



Cognitive Science 41 (2017) 1760–1803

Copyright © 2016 The Authors. Cognitive Science published by Wiley Periodicals, Inc. on behalf of Cognitive Science Society.

ISSN: 0364-0213 print / 1551-6709 online

DOI: 10.1111/cogs.12450

# First Language Attrition Induces Changes in Online Morphosyntactic Processing and Re-Analysis: An ERP Study of Number Agreement in Complex Italian Sentences

Kristina Kasparian,<sup>a,b</sup> Francesco Vespignani,<sup>c</sup> Karsten Steinhauer<sup>a,b</sup>

<sup>a</sup>*School of Communication Sciences & Disorders, McGill University*

<sup>b</sup>*Centre for Research on Brain, Language and Music (CRBLM), McGill University*

<sup>c</sup>*Department of Psychology and Cognitive Science, University of Trento*

Received 10 February 2016; received in revised form 12 June 2016; accepted 29 July 2016

---

## Abstract

First language (L1) attrition in adulthood offers new insight on neuroplasticity and the role of language experience in shaping neurocognitive responses to language. Attriters are multilinguals for whom advancing L2 proficiency comes at the cost of the L1, as they experience a shift in exposure and dominance (e.g., due to immigration). To date, the neurocognitive mechanisms underlying L1 attrition are largely unexplored. Using event-related potentials (ERPs), we examined L1-Italian grammatical processing in 24 attriters and 30 Italian native-controls. We assessed whether (a) attriters differed from non-attriting native speakers in their online detection and re-analysis/repair of number agreement violations, and whether (b) differences in processing were modulated by L1-proficiency. To test both local and non-local agreement violations, we manipulated agreement between three inflected constituents and examined ERP responses on two of these (subject, verb, modifier). Our findings revealed group differences in amplitude, scalp distribution, and duration of LAN/N400 + P600 effects. We discuss these differences as reflecting influence of attriters' L2-English, as well as shallower online sentence repair processes than in non-attriting native speakers. ERP responses were also predicted by L1-Italian proficiency scores, with smaller N400/P600 amplitudes in lower proficiency individuals. Proficiency only modulated P600 amplitude between 650 and 900 ms, whereas the late P600 (beyond 900 ms) depended on group membership and amount of L1 exposure within attriters. Our study is the first to show qualitative and

---

Correspondence should be sent to Kristina Kasparian and Karsten Steinhauer, School of Communication Sciences and Disorders, McGill University, 2001 McGill College Avenue, Unit 800, Montreal, Quebec H3A 1G1, Canada. E-mails: kristina.kasparian@mail.mcgill.ca (or) karsten.steinhauer@mcgill.ca

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

quantitative differences in ERP responses in attriters compared to non-attriting native speakers. Our results also emphasize that proficiency predicts language processing profiles, even in native-speakers, and that the P600 should not be considered a monolithic component.

*Keywords:* First language attrition; Neuroplasticity; Proficiency; Morphosyntactic processing; Event-related potentials; Number agreement

---

## 1. Introduction

### 1.1. Neuroplasticity and multilingualism

For more than half a century, multilingualism research has centered on whether and to what extent the brain mechanisms underlying language-related processes exemplify continued neuroplasticity over the lifespan. The long-standing view has been that maturational limits on neuroplasticity for language learning constrain second language (L2) acquisition, such that an L2 acquired in late childhood or adulthood must rely on different neurocognitive substrates and processes than those used for the native language (cf. “Critical Period Hypothesis [CPH]”: Lenneberg, 1967; Penfield & Roberts, 1959; also see Hahne & Friederici, 2001; Johnson & Newport, 1989; Kim, Relkin, Lee, & Hirsch, 1997; Weber-Fox & Neville, 1996).

In contrast, a growing body of evidence has argued against the view that the age at which a language is learned is the *limiting* factor in native-like language processing, and it has emphasized the decisive role of factors such as language proficiency (e.g., Bowden, Steinhauer, Sanz, & Ullman, 2013; Friederici, Steinhauer, & Pfeifer, 2002; McLaughlin, Osterhout, & Kim, 2004; Osterhout, McLaughlin, Pitkänen, Frenck-Mestre, & Molinaro, 2006; Perani et al., 1998; Rossi, Gugler, Friederici, & Hahne, 2006; Van Hell & Tokowicz, 2010; Wartenburger et al., 2003; White, Genesee, & Steinhauer, 2012), type of learning (Morgan-Short, Steinhauer, Sanz, & Ullman, 2012), first language background (Dowens, Guo, Guo, Barber, & Carreiras, 2011; Ojima, Nakata, & Kakigi, 2005; Tanner, Inoue & Osterhout, 2012; Tokowicz & MacWhinney, 2005), and socio-emotive factors such as motivation (Tanner, 2011) in predicting the degree of overlap in brain responses between native-monolingual and different types of multilingual speakers. Consistent with the notion of ongoing neuroplasticity for language learning well into adulthood, recent studies using advanced neuroimaging methods have measured changes in the brain’s anatomical structure with increasing language exposure and proficiency, such as increases in gray matter volume and cortical thickness in key language-related areas (Della Rosa et al., 2013; Klein, Mok, Chen, & Watkins, 2014; Mårtensson et al., 2012; Mechelli et al., 2004; Schlegel, Rudelson, & Tse, 2012; Stein et al., 2012).

Despite this gradual shift in perspective that the brain mechanisms underlying language remain more flexible beyond early childhood than was previously believed, the impact of age of acquisition (AoA)—independent from experiential factors which inarguably tend

to vary with age, such as exposure, proficiency, and motivation—remains controversial. Studies in support of maturational constraints on language learning and processing continue to pervade the literature on the neurocognition of bilingualism (e.g., Abrahamsson & Hyltenstam, 2009; Clahsen & Felser, 2006a,b; Moreno & Kutas, 2005; Pakulak & Neville, 2011).

### 1.2. *The contribution of first language (L1) attrition*

Unsurprisingly, much of the debate on neuroplasticity and language learning has led to investigations into speakers' L2 processing and performance. However, a unique multilingual experience that has the potential to shed new light on the neuroplasticity question is that of *L1 attrition*. Attrition has been defined as a gradual, non-pathological negative change or loss in one's L1 abilities following prolonged immersion into a new linguistic environment, usually after immigration to a new country (Köpke & Schmid, 2004). Attriters experience a clear shift in exposure and use from their native-L1 to the environmental-L2. Thus, contrary to typical L2 learners who continue to be L1-dominant, attriters not only experience increasing L2 exposure and proficiency over time in the new country, but the L2 becomes the predominant language while use of the L1 is reduced or interrupted. This change in how the L1 is used has been recently proposed as a more general and widely encompassing definition of the phenomenon of attrition (Schmid, 2011).

Investigating the neurocognitive aspects of L1 attrition logically complements research on L2 processing and is a key to inform theories on neuroplasticity and multilingualism. The corollary of the CPH—and related, milder claims of maturational constraints on language processes—is that one's L1 has a privileged status and is stable, as a result of having been hard-wired or “entrenched” in the brain with early exposure (Marchman, 1993; Penfield, 1965). The longer one is exposed to a given language, the more entrenched that language becomes (Van Hell & Tokowicz, 2010) and the less language mechanisms can be modified by L2 exposure (Pallier, 2007). This argument has been proposed to explain why cross-linguistic influence from the direction of the L1 to the L2 is so pervasive in late-learners, in whom L1 linguistic patterns and underlying neurocognitive mechanisms are deeply entrenched (e.g., Tokowicz & MacWhinney, 2005). In this view, adult attriters (who lived in an L1-dominant context until adulthood) should deviate less from non-attriting native speakers than young attriters (who experienced this shift in L1–L2 dominance at an earlier age). Behavioral research on attrition has supported this view (Ammerlaan, 1996; Bylund, 2009; Pelc, 2001; see reviews by Köpke, 2004, and Köpke & Schmid, 2004). Similar claims have been drawn from studies on L1 language loss following international adoption—albeit an extreme case of L1 attrition. Individuals who were adopted earlier than puberty did not show evidence of residual knowledge of their L1 and were indistinguishable in their brain activation responses from native-L2 controls (Pallier et al., 2003; Ventureyra et al. 2004; Ventureyra & Pallier, 2004; but see Pierce, Klein, Chen, Delcenserie, & Genesee, 2014). From such findings, it has been suggested that L1 attrition is subject to maturational constraints (see Pallier, 2007; Schmid, 2011) and that the L1 does indeed stabilize around 12 years of age (Bylund, 2008).

Including attriters in the mix of multilinguals that are recruited in neurolinguistic research therefore allows us to investigate whether early exposure to a language is key in guaranteeing native-like processing in the brain, or whether attriters' brain responses to either language are modulated by factors such as proficiency and exposure—even in the L1, and even if the shift in dominance took place in adulthood. While there has been a wealth of behavioral and anecdotal reports of L1 attrition, its neurocognitive correlates still remain to be explored.

This study uses event-related potentials (ERPs) to examine the neurocognitive correlates of real-time morphosyntactic processing in a group of adult attriters relative to non-attriting native speakers. We use “Attriters” to refer to a group of first-generation immigrants who moved from their native country (Italy) to a new country (Canada) in adulthood, and who report having experienced a shift in exposure/use and, in certain areas, in automaticity from the L1 to the L2. While behavioral studies have reliably shown that the domain of lexical-semantics is particularly vulnerable to attrition (De Bot, 1996; Hulsen, 2000; Köpke, 1999; Köpke & Schmid, 2004; Montrul, 2008; Opitz, 2011; see Paradis, 2003, 2007), the picture has been more mixed for in the domain of morphosyntax (Ammerlaan, 1996; Gürel, 2004, 2007; Kim, Montrul, & Yoon, 2010; Schmid, 2010; Schmid & Köpke, 2011; Sorace, 2011; Tsimpli, Sorace, Heycock, & Filiaci, 2004), and it remains an open question to what extent the adult grammar of L1 speakers immersed in an L2 may show attrition.

One research group led by Monika Schmid and colleagues (Bergmann et al., 2015) is currently conducting studies on L1 attrition using ERPs in German and Dutch attriters living in an English-speaking context, compared to monolingual native speakers of German and Dutch, as well as different groups of early/late L2 learners with diverse L1-backgrounds. Their auditory stimuli include non-finite verb form violations (e.g., *the rose has \*blossom*), and gender agreement violations between determiner-noun (e.g., *the<sub>neut</sub> \*garden<sub>masc</sub>*), or determiner-adjective-noun (*the<sub>masc</sub> fresh \*grass<sub>neut</sub>*). Their preliminary results indicate that L1 attriters perform like native speakers on behavioral tasks and elicit P600 effects of similar amplitude, distribution, and latency in response to all three kinds of violations, whereas late L2 learners deviate from native speakers (Bergmann et al., 2015; Schmid, 2013). The authors interpret these results in favor of maturational constraints on L2 processing, whereas the L1 is robust and remains native-like even despite a shift in dominance toward the L2.

Although our work developed independently from that of Schmid and colleagues, two distinctions between their study and ours are worth noting, as these differences contribute to the novelty of our work. First, in testing gender agreement between a determiner and a noun, it can be argued that the process being investigated is largely speakers' *lexicalized* knowledge of idiosyncratic associations between a noun and its gender (and its appropriate determiner). Given that, in German or Dutch, an adjective occurring between a gender-marked determiner and a gender-marked noun is not inflected for gender itself, it is not possible to test attriters' and L2 learners' sensitivity to morphosyntactic agreement rules independent of the more lexicalized process of matching a noun to its arbitrary determiner. Second, although Schmid and colleagues employed several proficiency

measures and matched their groups on overall proficiency, it was not one of their research goals to systematically examine whether/how individual differences in proficiency level modulated ERP responses.

In contrast, our study examines number agreement processing across three separate constituents within a sentence, each of which are inflected for number in Italian. Examining agreement processing in sentences with multiple inflected constituents also allows us to how morphosyntactic agreement is computed over a longer span within a sentence, as opposed to the detection and resolution of only local mismatches. As it has been shown that complex sentences and long-distance dependencies are a more reliable source of differences between highly proficient L2 learners and native speakers than salient morphosyntactic violations (e.g., Clahsen & Felser, 2006a,b), it is possible that these structures would reveal differences between attriters and non-attriter native speakers as well. A second aim was to examine the effect of L1-Italian proficiency on the neurocognitive processing patterns observed among native speakers. Contributing to findings from a growing number of L2 processing studies, several researchers have shown that L1 processing in native-monolinguals is also modulated by proficiency level (Pakulak & Neville, 2010; Prat, 2011). A further step is therefore to extend this line of inquiry to L1 attrition and to examine whether individual differences in L1 proficiency may relate to attrition effects at the processing level.

### *1.3. Number agreement processing*

Number agreement processing has been widely studied in monolingual native speakers of different languages (see review by Molinaro, Barber, & Carreiras, 2011). Given that agreement patterns are subject to crosslinguistic variation, it is of interest to extend this research to bilinguals whose two linguistic systems differ in their expression of number morphology. Italian is a language with a relatively free word-order and a rich morphological marking system where number agreement is salient and can often constrain the identification of a subject (Bates, McNew, MacWhinney, Devescovi, & Smith, 1982; MacWhinney & Bates, 1989). In contrast, number agreement in English is poorly signaled due to a less detailed system of morphological markers, and speakers instead rely on word-order for sentence interpretation.

Number agreement studies conducted with late L2 learners have shown evidence of cross-linguistic influence from the L1 onto the L2, revealing non-native-like processing profiles (i.e., missing LAN and/or P600) in cases where the L2 agreement properties in question did not exist in the speakers' L1 (e.g., Chen, Shu, Liu, Zhao, & Li, 2007; Ojima et al., 2005; Osterhout, McLaughlin, Kim, Greenwald, & Inoue, 2004; Tokowicz & MacWhinney, 2005). Several studies have emphasized continued neuroplasticity by showing that adult L2 learners converge on native speakers' processing patterns with continued learning and high proficiency levels (e.g., Hopp, 2010; Osterhout et al., 2006, 2008; Rossi et al., 2006). To date, online number agreement processing has not been investigated in L1 attrition, where the source of cross-linguistic influence is the L2 rather than the L1.

Due to their excellent temporal resolution, ERPs are particularly useful for investigating real-time processing of agreement patterns during sentence comprehension. Three components have generally been associated with the processing of number agreement violations: (a) a *left-anterior negativity* (LAN) elicited between 300 and 500 ms, reflecting the early detection of a morphosyntactic violation (e.g., Kaan, 2002; Molinaro, Vespignani et al., 2011; Osterhout & Mobley, 1995); (b) an *early frontal positivity* between 500 and 700 ms, argued to reflect difficulties integrating the mismatching constituent with the previous sentence context, particularly in ambiguous or complex sentences (Barber & Carreiras, 2005; Friederici, Hahne, & Saddy, 2002; Kaan & Swaab, 2003; Molinaro, Kim, Vespignani, & Job, 2008); and (c) a *posterior P600* between 700 and 1,000 ms, indexing morphosyntactic re-analysis and repair once the anomaly has been diagnosed, with larger and more prolonged P600s reflecting costlier repair (Carreiras, Salillas, & Barber, 2004; Hagoort & Brown, 2000; Molinaro, Kim et al., 2008; Silva-Pereyra & Carreiras, 2007). However, it is not the case that all three components are reliably elicited in response to all number agreement violations, nor have they been quantified in the same way across studies.

LAN-like negativities have been reported for subject-verb number agreement violations across languages, in contrast to gender and person agreement violations which typically elicit a broadly distributed N400 (see review by Molinaro, Barber et al., 2011). This dissociation has been explained in terms of strong *morphosyntactic* versus *conceptual* expectations of agreement which, when violated, elicit a LAN or N400, respectively (LAN: Molinaro, Barber et al., 2011; N400: Osterhout, 1997; Tanner, McLaughlin, Herschensohn, & Osterhout, 2013; Tanner, Inoue, & Osterhout, 2012). Other number agreement studies, however, did not find a LAN (e.g., Balconi & Pozzoli, 2005; Hagoort & Brown, 2000; Hagoort, Brown, & Groothusen, 1993; Nevins, Dillon, Malhotra, & Phillips, 2007; Osterhout, McKinnon, Bersick, & Corey, 1996) or reported a negativity with a bilateral-anterior focus (Kaan, 2002; Leinonen, Brattico, Jarvenpaa, & Krause, 2008), or rather a broad N400-like negativity (Coulson, King, & Kutas, 1998), or the effect only reached significance when a *t*-test was conducted on a small cluster of anterior electrodes (e.g., T3, C3, F3, F7), as was the case in several Italian studies of subject-verb number agreement (Angrilli et al., 2002; De Vincenzi et al., 2003; Mancini, Vespignani, Molinaro, Laudanna, & Rizzi, 2009; Molinaro, Barber et al., 2011). In Italian, the lack of a robust LAN has been attributed to the flexible word-order which allows a subject to follow a verb in a given sentence, thus weakening expectations of agreement between a sentence-initial noun and a subsequent verb (Molinaro, Barber et al., 2011).

