adjective was embedded in (in contrast to Keating's study (2009).

To observe any potential differences in first and second language gender agreement processing we used ERPs and eye-tracking methodologies. These methodologies provide different and complementary information regarding the moment when difficulties occur in a sentence and how they are processed (Frenck-Mestre, 2005; Mueller, 2005). More details about eye-tracking and the advantage of combining this methodology to ERPs will be presented in the introduction of Experiment 4. In Experiments 1, 2 and 3, we used ERPs which allow one to record brain activity triggered by linguistic processing. Different types of linguistic analyses can be identified through the ERP components, which reflect different types of processing. For instance, difficulty in lexical/semantic processing is usually reflected by an increased negative deflection in the waveform, which peaks at 400 ms, known as the N400 effect (Kutas & Hilyard, 1980). Furthermore, this component has recently been suggested to reflect distinct stages of L2 grammatical acquisition. Studies have shown that when learners are still in the process of acquiring their L2, they may show an N400 effect in response to syntactic violations. With increased proficiency, this N400 effect gradually shifts to a classic P600 effect (Inoue & Osterhout, 2005; McLaughlin et al., 2010; Osterhout, McLaughlin, Kim, & Inoue, 2004; Osterhout, McLaughlin, Pitkänen, Frenck-Mestre, & Molinaro, 2006). In contrast, the P600 effect (Osterhout & Holcomb, 1992) is an increased positive deflection with a maximum peak at 600 ms generated by morphosyntactic violations, generally assumed to reflect some stage of syntactic processing. This component has been consistently found in studies investigating gender agreement in sentence context (Barber & Carreiras, 2005; Deutsch & Bentin, 2001; Frenck-Mestre et al., 2008, 2009, 2010; Foucart & Frenck-Mestre, 2011; Gillon-Dowens et al., 2010; Gunter et al., 2000; Hagoort & Brown, 1999) as reported above. In the present study, both components were of interest. We expected a P600 effect to gender agreement violations, as usually found in response to these errors for both native speakers and late L2 learners. In addition, however, the presence of an N400 effect to these violations, specifically in the group of L2 learners could also reflect early stages of online syntactic processing, as has now been repeatedly found (see the work by Osterhout and colleagues, as well as Morgan-Short, Sanz, Steinhauer, & Ullman, 2010).

Experiment 1

Our first experiment involved gender agreement within the DP, between the noun and post-posed adjective which, as outlined above, is the most frequent position for adjectives in French. In line with previous studies examining gender agreement violations within the DP, we expected a P600 effect for French native speakers, either preceded or not by a LAN effect (Barber & Carreiras, 2005; Deutsch & Bentin, 2001; Foucart & Frenck-Mestre, 2011; Frenck-Mestre, 2005; Frenck-Mestre et al., 2009; Frenck-Mestre, Carrasco-ortiz, McLaughlin, Osterhout, & Foucart, 2010; Gillon-Dowens et al., 2010; Gunter et al., 2000; Hagoort & Brown, 1999). For L2 learners, the presence of a P600 effect elicited by gender agreement violations would reveal that English speakers can process gender in French and thus that late learners have the ability to acquire new features in their L2 even if they are not present in their L1, in line with the 'full access' theory (Schwartz & Sprouse, 1996). Moreover, such a result would indicate that gender agreement is computed rapidly and automatically as well as systematically enough to produce a robust electrophysiological response.

Participants

Fourteen French native speakers and 14 English–French L2 learners, students at a French university, took part in the four experiments reported below. Participants received monetary compensation. To reduce the variability due to participants' knowledge of French, the same L2 learners participated in all the ERP experiments (1–3) and the majority (12 out of 14) also participated in the eye-movement experiment. The same was true for the native French controls. Experiments were conducted at least 2 months apart from each other. All participants were between 18 and 29 years old (mean 21.6 years). Their vision was normal or corrected-to-normal. English speakers had started learning French at secondary school (mean age of start of instruction: 13.4 years old, ranging from 10 to 18 years; mean length of formal learning: 8 years, ranging from 5 to 11) and had been studying at a French university as Erasmus students for a mean of 3 months. They had passed the required exam to attend courses in a French university (Diplôme d'Etudes de Langue Française (DELF); individual results not available). Four English speakers had some knowledge of Spanish and one of them had studied Chinese; they reported that their proficiency was low to intermediate. French native speakers had learned foreign languages at school (mainly English and Spanish), but their proficiency was not high enough to affect processing in their L1. To check English speakers' knowledge of lexical gender in French, after the experiment, non-native participants had to complete an offline test which consisted in circling the correct gender marked article of the words presented during the experiment (errors: 5.7%, *SD*: 3.8). 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 (written comprehension, 4.7; oral comprehension, 4.4; written production, 3.6; oral production, 3.9). Based on both number of years of study and successful completion of the DELF (level 2 or better) as well as self-report, participants' proficiency level was considered to be advanced.

Materials

Ninety-six concrete French nouns, 48 masculine and 48 feminine, of low to medium frequency (mean frequency per million: 35.2, masculine: 40.6, feminine: 29.8; *Lexique 2*, New, Pallier, Brysbaert, & Ferrand, 2004) and in length from 3 to 8 letters (mean 5.8, masculine: 5.6, feminine: 6) were selected. Nouns were selected such that they had a unique, unambiguous translation in English. No homographs or cognates across languages were included. In addition to the 96 critical nouns, 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) according to their semantic fit with the nouns. The critical adjectives were both orthographically and phonologically modified when inflected for the feminine (e.g., *vert*$_{\text{masc}}$/vɛʀ/; vs. *verte*$_{\text{fem}}$/vɛʀt/'green'). These 40 adjectives were paired with the 96 nouns, with each adjective being presented between 1 and 6 times (twice on average).

These 96 noun–adjective pairs were presented in sentence pairs, with one grammatical and one ungrammatical sentence in each pair. Grammaticality was determined by gender agreement between the noun and post-posed adjective. All sentences followed the same pattern: adverb (or adverbial phrase), plural definite article, noun, critical adjective, copula and complement (see Table 1). The plural form of the article was used, such that no gender information was provided (in French, the plural form of the definite determiner is identical for masculine and feminine; e.g., *le*$_{masc}$ *livre*$_{masc}$ ('the book'), *la*$_{fem}$ *table*$_{fem}$ ('the table') → *les livres, les tables*). In addition to the experimental sentences, 96 filler sentences were created. Filler sentences were all syntactically correct. The structure of filler sentences was similar to experimental sentences. All sentences were semantically acceptable as confirmed by French native speakers who pre-checked the acceptability of all sentences (experimental stimuli and fillers) in an off-line task.

Two lists were created such that each experimental noun–adjective pair was seen in both conditions (gender agreement vs. violation) but in only one condition per participant. In each list, there were 48 sentences per condition, defined by Gender Agreement (agreement vs. violation) and Noun Gender (masculine vs. feminine nouns). The sentences were distributed in a fixed-random order, with the restriction that no list began with an ungrammatical sentence and no more than two ungrammatical sentences followed each other, there being six fixed random orders per list. Each list began with four training sentences. Each participant saw only one list.

Table 1
Example of all the conditions presented in Experiment 1. The critical word is underlined.

Conditions	Examples
Correct	*Depuis une semaine, les chaises*$_{fem}$ *<u>vertes</u>*$_{fem}$ *sont dans le jardin*
Incorrect	*Depuis une semaine, les chaises*$_{fem}$ *<u>*verts*</u>$_{masc}$ *sont dans le jardin*
	Since last week the <u>*green*</u>$_{\text{fem/masc*}}$ chairs$_{\text{fem}}$ are in the garden

Procedure

Sentences were presented visually word by word, at a rate of 650 ms per word (500 ms presentation followed by 150 ms blank screen) in a single block of 192 sentences preceded by four training sentences. Following each sentence, a 'yes/no' prompt was presented at the offset of the last word, and participants were requested to judge whether the sentence was correct or not. Half of the participants made positive responses with the left hand; the other half used their right hand. In all four experiments, responses to the questions were recorded as a control; however, they cannot be taken as fully indicative of sensitivity to syntactic violations since participants' attention was drawn to the general correctness of the sentence and not syntactic acceptability alone. Participants were seated 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 of rest periods. A short break was provided in the middle of the experiment. The entire session lasted approximately 45 min.

EEG recording

EEG activity was recorded continuously from 21 scalp locations using tin electrodes attached to an elastic cap (Electrocap International). 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). Horizontal eye-movements were monitored by means of an electrode placed at the outer canthus of the right eye while blinks and vertical eye-movements were monitored via an electrode beneath the left eye. An electrode was placed over both the left and right mastoid; all electrodes were referenced to the left mastoid during the recording and then re-referenced to the average of the mastoid electrodes for analyses. The EEG was amplified with a bandpass of 0.1–40 Hz (3 dB cut-off) by means of an SAI Bioamp 32 channel Model and was digitised online at 200 Hz. EEG were later filtered below 15 Hz. The electrode impedance threshold value was set to 3 kΩ for scalp electrodes and 10 kΩ for face electrodes. Epochs began 100 ms prior to stimulus onset and continued 1100 ms thereafter. Average ERPs were formed off-line from trials free of muscular and/or ocular artefact and amplifier blocking (rejection was performed by a

computerised 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 by calculating the mean voltage amplitudes and peak latencies, for four time windows: 80–180, 160–280, 300–500 and 500–800 ms post-presentation of the critical adjective. 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, N400 and/or LAN, and P600 components that are frequently observed in these studies. The main components of interest, based on prior studies of grammatical gender, were the P600 (defined as the mean positive amplitude 500–800 ms post-stimulus) the N400, and the LAN (defined as the mean negative amplitude 300–500 post-stimulus for each, but with differing scalp distributions). Analyses were also performed on the earlier time windows to ascertain whether any differences emerged between sentence conditions. Prior to analyses, trials with artefact were rejected (French: 5.7% and 6.7%, English: 6.1% and 7.1% for correct and incorrect conditions, respectively, with no significant difference between groups). At midline, a three-way ANOVA was performed with Group (L1 French vs. L1 English–L2 French) as a between participant factor and repeated measures for Agreement (gender agreement vs. violation) and Electrode (Fz, Cz and Pz). At lateral sites, four-way ANOVAs were performed with Group as a between participant factor and repeated measures on Agreement, 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). Noun gender was not included in analyses as there is an inherent confound in French for length and gender marking on the adjective, such that correct inflections are systematically shorter for masculine than feminine forms. To avoid this confound, it suffices to pool the two genders for the correct and incorrect conditions. The Greenhouse and Geisser (1959) correction was applied to all repeated measures with greater than one degree of freedom. All significant differences involving more than two conditions were confirmed by post-hoc comparisons.

Results

Behavioural results

Participants made end of sentence judgements, based on the overall acceptability of sentences; the responses were not, thus, only driven by syntactic considerations. Nonetheless, we examined the data to ascertain participants' likelihood to reject sentences that contained a grammatical gender agreement violation. The data were subjected to a 2 (Group: French natives vs. English–French learners) × 2 (Gender agreement vs. disagreement) × 2 (Response: positive vs. negative) ANOVA. A significant interaction was found by Group × Agreement × Response ($F(1,26) = 69.13$, $p < .001$, $\eta^2 = .72$) showing that French native speakers rejected more sentences with gender agreement violations than L2 learners (95% and 40%, respectively) but the two groups accepted a similar number of correct sentences (79% and 72%, respectively). Native speakers, thus, showed greater offline sensitivity to agreement violations, even when not specifically requested to judge sentences for syntactic grammaticality. This difference might be due to memory cost; indeed, it might be more difficult for L2 than for L1 speakers to store agreement features and recall violations after processing the rest of the sentence. Their online sensitivity to these violations shows, nonetheless, that they were indeed able to compute them.

