Confusing similar words: ERP correlates of lexical-semantic processing in first language attrition and late second language acquisition

Kristina Kasparian\textsuperscript{a,b,*}, Karsten Steinhauer\textsuperscript{a,b,*}

\textsuperscript{a} School of Communication Sciences & Disorders, McGill University, Montreal, Canada
\textsuperscript{b} Centre for Research on Brain, Language and Music (CRBLM), Montreal, Canada

ARTICLE INFO

Keywords:
Neuroplasticity
Lexical-semantic processing
First language attrition
Second language processing
Proficiency
Age-of-acquisition

ABSTRACT

First language (L1) attrition is a socio-linguistic circumstance where second language (L2) learning coincides with changes in exposure and use of the native-L1. Attriters often report experiencing a decline in automaticity or proficiency in their L1 after a prolonged period in the L2 environment, while their L2 proficiency continues to strengthen. Investigating the neurocognitive correlates of attrition alongside those of late L2 acquisition addresses the question of whether the brain mechanisms underlying both L1 and L2 processing are strongly determined by proficiency, irrespective of whether the language was acquired from birth or in adulthood. Using event-related-potentials (ERPs), we examined lexical-semantic processing in Italian L1 attriters, compared to adult Italian L2 learners and to Italian monolingual native speakers. We contrasted the processing of classical lexical-semantic violations (Mismatch condition) with sentences that were equally semantically implausible but arguably trickier, as the target-noun was “swapped” with an orthographic neighbor that differed only in its final vowel and gender-marking morpheme (e.g., cappello (hat) vs. cappella (chapel)). Our aim was to determine whether sentences with such “confusable nouns” (Swap condition) would be processed as semantically correct by late L2 learners and L1 attriters, especially for those individuals with lower Italian proficiency scores. We found that lower-proficiency Italian speakers did not show significant N400 effects for Swap violations relative to correct sentences, regardless of whether Italian was the L1 or the L2. Crucially, N400 response profiles followed a continuum of “nativelikeness” predicted by Italian proficiency scores – high-proficiency attriters and high-proficiency Italian learners were indistinguishable from native controls, whereas attriters and L2 learners in the lower-proficiency range showed significantly reduced N400 effects for “Swap” errors. Importantly, attriters and late L2 learners did not differ in their N400 responses when they belonged to the same proficiency subgroup. Attriters also showed an enhanced P600 response to both kinds of lexical-semantic anomalies, which we discuss as reflecting increased conflict-monitoring and conscious “second thought” processes specifically in attriters. Our findings provide some of the first ERP evidence of attrition effects, and are compatible with accounts of ongoing neuroplasticity for language in adulthood. Proficiency, rather than age-of-acquisition, seems to be the key factor in modulating certain neurocognitive responses, not only within L2 learners but also in L1 attriters.

1. Introduction

1.1. Neuroplasticity for language in adulthood: Contributions of L1 attrition

For decades, multilingualism research has sought to determine whether fundamental differences exist between first (L1) and second (L2) language learning, as a result of maturational limits on the plasticity of the neurocognitive mechanisms involved in language (“Critical Period Hypothesis (CPH)”: Lenneberg, 1967; Penfield and Roberts, 1959). Despite growing awareness that the brain mechanisms underlying language processes remain more flexible outside of early childhood than was previously believed to be the case, the view that age-of-acquisition (AoA) is the limiting factor in native-like language processing continues to pervade much of the literature on the neurocognition of bilingualism, particularly in morphosyntax (e.g., Abrahamsson and Hyltenstam, 2009; Clahsen and Felser, 2006a, 2006b; Hahne and Friederici, 2001; Pakulak and Neville, 2011; Weber-Fox and Neville, 1996). Others agree on the important role played by AoA but argue instead that factors such as language exposure...
and proficiency are those which ultimately determine the overlap in brain responses between native speakers and late learners (e.g., Bowden et al., 2013; Friederici et al., 2002; Osterhout et al., 2006; Rossi et al., 2006; Steinhauser et al., 2009; Van Hell and Tokowicz, 2010; Wartenburger et al., 2003). In the domain of lexical-semantic processing, reported differences between native speakers and late learners tend to be more quantitative than qualitative, and the influence of L2 proficiency level is generally agreed upon (e.g., Kotz and Elston-Güttler, 2004; Ojima et al., 2005; Wartenburger et al., 2003; see Slabakova, 2006). Still, it remains unclear what the progressive neurocognitive changes associated with changes in exposure and proficiency may be, whereas they are actually continuous or partly qualitative in more demanding sentence processing paradigms, and whether the driving factor of native-like neurocognitive responses is proficiency, given that some researchers continue to emphasize the independent contribution of AoA also in the lexical-semantic domain (e.g., Moreno et al., 2005; Newman et al., 2012).