Similarly, the distinction between the prototypical (posterior) P600 and an earlier (more fronto-central) positivity has not consistently been made, given that the majority of number agreement studies quantified their P600 *only* in the earlier time window between 500 and 750 ms where others quantified the frontal positivity (Angrilli et al., 2002; De Vincenzi et al., 2003; Hagoort et al., 1993; Hagoort, 2003; Osterhout & Mobley, 1995; Osterhout et al., 1996; Roehm, Bornkessel, Haider, & Schlesewsky, 2005). Only a few studies actually examined separate positivity windows, describing the effect between 500 and 700 ms as an early phase of the P600 with a more central (if not primarily frontal)

distribution, in contrast to a later P600 phase extending from about 700–1,000 ms and limited to posterior areas (Barber & Carreiras, 2005; Hagoort & Brown, 2000; Kaan, Harris, Gibson, & Holcomb, 2000; Kaan, 2002; Kaan & Swaab, 2003; Molinaro, Kim et al., 2008, 2011; Silva-Pereyra & Carreiras, 2007). Advocates for the claim that the P600 is not a monolithic component have taken such modulations in scalp distribution and timing as evidence that different positivities reflect distinct neurocognitive processes. In the agreement literature, the early/frontal positivity has been argued to represent the diagnosis of the incongruence while accessing non-syntactic, discourse-level information to detect the source of the error<sup>1</sup> (see Molinaro, Barber, et al., 2011). The late P600, in contrast, has been discussed as reflecting mechanisms of re-analysis and repair that are necessary to establish a well-formed sentence (see related “Diagnosis and Repair” theory by Fodor & Inoue, 1998, discussed in Friederici, Mecklinger, Spencer, Steinhauer, & Donchin, 2001 for garden-path sentences). The finding of larger “late P600s” in sentence contexts involving costlier repair supports this claim (Barber & Carreiras, 2005; Molinaro, Vespignani & Job, 2008; Silva-Pereyra & Carreiras, 2007).

Given that agreement studies have not uniformly investigated different positivities as reflecting potentially distinct processing stages, it remains unclear to date what factors may modulate these early/late positivities, and how consistent these ERP effects even are for different kinds of number agreement violations across languages or speakers.

#### 1.4. *This study*

Our study was based on a previous experiment conducted with Italian monolinguals by Molinaro, Vespignani, et al. (2011) (Experiment 1). Number agreement was manipulated between 3 sentence positions: (a) subject; (b) verb; and (c) an adjective modifying the subject-noun (e.g., *I lavoratori tornano dalla fabbrica sporchi di grasso/The workers<sup>(pl)</sup> return<sup>(pl)</sup> from the factory dirty<sup>(pl)</sup> with grease*). Four experimental conditions (Table 1), reflecting the four possible combinations of (dis-)agreement between the three sentence positions, were compared: (a) *Correct* (“xxx”); (b) *Inconsistent verb* (“xyx”); (c) *Inconsistent noun phrase* (“xyy”); and (d) *Inconsistent modifier* (“xxy”). Following these authors (but contrary to the majority of agreement studies), ERP correlates of morphosyntactic processing were examined on *two* target words: the *verb* and the *modifier*.

Although the original study did not test the fourth condition (xxy) and our stimuli were not identical, we expected processing patterns in our native-monolingual Italian speakers to largely replicate the findings of the previous study. On the verb, Italian native speakers had elicited a LAN<sup>2</sup> (350–450 ms) followed by an early (500–800 ms) and late (800–100 ms) P600 in response to subject-verb number mismatches. On the modifier, the authors found support of their “*Repair hypothesis*” which stipulates that the easiest way to process a number mismatch occurring early in a sentence is to repair this mismatch based on the number of the constituent on which it is detected (i.e., the verb), and to pursue this repaired/grammatical interpretation for the remainder of the sentence (i.e., integrating the modifier into the revised internal representation of the sentence). The modifier elicited a long-lasting P600 (with no preceding LAN) when it clashed with the verb

Table 1  
Experimental stimuli by condition

Condition	Subject-Noun	Verb	Intervening Phrase	Modifier	Prepositional Phrase
<b>xxx: Correct</b>	<b>x</b>	<b>x</b>		<b>x</b>	
Singular	Il lavoratore <i>The worker<sup>(sg)</sup></i>	<u>torna</u> <i><u>returns</u><sup>(sg)</sup></i>	dalla fabbrica <i>from the factory</i>	<u>sporco</u> <i><u>dirty</u><sup>(sg)</sup></i>	di grasso <i>with grease</i>
Plural	I lavoratori <i>The workers<sup>(pl)</sup></i>	<u>tornano</u> <i><u>return</u><sup>(pl)</sup></i>	dalla fabbrica <i>from the factory</i>	<u>sporchi</u> <i><u>dirty</u><sup>(pl)</sup></i>	di grasso <i>with grease</i>
<b>xyx: Inconsistent verb</b>	<b>x</b>	<b>y</b>		<b>x</b>	
Singular	Il lavoratore <i>The worker<sup>(sg)</sup></i>	* <u>tornano</u> <i>*<u>return</u><sup>(pl)</sup></i>	dalla fabbrica <i>from the factory</i>	<u>sporco</u> <i><u>dirty</u><sup>(sg)</sup></i>	di grasso <i>with grease</i>
Plural	I lavoratori <i>The workers<sup>(pl)</sup></i>	* <u>torna</u> <i>*<u>returns</u><sup>(sg)</sup></i>	dalla fabbrica <i>from the factory</i>	<u>sporchi</u> <i><u>dirty</u><sup>(pl)</sup></i>	di grasso <i>with grease</i>
<b>xxy: Inconsistent noun</b>	<b>x</b>	<b>y</b>		<b>y</b>	
Singular	Il lavoratore <i>The worker<sup>(sg)</sup></i>	* <u>tornano</u> <i>*<u>return</u><sup>(pl)</sup></i>	dalla fabbrica <i>from the factory</i>	<u>sporchi</u> <i><u>dirty</u><sup>(pl)</sup></i>	di grasso <i>with grease</i>
Plural	I lavoratori <i>The workers<sup>(pl)</sup></i>	* <u>torna</u> <i>*<u>returns</u><sup>(sg)</sup></i>	dalla fabbrica <i>from the factory</i>	<u>sporco</u> <i><u>dirty</u><sup>(sg)</sup></i>	di grasso <i>with grease</i>
<b>xyy: Inconsistent modifier</b>	<b>x</b>	<b>x</b>		<b>y</b>	
Singular	Il lavoratore <i>The worker<sup>(sg)</sup></i>	<u>torna</u> <i><u>returns</u><sup>(sg)</sup></i>	dalla fabbrica <i>from the factory</i>	* <u>sporchi</u> <i>*<u>dirty</u><sup>(pl)</sup></i>	di grasso <i>with grease</i>
Plural	I lavoratori <i>The workers<sup>(pl)</sup></i>	<u>tornano</u> <i><u>return</u><sup>(pl)</sup></i>	dalla fabbrica <i>from the factory</i>	* <u>sporco</u> <i>*<u>dirty</u><sup>(sg)</sup></i>	di grasso <i>with grease</i>

Notes. Number was counterbalanced such that the subject noun was either singular or plural. English translations are presented in italics. Target words (verb, modifier) are underlined. The asterisk marks the point of first violation.

(xyx), but not when it was congruent with the repaired version of the sentence (xxy). For both the verb and the modifier, the positivity showed a different scalp distribution depending on the processing stage; in the earlier time window (500–800 ms), the positivity was larger at fronto-central than posterior sites, whereas the later P600 (800–1,000 ms) was mainly posterior.

Extending this paradigm to the study of bilingualism and attrition, our goals were to examine the potential changes in the online detection and repair strategies in attriters' L1 morphosyntax. Specifically, we assessed whether (a) attriters differed from non-attriter native speakers (Controls) in their behavioral and/or ERP response patterns (e.g., amplitude and/or distributional differences in the negativity, frontal positivity and P600); and whether (b) differences in processing were modulated by L1-proficiency, L2-to-L1 transfer and/or characteristics specific to attriters' socio-linguistic circumstances (e.g., length of residence, age at immigration, amount of exposure to L1 relative to L2). For example, Attriters might elicit a more robust negativity in response to subject-verb agreement violations, as a result of L2-English influence (stronger expectations of subject-verb agreement in English than Italian). Differences in online morphosyntactic re-analysis and repair strategies might also be found in the P600 time-windows. Unlike native-



monolinguals who have been shown not to elicit a P600 effect in response to the repair condition (xxy), Attriters may process these sentences as morphosyntactic violations. Modulations in amplitude and duration of frontal and posterior positivities would reflect qualitatively different processing strategies, with larger and more prolonged P600s indicating more elaborated sentence repair (Molinaro et al., 2008).

Not only is this one of the earliest ERP investigations of L1 attrition, but also the first experiment to examine online morphosyntactic processing in attriters at multiple points in a given sentence, in an attempt to determine whether Attriters detect and recover from erroneous analyses in the same way as non-attributing Italian native speakers.

## 2. Methods

### 2.1. Participants

Participants in the attrition group consisted of 24 Italian native speakers (14 female;  $M_{\text{age}}$ : 36; range: 25–50) who had immigrated to Canada in adulthood ( $M_{\text{age}}$  at immigration (AoA of English)<sup>3</sup> = 28.2 years;  $M$  length of residence = 11 years). All participants reported having a very limited use of their native Italian, and having noticed changes or difficulties in their native-Italian fluency as a result of their predominant use of English. Based on these unanimous self-reports of changes to L1 exposure and L1 difficulties while immersed in an L2 context, we refer to these individuals as “Attriters.” Considering the cross-linguistic similarity between Italian and French, we only recruited attriters who had minimal knowledge or use of French (although living in Montreal).

Thirty Italian native speakers residing in Italy were recruited as a control group (17 female;  $M_{\text{age}}$ : 31; range = 25–54). Participants had minimal exposure to second languages (including English and Italian dialects), which we operationally defined as < 5 h per week.

All participants but one were right-handed and with no known history of neurological disorders.

### 2.2. Behavioral measures

Participants completed a background questionnaire pertaining to their demographic information (age, gender, education level), and language background. Attriters answered additional questions about their immigration history, first language exposure/use, motivation to maintain or achieve native-like proficiency in each language, and identity/attitudes toward each language and culture.

All participants (including Controls) completed four proficiency measures: (a) a *written self-report measure* where they were asked to rate their proficiency level on a scale from 1 to 7 in listening comprehension, reading comprehension, pronunciation, fluency, vocabulary, and grammatical ability; (b) a *written C-test* (Italian version: Kraš, 2008), where they were asked to fill in the blanks in five short texts in which twenty words in each text

had been partially deleted; (c) a *written error-detection test* designed specifically for this study, where participants had to detect and correct a number of errors in two separate texts; and lastly, (d) a timed *verbal semantic fluency task* where participants were asked to produce as many vocabulary items from two categories (“animals” and “fruits and vegetables”) as possible within 1 min. Participants also completed (a) a timed *reading fluency task* where they had to silently read and answer as many true-false statements as possible in 3 min (Woodcock, McGrew, Mather, & Schrank, 2003; adapted into Italian for this study); and (b) the *letter-number-sequencing task* from the Italian WAIS-IV as a measure of working memory (Orsini & Pezzuti, 2013). The purpose of these tasks was to ensure that group differences were not due to differences in reading speed and/or working memory capacity, given the rapid-serial-visual presentation mode of the sentence stimuli during the ERP experiment. Group means are provided in Table 2. Although Attriters scored numerically lower on all four proficiency measures, they did not differ significantly from Controls ( $p > .1$ ).

Subgroups of “*high*” and “*low*” proficiency were derived by median split. Note that, for the sake of brevity, we refer to individuals in the lower range as “low proficiency,” but it is obvious that their proficiency level is not “low” in the conventional sense of the word. *High* and *low* proficiency subgroups differed significantly on all measures except working memory (WM) span, but especially on the written measures (Table 3). High and low proficiency Attriters differed significantly on all measures (including WM), whereas high and low proficiency Controls only differed on the two written proficiency tests. Numerically, low proficiency Attriters scored lower than low proficiency Controls on all proficiency measures, but differences were not significant ( $ps > .1$ ). *High* and *low* subgroups of Attriters did not differ significantly on their age at testing ( $p = .2$ ), AoA ( $p = .8$ ), or length of residence ( $p = .1$ ).

Table 2

Group means (standard deviation) for proficiency and control tasks ( $ps > .1$ )

Behavioral Measures	Controls ( $n = 30$ )	Attriters ( $n = 24$ )
Self-report of proficiency (7-point scale)	7 (0)	6.87 (0.2)
Listening comprehension	7 (0)	7 (0)
Reading comprehension	7 (0)	7 (0)
Pronunciation	7 (0)	6.96 (0.2)
Fluency	7 (0)	6.79 (0.6)
Vocabulary	7 (0)	6.63 (0.7)
Grammar	7 (0)	6.83 (0.4)
C-test (%)	96.3 (4.4)	95.2 (4.6)
Error-detection test (%)	90.0 (5.1)	89.5 (5.9)
Verbal semantic fluency (average of two categories)	23.4 (5.5)	21.5 (3.9)
Reading fluency (no. correct in 3 min)	71.6 (13.0)	75.3 (15.0)
Working memory		
Correct	11.2 (2.7)	11.9 (2.6)
Span	5.4 (1.1)	5.7 (1.1)

Table 3  
 Proficiency subgroup means (standard deviation) for proficiency and control tasks. HP and LP are denotations for “higher” or “lower” proficiency, although L1 proficiency is not “low” in the conventional sense of the word

Behavioral Measures	Controls		Atrriers		All	
	HP	LP	HP	LP	HP	LP
Self-report of proficiency (7pt. scale)	7 (0)	7 (0)	6.97 (0.1)	6.74 (0.3)	6.9 (0.1)	6.8 (0.2)
Listening comprehension	7 (0)	7 (0)	7 (0)	7 (0)	7 (0)	7 (0)
Reading comprehension	7 (0)	7 (0)	7 (0)	7 (0)	7 (0)	7 (0)
Pronunciation	7 (0)	7 (0)	7 (0)	6.9	7 (0)	6.9 (0.2)
Fluency	7 (0)	7 (0)	7 (0)	6.5	7 (0)	6.8 (0.6)
Vocabulary	7 (0)	7 (0)	6.9	6.3	6.9 (0.2)	6.7 (0.6)
Grammar	7 (0)	7 (0)	6.9	6.7	6.9 (0.2)	6.8 (0.3)
C-test (%)	98.8 (1.1)	93.6 (2.8)	97.6 (2.7)	92.3 (4.7)	98.2 (2.1)	93.0 (3.7)
Error-detection test (%)	94.6 (3.4)	86.2 (4.3)	93.9 (3.6)	84.3 (3.4)	94.3 (3.6)	85.4 (3.9)
Verbal semantic fluency (average)	23.5 (6.9)	21.6 (5.6)	23.8 (3.0)	18.9 (3.3)	23.6 (5.3)	20.5 (4.9)
Reading fluency (No. correct)	73.4 (9.1)	70.0 (15.8)	82.4 (12.4)	66.8 (13.9)	77.7 (68.7)	68.7 (14.8)
Working memory						
Correct	11.5 (2.4)	10.9 (2.9)	13.2 (2.7)	10.5 (1.5)	12.3 (2.7)	10.7 (2.4)
Span	5.4 (1.1)	5.3 (1.3)	6.2 (1.2)	5.1 (0.7)	5.8 (1.2)	5.2 (1.1)

\*p < .05; \*\*p < .01; \*\*\*p < .005; NS = not significant (p > .1).

### 2.3. Stimuli

Sentence examples from all four conditions of the ERP study are provided in Table 1. The experimental stimuli consisted of eight-word sentences containing two target words: (a) a lexical *verb* (in third position); and (b) a *modifier* (in sixth position). Each sentence began with a masculine, animate subject noun phrase. The determiner matched the noun in number, as we were not interested in creating determiner-noun mismatches. Half of the subject-nouns were plural, and half singular. The verb and the modifier were separated by two constituents—a function word and an inanimate noun. This time lag was necessary to provide enough time for a possible structural re-analysis and to allow for slow-going ERP waves such as the P600 elicited by the verb to return to baseline prior to the presentation of the modifier. The intervening noun was inanimate and feminine in gender to eliminate ambiguity that would lead readers to attach the modifier to the intervening noun rather than to the subject-noun. In cases where the verb was intransitive (52.5%), the intervening phrase was a prepositional phrase, while for transitive sentences (47.5%), the intervening words consisted of a noun phrase (determiner + direct object noun). Sentences always ended with a prepositional phrase, in order for sentence wrap-up effects not to be confounded with effects on the modifier. Each target word contributed to each condition (counter-balanced across subjects), thus ruling out that effects in the grand-average were driven by contextual or lexical (frequency, length) differences between conditions. There were no repetitions of subject nouns, verbs, or modifiers across items.