ERP 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. Both French native speakers (Figs. 1 and 2) and English–French learners (Figs. 3 and 4) showed a positive deflection in the waveform for sentences containing gender agreement errors between 500 and 800 ms after the onset of the critical adjective, corresponding to a P600 effect. These differences were confirmed in ANOVAs performed on the mean voltages obtained per participant for each sentence condition as a function of time window and electrode site. No significant differences emerged as a function of experimental factors prior to the 500–800 ms window after the target word.

In the P600 time window, an effect of Agreement was found at midline ($F(1,26) = 10.76$, $p = .001$, $\eta^2 = .29$, Correct $M = 0.50$, $SD = 0.33$, Incorrect $M = 1.77$, $SD = 0.51$), at centro-parietal sites ($F(1,26) = 15.45$, $p < .001$, $\eta^2 = .37$; Correct $M = 0.59$, $SD = 0.31$, Incorrect $M = 1.79$, $SD = 0.43$) and at anterior lateral sites ($F(1,26) = 11.29$, $p < .001$, $\eta^2 = .30$; Correct $M = 0.24$, $SD = 0.34$, Incorrect $M = 1.13$, $SD = 0.40$). At midline, there was a significant interaction of Group × Agreement × electrodes ($F(2,52) = 7.18$, $p < .001$, $\eta^2 = .21$), at centro-parietal sites of Group × Agreement × Hemisphere × Electrodes ($F(3,78) = 3.58$, $p < .01$, $\eta^2 = .12$), and at anterior sites of Group × Agreement × Hemisphere ($F(1,26) = 4.55$, $p < .05$, $\eta^2 = .15$). Post-hoc comparisons (Scheffé) revealed that the P600 effect showed a central-parietal distribution, with a larger response over posterior sites for native speakers, but larger at anterior sites for L2 learners. Given the interactions with Group, subsequent ANOVAs were performed on the different time windows 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 target word onset. In the 500–800 ms time-window a main effect of Agreement was significant at midline ($F(1,13) = 7.81$, $p < .01$, $\eta^2 = .37$, Correct $M = 0.74$, $SD = 1.71$) and at centro-parietal sites ($F(1,13) = 7.5$, $p < .01$, $\eta^2 = .37$; Correct $M = 0.95$, $SD = 1.79$, Incorrect $M = 2.51$, $SD = 2.59$). This effect tended towards significance at anterior sites ($F(1,13) = 3.14$, $p = .09$, $\eta^2 = .19$). Adjectives that disagreed in gender with the previous noun provoked a positive deflection relative to those that agreed in gender. At anterior lateral sites a significant interaction was found between Agreement × Hemisphere

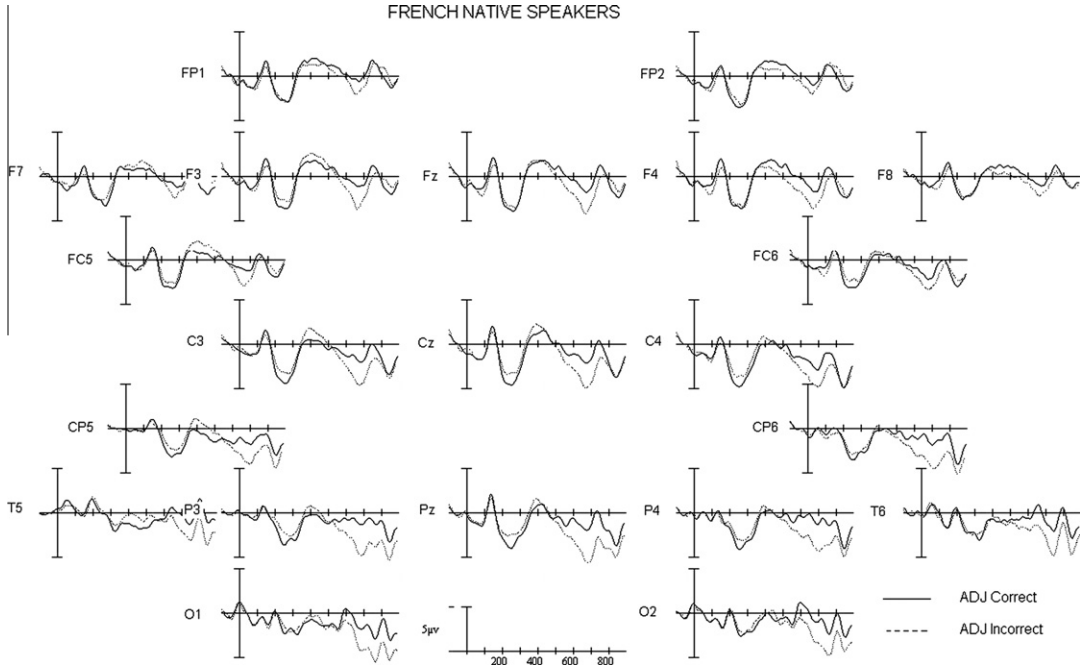

Fig. 1. Grand means – P600 effect for French native speakers in case of gender violations between the noun and the post-posed adjective.

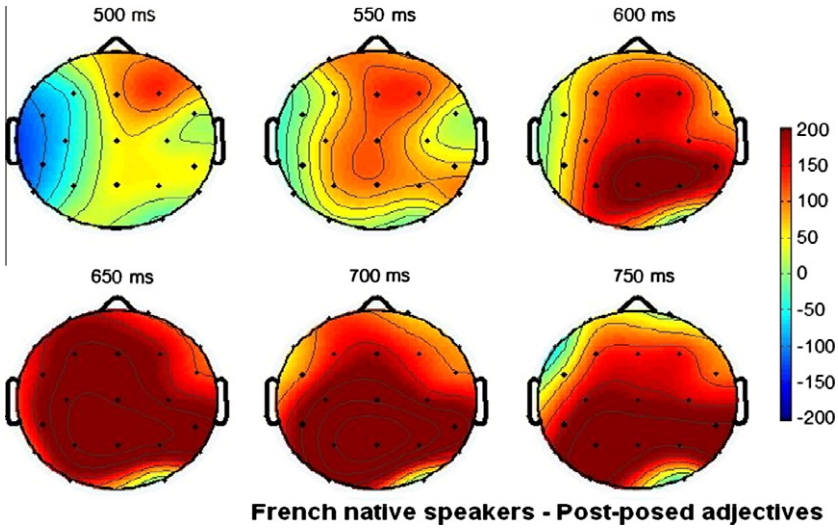

Fig. 2. Voltage map – P600 effect for French native speakers in case of gender violations between the noun and the post-posed adjective.

($F(1,13) = 6.47$, $p < .01$, $\eta^2 = .33$) showing a larger effect on the right than on the left anterior sites.

English–French learners. No significant effects emerged as a function of experimental factors prior to the 500–800 ms window after target word onset. In the 500–800 ms time-window a main effect of Agreement was found at anterior ($F(1,13) = 11.67$, $p < .001$, $\eta^2 = .47$, Correct $M = 0.26$, $SD = 0.62$, Incorrect $M = 1.26$, $SD = 0.66$) and centro-parietal sites ($F(1,13) = 15.38$, $p < .001$, $\eta^2 = .54$; Correct $M = 0.24$, $SD = 0.47$, Incorrect $M = 1.07$, $SD = 0.58$), and a trend at midline ($F(1,13) = 2.97$, $p = .10$, $\eta^2 = .18$), due to a P600 effect in response to agreement violations. At centro-parietal sites there was a significant interaction of Agreement × Hemisphere × Electrodes ($F(3,39) = 3.70$, $p < .01$, $\eta^2 = .22$) showing a larger effect on the left than on the right hemisphere.

Discussion

In the present experiment we manipulated gender agreement between the noun and the post-posed adjective, to investigate whether learners could process grammatical gender agreement online in their L2 despite the

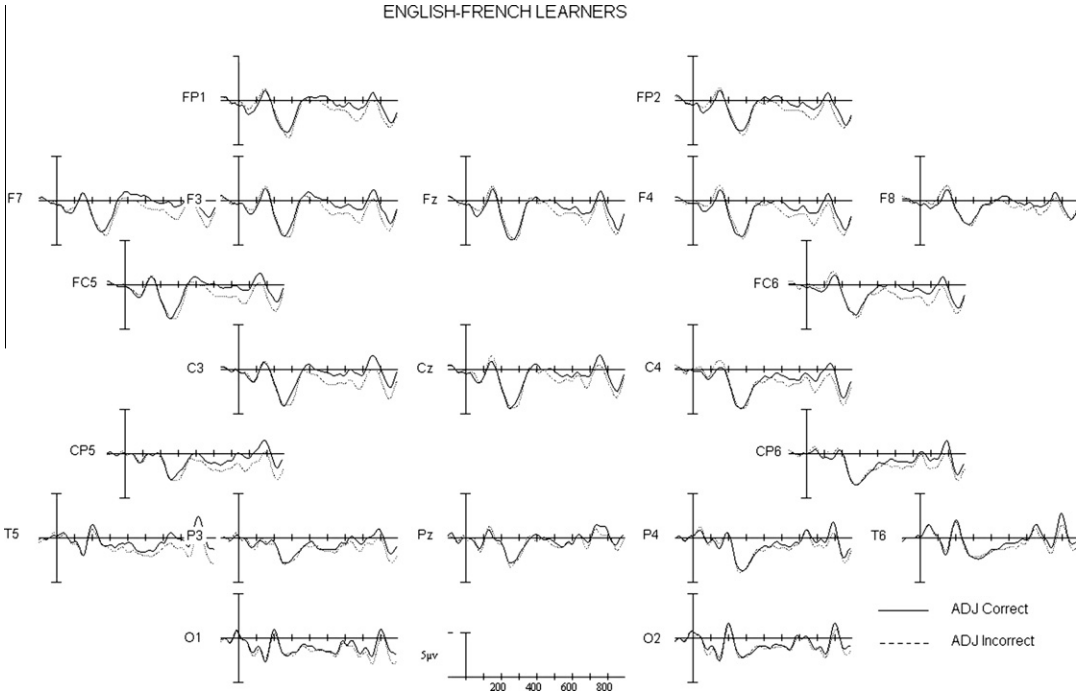

Fig. 3. Grand means – P600 effect for English–French learners in case of gender violations between the noun and the post-posed adjective.