The question of neuroplasticity for language learning has almost exclusively been explored from the angle of L2 processing and performance. However, a multilingual experience that can provide a complementary perspective in our study of neuroplasticity and language-processing is the phenomenon 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 and Schmid, 2004). Attriters experience a shift in exposure/use from their native-L1 to environmentally dominant L2. Thus, in contrast to L2 learners who continue to maintain predominant use of their native language and who therefore consider their L1 to be their dominant language, attriters are individuals for whom increasing L2 exposure and proficiency comes at a cost to their L1, exposure to which 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).

The logical corollary of claims in favor of AoA effects on language processes is that the L1, once been hard-wired or “entrenched” in the brain as a result of early exposure, is stable (Marchman, 1993; Penfield, 1965). The longer one is exposed to a given language, the more entrenched that language becomes (van Hell and Tokowicz, 2010) and the less language mechanisms can be modified by L2 exposure (Pallier, 2007). This argument has not only been proposed to explain why cross-linguistic influence from the direction of the L1 to the L2 is so pervasive in late learners (e.g., Tokowicz and MacWhinney, 2005), but also to account for the finding that attrition effects observed in children are far more pervasive than in older/adult Attriters (Ammerlaan, 1996; Bylund, 2008; Pec, 2001; see reviews by Köpke, 2004 and Köpke and Schmid, 2004; Montrul, 2008; adoption studies by Pallier et al., 2003; Ventureyra et al., 2004; Ventureyra and Pallier, 2004). From such findings, it has been suggested that L1 attrition – like L2 acquisition – is subject to maturational constraints (see Pallier, 2007; Schmid, 2011) and that the L1 stabilizes around puberty (Bylund, 2008). Although a multitude of studies have investigated the phenomenon of L1 attrition at the behavioral level, manifestations of attrition during real-time language processing using more sensitive temporal measures (such as eye-tracking or ERPs) are still relatively unexplored (but see Bergmann et al., 2015; Datta, 2010; Kasparian, 2015; Schmid, 2013).

Investigating the neurocognitive aspects of L1 attrition logically extends the work conducted on L2 processing, and is key to reaching a unified theoretical account of neuroplasticity in multilingual speakers. For one, examining attriters’ L1 allows us to determine whether early exposure is enough to guarantee native-like neurocognitive responses to a language, or whether attriters’ L1 processing is amenable to change, even when the shift in use and exposure has occurred well into adulthood. Secondly, comparing the L1 processing of attriters to the L2 processing of late learners allows us to assess whether the neurocognitive mechanisms underlying language processing are modulated by proficiency and exposure, irrespective of whether the language in question is the L1 or the L2 (i.e., regardless of AoA). In line with findings from the growing number of L2 processing studies, some researchers have shown that L1 processing in native monolinguals is also modulated by proficiency level (Pakulak and Neville, 2010; Prat, 2011). A further step – for which a direct comparison of L1 and L2 processing in native monolingual and several groups of bilingual speakers is necessary – would be to assess the likelihood that L1 and L2 processing fall along a continuum of more to less native-like processing, based on proficiency/exposure irrespective of whether the target language is the L1 or the L2 (i.e., AoA). In this vein, comparing L1 attrition and L2 acquisition allows us to examine whether processing changes with increasing attrition may be the reverse of those observed in L2 learning. To date, the notion of such a continuum has almost exclusively been explored within L2 learners relative to monolinguals, rather than systematically including native speakers who no longer predominantly use their L1.