Several modifications were made to the stimuli used by Molinaro, Vespignani et al. (2011; Experiment 1) which resulted in the creation of a number of new sentences. First, we balanced singular and plural versions of each sentence<sup>4</sup> to minimize the predictability of our agreement combinations and the possibility of identifying violations on a superficial level. Second, we changed the tense of the verbs from the remote past (“*passato remoto*”) to the present tense, as the remote past tense is subject to regional differences in Italy and is used somewhat infrequently in some regions. We additionally balanced transitivity, as the original stimuli contained an uneven proportion of intransitive and transitive sentence constructions. We also substantially reduced repetitions of non-target segments (intervening and sentence-final phrases). Finally, we replaced several nouns, verbs and modifiers that exceeded 10 letters in length, to ensure that words and their agreement inflections could be read in full without saccadic artifacts. Our stimuli were verified by two Italian native speakers.

A set of 120 different sentences were constructed and realized in each of the eight conditions (four main conditions × singular/plural). Eight experimental lists were created such that, across lists, each sentence contributed equally to each condition, while no sentence was repeated within any of the experimental lists. Each participant also saw 204 filler sentences, which were part of the larger study (testing Italian lexical-semantic processing and relative clause sentences) and will be reported in forthcoming papers (see Kasparian, Vespignani, & Steinhauer, 2013a,b, 2014a,b). Out of the total of 324 pseudorandomized stimuli (120 experimental and 204 fillers) per participant, 146 sentences (approx. 45%) were acceptable (grammatically and semantically), whereas

178 were expected to receive a rating of 3 or lower on a five-point rating scale (approx. 55%).

#### 2.4. Procedure

All participants provided informed consent prior to their participation in the study. After completing the questionnaires and behavioral tasks, participants were fitted with the EEG (electroencephalogram) cap and instructed that their task would be to rate the acceptability of various Italian sentences on a scale from 1 (severely ungrammatical and/or does not make sense) to 5 (perfect). Participants were encouraged to use the entire rating scale, rather than making a categorical judgment of “unacceptable” and “acceptable” using only 1 and 5. They were asked to decide based on their own intuition what types of errors would be considered more severe than others. The rationale for using a rating scale rather than a binary acceptability judgment task was threefold: (a) among native and highly proficient speakers, a rating scale may be more sensitive to individual variation and to fine-grained group differences than yes/no decisions; (b) to assess whether conditions showed graded response patterns (and whether these were related to graded ERP patterns); and (c) the filler sentences that were part of the larger study did not only contain outright violations but also tested infrequent/dispreferred word-order constructions or the likelihood of cross-linguistic co-activation where a binary judgment would not adequately capture the range of permissibility of these constructions.

Participants were seated in a comfortable chair in a dimly lit, sound-attenuated booth, at approximately 80 cm from the computer monitor with a *Cedrus* seven-button RB-740 response box placed in front of them (Cedrus Corporation, San Pedro, CA, USA). Subjects received a short demonstration to show how eye movement, blinks, and muscle movement create artifacts in the EEG signal, and they were encouraged to blink only between trials, as prompted by the image of an eye on the screen. They were instructed to carefully read each sentence until the end and to respond as quickly and as accurately as possible once the prompt appeared. A practice block of 20 sentences representative of those used in the actual experiment familiarized participants with the procedure.

Words were presented in white 40-font Arial characters, at the center of a black background. The sentence-final word was presented along with the period. Each trial began with the presentation of a white fixation cross for 500 ms, followed up for 200 ms by a blank screen (ISI). Each word then appeared one at a time for 300 ms (+ 200 ms ISI). A visual prompt (“???”) followed the offset of the sentence-final word, indicating the onset of the response interval. The prompt remained on the screen until participants pressed a button from 1 to 5. Immediately after their response, the image of the blue eye appeared at the center of the screen for a 2,000 ms interval for participants to blink their eyes. The next trial began after the blinking interval, with the presentation of another fixation cross. Each session lasted approximately 3 h, including setup, short breaks and cap removal. All consent forms, materials, and procedures were fully approved by the Ethics Review Board of each institution (Faculty of Medicine, McGill

University and Ethical Committee for Human Research, University of Trento) for the duration of the study.

### 2.5. EEG recording and analysis

The EEG was recorded continuously from 25 Ag/AgCl electrodes, 19 of which were electrodes mounted on a standard electro-cap according to the 10–20 system (Jasper, 1958), and six of which were external electrodes: four electro-oculogram (EOG) channels placed above and below the left eye (EOGV), and at the outer canthus of each eye (EOGH), as well as two reference electrodes placed on the mastoids (A1 and A2). All electrodes were referenced online to the left mastoid (A1). Impedances were kept strictly below 5 k $\Omega$  for scalp and reference electrodes, and below 10 k $\Omega$  for EOG electrodes. Signals were amplified using NeuroScan (Canada) and BrainVision (Italy) and filtered online with a band-pass filter of 0.1 to 100 Hz at a sampling rate of 500 Hz. Data pre-processing and analyses were carried out using EEProbe (ANT, Enschede, Netherlands).

Offline, EEG recordings were re-referenced to the average activity of the two mastoids<sup>5</sup> and filtered with a phase-true 0.3–40 Hz band-pass filter. Trials containing artifacts due to blinks, eye movements, and excessive muscle activity were rejected prior to averaging, using a moving-window (400 ms) standard deviation of 30 microvolts. On average, participants contributed 27/30 artifact-free trials per condition (range: 54%–100%), with no differences across conditions for either target word ( $ps > .1$ ).

ERPs were analyzed separately on the verb (–200 to 1,200 ms) and the modifier (–200 to 1,600 ms), and were time-locked to the onset of each target word with a baseline correction from –200 to 200 ms.<sup>6</sup> ERPs were quantified in time windows corresponding to each component of interest, based on previous agreement studies and on visual inspection of the grand average data for each participant group. On the verb, the time windows were (1) 300–500 (LAN/N400); (2) 550–650 (early frontal positivity); (3) 650–1,000 (P600); and (4) 1,000–1,200 (late P600). On the modifier, slightly different time-windows were selected based on visual inspection, especially to ensure that the negativity and positivity did not overlap in a given time window: (1) 300–500 (LAN/N400); (2) 500–600 (intermediate window); (3) 600–900 (P600); and (4) 1,000–1,300 (late P600). Note that the previous study by Molinaro, Vespignani et al. (2011) did not include a late P600 window extending beyond 1,000 ms, although figures in their paper indicate that the P600 had not yet returned to baseline at 1,000 ms.

To represent the four (dis-)agreement conditions in our ERP analyses, we crossed two factors: *Agreement 1* (= Ag1), describing (dis-)agreement between the *first* two sentence positions (i.e., subject-noun and verb), and *Agreement 2* (= Ag2), describing (dis-)agreement between the *last* two sentence positions (i.e., verb and modifier), each with two levels (*correct* and *violation*). Thus, conditions xxx and xyx were collapsed into Ag1-*correct* sentences, xyx and xyy were Ag1-*violation* sentences, xxx and xyy were Ag2-*correct* sentences, and finally xyx and xyy were collapsed into Ag2-*violation* sentences. Although, on the verb, only Ag1 is meaningful (as the third target word has not yet been

encountered), *Ag2* was included as a factor in the global ANOVA for the verb position to confirm that modulations in *Ag2* had no effect. This also allowed us to conduct identical ANOVAs on both target words. On the modifier, when interactions in the global ANOVA motivated follow-up comparisons by condition pairs (e.g., *xyy* vs. *xxx*), the factor “*Condition*” was used to describe the contrast (e.g., two levels: *xyy*, *xxx*). Note that statistical analyses of the correct condition in the same time windows revealed no significant group differences in ERP patterns between Controls and Attriters (all  $ps > .1$  for all verb and modifier time-windows of interest).

Repeated-measures ANOVAs were performed separately for four midline electrodes (Fz, Cz, Pz, Oz) and 12 lateral electrodes (F3/4, C3/4, P3/4, and F7/8, T3/4, T5/6). Global ANOVAs for the midline sites included within-subject factors *Ag1* (correct, violation), *Ag2* (correct, violation), *Ant-Post* (anterior, central, parietal, occipital), whereas lateral ANOVAs additionally included factors *Hemisphere* (left, right) and *Laterality* (lateral, medial). For all ANOVAs, *Group* (Controls, Attriters) and *Proficiency* (High, Low) were between-subjects’ factors.

Where appropriate, Greenhouse-Geisser correction was applied to analyses with more than two levels (e.g., *Ant-Post*). In these cases, the corrected  $p$  values but original degrees of freedom are reported. As a default, reported analyses are restricted to the midline only, except in cases where the lateral ANOVAs revealed additional effects (e.g., LAN).

### 3. Results

#### 3.1. Acceptability judgments

Acceptability ratings (on a scale from 1 to 5) for each sentence condition are shown in Fig. 1a. A repeated-measures ANOVA with within-subjects factor *Condition* (*xxx*, *xyx*, *xyy*, *xyy*) and between-subjects factor *Group* (Controls, Attriters) revealed a significant main effect of *Condition* ( $F(3, 156) = 146.99$ ,  $p < .001$  after Greenhouse-Geisser correction), but no effects or interactions with *Group* ( $ps > .10$ ). Follow-up analyses indicated that the correct condition *xxx* received a significantly higher rating than violation conditions *xyx* ( $F(1, 52) = 165.17$ ,  $p < .01$ ), *xyy* ( $F(1, 52) = 143.51$ ,  $p < .01$ ), and *xyy* ( $F(1, 52) = 162.54$ ,  $p < .01$ ), but that the violation conditions did not differ significantly from each other ( $p > .05$ ).

#### 3.2. Reaction times

Reaction times between the onset of the prompt and participants’ button press are shown in Fig. 1b. A repeated-measures ANOVA with within-subjects factor *Condition* (*xxx*, *xyx*, *xyy*, *xyy*) and between-subjects factor *Group* (Controls, Attriters) revealed a significant main effect of *Condition* ( $F(3, 156) = 6.958$ ,  $p < .01$ ), as well as a significant main effect of *Group* ( $F(3, 156) = 6.263$ ,  $p < .05$ ), but no significant interaction between *Condition*  $\times$  *Group* ( $p > .1$ ), indicating that Attriters took longer to respond

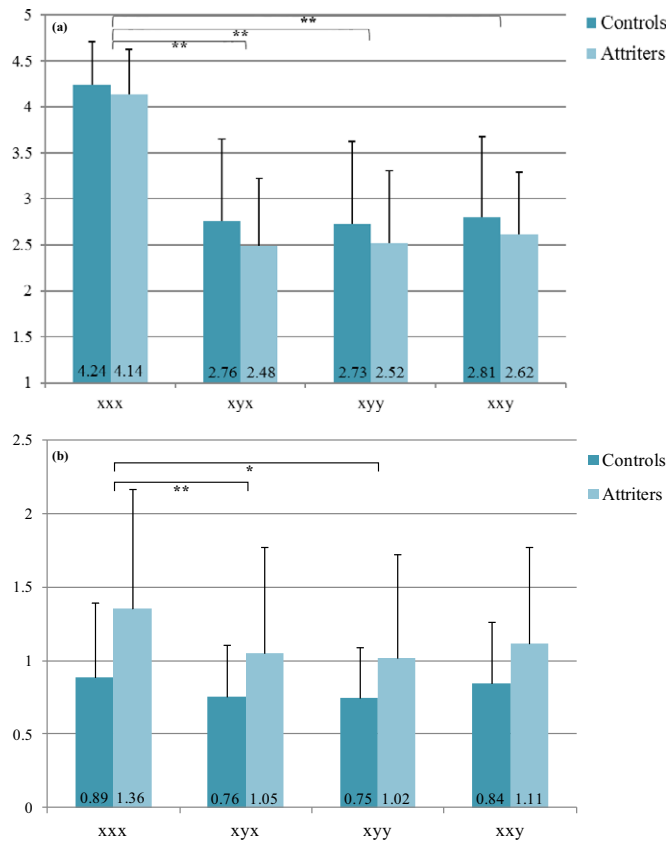


Fig. 1. (a) Group acceptability ratings on a scale from 1 (completely unacceptable) to 5 (perfect) by condition. Attriters do not differ overall from Controls.  $*p < .05$ ;  $**p < .01$ . Error bars represent standard deviation. (b) Group reaction times (in seconds) by condition. Attriters were consistently slower than Controls ( $p < .05$ ).  $*p < .05$ ;  $**p < .01$ . Error bars represent standard deviation.

overall than the Controls. Follow-up analyses indicated that response times to the correct *xxx* condition were significantly longer than response times for *xyx* violations ( $F(1, 52) = 10.03$ ,  $p < .01$ ) and *xyy* violations ( $F(1, 52) = 10.03$ ,  $p < .05$ ) but not significantly different from response times for *xxy* violations ( $p > .1$ ), suggesting that participants were faster in responding to sentences where a violation was already present on the verb than when the two-first constituents agreed in number. The *xxy* condition only marginally differed from *xyx* ( $F(1, 52) = 3.36$ ,  $p = .07$ ) and *xyy* violations ( $F(1, 52) = 3.47$ ,  $p = .07$ ).

### 3.3. ERPs elicited at the verb position

Grand average ERP waveforms for *Ag1* (subject-verb) conditions time-locked to the verb are presented in Fig. 2a (Controls) and Fig. 2b (Attriters).



In Controls, subject-verb violations elicited a small left-temporal negativity (LTN) localized primarily at T5 between 300 and 500 ms (Fig. 3a), followed by a frontal positivity between 550 and 650 ms and a large posterior P600 lasting until 1,200 ms. Attriters showed a prominent negativity between 300 and 500 ms which was primarily distributed over left and midline sites (Fig. 3b), followed by a frontal positivity between 550 and 650 ms, and a large P600 which appeared to be shorter in duration (lasting until 1,000 ms) and less focal (less posterior) than in Controls. These observations were confirmed by our statistical analyses.

### 3.3.1. Negativity between 300 and 500 ms

The global ANOVA in the 300–500 ms time window for midline electrodes revealed a significant main effect of *Ag1* ( $F(1, 50) = 16.63, p < .001$ ) and a significant *Ag1*  $\times$  *Group* interaction ( $F(1, 50) = 5.27, p < .05$ ). No interactions with *Ant-Post* or *Proficiency* reached significance ( $ps > .1$ ). Follow-ups by *Group* demonstrated a significant main effect of *Ag1* in Attriters ( $F(1, 23) = 18.54, p < .001$ ) but not in Controls ( $p > .1$ ), confirming that only Attriters elicited a broadly distributed negativity on the midline when processing a verb that mismatched in number with the preceding subject-noun.

The lateral global ANOVA revealed a significant main effect of *Ag1* ( $F(1, 50) = 15.31, p < .0001$ ) and a significant *Ag1*  $\times$  *Hemi* interaction ( $F(1, 50) = 7.95, p < .01$ ), reflecting a stronger negativity over left sites ( $F(1, 50) = 21.38, p < .0001$ ) rather than right ( $F(1, 50) = 5.83, p < .05$ ). Unlike at the midline, the interaction between *Ag1*  $\times$  *Group* was marginal ( $F(1, 50) = 2.89, p = .09$ ), as was the interaction between *Ag1*  $\times$  *Group*  $\times$  *Laterality* ( $F(1, 50) = 3.45, p = .07$ ). Follow-up analyses by *Laterality* supported the group differences observed during visual inspection: At medial sites, a significant main effect of *Ag1* ( $F(1, 50) = 15.93, p < .001$ ) was qualified by a significant *Ag1*  $\times$  *Group* interaction ( $F(1, 50) = 5.12, p < .05$ ), where Attriters elicited a significant negativity (*Ag1*:  $F(1, 23) = 14.78, p < .001$ ), but Controls did not ( $p > .1$ ). At lateral sites, however, the ANOVA revealed a significant main effect of *Ag1* ( $F(1, 50) = 10.96, p < .01$ ) and no significant interaction with *Group* ( $p > .1$ ), indicating that the negativity was shared by both groups at lateral electrodes. No significant differences were found between *Proficiency* subgroups ( $ps > .1$ ).

### 3.3.2. Frontal positivity between 550 and 650 ms

Between 550 and 650 ms, the positivity elicited by *Ag1* violations relative to correct sentences reached statistical significance across groups (*Ag1*:  $F(1, 50) = 12.31, p < .001$ ) and was qualified by a significant *Ag1*  $\times$  *Ant-Post* interaction ( $F(3, 150) = 7.34, p < .005$ ), reflecting that the positivity was most robust at Fz ( $F(1, 50) = 19.24, p < .0001$ ) than at more posterior electrodes (Cz:  $p < .005$ ; Pz:  $p < .01$ ; Oz:  $p > .1$ ). The frontal positivity did not statistically differ by *Group* or *Proficiency* ( $ps > .1$ ).