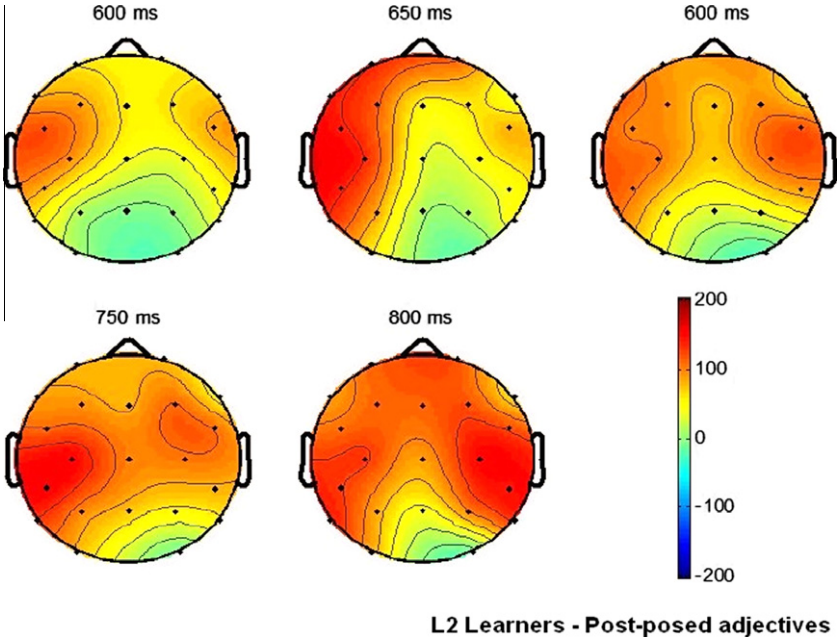

Fig. 4. Voltage map – P600 effect for English–French learners in case of gender violations between the noun and the post-posed adjective.

absence of this feature in their L1 and, if so, how similar their pattern of ERP response was to that of native speakers.

As predicted from the results obtained in previous studies (Hagoort & Brown, 1999; Osterhout & Mobley, 1995) native speakers displayed a P600 effect for gender agreement violations between the noun and the post-posed adjective. This effect was not preceded by any (early) differences in negativity as a function of gender agreement, similarly to previous studies (Foucart & Frenck-Mestre, 2011; Frenck-Mestre, 2005; Frenck-Mestre et al., 2008, 2009, 2010; Hagoort, Brown, & Groothusen, 1993;). The presence of a P600 effect suggests that grammatical gender is processed at the syntactic level.

In L2 learners, agreement errors also elicited a P600 effect with a similar latency to that found in native speakers, but with reduced amplitude and a more anterior distribution. No negativity preceded this effect. This is in line with the recent online evidence (Hahne, Mueller, & Clahsen, 2006; Rossi, Gugler, Friederici, & Hahne, 2006) that highly proficient L2 learners can reach native-like processing levels even if they learned their L2 late in life. These results will be further discussed in the framework of theories of syntactic processing in L2 in the general discussion.

To summarize, our L2 learners showed sensitivity to gender agreement violations between the noun and the post-posed adjective in the same manner as native speakers. However, as outlined in the introduction, the post-posed position is the most frequent order in French and hence learners are exposed to it more often than to the pre-posed position. Moreover, the post-posed position is a new surface order in L2 and therefore the syntactic rule related to that order (i.e., agreement between the noun and the post-posed adjective) does not conflict with any L1 rule. In contrast, pre-posed adjectives have the same surface order in French and English. Whether surface word order interacts with syntactic processing regarding gender agreement is an open question. To investigate whether such were indeed the case, we conducted a further experiment in which we manipulated agreement violations between the pre-posed adjective and the noun. For native speakers we expected the same pattern as for violations occurring between the noun and the post-posed adjective with, however, perhaps reduced amplitude of the P600 effect due to the lower frequency of this word order. For L2 learners, in view of the results of Experiment 1, we also expected a P600 effect. However, if the L2 learners are indeed less familiar with the pre-posed adjectives and/or if there is a conflict because the surface orders are similar in L1 and L2 but the syntactic rules are not due to differences in the grammatical features of the two languages (i.e., presence vs. absence of grammatical gender), then online computation of agreement may be more difficult thus leading to a reduction or even an absence of a P600 effect. Alternatively, and in line with the results of several ERP studies of late L2 learners, it is possible that we might observe an N400 effect to gender agreement errors in the L2 group if indeed this word order/agreement rule is less grammaticalized (Osterhout et al., 2004, 2008; McLaughlin et al., 2010).

Experiment 2

Method

Participants

See Experiment 1.

Materials

The same 96 nouns as in Experiment 1 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 letters, *Lexique 2*, New et al., 2004). The adjectives were selected such that they could be placed pre-nominally. Selection was also based on semantic fit with the critical nouns. The semantic acceptability of sentences was checked by French native speakers prior to the main experiment. Each adjective was seen between 1 and 6 times (twice on average).

Each of these 96 adjective–noun combinations was seen in a sentence pair, with one grammatical and one ungrammatical version in each pair. Grammaticality was determined by gender agreement between the pre-posed adjective and the noun. All sentences followed the same pattern: adverb (or adverbial phrase), plural definite article, pre-posed adjective, critical noun, copula and complement (see Table 2). The plural form of the article was used so that no gender information was provided by the determiner. Two lists were created such that each experimental adjective–noun pair was seen in both conditions (gender agreement vs. disagreement) but in only one condition per participant. In each list, 48 sentences were presented per condition, defined by Gender Agreement (gender agreement vs. mismatch) and Noun Gender (masculine vs. feminine). Two lists were created such that all critical nouns were seen in both gender agreement conditions but in only one condition per list. The sentences were presented in a fixed-random order, and six fixed random orders were created per list. The experiment was preceded by four warm-up sentences. Participants saw only one list.

Procedure and EEG recording

These were identical to Experiment 1.

Data analysis

This was identical to Experiment 1 with the exception of the P600 time window, which was reduced to a shorter period, i.e., from 500 to 700 ms, due to visual inspection which revealed both a shorter latency and smaller amplitude in comparison to Experiment 1. Prior to analyses, trials with artefact were rejected (French: 1.4% and 1.1%, English: 1.4% and 1.9% for correct and incorrect conditions, respectively and no significant difference emerged between groups). Grand averages and preliminary analyses for the factor Noun gender (masculine vs. feminine) revealed no differences ($F < 1$), hence, this factor was not included in the analyses.

Results

Behavioural results

As in Experiment 1, analyses were conducted on behavioural responses to sentence acceptability. No group

Table 2
Example of all the conditions presented in Experiment 2. The critical word is underlined.

Conditions	Examples
Correct	*De nos jours, les anciennes$_{fem}$ montres$_{fem}$ sont rares*
Incorrect	*De nos jours, les anciens$^{*}_{masc}$ montres$_{fem}$ sont rares*
	Nowadays old watches are rare

interaction was found – French native speakers and L2 learners rejected a comparable number of sentences that contained a grammatical gender agreement violation (41% and 30%, respectively) and accepted a roughly similar number of correct sentences (84% and 73%, respectively). Note, that when the agreement violation occurred in the non-canonical adjective order (Experiment 2), the number of rejected sentences that contain a violation was reduced for native speakers, as demonstrated by the significant interaction between Experiments 1 and 2 × Response for rejected ungrammatical sentences ($F(1,13) = 29.70$, $p < .001$, $\eta^2 = .69$).

ERP results

Gender agreement violations between the adjective and the following noun provoked differences in the waveforms that varied as a function of participant group and time window. Descriptively, the grand mean for French native speakers (Figs. 5 and 6) showed a positive deflection in the waveforms for agreement violations between 500 and 700 ms after the onset of the critical noun. English–French learners (Figs. 7 and 8) revealed a negative deflection in the waveform for agreement violations in the 80–180 ms and 300–500 ms time-windows. Subsequent ANOVAs confirmed these effects.

In the N100 (80–180 ms) time window, a main effect of Agreement was found at centro-parietal sites ($F(1,26) = 6.66$, $p < .01$, $\eta^2 = .20$, Correct $M = 0.87$, $SD = 0.37$), with a trend at anterior sites ($F(1,26) = 3.21$, $p = .08$, $\eta^2 = 11$). A trend towards significance emerged at centro-parietal sites for the interaction between Group × Agreement ($F(1,26) = 3.29$, $p = .08$. $\eta^2 = 11$).

In the N400 (300–500 ms) time window, a significant interaction of Group × Agreement was obtained at midline ($F(1,26) = 12.30$, $p < .001$, $\eta^2 = .27$) as well as at anterior ($F(1,26) = 6.31$, $p < .01$, $\eta^2 = .19$) and centro-parietal ($F(1,26) = 7.77$, $p < .001$, $\eta^2 = .23$) sites. This interaction was confirmed by post-hoc analyses (Scheffé test) which showed a wide spread negativity elicited by agreement violations for English–French learners at all sites ($p < .001$), but not for native speakers for whom no variation was found as a function of Agreement. Both the onset and distribution of the negativity observed in the learner group are suggestive of an N400 effect.

In the P600 time window (500–700 ms) a main effect of Agreement emerged at midline ($F(1,26 = 8.56$, $p < .001$, $\eta^2 = .25$; Correct $M = -0.09$, $SD = 2.11$, Incorrect $M = 0.70$, $SD = 2.70$) and centro-parietal sites ($F(1,26) = 4.69$, $p < .05$, $\eta^2 = .15$; Correct $M = 0.66$, $SD = 1.33$, Incorrect $M = 1.07$, $SD = 1.94$). A significant interaction of Group × Agreement was found at midline ($F(1,26) = 8.65$, $p < .001$, $\eta^2 = 25$), anterior ($F(1,26) = 6.04$, $p < .01$, $\eta^2 = .18$) and centro-parietal ($F(1,26) = 10.11$, $p < .001$, $\eta^2 = .27$) sites. Post-hoc analyses (Scheffé test) confirmed the presence of a P600 effect to agreement violations for French native speakers at all sites ($p < .001$) but not for English learners. No other effects or interactions were significant.

Given the interactions with Group, subsequent ANOVAs were performed on the ERP data for each group independently.

French native speakers. Gender violations between the adjective and the following noun provoked a P600 effect. A significant effect of Agreement appeared 500–700 ms after the onset of the critical noun at midline ($F(1,13) = 13.72$, $p < .001$, $\eta^2 = .51$; Correct $M = -0.09$, $SD = 1.79$, Incorrect $M = 1.48$, $SD = 3.13$), anterior ($F(1,13) = 4.72$, $p < .05$, $\eta^2 = .27$; Correct $M = -0.36$, $SD = 1.08$, Incorrect

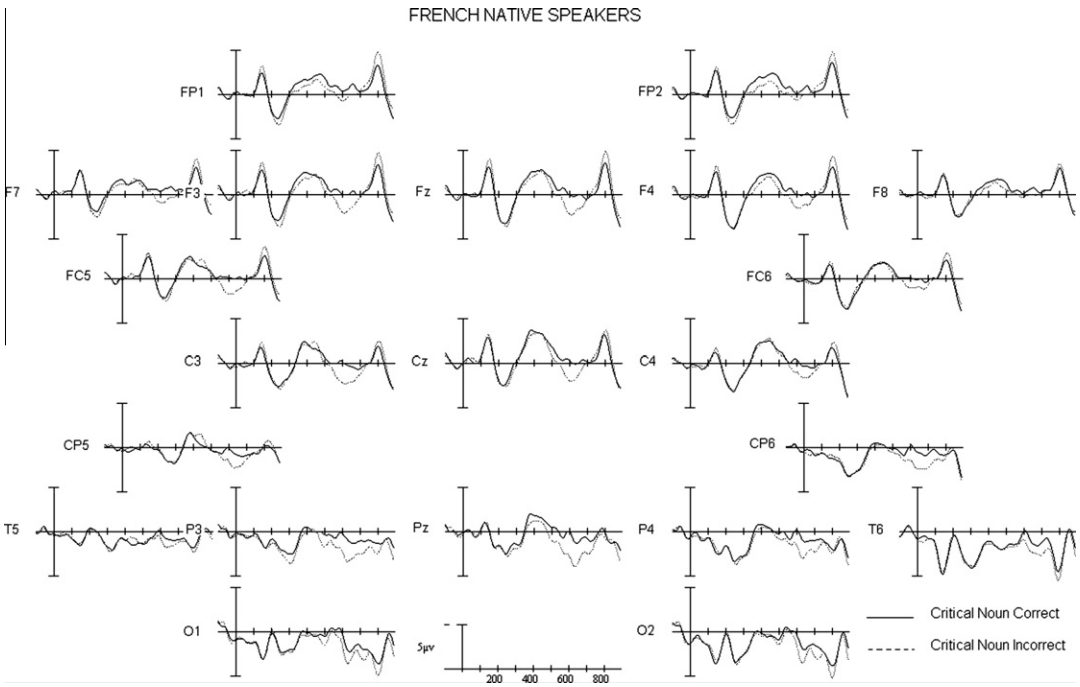

Fig. 5. Grand means – P600 effect for French native speakers in case of gender violations between the pre-posed adjective and the critical noun.