1.2. Lexical-semantic processing in L2 learning and L1 attrition

Lexical access during online comprehension has been widely explored in ERP studies of late L2 acquisition, where researchers have examined whether late L2 learners deviate from early learners and monolingual speakers in how they process single words or sentences where a word is either semantically congruent or incongruent with the preceding context (e.g., Hahne, 2001; McLaughlin et al., 2004; Mueller, 2005; Ojima et al., 2005; Weber-Fox and Neville, 1996). Such studies have found that processing patterns in high-proficiency learners converge on those of native speakers, whereas less proficient learners show ERP responses (typically N400 effects) that are delayed in onset or reduced in amplitude (e.g., Kotz and Elston-Güttler, 2004; Midgley et al., 2009; Moreno and Kutas, 2005; Ojima et al., 2005; but see Newman et al., 2012). One longitudinal ERP study of L2 word learning by McLaughlin et al. (2004) evaluated the impact of amount of L2 exposure. Modulations in N400 amplitude associated with word-pseudoword discrimination and semantic priming effects were found to be correlated with hours of L2 instruction – with increasing L2 exposure, learners showed N400 response patterns that converged on those observed in native speakers. Interestingly, these subtle early changes as a function of L2 exposure were better detectable with ERPs than with a behavioral lexical decision task.

Compatible with the notion of a continuum of ERP response patterns as a function of L2 exposure is a longitudinal ERP study that compared the trajectory of L2 learning to that of L2 attrition in the same learners when they ceased being exposed to the L2 (Pitkänen, 2010). In response to illegal non-word stimuli, L2 learners elicited ERP patterns that increasingly approximated the responses seen in native speakers (P600) with increasing L2 instruction. Following instruction, the amount of time that had elapsed since L2 instruction directly predicted learners’ ERP response patterns, with less native-like processing (i.e., more L2 attrition) associated with a larger number of days since instruction. Interestingly, the brain responses associated with increasing attrition mirrored the qualitative changes that had been observed during learning, but in the reverse direction. Similar to McLaughlin et al. (2004), participants’ behavioral judgments did not capture these changes in processing with increasing/decreasing exposure.

Lexical-semantic attributes have also received attention in the L1 attrition literature, where the particular vulnerability of vocabulary has been highlighted. Studies using production tasks such as picture-naming, verbal semantic fluency and free-speech have provided evidence that attriters are slower to access lexical items and may have smaller or less diverse L1 vocabularies (De Bot, 1996; Köpke, 1999; Linch et al., 2009; Schmid and Jarvis, 2014; Schmid and Keijzer, 2009; Waas, 1996; Yagmur, 1997; Yilmaz and Schmid, 2012). Others have found that...
attriters experience competition or transfer of lexical items from the L2 to the L1 – processes described as lexical borrowing or semantic intrusions when words are cross-linguistically similar (Pavlenko, 2000). Some theories have accounted for these difficulties by positing breakdowns at the level of stored lexical representations in memory (see Ecke (2004) for theories of forgetting), while others have claimed that attriters are rusty in accessing lexical items during production and that the difficulties have more to do with retrieval than with representation (e.g., Green, 1986; Paradis, 1989, 1997). It remains an open question whether lexical difficulties in attriters’ L1 are limited to production or if they also occur during comprehension. To our knowledge, the present study is the earliest ERP study of lexical-semantic processing in sentence contexts where we are able to study lexical access and integration as sentence comprehension unfolds in real-time.