### 3.3.3. P600 between 650 and 1,000 ms

A highly significant main effect of *Ag1* ( $F(1, 50) = 36.26, p < .0001$ ) was qualified by a significant interaction with factor *Ant-Post* ( $F(3, 150) = 17.56, p < .0001$ ), reflecting

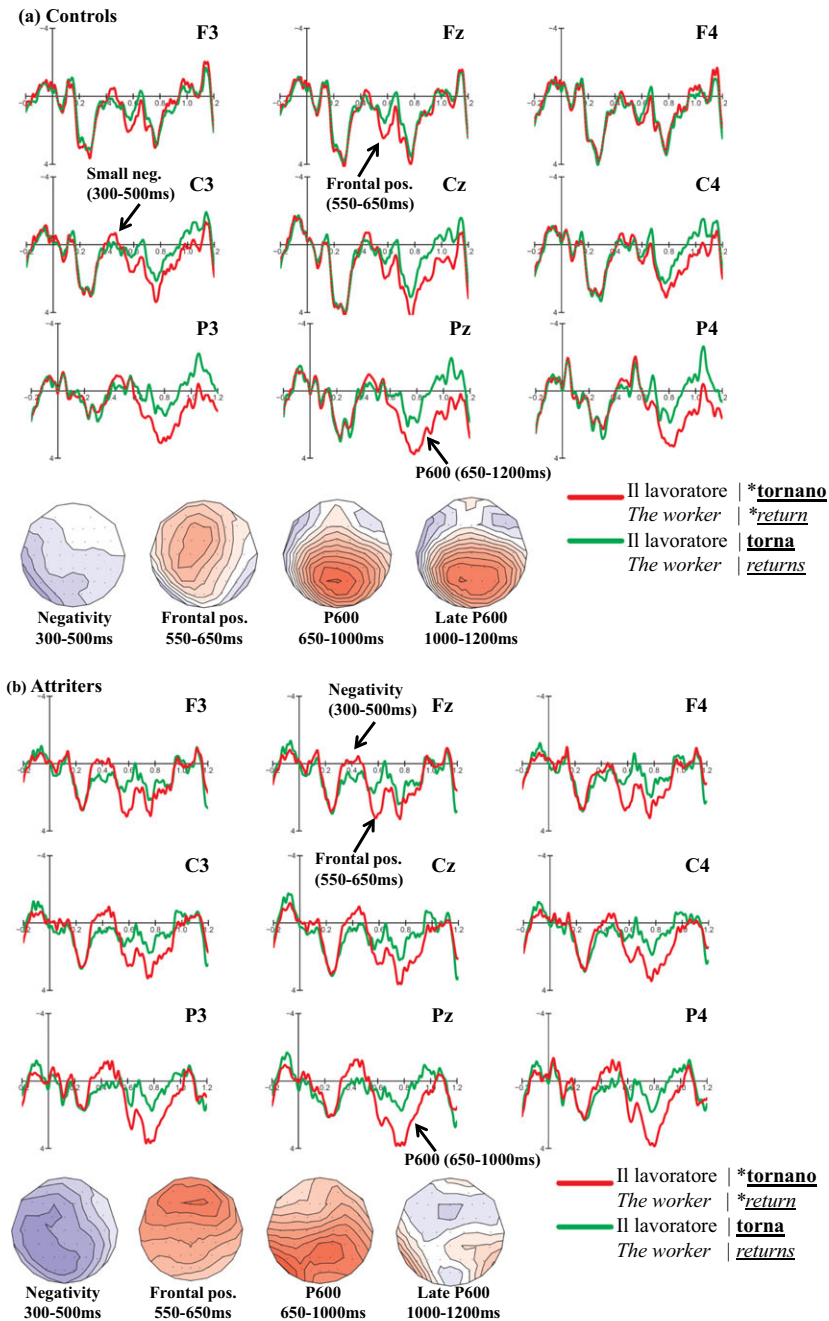


Fig. 2. Event-related potentials (ERPs) elicited by the verb in response to *AgI* violations (red) compared to *AgI* correct (green) in Controls (a) and Attriters (b). Time ranges (in milliseconds) depicted on the *x*-axis are relative to the onset of the verb (0 ms). Negative values are plotted up. Voltage maps illustrate the scalp distribution of the effects observed for the time windows of interest.

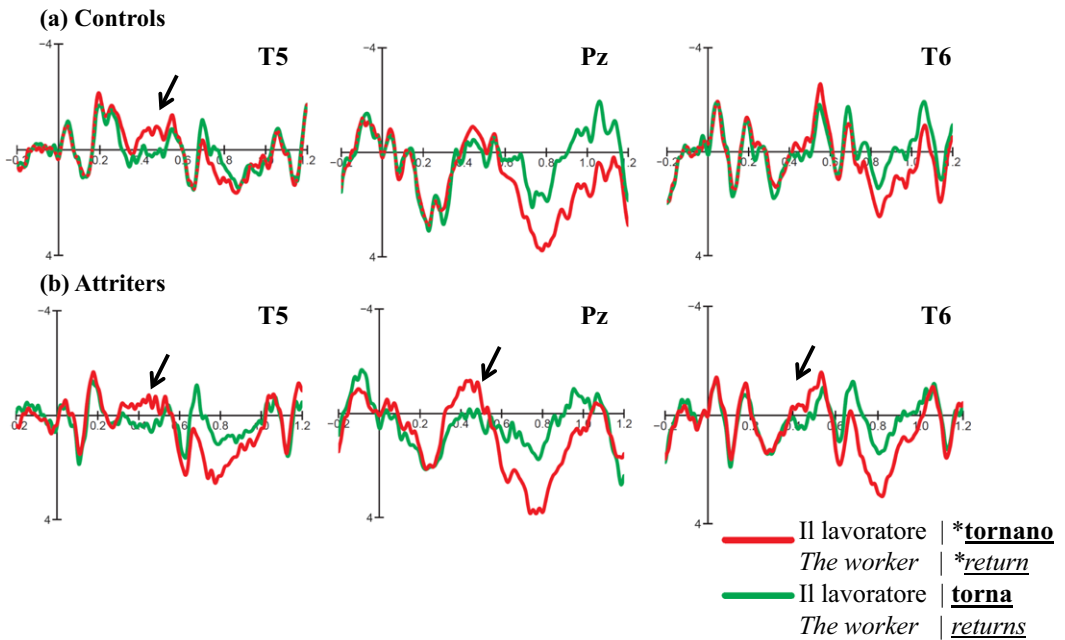


Fig. 3. A comparison of electrodes T5 (left), Pz (midline), and T6 (right) for *AgI* violations (red) relative to *AgI* correct (green) in Controls and Attriters. Time ranges (in milliseconds) depicted on the *x*-axis are relative to the onset of the verb (0 ms). Negative values are plotted up. The negativity elicited in Controls is focused at T5, whereas it is broadly distributed for Attriters.

the posterior distribution of the P600 (Fz:  $p = .05$ ; Cz:  $p < .0001$ ; Pz:  $p < .0001$ ; Oz:  $p < .0001$ ). Surprisingly, the expected interaction with factor *Proficiency* did not reach significance ( $ps > .1$ ) in the global ANOVA (but see correlational analyses below). The lack of a significant *Group* interaction points to a P600 effect of similar amplitude and scalp distribution for Controls and Attriters in this time window.

Correlations revealed that P600 amplitude at Pz (where the effect was maximal) was positively correlated with the C-test ( $r = .320$ ,  $p < .05$ ) and verbal semantic fluency ( $r = .345$ ,  $p < .01$ ), such that individuals with higher Italian proficiency scores elicited a larger P600 effect in response to subject-verb number agreement mismatches.<sup>7</sup>

### 3.3.4. Late P600 between 1,000 and 1,200 ms

In this late interval, *AgI* violations elicited a posterior P600 (*AgI*:  $F(1, 50) = 27.44$ ,  $p < .0001$ ; *AgI*  $\times$  *Ant-Post*:  $F(3, 150) = 26.63$ ,  $p < .0001$ ) compared to correct sentences. A significant interaction with factor *Group* ( $F(1, 50) = 10.33$ ,  $p < .005$ ) indicated that this effect was present in the Controls ( $F(1, 29) = 21.79$ ,  $p < .0001$ ) but not the Attriters ( $p > .1$ ). *Proficiency* was not a meaningful factor in this late P600 time window ( $ps > .1$ ), in contrast to the previous P600 window (Fig. 4). Correlational analyses confirmed that proficiency scores did not modulate the P600 in this late interval ( $ps > .1$ ).

### 3.4. ERPs elicited at the modifier position

On the modifier, condition *xyy* was expected to elicit a large negativity followed by a large P600, as it constituted the most salient number agreement violation out of the four conditions. To replicate Molinaro et al.'s reports that subject-noun number mismatches are repaired on the basis of verb number before the modifier is subsequently integrated into the sentence, we would expect that condition *xyy* would not differ from condition *xxx* on the modifier, but that condition *xyx* would elicit violation effects. Finally, we expected ERP patterns to be affected by group membership and/or by Italian proficiency.

ERP waveforms for the *xyy* condition versus *xxx* for each group (Fig. 5) illustrate that Controls and Attriters show a similar pattern, namely a large, broadly distributed N400-like negativity (most prominent at medial electrodes) followed by a large parietal P600. In the Attriters, the P600 is also present at frontal sites, while no frontally distributed positivity is discernible in Controls. The scalp distribution and duration of the negativity seem to be influenced by proficiency (Fig. 6), with lower proficiency individuals eliciting a less left-lateralized and longer lasting negativity (until 600 ms). The P600 appears to be differentially modulated both proficiency level (600–900 ms) and by group (1,000–1,300 ms), with lower proficiency individuals eliciting a less focal P600 of smaller amplitude in the earlier time window, and Attriters eliciting a shorter lived P600 than Controls.

This differential proficiency/group effect on different slices of the P600 is further illustrated with difference waves in Fig. 7 and is reminiscent of the pattern we had observed on the verb in Fig. 4.

Comparisons of violation conditions *xyx* and *xyy* with the correct condition *xxx* (Fig. 8) show that, in Controls, the *xyx* condition elicits a P600 starting around 650 ms, whereas

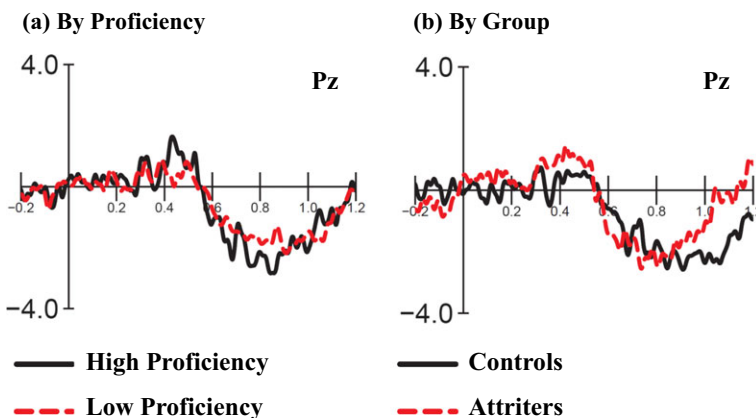


Fig. 4. P600 difference waves (*AgI* violation—correct) at Pz emphasizing proficiency differences (LP in red < HP in black) in the early time window (a) but group differences (Attriters in red < Controls in black) in the late time window (b).

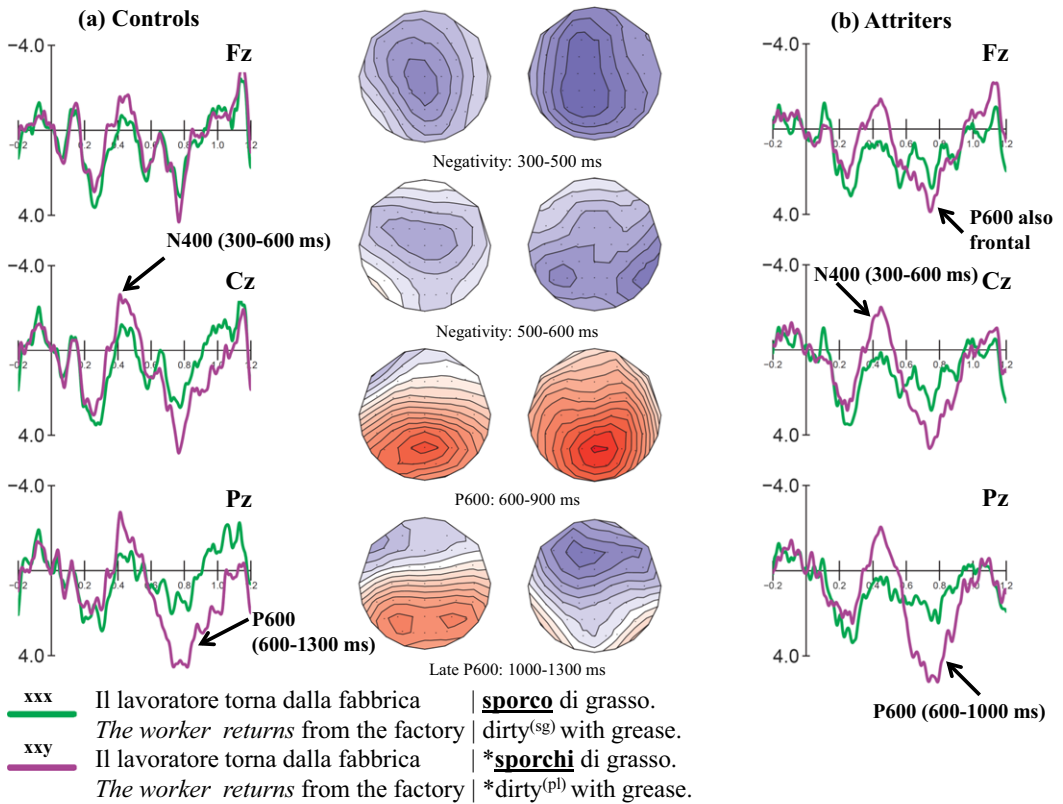


Fig. 5. Event-related potentials (ERPs) elicited by the modifier in response to xxy violations (purple) compared to xxx (green) shown at the midline for Controls (a) and Attriters (b). Voltage maps illustrate that the topography of the effects are similar in both groups, but the P600 in Attriters appears less focal and shorter in duration.

the xyy condition seem to largely overlap with the correct control condition in the P600 interval. In contrast, Attriters seem to show a clear P600 effect for both violation conditions, relative to correct sentences.

3.4.1. Negativity (LAN/N400) between 300 and 500 ms

The global ANOVA in the 300–500 ms time window for midline electrodes revealed a significant main effect of Ag2 (i.e., agreement between verb and modifier;  $F(1, 50) = 11.92, p < .001$ ) as well as a main effect of Ag1 (i.e., agreement between subject and verb;  $F(1, 50) = 6.14, p < .05$ ), which were qualified by a significant interaction between  $Ag1 \times Ag2$  ( $F(1, 50) = 17.57, p < .0001$ ).  $Ag1 \times Ant-Post$  was also significant ( $F(3, 150) = 4.74, p < .005$ ), as was  $Ag2 \times Ant-Post \times Proficiency$  ( $F(3, 150) = 2.68, p < .05$ ) and  $Ag2 \times Ag1 \times Ant-Post$  ( $F(3, 150) = 3.80, p < .05$ ). Interactions with *Group* did not reach significance, indicating that the negativity was shared between Controls and Attriters.

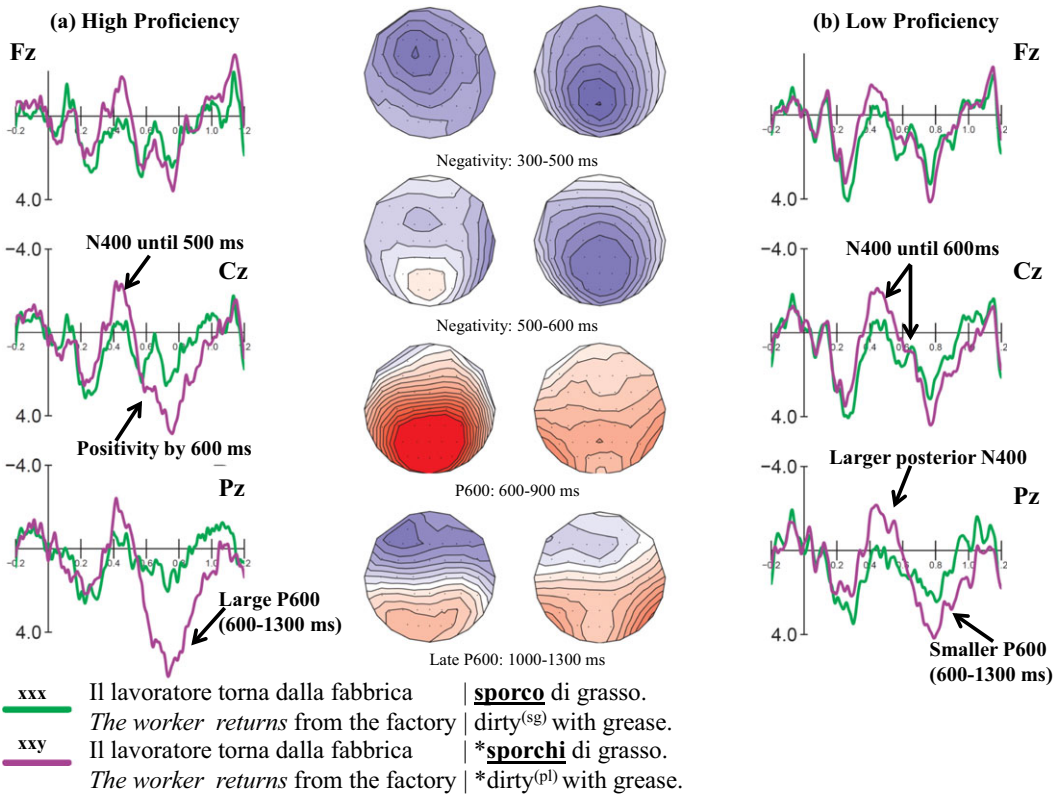


Fig. 6. Event-related potentials (ERPs) elicited by the modifier for *xxy* violations (purple) compared to *xxx* (green) in High Proficiency (a) versus Low Proficiency (b) groups. Low proficiency individuals show a less frontal and longer lasting negativity, as well as a weaker P600 effect relative to High Proficiency speakers.