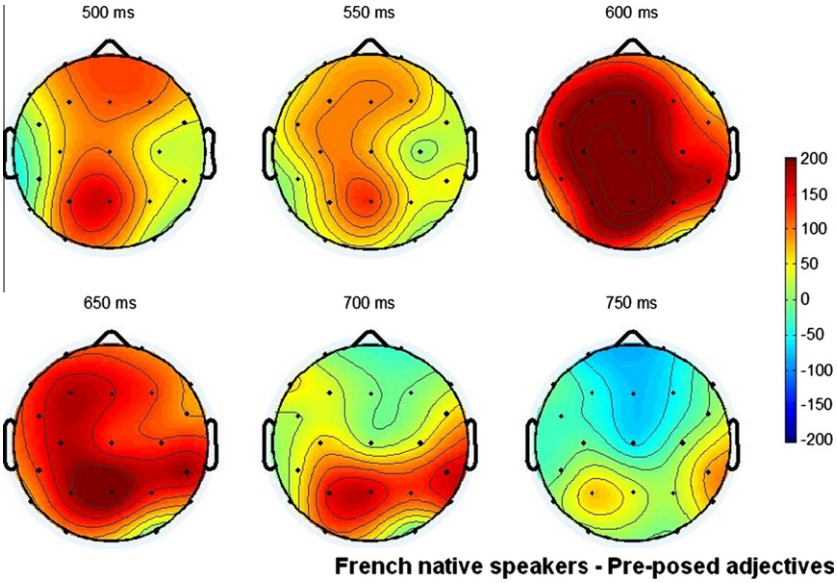

Fig. 6. Grand means – P600 effect for French native speakers in case of gender violations between the pre-posed adjective and the critical noun.

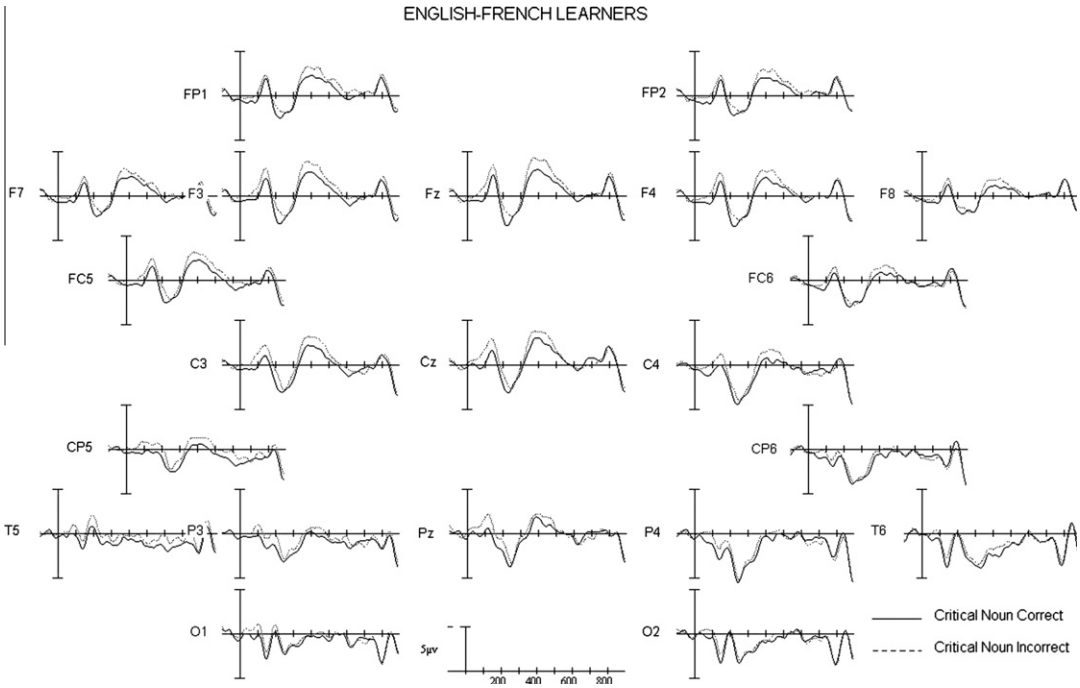

Fig. 7. Grand means – English–French learners' ERP response to gender agreement violations between the pre-posed adjective and the critical noun.

$M = 0.58$, $SD = 2.12$) and centro-parietal ($F(1,13) = 10.54$, $p < .001$, $\eta^2 = .45$; Correct $M = 0.52$, $SD = 1.43$, Incorrect $M = 1.53$, $SD = 2.45$) sites. A significant interaction of Agreement and Electrode emerged at anterior ($F(2,26) = 3.82$, $p < .05$, $\eta^2 = .22$) and centro-parietal ($F(3,39) = 3.98$, $p < .01$, $\eta^2 = .23$) sites. Post-hoc analyses (Scheffé) revealed that the effect of agreement was larger at posterior and left hemisphere sites.

English–French learners. Gender violations between the pre-posed adjective and the noun provoked negative deflections in the waveform in two time windows: 100–180 ms (N100) and 300–500 ms (N400) after the onset of the critical noun. In both time windows, a significant effect of Agreement was found at all sites: N100 (midline ($F(1,13) = 11.82$, $p < .001$, $\eta^2 = .48$; Correct $M = -0.36$, $SD = 2.17$, Incorrect $M = -1.68$, $SD = 2.38$); anterior

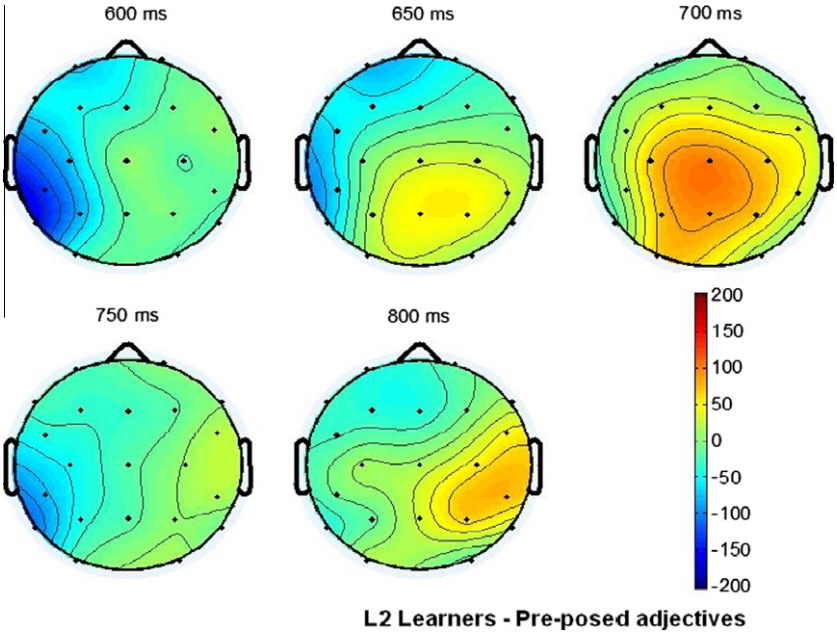

Fig. 8. Voltage map – English–French learners' ERP response to gender agreement violations between the pre-posed adjective and the critical noun.

($F(1,13) = 14.11$, $p < .001$, $\eta^2 = .52$; Correct $M = -0.36$, $SD = 1.95$, Incorrect $M = -1.21$, $SD = 1.86$); centro-parietal (($F(1,13) = 10.47$, $p < .001$, $\eta^2 = .45$; Correct $M = 0.93$, $SD = 2.22$, Incorrect $M = -0.08$, $SD = 1.90$)); N400 (midline ($F(1,13) = 12.92$, $p < .001$, $\eta^2 = .50$; Correct $M = -1.55$, $SD = 2.62$, Incorrect $M = -2.42$, $SD = 2.42$), anterior ($F(1,11) = 5.87$, $p < .05$, $\eta^2 = .31$; Correct $M = -1.12$, $SD = 2.10$, Incorrect $M = -1.83$, $SD = 1.91$) and centro-parietal sites ($F(1,13) = 12.8$, $p < .001$, $\eta^2 = .50$; Correct $M = 0.72$, $SD = 2.23$, Incorrect $M = 0.02$, $SD = 2.14$).

Discussion

In the second experiment, we manipulated gender agreement violations between the pre-posed adjective and the noun to investigate whether L2 learners' sensitivity to agreement violations was similar when violations occurred on adjectives in canonical word order (i.e., post-posed) and when they occurred on adjectives in a less frequent position (i.e., pre-posed). Moreover, we examined whether the fact that the surface order was shared in both languages but the syntactic rules differed in French and in English due to the language-specific properties affected online gender processing in L2; in other words, whether the fact that adjectives can be positioned before the noun in both languages but only agree in gender in French affected agreement processing in L2.

Our results for French native speakers suggest that gender agreement is processed in the same way independently of the position of the adjective. Indeed, French native speakers displayed a P600 effect in response to gender agreement violations between the pre-posed adjective and the noun similar to that obtained in response to violations between the noun and the post-posed adjective. To ascertain whether any differences were obtained as a function of word order, we compared the ERP data for the French participants from the two experiments. The results from this analysis showed no interaction between agreement and experiment (at midline for the interaction between factors Experiments × Agreement, $F(1,26) = 0.16$, p = n.s).

Concerning the L2 learner group, our results suggest that despite the similarity in surface word order across French and English for pre-posed adjectives our participants had greater difficulty processing gender agreement for this word order. Rather than the typical P600 effect, our native English speakers showed an N400 effect for gender agreement violations between the pre-posed adjective and the noun. This effect was preceded by a very early negativity, the implication of which is not clear (note that this early effect was not due to a carryover effect on the adjective preceding the critical noun). These two effects will be discussed independently.

Recently, several studies investigating native processing have reported very early ERP effects in response to syntactic manipulations; these early effects were not always followed by a later positivity (Hasting & Kotz, 2008; Hasting, Kotz, & Friederici, 2007; Pulvermüller & Shtyrov, 2003; Pulvermüller, Shtyrov, Hasting, & Carlyon, 2008). It is important to note, however, that these effects have been obtained mainly for verbal agreement manipulations (but see Malaia, Wilbur, & Weber-Fox, 2009) which may indeed differ from the gender agreement manipulations investigated here and, importantly, these early effects have usually been observed in less complex contexts involving only a pronoun and an inflected verb and in a mismatch paradigm (but see Hasting & Kotz, 2008; Malaia et al., 2009). Thus, it may be less surprising to find such a rapid response to violations under those conditions than here. Whether the early negativity we observed in the group of L2 learners in response to gender agreement violations reflects sensitivity to syntactic agreement is questionable

since no such response was found in our control group of native French speakers. Further work is necessary to elucidate this effect.