1.3. Confusable words

The present study examines the processing of “confusable words” within the target language (Italian) to test whether an aspect of lexical-semantics that is challenging for L2 learners would also pose a processing challenge for attriters who were minimally exposed to their native Italian.

In Italian, pairs of words that differ in their final vowel and their semantic meaning (e.g., cappello (hat) vs. cappa (hearse)) are notoriously difficult for Italian learners, and are the subject of many cautionary notes in Italian textbooks (e.g. Fogliato and Testa, 2006, p. 41) and online language blogs (e.g., “Tricky Words,” 2012; “Maschino o Femminile”, 2009). These lexical items may also be difficult to master due to their interface with gender; a word-final vowel change in Italian typically corresponds to a change in gender, such that one word in the minimal pair is feminine and the other masculine. Learning gender in the L2 has been shown to be especially challenging to learners from L1 backgrounds in which gender is not morphologically marked or where gender rules operate differently than in the L2 (Dowens et al., 2010; Foucart and Frenck-Mestre, 2012; Hawkins, 2001; Sabourin and Stowe, 2006; Tokowicz and MacWhinney, 2005). In the present case, the absence of morphological gender marking on nouns in English may make it additionally challenging for English-Italian learners to semantically distinguish between the confusable words based on their gender.

Studies with L2 learners conducted in both visual and auditory modalities have shown that word recognition is slower and/or more difficult when words can be confounded with orthographically/phono-logically similar words within the L2 (Carreiras et al., 1997; Rüschemeyer, 2005; Rüschemeyer et al., 2008). A behavioral and ERP study by Rüschemeyer et al. (2008) showed that German L2 learners were more likely than native speakers to judge auditorily-presented semantically-anomalous sentences (e.g., Der Tisch wurde am Nachmittag geangelt=The table was caught in the afternoon) as semantically acceptable if replacing the incongruent noun by its phonological neighbor (e.g., Fisch=fish) would have rendered the sentence plausible, compared to anomalous sentences with no confusable noun. In contrast, there were no differences between conditions for native speakers. In their ERP experiment, L2 learners showed N400 amplitude differences on a visually-presented prime between neighbor vs. non-neighbor conditions. The authors concluded that L2 learners “are more likely to be tricked by the presence of a phonological neighbor than are native speakers” (p. 134) and that they experience interference from semantically unrelated orthographic/phonological neighbors during L2 word-recognition. These results are in line with other claims that L2 learners are less efficient at inhibiting competing activation of intralingual competitors (Elston-Güttler and Friederici, 2007; FitzPatrick and Indefrey, 2010; Sebastián-Gallés et al., 2005; Weber and Cutler, 2004).

Research on parallel lexical-activation and competition has shown that proficiency and exposure in the target language have a role to play in regulating the selectivity of lexical access (Dijkstra and Van Heuven, 2002; Gollan et al., 2008; Kroll and Stewart, 1994). Increasing proficiency leads to stronger form-meaning links and, thus, better efficiency in inhibiting non-target semantic meanings (Jared and Kroll, 2001). Language dominance influences the ease of lexical-access due to frequency of use, with words in the dominant language at a higher activation baseline level than words in the non-dominant language (Gollan et al., 2008). With such findings in mind, a comparison of native monolinguals, L1 attriters and L2 learners, with considerations of language dominance, proficiency and exposure levels, could be beneficial for our understanding of lexical-access during comprehension, as well as of the factors that drive similarities or differences in processing in these individuals.

The orthographic neighbors tested in studies by Rüschemeyer et al. (2008) and Laszlo and Federmaneier (2009) were words that differed in their initial phoneme (e.g. Tisch/Fisch or zisch/disich respectively), which may potentially make them less confusable than Italian nouns that are identical until their final vowel and gender marker (see Cohort model, Marslen-Wilson, 1987). Arguably, the point of disambiguation being a word-final phoneme (an inflectional morpheme for gender) may make the distinction between these items less salient in the input and may add an additional dimension of confusability (form-concept-gender). These inherent difficulties may result in less robust form-meaning links in the lexicon, at least at lower levels of proficiency or exposure in the target language.