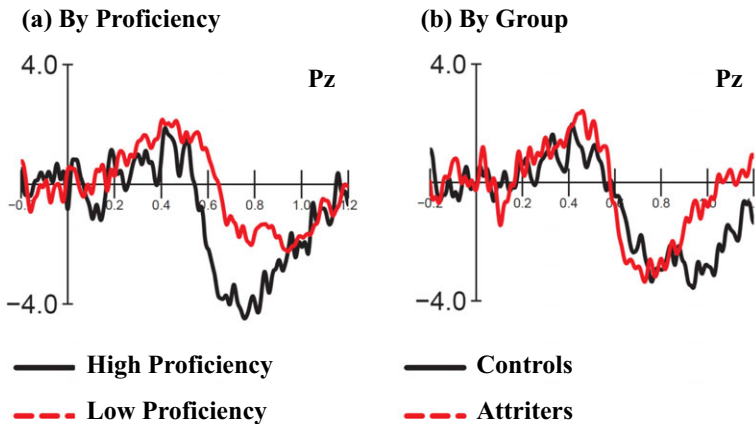


Fig. 7. P600 difference waves (*xxy*-*xxx*) at Pz emphasizing proficiency differences (LP in red < HP in black) in the early time window (a) but group differences (Attriters in red < Controls in black) in the late time window (b).

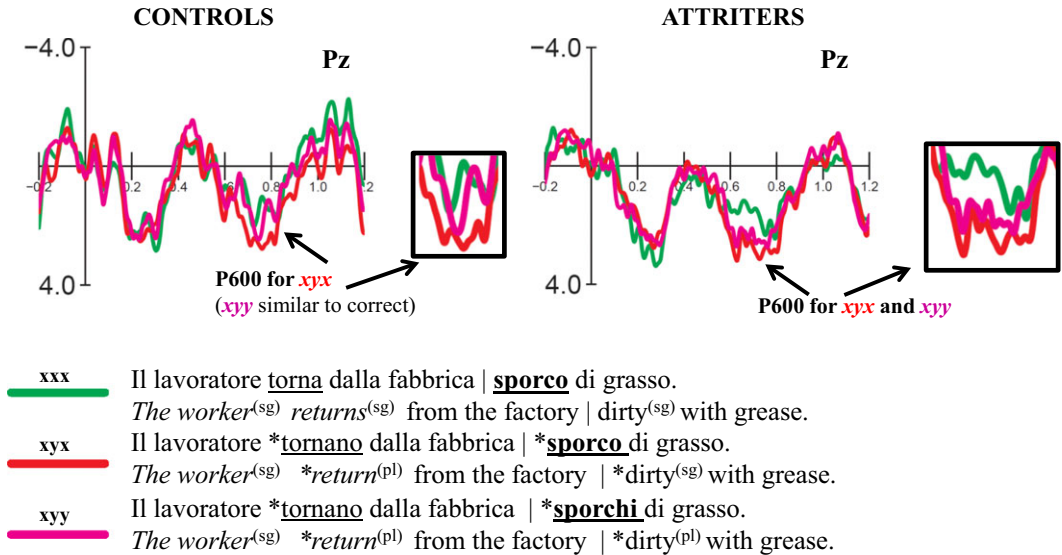


Fig. 8. Event-related potentials (ERPs) elicited by the modifier in response to *xyx* (red) and *xyy* (pink) violations relative to *xxx* at the most representative electrode (Pz). In Controls, *xyy* seems to elicit a smaller P600 than *xyx* violations, consistent with the “Repair hypothesis.” In Attriters, both violations seem to overlap and to elicit a P600 effect.

Given that our predictions as well as visual inspection of the data pointed to the negativity being primarily driven by the *xyy* condition, we proceeded directly to investigating follow-up analyses by condition pairs (which was motivated by significant  $Ag2 \times Ag1$  interactions). The only comparison that revealed significant effects in the negativity time-window was *xyy* versus *xxx*. *Condition* was found to be highly significant ( $F(1, 50) = 25.31, p < .0001$ ), as was  $Condition \times Ant-Post$  ( $F(1, 50) = 4.58, p < .005$ ). These effects were qualified by a significant  $Condition \times Ant-Post \times Proficiency$  interaction ( $F(3, 150) = 3.14, p < .05$ ). Follow-up analyses by *Proficiency* indicated that the negativity was present in both subgroups (High:  $Cond: F(1, 26) = 10.01, p < .005$ ;  $Cond \times Ant-Post: F(3, 78) = 3.30, p < .05$ ; Low:  $Cond: F(1, 26) = 15.67, p < .001$ ;  $Cond \times Ant-Post: F(3, 78) = 4.84, p < .001$ ) but that it differed in its scalp distribution. While the effect was frontally predominant in the higher proficiency subgroup (Fz:  $p < .05$ ; Cz:  $p = .05$ ; Pz and Oz:  $p > .1$ ), it was predominant at Pz in lower proficiency individuals (Fz:  $p > .05$ ; Cz:  $p < .05$ ; Pz:  $p < .001$ ; Oz:  $p < .01$ ).

Correlational analyses supported the ANOVA results that the scalp distribution of the negativity in the *xyy* condition was influenced by Italian proficiency. Proficiency scores positively correlated with *xyy* versus *xxx* amplitude at Pz (C-test:  $r = .293, p < .01$ ) but negatively with the effect at Fz (*Error-detection test*:  $r = -.210, p < .05$ ), demonstrating that the negativity was more enhanced at frontal sites in higher proficiency individuals.

### 3.4.2. Intermediate time window between 500 and 600 ms

This intermediate time window was selected to corroborate the grand-average data depicting an ongoing negativity only for lower proficiency individuals. The midline ANOVA revealed a significant main effect of *Ag1* ( $F(1, 50) = 5.61, p < .05$ ), but not of *Ag2* ( $p > .1$ ), as well as a significant interaction between  $Ag2 \times Ag1$  ( $F(1, 50) = 7.38, p < .01$ ) and  $Ag1 \times Ant-Post$  ( $F(3, 150) = 5.26, p < .01$ ). The interaction between  $Ag2 \times Proficiency$  was marginal ( $F(1, 50) = 4.04, p = .05$ ), but the three-way interaction with factor *Ant-Post* reached significance ( $Ag2 \times Ant-Post \times Proficiency: F(3, 150) = 3.57, p < .05$ ).

Follow-up analyses by *Condition* pairs were then performed, motivated by the significant  $Ag2 \times Ag1$  interaction. None of the condition pairs revealed any significant effects except *xy* versus *xx*, where we expected the *Proficiency* interaction (*Condition*:  $F(1, 50) = 5.33, p < .05$ ; *Condition*  $\times$  *Proficiency*:  $F(1, 50) = 3.15, p = .08$ ; *Condition*  $\times$  *Ant-Post*  $\times$  *Proficiency*:  $F(3, 150) = 2.28, p = .08$ ). Despite marginal interactions with factor *Proficiency*, follow-up analyses within each proficiency subgroup clearly supported the trend seen in the data: The negativity persisted from 500 to 600 ms for lower proficiency (Condition:  $F(1, 26) = 15.47, p < .001$ ) but not higher proficiency individuals ( $p > .1$ ).

Correlational analyses confirmed this trend and revealed that individuals with higher proficiency scores elicited a larger positive amplitude at Pz for *xy* versus *xx* in the 500–600 ms range (C-test:  $r = .337, p < .01$ ; *Error-detection test*:  $r = .274, p < .05$ ).

### 3.4.3. P600 between 600 and 900 ms

On the midline in the prototypical P600 window, a significant main effect of *Ag1* ( $F(1, 50) = 10.95, p < .005$ ) and of *Ag2* ( $F(1, 50) = 30.86, p < .0001$ ) were qualified by a significant interaction between the two factors ( $F(1, 50) = 11.72, p < .005$ ). Interactions with *Ant-Post* were also significant ( $Ag2 \times Ant-Post: F(3, 150) = 11.90, p < .0001$ ;  $Ag2 \times Ag1 \times Ant-Post: F(3, 150) = 12.08, p < .0001$ ), reflecting the posterior prominence of the positivity. Interactions with *Proficiency* were significant ( $Ag2 \times Proficiency: F(1, 50) = 5.77, p < .05$ ;  $Ag2 \times Ant-Post \times Proficiency: F(3, 150) = 5.12, p < .05$ ;  $Ag2 \times Ag1 \times Ant-Post \times Proficiency: (F(3, 150) = 5.34, p < .005$ ). No interactions with factor *Group* were statistically significant ( $ps > .1$ ), suggesting that the P600 in this time window was modulated by Italian proficiency level irrespective of group membership.

Follow-up analyses by *Condition* (motivated by the  $Ag2 \times Ag1$  interactions) revealed the most significant difference to be between *xy* and *xx* conditions (*Condition*:  $F(1, 50) = 29.77, p < .0001$ ; *Condition*  $\times$  *Ant-Post*:  $F(3, 150) = 16.01, p < .0001$ ). Interactions with *Proficiency* also reached significance (*Condition*  $\times$  *Proficiency*:  $F(1, 50) = 6.98, p < .01$ ; *Condition*  $\times$  *Ant-Post*  $\times$  *Proficiency*:  $F(3, 150) = 7.10, p < .0005$ ). ANOVAS within each proficiency subgroup indicated a significant P600 effect only in the higher proficiency individuals (*Condition*:  $F(1, 26) = 35.67, p < .0001$ ; *Condition*  $\times$  *Ant-Post*:  $F(3, 78) = 16.20, p < .0001$ ; significant at Cz, Pz, Oz at  $p < .0001$ ). In lower proficiency individuals, the P600 effect was only marginally significant (*Condition*:  $F(1,$



26) = 3.39,  $p = .08$ ; *Condition*  $\times$  *Ant-Post*:  $F(3, 78) = 2.28, p = .08$ , only at Pz:  $p < .05$ ). Correlational analyses provided further support that higher Italian scores were associated with a larger P600 amplitude at Pz (C-test:  $r = .381, p < .01$ ; *Error-test*:  $r = .358, p < .01$ ; *Semantic fluency*:  $r = .239, p < .05$ ).<sup>8</sup>

The comparison between *xyx* and *xxx* conditions yielded a significant main effect of *Condition* ( $F(1, 50) = 5.19, p < .05$ ) and an interaction with factor *Ant-Post* ( $F(3, 150) = 5.03, p < .05$ ) as well as a marginal three-way interaction between *Condition*  $\times$  *Ant-Post*  $\times$  *Proficiency* ( $F(3, 150) = 2.47, p = .06$ ). Despite the marginal significance, follow-up analyses within each proficiency subgroup supported the trend, such that only higher proficiency individuals showed a significant P600 for the *xyx* condition (*Condition*:  $F(1, 26) = 4.54, p < .05$ ; *Condition*  $\times$  *Ant-Post*:  $F(3, 78) = 6.09, p < .01$ ) which was significant at Pz ( $p < .005$ ) and Oz ( $p < .05$ ). In lower proficiency individuals, no effects approached significance ( $ps > .10$ ).

The comparison between *xyy* and *xxx* also revealed a significant *Condition*  $\times$  *Ant-Post* interaction ( $F(3, 150) = 4.49, p < .05$ ), indicating a posterior P600 effect, contrary to what would be predicted by the repair hypothesis. In contrast to the pattern observed in the plots, the difference between the two conditions was neither modulated by *Group* nor *Proficiency* ( $ps > .1$ ) in this P600 time window. This was surprising given that, in the monolingual Controls, the *xyy* condition seemed to overlap with the correct condition. Finally, comparing *xyy* and *xyx* violation conditions revealed a marginal main effect of *Condition* ( $F(1, 50) = 3.88, p = .06$ ) but no interactions with *Ant-Post* nor with between-subject factors such as *Group* nor *Proficiency* ( $ps > .1$ ).

To verify whether our hypotheses for group differences in repair strategies were partially supported (as visual inspection of the plots suggested), we took a closer look within each group. For the repair condition “*xyy*” relative to the correct condition “*xxx*,” Attriters showed a significant P600 effect (*Cond*  $\times$  *Ant-Post*:  $F(3, 69) = 5.63, p < .01$ ), whereas Controls did not ( $ps > .3$ ). In other words, while data from our non-attributing native speakers replicated the online repair strategies reported in the original study with Italian monolinguals (Molinaro et al., 2011), Attriters seemed to treat these sentences as morphosyntactic violations. Further support of this pattern was provided by the comparison between the two violation conditions (*xyy* vs. *xyx*), where we expected significant differences in Controls (*Condition*:  $F(1, 29) = 4.13, p = .05$ ) but not in Attriters ( $ps > .4$ ).

#### 3.4.4. Late P600 between 1,000 and 1,300 ms

The global ANOVA on the midline confirmed the pattern observed in the data (Fig. 7), namely that *Group* (but not *Proficiency*) was the meaningful factor that modulated P600 effects in this very late time window. The interaction between *Ag2*  $\times$  *Group* was significant ( $F(1, 50) = 5.91, p < .01$ ), as were the interactions with factor *Ant-Post* (*Ag2*  $\times$  *Ant-Post*:  $F(3, 150) = 13.31, p < .0001$ ; *Ag2*  $\times$  *Ag1*  $\times$  *Ant-Post*:  $F(3, 150) = 13.95, p < .0001$ ), as the effect was visibly strongest at posterior electrodes.

*Group* interactions (but not *Proficiency* interactions) were also found in follow-up analyses performed by *Condition* pairs. The comparison between *xyy* versus *xxx* yielded a significant *Condition*  $\times$  *Group* interaction ( $F(1, 50) = 8.75, p < .005$ ), which,

when followed up within each group, revealed that Controls showed a significant P600 effect in response to *xyy* violations ( $F(1, 29) = 8.39, p < .01$ ) but Attriters did not ( $p > .1$ ). In the comparison between condition *xyx* versus *xxx*, the *Condition*  $\times$  *Group* interaction also reached significance ( $F(1, 50) = 5.21, p < .05$ ), once again reflecting the presence of the P600 effect in Controls ( $F(1, 29) = 6.93, p < .05$ ) but not in Attriters ( $ps > .1$ ). Contrasting conditions *xyy* versus *xxx* yielded a significant *Condition*  $\times$  *Ant-Post* interaction ( $F(3, 150) = 3.19, p < .05$ ), but violation conditions *xyy* versus *xyx* were not statistically different from one another in this late P600 window ( $ps > .1$ ).

### 3.5. Experiential factors and ERP patterns in Attriters

Although, on average, age at testing was superior in Attriters than in native-speaker Controls, there were no modulations of behavioral or ERP responses by age of testing or level of education (all  $ps > .1$ ).

We assessed the role of background factors (such as age, education, age at immigration, and length of residence) as well as factors related to language use (such as amount of L1 and L2 exposure) on proficiency scores and ERP patterns. Within Attriters, length of residence was found to negatively correlate with scores on the written Error-detection test only ( $r = -.49, p < .01$ ). Amount of L1 exposure (in terms of hours/week) was positively correlated with Attriters' overall proficiency scores ( $r = .43, p < .01$ ) as well as with their performance on the semantic fluency task ( $r = .35, p < .05$ ).

With respect to ERP patterns, amount of daily L1 exposure (% relative to L2 exposure) was positively correlated with the late P600 elicited by the modifier in the *xyy* conditions, both in the 900–1,000 ms time window ( $r = .40, p < .005$ ) as well as the 1,000–1,300 ms window ( $r = .45, p < .000$ ). Thus, Attriters with more L1 exposure were more similar to native-controls in showing a late P600 effect.

No correlations were found between ERP patterns and Attriters' length of residence or age at immigration ( $ps > .1$ ).