The N400 effect found in the L2 group differed from the classic P600 effect triggered by agreement violations and which we indeed found in the group of native speakers. It is possible that the N400 reflects early stages of syntactic processing. As mentioned in the introduction, recent studies have shown that when learners are still in the process of acquiring their L2, they produce an N400 in response to syntactic violations that gradually shifts to a classic P600 effect as proficiency increases (Inoue & Osterhout, 2005; McLaughlin et al., 2010; Osterhout et al., 2004, 2006). Hence, it is possible that our participants were still in the process of acquiring agreement between the preposed adjective and the noun, despite their advanced proficiency. This possibility will be further discussed in the general discussion.

Experiment 3

Overall, the results of Experiments 1 and 2 suggest that English–French learners can acquire gender and process gender agreement in local contexts, with varying sensitivity to agreement violations depending on the position of the adjective within the DP. In Experiment 3, we examined whether L2 learners could process gender agreement across a syntactic boundary, i.e., for predicative adjectives. Even in monolinguals, it has been demonstrated that processing agreement between a noun and a predicative adjective is more costly than within the DP (Barber & Carreiras, 2005). The SSH (Clahsen & Felser, 2006) suggests that L2 learners should experience greater difficulty processing gender agreement between distanced elements due to limited processing resources when reading in the L2 and the inability to retain information related to gender in working memory while processing other lexical items/relations between the noun and the predicative adjective. If such is the case, L2 learners should show a reduced or even null response to agreement violations for predicative adjectives. This is indeed what was obtained for English–Spanish learners in Keating's (2009) study. However, in Keating's study, the noun and predicative adjective were separated by an adverb in addition to the copula, and the grammatical structure in which the adjective was embedded was more complex. The present experiment tested whether L2 learners would show online sensitivity to violations of gender agreement within the simplest case, i.e., where the noun and predicative adjective were separated only by the copula and were within the same main clause.

Method

Participants

See Experiment 1.

Materials

Stimuli consisted of the same 96 nouns used in Experiments 1 and 2. Thirty-nine adjectives were selected such that they matched semantically with the noun (mean frequency per million: 105.4, mean length: 5.4, range between 3 and 8, *Lexique 2*, New et al., 2004). Each adjective was seen between 1 and 6 times (twice, on average). These noun adjective pairs were embedded in predicative adjective structures in short sentence contexts. All sentences followed the same pattern: adverb (or adverbial phrase), plural definite article, noun, copula, critical predicative adjective and complement. Each noun–predicative adjective combination was seen in a sentence pair, with one grammatical and one ungrammatical sentence in each pair. Grammaticality was determined by gender agreement between the noun and predicative adjective (see example in Table 3). The semantic acceptability of the sentences was checked by French native speakers prior to the experiment. Forty-eight sentences were presented per condition, defined by Gender Agreement (gender agreement vs. mismatch between the noun and the critical predicative adjective) and Noun Gender (masculine vs. feminine nouns). A Latin square design was used such that the same sentence was seen in both correct and incorrect conditions but not by the same participant. In addition to the 96 experimental sentences, 96 syntactically correct and anomalous filler sentences were created. The order of presentation of sentences was randomised for each participant. The experiment proper was preceded by four training sentences.

Procedure and EEG recording

These were identical to Experiments 1 and 2.

Data analysis

This was identical to Experiment 2. Prior to analyses, trials with artefact were rejected (French: 1.2% and 0.7%, English: 0.7% and 1.2% for correct and incorrect conditions, respectively; no significant difference emerged between conditions or groups).

Results

Behavioural results

Similarly to Experiment 1, analyses of the sentence acceptability task revealed a significant interaction of Agreement × Response × Group ($F(1,26) = 4.33$, $p < .05$, $\eta^2 = .14$), showing that French native speakers and L2 learners accepted a comparable number of correct sentences (87% and 81%, respectively), but that natives speakers rejected more sentences with gender agreement violations than L2 learners (44% and 25%, respectively). These results confirm that, even though the task intentionally did not

Table 3
Example of all the conditions presented in Experiment 3. The critical word is underlined.

Conditions	Examples
Correct	*Au printemps, les pommes$_{fem}$ sont <u>vertes</u>$_{fem}$ sur cet arbre*
Incorrect	*Au printemps, les pommes$_{fem}$ sont <u>verts</u>*$_{masc}$ sur cet arbre*
	In spring apples are <u>green</u> on this tree

focus on grammaticality, native speakers showed greater offline sensitivity to agreement violations.

ERP results

The grand means revealed differences in the waveforms for predicative adjectives that agreed in gender with the preceding nouns compared to those that did not. Descriptively, waveforms for French native speakers (Figs. 9 and 10) revealed a positive deflection between 500 and 700 ms after the onset of the critical adjective in the case of gender mismatch, whereas English–French learners (Figs. 11 and 12) showed no apparent differences based on agreement. ANOVAs were performed on these data. No main effects or interactions obtained prior to the P600 (500–700 ms) time-window in which a main effect of Agreement emerged at midline ($F(1,26) = 4.39$, $p < .05$, $\eta^2 = .14$; Correct $M = 0.63$, $SD = 2.43$, Incorrect $M = 1.33$, $SD = 3.08$), anterior ($F(1,26) = 3.14$, $p = .09$, $\eta^2 = .10$) and centro-parietal ($F(1,26) = 0.52$, $p = .06$, $\eta^2 = .12$) sites. The analysis revealed a trend for an interaction with Group at midline $F(1,26) = 3.11$, $p = .09$. This was due to individual differences in the L2 groups showing that some participants revealed a positivity to violations that started early and continued but did not resemble the classic P600 effect in distribution and latency.

Discussion

In Experiment 3, we manipulated gender agreement between the noun and the predicative adjective to investigate whether L2 learners were able to process gender agreement when the noun and the adjective are separated by a syntactic boundary. The results for the L2 group were mixed; some participants showed no consistent cortical response to gender agreement for predicative adjectives while others produced an early and sustained positivity that might be a transitional response, which should become a true P600 effect with greater exposure and/or experience. Overall, the results for the L2 group showed no consistent or language-related ERP response to gender agreement, in contrast to those of native control group, showing a small but robust P600 effect.

The pattern of results obtained for the L2 learners contrasts with that reported in Gillon-Dowens and colleagues' (2010) study, in which a small but significant P600 effect was found for English–Spanish learners for gender agreement violations. It should be underlined, however, that the number of years of exposure to the L2 was vastly greater in the latter study. Our results are more in line with those reported by Keating (2009), in his eye movement study of gender agreement processing for different levels of English–Spanish learners. The overall pattern of results was complex however, as in our study, no clear online sensitivity to gender agreement in a late learned L2 was found for predicative adjectives although the inspection of individual results suggested that some participants were indeed responsive.

It is difficult to draw definite conclusions from the present ERP results given the large variability in response we observed in the L2 group. Given such, we further investigated gender agreement in predicative adjectives with a different methodology. We ran the same experiment with eye-tracking, which allowed us both to directly compare the sensitivity of ERPs and eye-movements as a measure

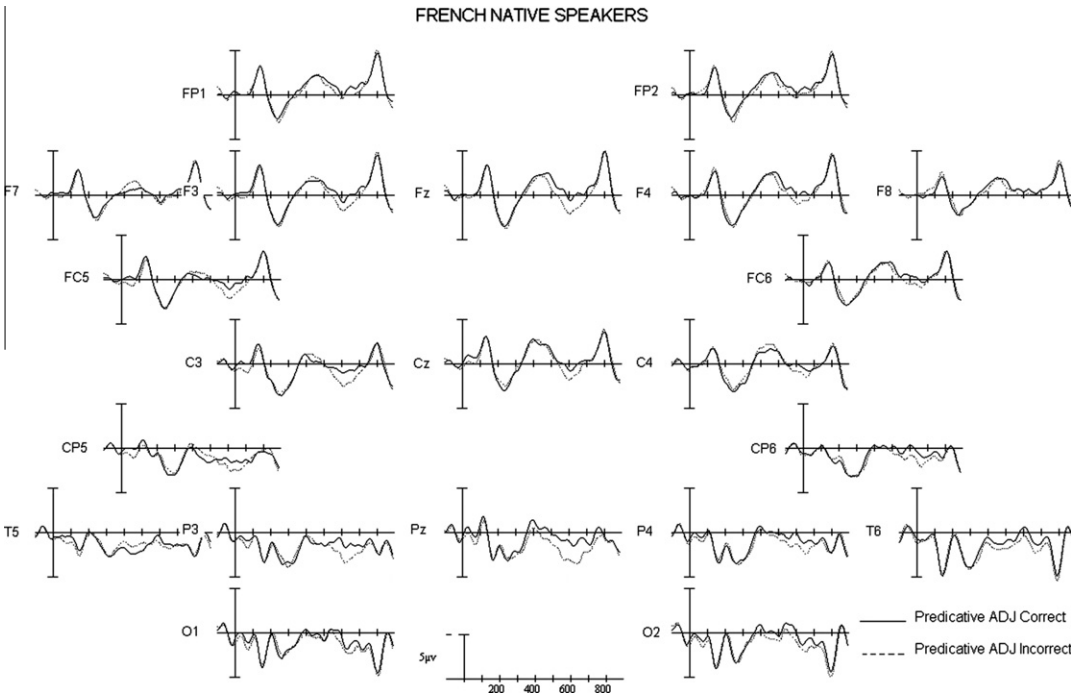

Fig. 9. Grand means – P600 effect for French native speakers in case of gender violations between the noun and the predicative adjective.

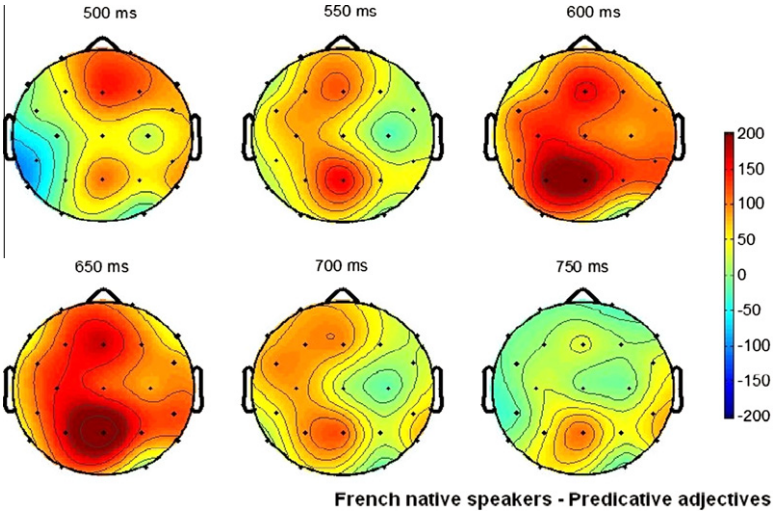

Fig. 10. Voltage map – P600 effect for French native speakers in case of gender violations between the noun and the predicative adjective.