Testing lexical access and semantic integration of confusable nouns in Italian (L1) attriters and Italian (L2) learners gives rise to two possibilities: (1) processing patterns will fall along a continuum from native-like to less native-like based on Italian proficiency / exposure irrespective of group membership, such that attrition parallels L2 acquisition in reverse, or (2) these form-gender-concept mappings are not affected by reduced L1 exposure or modulated by proficiency, as this information is robust when acquired as an L1 early in life.

1.4. Beyond N400 effects

Although ERP investigations of lexical-semantic processing in sentence contexts have mainly focused on modulations of the N400 component (Kutas and Hillyard, 1980, 1984; Kutas and Federmeier, 2011 for a review), a systematic review by Van Petten and Luka (2012) identified a number of studies that revealed post-N400 parietal positivities consistent with a P600 effect.

It has been shown that the P600 is not merely elicited in response to sentences that are syntactically anomalous or disregarded (Friederici et al., 1996; Hagoort et al., 1993; Kaan and Swaab, 2003; Osterhout and Holcolm, 1992; Osterhout et al., 1994; Phillips et al., 2005), but also when semantically-based predictions in sentence or discourse contexts are disconfirmed (DeLong et al., 2011; Federmaneier et al., 2007; Kuperberg, 2007; Otten and Van Berkum, 2008; Steinhauer et al., 2010). Van Petten and Luka describe the P600 as a reflection of a “processing cost” associated with having to review the preceding context to diagnose the problem and repair it.

A related and perhaps even more general explanation of P600 effects elicited in response to lexical-semantic violations is that the P600 reflects an underlying “conflict monitoring” process (e.g., Kolk et al., 2003; Kolk and Chwilla, 2007; Van Herten et al., 2005, 2006; Van de Meerdondonk et al., 2009; Vissers et al., 2006, 2008). A monitoring process involves double-checking all forms of the input, including orthographic/phonological features, for possible processing errors once an anomaly has been detected (Van de Meerdondonk et al., 2009). According to this view, the P600 depends on the severity of the conflict during lexical integration; a mild conflict leads to integration difficulties that are resolved without having to re-analyze the input (N400 without subsequent P600), whereas strong conflicts trigger a re-analysis process, thereby eliciting a P600.

In addition to examining potential P600 differences between semantically-anomalous sentences where the target-word is either an
orthographic/phonological neighbor or an entirely unrelated word, it is
of theoretical interest to determine whether L1 attriters and/or late L2
learners show processing differences from non-attriter native speakers
at later stages of lexical-semantic processing, depending on the degree
of the perceived conflict.

1.5. The present study

The critical words in our experiment involved Italian minimal pairs
that differed in their final vowel (and gender), but also in their lexical-
semantic meaning. These words, therefore, were orthographic and
phonological neighbors of the words that were intended to fit the
sentence context. We examined ERP effects in response to sentences
where the intended target word was switched with its minimal pair
(“Swap” condition), as well as to sentences where an orthographically and
semantically unrelated word was inserted into the sentence context
(“Mismatch” condition). Both of these semantic violations were
compared to semantically-correct sentences (see Table 1).

While the “Mismatch” sentences were similar to classical violations
tested in ERP paradigms of lexical-semantic processing (e.g. “He
spread the warm bread with socks”; Kutas and Hillyard, 1980, 1984), the “Swap” sentences – though equally semantically implausible – arguably consisted of a less salient type of lexical-semantic violation,
given the confusability of the target words.