## 4. Discussion

This study compared attriters and non-attributing native speakers in their L1-Italian morphosyntactic processing, to determine whether online error detection and re-analysis mechanisms involved in number agreement computation are vulnerable to attrition, behaviorally and/or at the neurocognitive level. Moreover, we examined whether processing differences were driven by L1 proficiency level (even among two groups of *native* speakers), and whether additional factors (L2 influence, amount of L1 exposure, length of residence, age of immigration) played a role in modulating the degree of native-like-ness of attriters' morphosyntactic processing patterns.

#### 4.1. *Qualitative and quantitative differences in Attriters' L1 morphosyntactic processing*

##### 4.1.1. *Verb*

On the verb, we showed that Attriters and Controls differed in the early negativity (300–500 ms) elicited by subject-verb number mismatches, both in terms of its amplitude and its scalp distribution. Controls showed a weak negativity that was focused at left-temporal sites (T3, T5) and only reached significance at the most lateral electrodes. This effect is consistent with a LTN, which has been previously reported to occur in response to morphosyntactic violations instead of a LAN in some reading studies (see Steinhauer, White, & Drury, 2009; Neville, Nicol, Barss, Forster, & Garrett, 1991; Newman et al., 2007; Weber-Fox & Neville, 1996). In contrast, Attriters showed a more robust, broadly distributed negativity that extended from midline to lateral sites (though larger over the left hemisphere). The negativity was followed by a frontal positivity (550–650 ms) similar to the effect reported in the original study by Molinaro et al. (2011). Although the frontal positivity was numerically larger in the Attriters, it was statistically indistinguishable from the effect elicited in Controls. Subject-verb number mismatches also elicited a large posterior P600 (as of 650 ms) relative to correct sentences. This posterior positivity was divided into two distinct phases—a first phase (650–900 ms), shared by Controls and Attriters but modulated by L1 proficiency (with larger P600 amplitudes in higher proficiency individuals), and a second phase (1,000–1,200 ms) where only the Controls showed a late, ongoing P600, whereas the P600 in Attriters returned to baseline by 1,000 ms.

One possible view is that the negativity was shared by both groups but that amplitude and distributional differences were merely caused by component overlap with the subsequent positivity. According to arguments made by Tanner and colleagues (Tanner, 2015; Tanner et al., 2013; Tanner & Van Hell, 2014; Tanner et al., 2014; see response by Molinaro, Barber, Caffarra, & Carreiras, 2015), extending previous suggestions by Osterhout and colleagues (Osterhout, 1997; Osterhout & Mobley, 1995), left-lateralized negativities are the result of N400s that have been altered by the onset of the following positivity, which cancels out the negativity at sites where both effects overlap in time. However, we question the ability of this account to explain our data, given that the frontal positivity (550–650 ms) was shared across both groups, and that neither its amplitude nor scalp distribution differed significantly between Controls and Attriters. In fact, the frontal positivity was numerically larger in the *Attriters*, that is, in the group with the larger negativity. It is not conceivable that a larger negativity would survive when followed by a numerically larger positivity in an adjacent time window. Secondly, even if component overlap were to explain the resulting difference in scalp distribution of an otherwise shared N400 across groups, Controls and Attriters would have to have differed on the frontal positivity and, thus, our data would have still revealed a group difference.

Rather, the group differences in the negativity time window seem to point to qualitative differences in expectations of agreement between a sentence-initial subject-noun and a subsequent verb. The absence of a strong LAN in monolingual Italian native speakers is not an entirely surprising finding, given that several Italian studies have previously struggled to detect a significant LAN in overall statistical analyses (Angrilli et al., 2002;

De Vincenzi et al., 2003; Mancini et al., 2009; Molinaro et al., 2011). It has been argued that, due to the relatively free word order of Italian and the grammaticality of post-verbal subject constructions, the expectation of number agreement between a sentence-initial noun and a subsequent verb is weaker in Italian than in languages such as English, where post-verbal subjects are syntactically unacceptable and the verb must therefore agree with its preceding noun (Molinaro, Barber et al., 2011). Given that the subject may follow the verb in Italian, a mismatch in number agreement detected on the verb does not necessarily signal a grammatical violation and thus may fail to elicit a robust LAN compared to, for example, determiner-noun number agreement violations (Vespignani, Molinaro, & Job, unpublished data). One study by Mancini, Molinaro, Rizzi, and Carreiras (2011b) on Spanish found a similar “left-posterior negativity” between 300 and 500 ms in monolingual Spanish speakers for a condition called “unagreement” where a person mismatch between subject and verb nonetheless produces a grammatical pattern. The authors interpreted this negativity as reflecting a violation of participants’ semantic/pragmatic expectations, rather than the detection of a morphosyntactic error. Attriters, on the other hand, seemed more likely than monolingual-controls to immediately process subject-verb disagreement as a morphosyntactic violation, likely reflecting influence of their dominant English grammar. In this view, Attriters’ reliance on word-order cues (subject precedes verb) led to a strong morphosyntactic and conceptual (LAN/N400) expectation of agreement between subject-noun and verb, rather than exploring the possibility of a post-verbal subject construction as a solution to the agreement mismatch.

The group differences in the duration of the posterior P600 are also indicative of processing differences between Attriters and non-attriter Controls. As later stages of the P600 have been associated with re-analysis and repair processes (Carreiras et al., 2004; Hagoort & Brown, 2000; Mancini, Molinaro, Rizzi, & Carreiras, 2011a,b; Molinaro et al., 2008; Silva-Pereyra & Carreiras, 2007), we interpret these results as suggesting that Controls engage in more extensive/elaborated repair than Attriters do for the same number-agreement violations. This pattern, replicated on both target words (verb and modifier), together with Attriters’ longer response latencies in providing acceptability judgment ratings, seems to support the view of less efficient online processing in Attriters. This interpretation will be further discussed below.

#### 4.1.2. *Modifier*

On the modifier, in response to *xy* sentences where the modifier marks the first point of violation (after both the subject-noun and the verb agree in number), we showed that Attriters and Controls elicited a similar biphasic pattern consisting of an N400 (300–500 ms) followed by a P600 (600–900 ms) without a preceding frontal positivity. The finding of an N400-like negativity is consistent with the view that computing certain types of agreement information requires access to lexical-semantic or discourse-level information (Barber, Salillas, & Carreiras, 2004; Deutsch & Bentin, 2001; Molinaro et al., 2008; see Molinaro, Barber et al., 2011), given that, at this point in the sentence, readers must determine the antecedent of the modifier (the subject-noun). Interestingly, our data showed that a biphasic pattern only occurred on target words which marked the

first point of violation in the sentence (i.e., on verb in violation conditions and on modifier in *xyy* condition, *but not* on modifier in *xyx* or *xyy* conditions). The finding that the negativity is more robust for the modifier than on the verb may be in line with the notion that subject-verb number mismatches in Italian are less likely to elicit robust LAN/N400 effects due to the possibility of a post-verbal subject noun, whereas encountering a modifier must unambiguously agree with its preceding antecedent. Although the amplitude and scalp distribution of the N400 on the modifier were modulated by L1-proficiency, there were no group differences in this time window. However, the groups differed once again on the duration of the P600, which persisted into the 1,000–1,300 ms time-range in native-Controls but not in Attriters. Interestingly, *within* Attriters, correlational analyses showed that the more frequent the L1-Italian exposure, the larger the late P600. In other words, the more Attriters continued to use their L1, the more native-like they were in the P600 they elicited on the modifier.

Comparing the two other violation conditions (*xyy* and *xyx*) relative to correct *xxx* sentences revealed that Attriters elicited a numerically larger P600 effect for *xyy* sentences (i.e., the “repair” condition) compared to non-attriter Controls, for whom the *xyy* condition largely overlapped with the correct condition and only the *xyx* condition seemed to elicit a P600 effect (Fig. 8). This trend was in the direction of our hypotheses, namely that Attriters would be less efficient than Controls in their online repair/re-analysis strategies upon encountering an initial violation on the verb. However, group differences in this graded pattern of ERP responses ( $xyx > xyy$ ) did not emerge as statistically significant. In hindsight, it is possible that our choice of task weakened the potential for repair to be as robustly pursued as in the original study, where the task involved reading for comprehension rather than rating acceptability (Molinaro et al., 2011; Molinaro et al., 2008). It is possible that repair is not mandatory or fully pursued in an acceptability judgment task as compared to reading for comprehension. In line with this speculation, behavioral response times were longest for conditions where the verb agreed with the preceding subject, indicating that participants were faster to make up their mind about the acceptability of the sentence if the violation occurred early on (on the verb). A comprehension question assessing readers’ interpretation of the number value of the sentences (e.g., was the sentence about one or two workers?) may have proven more sensitive to processing differences related to input-revision and repair.

One may argue the negativity observed on the modifier may be driven by the potential structural ambiguity of the sentence constructions, where readers may have initially attempted to attach the modifier to the immediately preceding word (which was feminine, inanimate, and singular across all trials). An attachment attempt of the modifier with the preceding noun would result in a Gender+Plausibility+Number mismatch, and give rise to an N400 effect. It is conceivable that such a parsing preference could exist and could be modulated by factors related to attrition or proficiency. However, we may rule out this possibility, as the mismatch in gender and plausibility between the modifier and the intervening noun should have affected all conditions equally (including the correct condition), whereas only the *xyy* condition elicited a large N400 effect. Moreover, if the N400 were driven by the number mismatch between the modifier and preceding noun (rather than its

mismatch with the sentence-initial subject-noun), then we would expect differences between singular and plural trials, given that the intervening noun was consistently singular. We investigated this possibility but did not find any evidence that the ERP violation effects were largest in subconditions where the modifier was plural (and clashed with the preceding singular noun). We can therefore rule out that any processing differences we describe in our data were due to potential differences in modifier-antecedent attachment preferences.

The pattern of a shorter P600 response in Attriters than in Controls (as well as to Attriters with greater L1 use), together with Attriters' longer response-times, as well as their tendency to elicit larger P600 responses to the "repair" condition ( $xyy$ ) than native-monolinguals, suggest overall that Attriters may engage in shallower online repair/re-analysis processes during real-time comprehension. Although they did detect the number agreement violations at both points within the sentence and reached similar acceptability judgment ratings (albeit at significantly slower rates), Attriters differed in how they computed syntactic relations online. Importantly, although one could argue that the shorter P600 in Attriters is a trivial effect that reflects unspecific inter-individual differences that happened to manifest as a group difference, data we collected from the same participants over the same testing session but with different experimental stimuli did not show this pattern (Kasparian & Steinhauer, 2016). This rules out the possibility that Attriters show shorter P600 effects across the board. Lastly, Attriters were not slower in their response-times compared to native-Controls across the board but specifically in the present experiment, which argues against the possibility that Attriters are slower in their processing overall.

Interestingly, the P600 is the effect that Monika Schmid and collaborators focused on in their large-scale study of gender agreement processing (Bergmann et al., 2015; Schmid, 2013) and on which they reported no significant amplitude or latency/duration differences between Attriters and non-attributing monolingual controls. However, as discussed earlier, determiner-noun agreement (without an inflected intervening adjective in Dutch and German) may not be sensitive to group differences in processing in native speakers. An interesting avenue for future research would be to continue investigating potential P600 modulations in Attriters and how these effects may depend on the structures that are investigated, the specific language-pairings and factors related to attrition and to proficiency level.

#### *4.2. The role of proficiency level in modulating ERP response patterns*

In line with research emphasizing the impact of proficiency on native-like L2 processing patterns (e.g., Bowden et al., 2013; Friederici, Steinhauer, et al., 2002; Morgan-Short et al., 2012; Newman, Tremblay, Nichols, Neville, & Ullman, 2012; Osterhout et al., 2006; Rossi et al., 2006; Steinhauer et al., 2009), our findings confirmed that proficiency scores predicted the amplitude, scalp distribution, latency and/or duration of ERP correlates of language processing, even in native speakers processing their *L1*.

Both on the verb and on the modifier, L1-proficiency modulated the amplitude of the earlier portion of the posterior P600 (verb: 650–1,000 ms; modifier: 600–900 ms), such that native speakers (Controls and Attriters) with higher proficiency scores elicited a P600 effect of a larger amplitude, indicating that they were better able to diagnose the ungrammaticality of the sentence than native speakers with lower scores on the Italian proficiency measures (see Friederici et al., 2001). This finding is reminiscent of a number of previous studies reporting reduced P600s in less proficiency L2 learners (e.g., Rossi et al., 2006). Thus, proficiency of the target language modulates processing patterns, regardless of the status of that language (L1 or L2). Crucially, and consistently for both target-words, proficiency only modulated the P600 amplitude in the early time window, whereas the P600 effect extending beyond 1,000 ms was dependent on group membership rather than on proficiency level, with a significant P600 persisting only in Controls.

Furthermore, the N400 effect observed on the modifier (*xy* condition) was also predicted by proficiency scores; lower L1-proficiency speakers (Controls as well as Attriters) elicited a smaller, less frontal, and longer lasting (until 600 ms) N400 in response to *xy* violations, whereas the N400 was larger and more anterior in higher proficiency individuals. Although the voltage maps in Fig. 7 suggest that the negativity in the high-proficiency subgroup resembles a LAN, whereas the negativity in the low-proficiency subgroup has an N400-like distribution, the lateral ANOVA did not reveal a significant *Hemisphere* × *Group* interaction. Proponents of the component overlap view proposed by Tanner and colleagues (Osterhout, 1997; Tanner & Van Hell, 2014; Tanner et al., 2013, 2014) would argue that both subgroups show an N400 effect, but that the left-anterior distribution of the negativity in the high-proficiency individuals is the result of the larger and earlier onsetting parietal P600 that cancels out the N400 at electrodes sites where the two effects overlap. To investigate this possibility, we visually examined the distribution of the negativity with a 50 ms moving window and determined that the negativity in high-proficiency individuals originated at left-frontal electrodes as early as 250 ms. Thus, the steepness of the P600 in the high-proficiency group is not the reason for the frontal distribution of the negativity in these individuals. Given that, based on the literature, a LAN + P600 pattern may be the expected response to number agreement violations, it seems intuitive for the negativity to be more frontally localized (and more left-lateralized, at least qualitatively) in higher rather than lower proficiency individuals.

Our study is consistent with similar accounts of proficiency effects on L1 morphosyntactic processing (Newman et al., 2012; Pakulak & Neville, 2010) and among the few ERP studies to systematically measure proficiency and assess its impact on the ERP patterns of native speakers, thus moving away from the assumption of “ceiling effects” in native speakers without any individual variability in proficiency. In sum, our results emphasize that proficiency is a key factor in modulating native-like neurocognitive responses, regardless of whether the language being processed is the L2 or the L1.

### 4.3. What qualifies as “first language attrition”?

We have argued that attriters differed from non-attriting native speakers in their L1 processing of number agreement, showing evidence of crosslinguistic influence from L2-English in response to subject-verb agreement mismatches, as well as less thorough online repair or re-analysis mechanisms as reflected by differences in the late P600 (and longer behavioral response times). However, it is important to rule out the possibility that other differences between groups account for these patterns. First, although Attriters were more advanced in age than Controls, there were no significant correlations between ERP patterns and age at testing. The groups also did not differ significantly in WM and/or reading speed.

One possibility is that attentional differences contribute to group differences in ERP response patterns. Given that we explicitly set out to recruit individuals who were reporting changes/difficulties in their L1, participants in our attrition group were aware that they were being selected for a study assessing their native language. We chose not to hide this information from them to ensure that our candidates reported experiencing a clear change in the way they used their L1 relative to the L2. However, this also inherently meant that attention, focus, and motivation to show maintenance of the L1 may have played a greater role in these individuals than in native speakers who were tested in Italy. Attriters may be more likely to be self-conscious and to want to perform well in their native language. That said, the potentially heightened attention in these individuals may be an intrinsic characteristic of attrition and an influencing factor of its own right.

Is it fair to label these differences as “attrition”? It may be argued that differences in amplitude, distribution, or duration of otherwise qualitatively similar ERP components may not constitute compelling evidence in favor of attrition. Such a criticism raises important issues relevant to the operationalization of “attrition” as well as for ERP research in general. Qualitative and/or quantitative differences in amplitude, scalp distribution, latency, and duration have long been interpreted in L2-processing studies as differential processing mechanisms between (late) L2 learners and native speakers (e.g., Hahne & Friederici, 2001; Osterhout et al., 2006; Rossi et al., 2006; Weber-Fox & Neville, 1996). Moreover, it is important to note that the only other ERP study conducted with attriters failed to show any such differences—be it qualitative *or* quantitative—in the L1 processing patterns of Attriters compared to native-monolingual Controls.