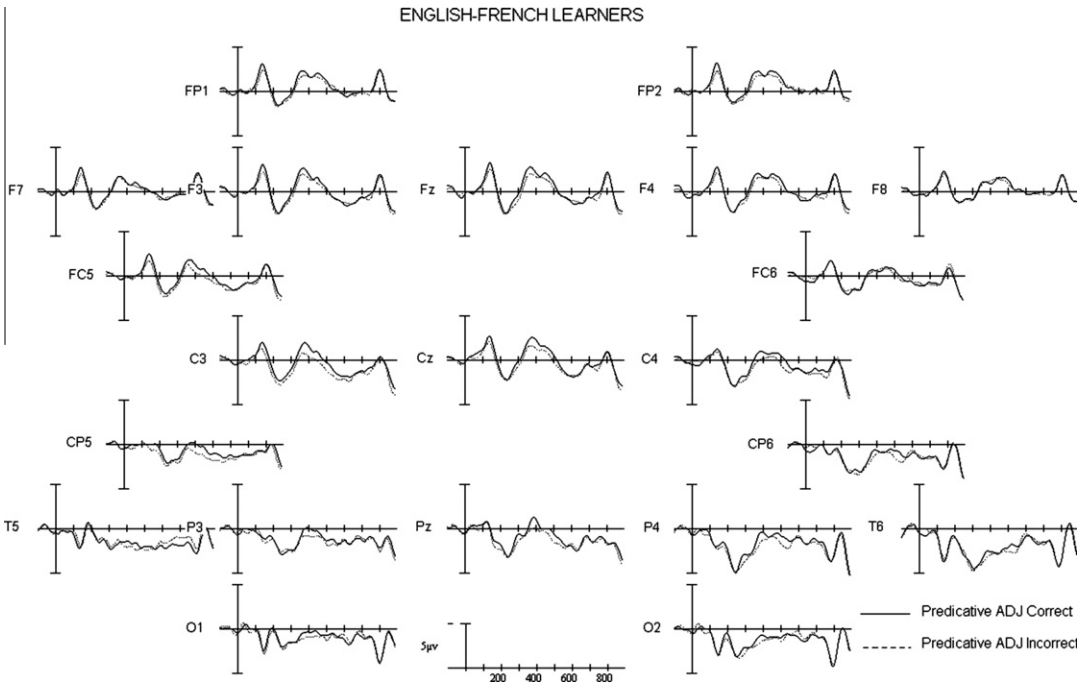

Fig. 11. Grand means – English–French learners' ERP response to gender agreement violations between the noun and the predicative adjective.

of online processing and to compare our findings with recent eye-movement studies of L2 gender processing (Keating, 2009).

Experiment 4

Eye-tracking is a valuable methodology for psycholinguistics as it provides an online record of the processes involved in 'natural' reading. In contrast to the typical ERP paradigm where sentences are presented word by word, eye-tracking allows complete sentences to be presented and reading speed to be controlled by the reader. Hence, readers can both control the amount of time they spend on a particular word and re-read a region, neither of which is possible in the ERP paradigm. Both ERP and eye-movement methodologies can detect very subtle effects; they reveal where difficulties occur during syntactic processing and both the extent (for eye-tracking) as well as the nature (for ERPs) of the difficulty. However, studies using either one or the other methodology have shown different results for L2 processing (Frenck-Mestre, 2005). While

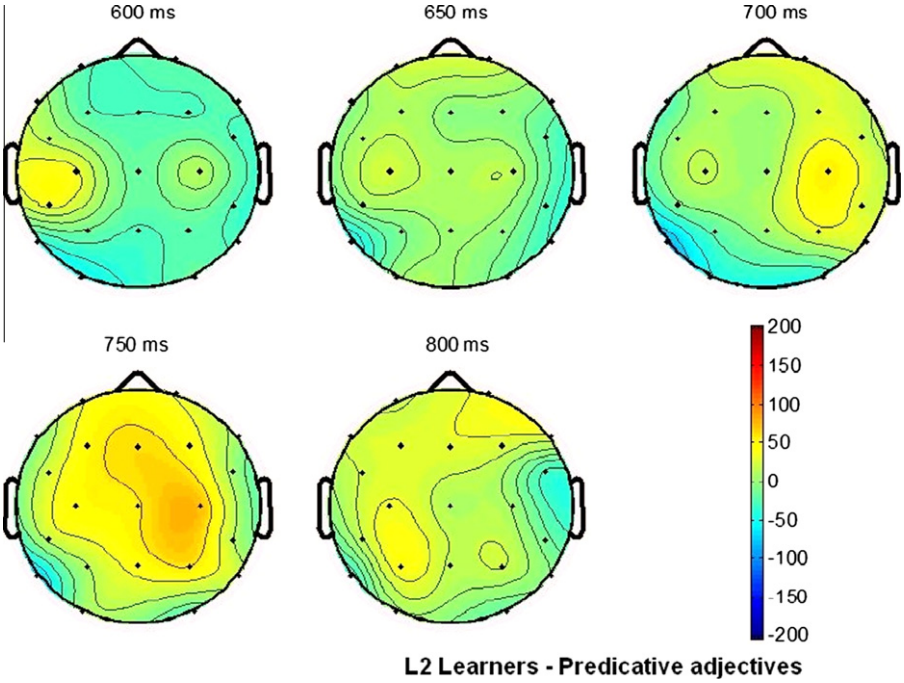

Fig. 12. Voltage map – English–French learners' ERP response to gender agreement violations between the noun and the predicative adjective.

eye-tracking studies have suggested similar syntactic parsing in L1 and L2, ERPs have quite often revealed different effects for native and non-native speakers, which are nonetheless attenuated by increased L2 proficiency, independent of age of acquisition. These discrepancies can be accounted for by several differences across the ERP and eye-movement studies to date. First, the types of results obtained with the methodologies are not similar. Indeed, both ERPs and eye-movements provide information on *when* and *where* in the sentence the difficulty is processed however, only ERP components can reveal the *nature* of the processing involved in the resolution of the difficulty via the type of component that is evoked (even though the interpretation of ERP components remains a question of debate). Second, the types of syntactic processing investigated were not the same for the two methodologies. ERP studies have mainly examined syntactic anomalies, whereas eye-movement studies have predominantly observed syntactic ambiguities (but see Keating, 2009).

These differences underline the need to combine ERPs and eye-movements to study syntactic processing in a second language. To our knowledge, only one monolingual study has used these two methodologies with the same materials (Deutsch & Bentin, 2001). The study examined the possible interaction between syntactic and semantic processing by examining gender agreement in Hebrew. The results were consistent across methodologies. The effects obtained were complementary and allowed the authors to draw conclusions and discuss their findings in relation to interactive, constraint-based models of online sentence processing (MacDonald & Seidenberg, 2006). In the present experiment, we used the same materials as in Experiment 3 but with eye-tracking. Our aim was to further explore gender agreement processing in the L2 when the noun and the adjective were separated by a syntactic boundary, and to determine whether eye-tracking might reveal a different pattern of results than what we obtained in our ERP study (Experiment 3).

Method

Participants

The majority of the participants that took part in this experiment had also participated in Experiment 3 (12 of the 14 English–French learners and 11 of the French native controls).

Materials

The materials were those used in Experiment 3. Both feminine and masculine nouns were present in the materials however, noun gender was not taken into account in the analyses of predicative adjectives given the inherent confound between length and gender marking in French. Adjectives are systematically longer when inflected for the feminine than the masculine, by at least one letter (e.g. *petite* vs. *petit* "small", *longue* vs. *long* "long"). As such, for the masculine the mean length of correctly inflected adjectives is shorter than for incorrectly inflected adjectives while the inverse is true for the feminine (for our materials, mean length was 5.92 and 7.54 for correct and incorrect masculine, respectively, and 7.01 and 5.88 for correct and incorrect feminine, respectively. When the masculine and feminine are considered together, this confound with length no longer holds (for our materials: 6.5 and 6.7 for correctly and incorrectly inflected cases,

respectively). One pair of adjectives (*puissants/puissantes* "powerful") was not included in analyses as it was the only case in which the adjective surpassed nine letters, specifically in the feminine.

Apparatus and procedure

Eye movements were recorded using a Dr. Bouis eye-tracker. Horizontal signals from the right eye of the participant were sampled every 5 ms. Participants sat with their heads restrained by means of a bite-bar and adjustable head and chin rests 60 cm from a CRT screen, such that 3.25 character spaces subtended 1° of visual angle. Initial calibration was achieved via five fixation points placed equidistantly across the screen. Thereafter, re-calibrations were performed every four sentences. Participants were requested not to move during reading. Sentences were presented individually on a single line in the centre of the screen. Prior to sentence presentation, participants fixated a point on the left of the screen; once the eye was detected, the sentence was presented. Participants read the sentences silently and assessed their syntactic and/or semantic acceptability. Data analysis was performed off-line and all trials contaminated by head movement were discarded. The session took approximately 35 min.

Data analysis

For purposes of analysis, sentences were divided into six regions of interest: (R1) an adverb plus the plural definite article, (R2) the noun, (R3) the copula, (R4) the critical adjective, (R5) the spill over region and (R6) sentence end. An example is provided below.

R1	R2	R3	R4	R5	R6
(1) Au printemps les/pommes/sont/vertes/sur cet/arbre.					
In spring apples are green on this tree.					

Each region was a minimum of five characters spaces to reduce the likelihood that the region was skipped. Despite such, the copula (R3) was skipped on more than 80% of trials; hence, the copula was collapsed with the following region (R3 + R4). When either the critical noun (R2) or the critical region of interest, i.e., the predicative adjective (R4), was skipped, the sentence was excluded from analyses. First pass (first fixation and gaze duration) and total reading times for each of the above regions are reported in Table 4.

Independent analyses were performed on the data for each region of interest. Two-way ANOVAs were conducted on both reading times and number of regressions for each region, with repeated measures for Agreement (gender agreement vs. gender mismatch) and Group (native French vs. English–French learners) as a between participant factor. No differences emerged as a function of experimental factors prior to the critical predicative adjective, nor for the spill-over region (R5) or sentence end (R6). The only region that revealed significant effects was the predicative adjective (R4, combined with the 20% of fixations in R3), Results are reported for first fixation durations and first pass gaze durations (the sum of all fixations from when the reader's eye initially entered a region from the left until it exited the region) as well as for total reading times (i.e., all fixations in a region) for this region, below. In addition, we analysed both first pass and second pass regressions, from the critical adjective to prior regions of the sentence.

Results

Behavioural results

As in the ERP experiments, analyses were conducted on behavioural responses to sentence acceptability. No group interaction was found – comparable percentages were found for both groups for acceptance of correct sentences (86% and 88%, for French and L2 learners, respectively) and rejection of sentences that contained a grammatical gender agreement violation (77% and 82%, for French and L2 learners, respectively). Note, that the correct rejection rate for ungrammatical sentences is higher than in Experiment 3 undoubtedly in part because with eye-tracking, participants had the possibility both to dwell on any given region as well as re-read the sentence before making a decision, which was not the case with ERPs.

Reading times

First pass reading times

First fixation duration. Mean first fixation durations are reported for the critical predicative adjective in Table 4. No significant effects or interactions were observed for this measure (Group: $F_1 < 1$; $F_2(1,46) = 1.49$, $p = .22$) Agreement: $F_1(1,22) = 2.28$, $p = .14$; $F_2(1,46) = 2.36$, $p = .13$; Group × Agreement: F_1 and $F_2 < 1$).

Gaze duration. Mean gaze durations are reported for the critical predicative adjective in Table 4. Analyses revealed an effect of Agreement ($F_1(1,22) = 13.86$, $p < .001$, $\eta^2 = .39$; $F_2(1,46) = 4.51$, $p < .04$, $\eta^2 = .09$) and an effect Group by items ($F_1(1,22) = 1.61$, $p = .22$, $\eta^2 = .07$; $F_2(1,46) = 18.20$, $p < .001$, $\eta^2 = .28$), with no interaction effect ($F_1(1,22) = 2.68$, $p = .12$, $\eta^2 = .11$; $F_2(1,46) = 1.66$, $p = .24$). Mean gaze durations were shorter when the predicative adjective agreed in gender with the preceding noun than when it did not (344 ms and 371 ms, respectively).

Total reading times. Mean total reading times are reported for the critical predicative adjective in Table 4. Analyses of this region revealed an effect of Agreement ($F_1(1,22) = 15.39$, $p < .001$, $\eta^2 = .41$; $F_2(1,46) = 8.66$, $p < .005$, $\eta^2 = .16$) and an effect of Group by items ($F_1(1,22) = 1.70$, $p = .21$; $F_2(1,46) = 10.84$, $p < .002$, $\eta^2 = .19$) with no interaction effect (F_1 and $F_2 < 1$). Mean total reading times were longer when adjectives were incorrectly than correctly inflected for gender (627 vs. 544 ms, respectively).

Regressions. Table 5 shows the percentage of first pass and second pass regressions launched from the critical adjective as a function of Group (native French vs. English–French learners) and Agreement (gender agreement vs.

Table 4
Means and standard deviations for first pass, gaze duration and total reading time of all regions as a function of Group (French and English speakers) and Agreement (correct vs. incorrect).

	First fixation		Gaze duration		Total reading times	
	Cor.	Inc.	Cor.	Inc.	Cor.	Inc.
Initial (R1)						
French	452	452	763	780	789	793
SD	132	148	203	228	201	199
English	451	464	715	775	806	824
SD	104	104	162	276	170	212
Noun (R2)						
French	267	275	326	321	443	434
SD	56	63	56	82	133	171
English	291	288	396	373	607	615
SD	53	64	106	86	195	208
*Adjective (R4)**						
French	275	285	338	352	517	600
SD	54	60	75	77	156	157
English	279	289	363	398	585	664
SD	52	65	99	124	220	269
Final (R5 + R6)						
French	335	296	419	389	663	663
SD	100	91	158	153	200	203
English	330	326	455	420	790	779
SD	89	61	163	132	268	232

* Main region of interest.

Table 5
First pass and second pass regressions (in percentages) launched from the critical predicative adjective region, as a function of participant group and gender agreement.

	Correct	Incorrect
Native group		
First pass	16	23
Second pass	30.5	35.5
L2 group		
First pass	15	10.5
Second pass	30	44

violation). Analyses of first pass regressions launched from the critical adjective revealed no main effects and only a trend for an interaction of Group × Agreement by participants ($F_1(1,22) = 3.45$, $p = .07$)[1]. No effects as a function of experimental factors were found for regressions launched during subsequent readings of sentences.

Discussion

In Experiment 4, we used eye-tracking to re-examine the processing of gender agreement for predicative adjectives, using the same materials as in Experiment 3 in which ERPs were recorded. In both experiments, the primary aim was to determine whether L2 learners can process gender agreement online in non-local contexts; however, our results also provide valuable information regarding native language processing and the sensitivity of the two techniques.

Both French native speakers and English–French learners showed longer reading times when the gender of the predicative adjective and the gender of the noun mismatched than when they matched. In contrast to Keating's (2009) study in which the effect of agreement emerged only for total reading times, and only clearly for the group of native speakers, the effects we obtained were significant during first pass reading times, for gaze durations, as well as for total reading times. Moreover, this was true for both participant groups. Our study thus shows early and equivalent detection of gender agreement violations during online sentence processing for native and non-native speakers.

In sum, this eye-tracking experiment suggests that L2 learners can process gender agreement even when it occurs in non-local contexts as recently demonstrated in an ERP study (Gillon-Dowens et al., 2010), at least under conditions where they can control the amount of time they allocate to processing. English–French learners displayed a similar pattern to that of native speakers. There was no apparent delay in processing in the learner group, in contrast to data from ERP studies, which suggest that L2 syntactic processing is often delayed compared to L1 processing (Ardal, Donald, Meuter, Muldrew, & Luce, 1990; Hahne, 2001). Our results, once again, suggest that English speakers are not handicapped by the absence of a grammatical gender system in their L1. These results will further discussed in relation to theories of syntactic processing in L2 in the general discussion.

General discussion

In the present paper we report three ERP experiments and one eye-tracking experiment, which investigated

[1] Item analyses were not undertaken for regressions given the relative scarcity of data.

whether late English–French learners are able to acquire gender and process agreement online in their L2. In all four experiments we compared the performance of French native speakers to that of English–French learners during sentence processing. Experimental sentences contained gender agreement violations either within the DP, between the noun and the post-posed adjective (Experiment 1), the pre-posed adjective and the noun (Experiment 2), or within the verbal phrase for predicative adjectives (Experiments 3 and 4). The use of various types of adjectives allowed us to examine (a) whether L2 speakers can acquire gender and process it online; (b) whether gender agreement processing is hindered when the surface order is similar in L1 and L2 but the grammatical properties of the two languages differ; and (c) whether L2 learners can process gender in non-local contexts.

Results for late L2 learners revealed sensitivity to gender agreement violations between the noun and the post-posed adjective similar to that found in native speakers, as revealed by a P600 effect. This suggests that English speakers can acquire grammatical gender in French and compute agreement online even when they acquire their L2 later in life. These results contrast with those obtained in previous studies for English–Dutch learners (Sabourin & Stowe, 2008) but are in line with those obtained for English–Spanish learners (Gillon-Dowens et al., 2010; Keating, 2009; Tokowicz & MacWhinney, 2005; White et al., 2004). In view of our findings and those of previous studies, our conclusion converges with the assumption made by the 'full access' model presented in the introduction, since the results for English–French learners reflected their ability to assign and process gender online, which confirms the possibility for L2 learners to acquire new features and grammaticalize agreement rules.

The sensitivity of our English–French learners to agreement violations was apparently affected, however, by the difference of syntactic rules in L1 and L2 due to language properties (i.e., presence vs. absence of gender). Indeed, when the surface order of the adjective–noun pair was similar in English and French (i.e., pre-posed adjective), English speakers showed some sensitivity to agreement violations but displayed a different effect from native speakers; while French native speakers displayed a classic P600 effect for gender agreement violations, English speakers showed an N400 effect which has been associated with L2 learning in recent studies (Inoue & Osterhout, 2005; McLaughlin et al., 2010; Osterhout et al., 2004). This effect will be further discussed below. Our results are line with Tokowicz and MacWhinney's (2005) proposal that structures are more rapidly acquired and processing is more accurate when the features and/or grammatical rules of the L1 and L2 are not in conflict than when they are (Experiment 1 vs. Experiment 2 in the present paper). This proposal is also consistent with the results obtained in Foucart and Frenck-Mestre (2011) for German–French learners processing the same kind of violations. These learners showed a P600 effect in response to agreement violations when the grammatical rules were similar in both languages (i.e., determiner–noun agreement) but not when the rules differed in French and German (i.e., adjective–noun agreement for plural DPs). In contrast, our results conflict with the proposal put forward by Sabourin and colleagues (Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008; Sabourin, Stowe, & de Haan, 2006) that L2 learners can reach native-like processing only when syntactic structures are similar in L1 and L2. In addition, these authors suggested 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. Our results as well as those obtained in other studies clearly contradict this proposal (Bruhn de Garavito & White, 2002; Gillon-Dowens et al., 2010; Keating, 2009; Tokowicz & MacWhinney, 2005; White et al., 2004).

Our ERP data revealed different patterns of effects for native speakers and L2 learners when processing gender agreement for pre-posed adjectives in French. A possible explanation for this is linked to our learners' expectations concerning adjective placement. Recently, Anderson (2007, 2008) investigated the acquisition of interpretative correlates of variable adjective position in French. Overall, his results showed a gradual development in the acquisition of the interpretation of adjectives by English learners of French, with the interpretation of post-posed adjective being mastered earlier in the acquisition process than that of pre-posed adjectives. Anderson notes that when taught adjectives in the French classroom, learners are told that adjectives are predominantly post-nominal in French with a few exceptions being pre-nominal or both pre- and post-nominal. Hence, English learners first tend to assess pre-nominal adjectives as ungrammatical and it is only when they are proficient enough to understand the interpretation of the adjective that they consider pre-nominal adjectives as grammatical. As such, the N400 we observed in the L2 learner group may reflect difficulties in lexical-semantic processing rather than syntactic processing.

Gender processing is apparently more challenging when occurring in non-local contexts, i.e., for predicative adjectives, both for monolingual and L2 learners (Barber & Carreiras, 2005; Deutsch & Bentin, 2001). It is important to note, nonetheless, that our results showed different patterns for native and non-native speakers depending upon the measure we used to test for online sensitivity to gender agreement. French native speakers showed sensitivity to gender agreement violations between the noun and the predicative adjective independent of the measure, i.e., whether eye movements or ERPs. In contrast, the L1 English–L2 French participants, showed a consistent response to agreement violations only when they could read under naturalistic situations, i.e., where the whole sentence was presented, different agreeing elements could be gleaned at least in the periphery, and where it was possible to control the amount of time allocated to processing, i.e., by either pausing or re-reading the sentence. Such is possible with eye-tracking but not in the word-by-word paradigm generally employed in visual ERP designs. Hence, our late bilinguals showed sensitivity in the eye-tracking record but not systematically enough to produce a robust pattern in the ERP record. Indeed, the group of L2 learners showed heterogeneous ERP responses with no clear language-related component observed. This was also reflected in the behavioural responses; L2 participants rejected a higher number of sentences that contained a gender agreement

violation in the eye-tracking experiment than in the ERP experiment. The differences in results across methods specifically in the L2 group can possibly be accounted for by memory load. McDonald (2006) has suggested that differences in native and second language processing can be attributed to the capacity to allocate resources and working memory. In this view, under conditions where memory is taxed, due to having to hold elements in memory in order to compute agreement, as was the case for word-by-word presentation and predicative adjectives, L2 processing should indeed be less efficient. When memory is less taxed, as is the case for whole sentence presentation, L1 and L2 processing should be more similar. This is exactly the pattern we obtained, with more similar processing across native and L2 processing for eye-movements than ERPs.

It is important to note that the results we obtained with eye-tracking showed virtually indistinguishable patterns for L2 learners a and native speakers. Our results suggest thus that L2 learners are sensitive to gender agreement violations even in non-local contexts, at least, when they have the possibility to use peripheral information, re-read the sentence and allocate further resources for agreement checking. They converge with the ERP results obtained for English–Spanish late learners who had extended exposure to their L2 (Gillon-Dowens et al., 2010) but contrast with the null-effect obtained by Keating (2009) for English–Spanish advanced learners in an eye-tracking paradigm. The contrast with the latter study, despite converging methodology, may stem from the fact that in our experiment the noun and the adjective were only separated by the copula, as opposed to several words and complex grammatical structures in Keating's study. Hence, it is conceivable that the difficulty of processing agreement may increase with the number of words as well as number of syntactic boundaries between the noun and the adjective.