If these differences qualify as “attrition,” why is this decline not reflected in our Attriters’ L1 proficiency scores which, though numerically lower, were not statistically different from those of native speakers still residing in Italy? One possibility is that, at least for the written measures, Attriters arrived at the same responses but after spending more time on the task. Longer response latencies were indeed found for the end-of-sentence acceptability judgment task, in the absence of significant differences in rating values for each condition. Furthermore, it has been shown that attrition effects can be found in some tasks but not in others (Schmid, 2011). It is also conceivable that differences begin to appear at the neurocognitive processing level *before* they appear in behavior, as has been



demonstrated in longitudinal studies of L2 learning such as in McLaughlin et al. (2004). The relationship between online processing and linguistic behavior, and the time course of attrition effects as reflected by both methods, is still an interesting question that is open to investigation. It is important to keep in mind that *all* attriters who participated in this study were recruited because of their self-reported experiences of attrition effects. Thus, attrition effects must not be entirely absent in their linguistic behavior, given that Attriters (and sometimes their friends or family) have pinpointed changes in their language. Rather, our behavioral tasks were less sensitive to these effects than were real-time ERP measures.

One may argue that the attriters tested in this study were only at the earliest stages of attrition. Proficiency scores were indeed lower in individuals with a longer length of residence and with less L1 exposure. However, it may be that we need to move away from the expectation that proficiency and behavioral differences are the yardstick with which we define and quantify attrition. As we have shown in this study, several group differences between Attriters and Controls in amplitude, timing, and scalp distribution of ERP patterns were found to be independent from proficiency effects, indicating that a normal degree of proficiency variation within native speakers cannot be the whole story. Attrition does not seem to manifest itself as a generalized reduction in proficiency, neither behaviorally in individuals' proficiency profiles, nor in terms of ERP responses that are proficiency-dependent. A broader definition such as the one proposed by Schmid (2011) where "attrition" is generally used to describe a *change* to the L1 as a result of predominant L2 exposure/use may be more appropriate to understand the neurocognitive and linguistic underpinnings of L1 attrition.

In our view, attrition effects should more broadly encompass (a) *proficiency effects*, where L1 processing patterns are more native-like at higher ranges of L1 proficiency and L2 processing is more native-like at higher levels of L2 proficiency; (b) *crosslinguistic transfer effects*, where increased attrition is characterized by an increase in L2-to-L1 influence and a decrease in L1-to-L2 influence; and (c) *special experiential circumstances*, such as increased attention, motivation to perform well, amount of L1 versus L2 exposure, length or residence, etc.

A final point is whether group differences are merely the result of comparing bilingual and monolingual-speakers. Although we also ran the present experiment with English-Italian L2 learners of intermediate and advanced proficiency levels, the focus of the current paper was to assess the impact of L1 dominance and to compare attriters to non-attriter native speakers on their *L1*, rather than also comparing them to a group of learners for whom Italian was the *L2*. Second, in addition to differences between groups, we found evidence of more or less native-like processing patterns even *within* the bilingual group of Attriters. For example, amount of L1 exposure modulated the amplitude of the late P600 which has been associated with structural re-analysis processes. However, given that cross-linguistic influence from English to Italian is a likely source of processing differences (as argued in the context of the robust N400 observed in Attriters in response to subject-verb agreement violations), it would be highly beneficial as a next step to examine not only L2 learners of Italian, but especially to test a

group of L1-dominant Italian-English bilinguals residing in Italy, who are highly proficient in the English-L2 but who do not experience a change in the way the L1 is used in the environment.

In sum, this research is only the first step in the direction of understanding the neurocognitive correlates of L1 attrition, but it has highlighted some of the first differences in the real-time processing of L1 morphosyntax in attriters relative to non-attriting native speakers.

#### 4.4. *Beyond a monolithic account of the P600*

As one of the broader aims of our study was to replicate previous studies on number agreement processing and investigate ERP responses as reflecting distinct processing stages, a brief mention of the functional significance of the distinct phases of positive-going ERP components is of interest for ERP research on language processing.

In line with several number agreement studies, including the original study by Molinaro et al. (2011), we replicated the early positivity that was prominent over fronto-central areas of the scalp (Barber & Carreiras, 2005; Hagoort & Brown, 2000; Kaan, 2002; Kaan & Swaab, 2003; Kaan et al., 2000; Molinaro, Kim et al., 2008; Molinaro, Barber, & Carreiras, 2011; Silva-Pereyra & Carreiras, 2007). This effect was elicited by the verb when it clashed in number with the sentence-initial subject-noun. Previous reports have attributed the frontal positivity to an increased difficulty in integrating the mismatching constituent with the previous sentence fragment, and having to override the preferred structural representation of the sentence (Hagoort, Brown, & Osterhout, 1999). However, one should also consider the literature where frontal positivities elicited in a variety of paradigms have been referred to as P3a components. The P3a is often driven by surprise (Polich, 2007; Squires, Squires, & Hillyard, 1975) and is viewed as part of an orientation response allocating special attention to the stimulus (Näätänen & Galliard, 1983). This interpretation would explain why the early frontal positivity in our study was present for violations realized early on in the sentence (i.e., on the verb but not on the modifier) and was larger for trials in the first half of the experiment, compared to the second half (see Appendix S1). We preliminarily argue that the early frontal positivity we observed is a P3a, driven by a violation that occurs early on in a sentence without much context, and is absent when the violation is more predictable. This hypothesis could be further investigated in future studies.

Our study also replicated the prototypical, posterior P600 effect that has been elicited in response to morphosyntactic violations. In terms of its amplitude, we found large P600s in response to subject-verb agreement violations, as well as in response to the modifier when it consisted of the first and only violation in the sentence (*xyx* condition). In contrast, P600 effects were significantly smaller in those conditions where the modifier marked the second point of violation within the sentence (*xyx* and *xxy*). Interestingly, this graded P600 pattern was strikingly similar to the P600 pattern reported in the original study by Molinaro and colleagues (2011) as well as in an English study with a similar experimental design by Molinaro et al. (2008), despite

our use of an acceptability judgment task rather than a comprehension task (see discussions of task effects on the P600 in Bornkessel-Schlesewsky et al., 2011; Coulson et al., 1998; Friederici et al., 2001; Frisch, Kotz, von Cramon, & Friederici, 2003; Osterhout & Hagoort, 1999; Royle, Drury, & Steinhauer, 2013; Steinhauer, Mecklinger, Friederici, & Meyer, 1997).

An important contribution of our study was the finding of functionally distinct portions of the posterior P600. In line with the literature, we interpreted the first window of the posterior P600 as reflective of the diagnosis of a violation (see Fodor & Inoue, 1998, discussed in Friederici et al., 2001 for garden-path sentences), whereas we associated the later window with processes related to morphosyntactic repair (Carreiras et al., 2004; Hagoort & Brown, 2000; Mancini et al., 2009; Molinaro et al., 2008; Silva-Pereyra & Carreiras, 2007). However, the majority of agreement studies had not extended the time window of the P600 beyond 900 or 1,000 ms. It is important to emphasize that the later P600 does not seem to reflect a mere continuation (longer duration) of the P600 elicited 650–1,000 ms, as topographical differences were found in Controls in an additional analysis<sup>9</sup> comparing the two time windows ( $TW \times Agr1 \times Ant-post$ :  $(F(3, 87) = 3.67, p < .05)$ ). These distributional differences between the earlier and late P600 window provides further support that the underlying processes (diagnosis vs. repair) are functionally distinct.

An unexpected and novel finding was that these P600 effects in different time windows were *differentially* affected by proficiency (early phase) and group membership (late phase). This differential impact of L1 proficiency and group emphasizes that different stages of the P600 reflect different underlying processes and therefore it would be much too simplistic to consider the P600 as a monolithic component.

## 5. Conclusions

In one of the first ERP investigations of L1 attrition in morphosyntactic processing, we showed that adult attriters and non-attriting native speakers differed in the neurocognitive correlates underlying real-time comprehension of Italian sentences that required online re-analysis. This was the first ERP study to show qualitative and quantitative differences in ERP components of interest in the absence of robust differences offline, at the behavioral level. Our results also emphasized that proficiency modulates native-like processing patterns, even in one's L1, but that attrition cannot simply be characterized as "lower L1 proficiency" profiles, given the additive effects of group and proficiency in our study. The finding that even the "entrenched" L1 grammar of individuals who lived in an exclusively monolingual context up until adulthood is subject to change after a period of predominant L2 exposure/use in a non-L1-dominant environment corroborates the view of ongoing neuroplasticity in adulthood, where language experience is able to alter neurocognitive mechanisms beyond an early maturational window. Our study also highlighted and addressed a number of theoretical and methodological questions that could be considered as avenues for impending research in both L1 and L2 processing.

## Acknowledgments

K.K. was supported by a Vanier Canada Graduate Scholarship from the Canadian Institutes of Health Research, as well as a Richard Tomlinson Doctoral Fellowship awarded by the Faculty of Medicine of McGill University. K.K.'s research stay in Italy for data collection was made possible by a Michael Smith Foreign Study Supplement (MSFSS-CGS). This research was supported by grants awarded to K.S. by the Canada Research Chair program and the Canada Foundation for Innovation (CRC/CFI; project No. 201876), the Natural Sciences and Engineering Research Council of Canada (NSERC; RGPGP 312835-05 and RGPIN 402678-11), the Social Sciences and Humanities Research Council of Canada (SSHRC; No. 435-2013-2052), and the Fonds de Recherche Société et Culture, Québec (FQRSC; No. 2010-SE-103727). The authors thank Nicola Molinaro for valuable feedback on the experimental design and stimuli, as well as John E. Drury and Stefanie Nickels for comments on an earlier draft of this manuscript. We are also indebted to research assistants Linna Jin, Kristina Maiorino, Filippo Vicari, and Paolo Zandomenighi for their invaluable help with data collection.

## Notes

1. It should be noted that the early (fronto-central) stage of the positivity is not limited to ambiguous/complex sentences or to agreement errors, but has been elicited in response to other morphosyntactic violations such as phrase structure violations (e.g., Bowden et al., 2013). The early frontal positivity has also been described as a member of the P300 component family (P3a: Squires, Squires, & Hillyard, 1975; in Italian morphosyntactic violations: Mueller, Oberecker, & Friederici, 2009; see review in Osterhout & Hagoort, 1999, and Polich, 2007), although many studies have termed it a “frontal P600” (e.g., Kaan & Swaab, 2003).
2. The LAN was only significant in the authors' direct comparison of condition pairs, rather than in a global ANOVA.
3. Participants unanimously considered their AoA of English to coincide with their age of immersion into English (i.e., immigration), given that their exposure to English within the Italian school system was only minimal.
4. Molinaro et al., (2011) only presented participants with sentences where the subject-noun was plural as their main goal was to compare the processing of morphosyntactic plurals and coordinate plurals (joined by “and”).
5. Following an observation made by Molinaro, Barber, et al. (2011) that LAN effects were more robust in agreement studies using an offline reference of average mastoids rather than the left mastoid, we ran our analyses with each reference choice and found that the pattern of our results was identical whether we used the left mastoid or the average mastoids as our offline reference.

6. Our original baseline of  $-200$  to  $0$  ms was adjusted to compensate for early differences triggered by the subject-noun in Attriters but also in L2 learners; to ensure comparability between the L2 data (reported elsewhere) and data from Controls and Attriters, maintaining a consistent baseline interval was considered important. Crucially, all ERP effects reported here were consistent with either baseline correction (see example in Appendix S1).
7. These correlations between P600 amplitude and proficiency were significant even with group partialled out (C-test:  $r = .296$ ,  $p < .05$ ; Semantic fluency:  $r = .344$ ,  $p < .05$ ), indicating that proficiency differences cannot simply be reduced to group differences between Attriters and Controls on the P600. We thank an anonymous reviewer for this suggestion.
8. These correlations between P600 amplitude and proficiency held when group was partialled out (C-test:  $r = .404$ ,  $p < .01$ ; Error-detection test:  $r = .360$ ,  $p < .01$ ; Semantic fluency test:  $r = .246$ ,  $p = .08$ ), indicating that proficiency effects cannot be reduced to group differences between Attriters and Controls. We thank an anonymous reviewer for this suggestion.
9. We thank an anonymous reviewer for this suggestion.

## References

- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59(2), 249–306. doi:10.1111/j.1467-9922.2009.00507.x
- Ammerlaan, T. (1996). *You Get a Bit Wobbly...: Exploring Bilingual Lexical Retrieval Processes in the Context of First Language Attrition*. PhD dissertation, Nijmegen University.
- Angrilli, A., Penolazzi, B., Vespignani, F., De Vincenzi, M., Job, R., Ciccirelli, L., Palomba, D., & Stegagno, L. (2002). Cortical brain responses to semantic incongruity and syntactic violation in Italian language: An event-related potential study. *Neuroscience Letters*, 322(1), 5–8. doi:10.1016/S0304-3940(01)02528-9
- Balconi, M., & Pozzoli, U. (2005). Comprehending semantic and grammatical violations in Italian. N400 and P600 comparison with visual and auditory stimuli. *Journal of Psycholinguistic Research*, 34(1), 71–98. doi:10.1007/s10936-005-3633-6
- Barber, H. A., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, 17(1), 137–153. doi:10.1162/0898929052880101
- Barber, H., Salillas, E., & Carreiras, M. (2004). Gender or genders agreement. On-line study of sentence comprehension, 309–328.
- Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A crosslinguistic study. *Cognition*, 11(3), 245–299. doi:10.1016/0010-0277(82)90017-8
- Bergmann, C., Meulman, N., Stowe, L. A., Sprenger, S. A., & Schmid, M. S. (2015). Prolonged L2 immersion engenders little change in morphosyntactic processing of bilingual natives. *NeuroReport*, 26(17), 1065–1070.
- Bornkessel-Schlesewsky, I., Kretschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., Roehm, D., & Schlewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language*, 117, 133–152. doi:10.1016/j.bandl.2010.09.010

- Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, *51*(13), 2492–2511. doi:10.1016/j.neuropsychologia.2013.09.004
- Bylund, E. (2009). Maturation constraints and first language attrition. *Language Learning*, *59*(3), 687–715.
- Carreiras, M., Salillas, E., & Barber, H. (2004). Event-related potentials elicited during parsing of ambiguous relative clauses in Spanish. *Cognitive Brain Research*, *20*(1), 98–105. doi:10.1016/j.cogbrainres.2004.01.009
- Chen, L., Shu, H., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject-verb agreement in L2 learning. *Bilingualism: Language and Cognition*, *10*, 161–174. doi: 10.1017/S136672890700291X
- Clahsen, H., & Felser, C. (2006a). How native-like is non-native language processing? *Trends in Cognitive Sciences*, *10*(12), 564–570. doi:10.1016/j.tics.2006.10.002
- Clahsen, H., & Felser, C. (2006b). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(01), 3–42. doi:10.1017/S0142716406060024
- Coulson, S., King, J., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, *13*(1), 21–58. doi:10.1080/016909698386582
- De Bot, K. (1996). Language loss. In H. Goebel, P. Nelde, Z. Stary, & W. Wölk (Eds.), *Contact linguistics: An international handbook of contemporary research* (Vol. 1, pp. 579–585). Berlin: Walter de Gruyter.
- De Vincenzi, M., Job, R., Di Matteo, R., Angrilli, A., Penolazzi, B., Ciccirelli, L., & Vespignani, F. (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain and Language*, *85*(2), 280–296. doi:10.1016/S0093-934X(03)00055-5
- Della Rosa, P. A., Videsott, G., Borsa, V. M., Canini, M., Weekes, B. S., Franceschini, R., & Abutalebi, J. (2013). A neural interactive location for multilingual talent. *Cortex*, *49*(2), 605–608. doi:10.1016/j.cortex.2012.12.001
- 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*(2), 200–224. doi:10.1006/jmla.2000.2768
- Dowens, M. G., 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*(7), 1651–1659. doi:10.1016/j.neuropsychologia.2011.02.034
- Fodor, J. D., & Inoue, A. (1998). Attach anyway. In *Reanalysis in sentence processing* (pp. 101–141). Springer Netherlands.
- Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research*, *31*(1), 45–63. doi:10.1023/A:1014376204525
- Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: A spatio temporal analysis of event-related brain potentials. *Cognitive Brain Research*, *11*, 305–323. doi:10.1016/S0926-6410(00)00065-3
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, *99*(1), 529–534. doi:10.1073/pnas.012611199
- Frisch, S., Kotz, S. A., von Cramon, D. Y., & Friederici, A. D. (2003). Why the P600 is not just a P300: The role of the basal ganglia. *Clinical Neurophysiology*, *114*, 336–340. doi:10.1016/S1388-2457(02)00366-8
- Gürel, A. (2004). Selectivity in L2-induced L1 attrition: A psycholinguistic account. *Journal of Neurolinguistics*, *17*(1), 53–78. doi:10.1016/S0911-6044(03)00054-X
- Gürel, A. (2007). Linguistic determinants of L1 attrition. *Language Attrition. Theoretical perspectives*, 99–120.
- Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, *38*(11), 1531–1549. doi:10.1016/S0028-3932(00)00053-1