Our results are in partial contradiction with the prediction made by the SSH (Clahsen & Felser, 2006) that late L2 learners cannot compute agreement in non-local contexts because of a lack of processing resources. The SSH argues that processing in non-local contexts is more demanding for working memory; hence, L2 learners cannot hold the information required to compute agreement between the noun and the predicative adjective while computing agreement of intervening elements (e.g., the verb). Our eye-tracking results, in contrast, suggest that L2 learners can process agreement in non-local contexts. Indeed, the L2 learners showed immediate sensitivity to agreement violations, as revealed by longer reading times, in like manner to native speakers, even on first pass measures. Moreover, the L2 learners did not show excessively long reading times in comparison to native speakers, thus suggesting that computation was not delayed. Nonetheless, the difference between our ERP and eye-tracking results for the L2 learners suggests that under conditions where one must hold items in memory and cannot self-regulate reading time, L2 processing will be less efficient when non-local dependencies are in question. It is an open question, indeed, whether the differences we found across eye-tracking and ERPs are truly the result of lesser working memory in the L2, as suggested by some (Clahsen & Felser, 2006; McDonald, 2006) or if indeed the two methodologies in fact tap into different types of processing.

Overall, our results converge with other studies suggesting that highly proficient L2 learners can reach native-like processing levels even if they learned their L2 late in life (Gillon-Dowens et al., 2010; Herschensohn, 2006; Hopp, 2007; Rossi et al., 2006; Tokowicz & MacWhinney, 2005; White et al., 2004). However, syntactic processing seems to go through different stages during L2 acquisition. Recently, Osterhout and collaborators (Osterhout et al., 2004, 2006) have investigated the neurological changes that occur in learners' brain over time with increasing proficiency and/or exposure to their L2. Although their findings are preliminary, they claim that it may be possible to establish *which* aspects of the L2 have been learned and *when* they have been learned. In several longitudinal studies, they showed that L2 learners display a typical N400 effect in response to semantic violations, and this effect is consistent throughout acquisition. In contrast, it seems that in early stages of acquisition, the processing of syntactic violations is reflected by an N400 effect and not a P600 effect as it is usually expected for this type of violations. However, when proficiency increases, the N400 effect gradually evolves into a more classic P600 effect. To account for this evolution, the authors proposed that L2 learners proceed through various stages of syntactic learning; they first memorise combinations of words and morphemes, and only later do they integrate syntactic rules, in like manner to children during L1 acquisition (Tomasello, 2000; Ullman, 2001, 2005). However, it seems unlikely that L2 learners memorise all adjective–noun combinations and the same authors have recently suggested that learners memorise probabilistic dependencies between non-adjacent morphemes rather than the whole words (McLaughlin et al., 2010). For example, they memorise that the French feminine article *la* is likely to be followed by an adjective ending in 'e' in contrast to the article *le* (e.g., *la chaise verte*, the green chair vs. *le livre*$_{masc}$ *vert*$_{masc}$, the green book). Once proficiency increases, L2 learners integrate the gender agreement rule, and an agreement violation between the noun and the adjective provokes a P600 effect as typically found for syntactic violations. In the case of our English–French learners, they revealed a P600 effect for agreement violations between the noun and the post-posed adjective, an N400 effect for violations between the pre-posed adjective and the noun, and only a minimal effect in the ERP trace for the predicative adjective. Thus, it implies that our L2 learners had integrated the agreement rule for post-posed adjectives, but not completely for pre-posed adjectives, and not consistently for predicative adjectives. Again, this hypothesis converges with Tokowicz and MacWhinney's (2005) proposal that syntactic processing is more rapidly acquired in structures that are unique in L2 (i.e., here, post-posed adjectives). It is important to recall that the canonical order for adjectives in French is post-posed, hence, L2 learners encounter them more often and earlier in their L2 instruction (Anderson, 2008). In line with these studies and in view of the results for post-posed adjectives, we suggest that processing of pre-posed adjectives may become native-like with increased proficiency.

Last, a word is in order regarding the topographical differences we found across the native and L2 learner groups for the reported P600 effect. For our native control group, we reported a consistent posterior distribution for the P600 elicited by gender agreement violations, independent of both the position of the adjective within the determiner phrase (post-posed or pre-posed) and type (attributive or predicative). The only difference was in magnitude, with progressively smaller effects from canonical order of attributive adjectives to predicative adjectives. In the L2 group, a P600 effect was only found for gender agreement violations of attributive adjectives in canonical word order (adjective–noun, in French) and its distribution was rather frontal. It has been suggested that differences in distribution of the P600 effect reveal qualitative differences in processing, with more frontal P600s linked to revision, for dispreferred structures, rather than to actual repair of outright violations, which would elicit a more posterior distribution (Hagoort, Brown, & Osterhout, 1999). Others have claimed that this distinction is less clear cut, with posterior distributions being indistinguishable for "revision" and "repair" processes, and frontal P600s linked, rather, to level of syntactic complexity (Kaan & Swaab, 2003). In our study, the factor that drove differences in the P600 distribution was language proficiency, with native speakers consistently showing a posterior distribution and L2 learners a more anterior one. It is conceivable that, in general, L2 online syntactic processing is more costly and/or "complex" that native language processing, which might then predict more frontal P600s in the former case.

To conclude, we have shown that English–French learners are able to acquire gender and to process it online in their L2. Hence, it seems to be possible for L2 learners to acquire new features in their L2 even when learned after early childhood. We claim that English speakers can process gender in a similar way to native speakers, but attaining native-like processing may be faster for structures that are unique in L2 and do not conflict with L1 structures. Further longitudinal studies and studies involving near-native L2 speakers are required to investigate the different stages of L2 acquisition.

References

Anderson, B. (2007). Learnability and parametric change in the nominal system of L2 French. *Language Acquisition, 14*, 165–214.

Anderson, B. (2008). Forms of evidence and grammatical development in the acquisition of adjectives position in L2 French. *Studies in Second Language Acquisition, 30*, 1–29.

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.

Ayoun, D. (2007). The second language acquisition of grammatical gender and agreement. In Dalila Ayoun (Ed.), *French applied linguistics* (pp. 130–170). Johns Benjamins.

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

Bernstein, J. (1991). DPs in French and Walloon: Evidence for parametric variation in nominal head movement. *Probus, 3*, 101.

Bolte, J., & Connine (2004). Grammatical gender in spoken word recognition in German. *Perception & Psychophysics, 66*, 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). Doradrecht: Kluwer.

Carroll, S. (1989). Second language acquisition and the computational paradigm. *Language Learning, 39*, 535–594.

Carstens, V. (2000). Concord in minimalist theory. *Linguistics Inquiry, 31*, 319–355.

Chomsky, N. (1995). Categories and transformations. In *The minimalist program* (pp. 219–394). Cambridge, Mass.: MIT Press.

Christianson, K., Hollingworth, A., Halliwell, J., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*, 368–407.

Cinque, G. (1994). On the evidence for partial N-movement in the Romance DP. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, & R. Zanuttini (Eds.), *Paths towards universal grammar* (pp. 85–110). Washington, DC: Georgetown University Press.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3–42.

Clark, E. (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.

Deprez, V. (2005). Morphological number, semantic number and bare nouns. *Lingua, 115*, 857–883.

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.

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition, 14*(3), 379–399.

Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research, 17*, 213–247.

Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of grammatical gender*. Amsterdam: J. Benjamins.

Frenck-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.

Frenck-Mestre, C., Carrasco-Ortiz, H., McLaughlin, J., Osterhout, L., & Foucart, A. (2010). Linguistic input factors in native and L2 processing of inflectional morphology. Evidence from ERPs. In D. V. Maya Hickmann (Ed.), *Language, Interaction and Acquisition* (pp. 206–228). Amsterdam / Philadelphia: John Benjamins.

Frenck-Mestre, C., Foucart, A., Carrasco, H., & Herschensohn, J. (2009). *Processing of grammatical gender in French as a first and second language: Evidence from ERPs. EUROSLA yearbook* (pp. 76–106).

Frenck-Mestre, C., Osterhout, L., McLaughlin, J., & Foucart, A. (2008). The effect of phonological realization of inflectional morphology on verbal agreement in French: Evidence from ERPs. *Acta Psychologica, 128*, 528–536.

Gess, R., & Herschensohn, J. (2001). Shifting the DP parameter: A study of anglophone French L2ers. In Joaquim Camps & Caroline R. Wiltshire (Eds.), *Romance syntax, semantics and their L2 acquisition* (pp. 105–119). Amsterdam: John Benjamins.

Gillon Dowens, M., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in chinese learners of spanish–evidence from event related potentials. *Neuropsychologia, 49*, 1651–1659.

Gillon-Dowens, M., Vergara, M., Barber, H. A., & Carreiras, M. (2010). Morpho-syntactic processing in late L2 learners. *Journal of Cognitive Neuroscience, 22*, 1870–1887.

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.

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

Hagoort, P., Brown, C., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 273–316). Oxford, UK: Oxford University Press.

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., 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*, 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*, 386–400.
Hawkins, R. (1998). Explaining the difficulty of French gender attribution for speakers of English. In *Paper presented at the EUROSLA conference*, Paris.
Hawkins, R. (2001). The theoretical significance of Universal Grammar in second language acquisition. *Second Language Research, 17*, 345–367.
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.
Hawkins, R., & Franceschina, F. (2004). Explaining the acquisition and non-acquisition of determiner–noun gender concord in French and Spanish. In P. Prevost & J. Paradis (Eds.), *The acquisition of French in different contexts. Focus on functional categories* (pp. 175–207). Amsterdam: John Benjamins.
Herschensohn, J. (2006). Français langue seconde: From functional categories to functionalist variation. *Second Language Research, 22*, 95–113.
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. In *Annual CUNY sentence processing conference*, Tucson, March 2005.
Jasper, H. H. (1958). The ten twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 5*, 196–214.
Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience, 15*, 98–110.
Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning, 59*, 503–535.
Kutas, M., & Hilyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.
Laenzlinger, C. (2005). French adjective ordering: Perspectives on DP-internal movement types. *Lingua, 115*, 645–689.
MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 581–611). London: Elsevier Inc.
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.
McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*, 381–401.
McLaughlin, J., Tanner, D., Pitkanen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., et al. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning*.
Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Acquisition of gender agreement in second language learners: An event-related potential study. *Language Learning, 60*, 154–193.
Mueller, J. (2005). Electrophysiological correlates of second language processing. *Second Language Research, 21*, 152–174.
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.
New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, and Computers, 36*, 516–524.
Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785–806.
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). 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.
Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., et al. (2008). Second language learning and changes in the brain. *Journal of Neurolinguistics, 21*, 509–521.
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. L., & Haverkort, M. (2003). Neural substrates of representation and processing of a second language. In R. van Hout, A. Hulk, F. Kuiken, & 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*, 397–430.
Sabourin, L., Stowe, L. A., & de Haan, G. J. (2006). Transfer effects in learning an L2 grammatical gender system. *Second Language Research, 22*, 1–29.
Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research, 12*, 40–72.
Spinelli, E., Meunier, F., & Seigneuric, A. (2006). Does gender information influence early phases of spoken word recognition? *The Mental Lexicon, 1*, 277–297.
Taner, D. (2008). In Melissa Bowles et al. (Eds.), *Selected proceedings of the 2007 second language research forum* (pp. 22–34). Somerville, MA: Cascadilla Proceedings Project.
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.
Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*, 209–253.
Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience, 2*, 717–726.
Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition* (pp. 141–178). Washington, DC: Georgetown University Press.
White, L. (1989). *Universal grammar and second language acquisition*. Amsterdam: John Benjamins.
White, L. (2003). *Second language acquisition and universal grammar*. Cambridge: Cambridge University Press.
White, L., Valenzuela, E., Kozlowska-Macregor, M., & Leung, Y. I. (2004). Gender and number agreement in non-native Spanish. *Applied Psycholinguistics, 25*, 105–133.