- 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(4), 439–483. doi:10.1080/01690969308407585
- Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. M. Brown & P. Hagoort (Eds.) *The neurocognition of language* (pp. 273–317). Oxford, UK: Oxford University Press. doi: 10.1093/acprof:oso/9780198507932.001.0001
- Hagoort, P. (2003). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *Neuroimage*, 20, S18–S29.
- Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, 4(2), 123–141. doi: 10.1017/S1366728901000232
- Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, 120(4), 901–931. doi:10.1016/j.lingua.2009.06.004
- Hulsen, M. (2000). Language loss and language processing: Three generations of Dutch migrants in New Zealand. Doctoral Dissertation, Nijmegen: Katholieke Universiteit Nijmegen.
- Jasper, H. H. (1958). Report to the committee on methods of clinical examination in electroencephalography. Appendix: The ten-twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 10, 371–375. doi:10.1016/0013-4694(58)90053-1
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99. doi:10.1016/0010-0285(89)90003-0
- Kaan, E. (2002). Investigating the effects of distance and number interference in agreement processing: An ERP study. *Journal of Psycholinguistic Research*, 31(2), 165–193. doi:10.1023/A:1014978917769
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15(2), 159–201. doi:10.1080/016909600386084
- Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, 15(1), 98–110. doi:10.1162/089892903321107855
- Kasparian, K., & Steinhauer, K. (2016). Confusing similar words: ERP correlates of lexical-semantic processing in first language attrition and late second language acquisition. *Neuropsychologia*. doi: 10.1016/j.neuropsychologia.2016.10.007. [Epub ahead of print]
- Kasparian, K., Vespignani, F., & Steinhauer, K. (2013a). When the second language takes over: ERP evidence of L1-attrition in morphosyntactic processing. Talk presented at the International Conference on Multilingualism, Montreal, Canada, October 24–25, 2013.
- Kasparian, K., Vespignani, F., & Steinhauer, K. (2013b). My Italian is not what it used to be: Investigating the neural correlates of L1 attrition and late L2 acquisition. Poster presented at the Workshop on Neurobilingualism, Groningen, The Netherlands, August, 25–27, 2013.
- Kasparian, K., Vespignani, F., & Steinhauer, K. (2014a). The case of a non-native-like first language: ERP evidence of first language (L1) attrition in lexical and morphosyntactic processing. *International Journal of Psychophysiology*, 2(94), 159–160.
- Kasparian, K., Vespignani, F., & Steinhauer, K. (2014b). Neurophysiological correlates of L1 attrition and L2 acquisition: A continuum based on proficiency. Poster presented at the Society for the Neurobiology of Language, Amsterdam, Netherlands, August 27–29, 2014
- Kim, J. H., Montrul, S., & Yoon, J. (2010). Dominant language influence in acquisition and attrition of binding: Interpretation of the Korean reflexive caki. *Bilingualism: Language and Cognition*, 13(01), 73–84. doi: 10.1017/S136672890999037X
- Kim, K. H., Relkin, N. R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, 388, 171–174. doi:10.1016/S0002-9394(14)71720-9
- Klein, D., Mok, K., Chen, J.-K., & Watkins, K. E. (2014). Age of language learning shapes brain structure: A cortical thickness study of bilingual and monolingual individuals. *Brain and Language*, 131, 20–24. doi:doi:10.1016/j.bandl.2013.05.014

- Köpke, B. (1999). L'attrition de la première langue chez le bilingue tardif: Implications pour l'étude psycholinguistique du bilinguisme. Unpublished Doctoral Dissertation, Université de Toulouse-Le Mirail. doi: 10.1515/9783111594149.115
- Köpke, B. (2004). Neurolinguistic aspects of attrition. *Journal of Neurolinguistics*, 17(1), 3–30. doi:10.1016/S0911-6044(03)00051-4
- Köpke, B., & Schmid, M. S. (2004). Language Attrition: The Next Phase. In M. S. Schmid, B. Köpke, M. Keijzer & L. Weilemar (Eds.), *First language attrition: Interdisciplinary perspectives on methodological issues* (pp. 1–43). Amsterdam: John Benjamins. doi: 10.1075/sibil.28
- Kraš, T. (2008). L2 acquisition of the lexicon-syntax interface and narrow syntax by child and adult Croatian learners of Italian. Unpublished doctoral dissertation, University of Cambridge, 292. doi:10.1075/eurosla.10.12kra
- Leinonen, A., Brattico, P., Jarvenpaa, M., & Krause, C. M. (2008). Event-related potential responses to violations of inflectional and derivational rules of Finnish. *Brain Research*, 1218, 181–193. doi:10.1016/j.brainres.2008.04.049
- Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.
- MacWhinney, B., & Bates, E. (1989). *The crosslinguistic study of sentence processing*. New York: Cambridge University Press. doi:10.2307/1423201
- Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011a). A person is not a number: Discourse involvement in subject-verb agreement computation. *Brain Research*, 1410, 64–76. doi:10.1016/j.brainres.2011.06.055
- Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011b). When persons disagree: An ERP study of Unagreement in Spanish. *Psychophysiology*, 48, 1361–1371. doi:10.1111/j.1469-8986.2011.01212.x
- Mancini, S., Vespignani, F., Molinaro, N., Laudanna, A., & Rizzi, L. (2009). Number agreement processing with different persons: An ERP study. Poster presented at the 15th AMLAP Conference. Barcelona, Spain, September 7–9, 2009. doi:10.1016/j.lingua.2014.04.014
- Marchman, V. A. (1993). Constraints on plasticity in a connectionist model of the English past tense. *Journal of Cognitive Neuroscience*, 5(2), 215–234. doi:10.1162/jocn.1993.5.2.215
- Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage*, 63(1), 240–244. doi:10.1016/j.neuroimage.2012.06.043
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, 7, 703–704. doi:10.1038/nn1264
- Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Neurolinguistics: structural plasticity in the bilingual brain. *Nature*, 431(7010), 757. doi:10.1038/431757a
- Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2015). On the left anterior negativity (LAN): The case of morphosyntactic agreement: A Reply to Tanner et al. *Cortex*, 66, 156–159.
- Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, 47(8), 908–930. doi:10.1016/j.cortex.2011.02.019
- Molinaro, N., Kim, A., Vespignani, F., & Job, R. (2008). Anaphoric agreement violation: An ERP analysis of its interpretation. *Cognition*, 106(2), 963–974. doi:10.1016/j.cognition.2007.03.006
- Molinaro, N., Vespignani, F., Zamparelli, R., & Job, R. (2011). Why brother and sister are not just siblings: Repair processes in agreement computation. *Journal of Memory and Language*, 64(3), 211–232. doi:10.1016/j.jml.2010.12.002
- Montrul, S. (2008). *Incomplete acquisition in bilinguals: Re-examining the age factor*. Amsterdam: John Benjamins. doi:10.1017/S0272263109990088
- Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: An electrophysiological exploration in both languages of Spanish-English bilinguals. *Cognitive Brain Research*, 22(2), 205–220. doi:10.1016/j.cogbrainres.2004.08.010



- Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, 24(4), 933–947. doi:10.1162/jocn\_a\_00119
- Näätänen, R., & Galliard, A. W. K. (1983). The orienting reflex and the N2 deflection of the ERP. In A. W. K. Galliard & W. Ritter (Eds.), *Tutorials in event related potential research: Endogenous components* (pp. 119–141). Amsterdam: Elsevier/North-Holland. doi:10.1016/S0166-4115(08)62036-1
- Neville, H. J., Nicol, J., Barss, A., Forster, K., & Garrett, M. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, 3, 155–170. doi:10.1162/jocn.1991.3.2.151
- Nevins, A., Dillon, B., Malhotra, S., & Phillips, C. (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research*, 1164, 81–94. doi:10.1016/j.brainres.2007.05.058
- Newman, A. J., Ullman, M. T., Pancheva, R., Waligura, D. L., & Neville, H. J. (2007). An ERP study of regular and irregular English past tense inflection. *NeuroImage*, 34(1), 435–445.
- Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The influence of language proficiency on lexical semantic processing in native and late learners of English. *Journal of Cognitive Neuroscience*, 24(5), 1205–1223. doi:10.1162/jocn\_a\_00143
- Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience*, 17(8), 1212–1228. doi:10.1162/0898929055002436
- Opitz, C. (2011). First language attrition and second language acquisition in a second language environment. PhD dissertation, Trinity College Dublin. doi: 10.1017/S0272263100006240
- Orsini, A., & Pezzuti, L. (2013). *WAIS-IV*. Giunti OS, Firenze: Manuale.
- Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language*, 59(3), 494–522. doi:10.1006/brln.1997.1793
- Osterhout, L., & Hagoort, P. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes*, 14(1), 1–14. doi:10.1080/016909699386356
- Osterhout, L., McKinnon, R., Bersick, M., & Corey, V. (1996). On the language-specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family? *Journal of Cognitive Neuroscience*, 8(6), 507–526. doi:10.1162/jocn.1996.8.6.507
- Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond*. London: Psychology Press. doi: 10.4324/9780203509050
- Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: A means for exploring the neurocognition of second language processing. *Language Learning*, 56(s1), 199–230. doi:10.1111/j.1467-9922.2006.00361.x
- Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34(6), 739–773. doi:10.1006/jmla.1995.1033
- Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., Frenck-Mestre, C., & Hirschensohn, J. (2008). Second-language learning and changes in the brain. *Journal of Neurolinguistics*, 21(6), 509–521. doi:10.1016/j.jneuroling.2008.01.001
- Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, 22(12), 2728–2744. doi:10.1162/jocn.2009.21393
- Pakulak, E., & Neville, H. J. (2011). Maturational constraints on the recruitment of early processes for syntactic processing. *Journal of Cognitive Neuroscience*, 23(10), 2752–2765. doi:10.1162/jocn.2010.21586

- Pallier, C. (2007). Critical periods in language acquisition and language attrition. In B. Kopke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language attrition: Theoretical perspectives* (pp. 156-167). Amsterdam: John Benjamins. 155–168. doi: 10.1075/sibil.33.11pal
- Pallier, C., Dehaene, S., Poline, J. B., LeBihan, D., Argenti, A. M., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex*, 13, 155161. doi:10.1016/S1053-8119(01)91925-1
- Paradis, M. (2003). The bilingual Loch Ness Monster raises its non-asymmetric head again- or, why bother with such cumbersome notions as validity and reliability? Comments on Evans et al. (2002). *Brain and Language*, 87, 441–448. doi:10.1016/S0093-934X(03)00136-6
- Paradis, M. (2007). L1 attrition features predicted by a neurolinguistic theory of bilingualism. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language attrition: Theoretical perspectives* (pp. 9–37). Amsterdam: John Benjamins. doi: 10.1075/sibil.33.09par
- Pelc, L. (2001). L1 Lexical, Morphological and Morphosyntactic Attrition in Greek–English Bilinguals. Doctoral dissertation, New York, NY: City University of New York.
- Penfield, W. (1965). Conditioning the uncommitted cortex for language learning. *Brain*, 88, 787798.
- Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. New York: Athenaeum.
- Perani, D., Paulesu, E., Sebastian-Galles, N., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S. F., Fazio, F., & Mehler, J. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain*, 121, 1841–1852. doi: 10.1093/brain/121.10.1841
- Pierce, L. J., Klein, D., Chen, J. K., Delcenserie, A., & Genesee, F. (2014). Mapping the unconscious maintenance of a lost first language. *Proceedings of the National Academy of Sciences*, 111(48), 17314–17319. doi:10.1073/pnas.1409411111
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128–2148. doi:10.1016/j.clinph.2007.04.019
- Prat, C. S. (2011). The brain basis of individual differences in language comprehension abilities. *Language & Linguistics Compass*, 5, 635–649. doi:10.1111/j.1749-818X.2011.00303.x
- Roehm, D., Bornkessel, I., Haider, H., & Schlesewsky, M. (2005). When case meets agreement: Event related potential effects for morphology-based conflict resolution in human language comprehension. *NeuroReport*, 16(8), 875–878.
- Rossi, S., Gugler, M., Friederici, A., & 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(12), 2030–2048. doi:10.1162/jocn.2006.18.12.2030
- Royale, P., Drury, J. E., & Steinhauer, K. (2013). ERPs and task effects in the auditory processing of gender agreement and semantics in French. *The Mental Lexicon*, 8(2), 216–244. doi:10.1075/ml.8.2.05roy
- Schlegel, A. A., Rudelson, J. J., & Tse, P. U. (2012). White matter structure changes as adults learn a second language. *Journal of Cognitive Neuroscience*, 24(8), 1664–1670. doi:10.1162/jocn\_a\_00240
- Schmid, M. S. (2010). Languages at play: The relevance of L1 attrition to the study of bilingualism. *Bilingualism: Language and Cognition*, 13(01), 1. doi:10.1017/S1366728909990368
- Schmid, M. S. (2011). *Language attrition*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511852046
- Schmid, M. S. (2013). First language attrition as a window to constraints on bilingual development. Keynote lecture at the International Symposium on Bilingualism (ISB9), Singapore, June 10, 2013.
- Schmid, M. S., & Köpke, B. (2011). L'attrition de la première langue en tant que phénomène psycholinguistique. *Language, Interaction and Acquisition: Special Issue on L1 attrition*, 2(2), 197–220. doi: 10.1075/lia.2.2.02kop
- Silva-Pereyra, J. F., & Carreiras, M. (2007). An ERP study of agreement features in Spanish. *Brain Research*, 1185(14), 201–211. doi:10.1016/j.brainres.2007.09.029
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1–33. doi:10.1016/0013-4694(75)90263-1

- Squires, N., Squires, K., & Hillyard, S. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, 38(4), 387–401. doi:10.1016/0013-4694(75)90263-1
- Stein, M., Federspiel, A., Koenig, T., Wirth, M., Strik, W., Wiest, R., & Dierks, T. (2012). Structural plasticity in the language system related to increased second language proficiency. *Cortex*, 48(4), 458–465. doi:10.1016/j.cortex.2010.10.007
- Steinhauer, K., Mecklinger, A., Friederici, A. D., & Meyer, M. (1997). Probability and strategy: An event-related potential study of processing syntactic anomalies. *Zeitschrift fuer Experimentelle Psychologie*, 2, 305–331. doi:10.3758/BF03197249
- Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research*, 25(1), 13–41. doi:10.1177/0267658308098995
- Tanner, D. (2011). Agreement mechanisms in native and nonnative language processing: Electrophysiological correlates of complexity and interference. ProQuest LLC, 789 East Eisenhower Parkway, PO Box 1346, Ann Arbor, MI 48106.
- Tanner, D., (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: A Commentary on “Grammatical agreement processing in reading: ERP findings and future directions” by Molinaro et al., 2014. *Cortex*, 66, 1–7. doi:10.1016/j.cortex.2014.04.007
- Tanner, D., Inoue, K., & Osterhout, L. (2012). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition*, 17(02), 277–293. doi: 10.1017/S1366728913000370
- Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, 16(02), 367–382. doi: 10.1017/s1366728912000302
- Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, 56, 289–301.
- 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*, 27, 173–204. doi:10.1017/S0272263105050102
- Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism*, 8(3), 257–277. doi:10.1177/13670069040080030601
- Van Hell, J. G., & Tokowicz, N. (2010). Event-related brain potentials and second language learning: Syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research*, 26, 43–74. doi:10.1177/0267658309337637
- Ventureyra, V., & Pallier, C. (2004). In search of the lost language: The case of adopted Koreans in France. In M. S. Schmid, B. Köpke, M. Keijzer, & L. Weilemar (Eds.), *First language attrition: Interdisciplinary perspectives on methodological issues* (pp. 207–221). Amsterdam: John Benjamins. doi: 10.1075/sibil.28.12ven
- Ventureyra, V. A., Pallier, C., & Yoo, H. Y. (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics*, 17(1), 79–91.
- Vespignani, F., Molinaro, N., & Job, R. (unpublished data). ERP study of parsing strategies during repair.
- Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron*, 37(1), 159–170. doi:doi:10.1016/S0896-6273(02)01150-9
- Weber-Fox, C. M., & Neville, H. J. (1996). Maturation constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8(3), 231–256. doi:10.1162/jocn.1996.8.3.231

White, E. J., Genesee, F., & Steinhauer, K. (2012). Brain Responses before and after intensive second language learning: Proficiency based changes and first language background effects in adult learners. *PLoS ONE*, 7(12), 52318. doi:10.1371/journal.pone.0052318

Woodcock, R. W., McGrew, K. S., Mather, N., & Schrank, F. A. (2003). *Woodcock-Johnson III diagnostic supplement to the tests of cognitive abilities*. Itasca, IL: Riverside. doi:10.1177/003435520104400407

### Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

**Appendix S1.** An example of data patterns with the original baseline correction (−200 to 0 ms). The N400, frontal positivity, and P600 effects we reported on the verb for Attriters (with a baseline of −200 to 200 ms) were also reliable with this original baseline.

**Appendix S2.** An illustration of the larger frontal positivity / P3a amplitudes on the verb during the first half of the experiment compared to the second half (for all subjects).