In terms of condition effects, we expected non-attriters native
speakers (Controls) to elicit large N400 effects for both “Mismatch”
and “Swap” anomalies relative to “Correct” sentences. A graded N400
pattern (Mismatch > Swap > Correct) would be conceivable, as re-
duced N400s have been reported for conditions where the presented
target word overlaps with the contextually appropriate word in ortho-
graphic/phonological form (Laszlo and Federmeier, 2009) or in
semantic meaning (e.g., The pizza was too hot to eat <*drink
< **cry; Kutas and Hillyard, 1984; see also Federmeier and Kutas,
1999; Federmeier et al., 2007; Wlooto and Federmeier, 2016). Crucially,
however, we expected Controls to significantly distinguish between
“Swap” and “Correct” sentences in the N400 time-window. Further
downstream, in the P600 interval, we expected that the semantically anomalous sentences could potentially elicit P600 effects.

Table 1

<table>
<thead>
<tr>
<th>CONDITION</th>
<th>Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>Per coprire la testa, il pescatore porta il cappello di lana.</td>
</tr>
<tr>
<td></td>
<td>To cover his head, the fisherman wears the hat of wool.</td>
</tr>
<tr>
<td></td>
<td>Per assister alla messa, la coppia frequenta il cappello ogni giorno.</td>
</tr>
<tr>
<td></td>
<td>To assist the mass, the couple attends the chapel every day.</td>
</tr>
<tr>
<td></td>
<td>Per indicare gli orari, il negoziante lascia il cartello in vitrina.</td>
</tr>
<tr>
<td></td>
<td>To indicate the opening-hours, the shopkeeper leaves the sign in the window.</td>
</tr>
<tr>
<td></td>
<td>Per nascondere i documenti, la spia mette il cartello in cassaforte.</td>
</tr>
<tr>
<td></td>
<td>To hide the documents, the spy puts the briefcase in the safe.</td>
</tr>
<tr>
<td>Swap</td>
<td>Per coprire la testa, il pescatore porta la “cappella” di lana.</td>
</tr>
<tr>
<td></td>
<td>To cover his head, the fisherman wears the “chapel” of wool.</td>
</tr>
<tr>
<td></td>
<td>Per assister alla messa, la coppia frequenta il “cappello” ogni giorno.</td>
</tr>
<tr>
<td></td>
<td>To assist the mass, the couple attends the “hat” every day.</td>
</tr>
<tr>
<td></td>
<td>Per indicare gli orari, il negoziante lascia la “cartella” in vitrina.</td>
</tr>
<tr>
<td></td>
<td>To indicate the opening-hours, the shopkeeper leaves the “briefcase in the window.”</td>
</tr>
<tr>
<td></td>
<td>Per nascondere i documenti, la spia mette il “cartello” in cassaforte.</td>
</tr>
<tr>
<td></td>
<td>To hide the documents, the spy puts the “sign in the safe.</td>
</tr>
<tr>
<td>Mismatch</td>
<td>Per coprire la testa, il pescatore porta la “cartella” di lana.</td>
</tr>
<tr>
<td></td>
<td>To cover his head, the fisherman wears the briefcase of wool.</td>
</tr>
<tr>
<td></td>
<td>Per assister alla messa, la coppia frequenta il “cartello” ogni giorno.</td>
</tr>
<tr>
<td></td>
<td>To assist the mass, the couple attends the “document every day.</td>
</tr>
<tr>
<td></td>
<td>Per indicare gli orari, il negoziante lascia la “cappella” in vitrina.</td>
</tr>
<tr>
<td></td>
<td>To indicate the opening-hours, the shopkeeper leaves the “chapel in the window.”</td>
</tr>
<tr>
<td></td>
<td>Per nascondere i documenti, la spia mette il “cappello” in cassaforte.</td>
</tr>
<tr>
<td></td>
<td>To hide the documents, the spy puts the “hat in the safe.</td>
</tr>
</tbody>
</table>

orthographic/phonological neighbor or an entirely unrelated word, it is
of theoretical interest to determine whether L1 attriters and/or late L2
learners show processing differences from non-attriter native speakers
at later stages of lexical-semantic processing, depending on the degree
of the perceived conflict.

The critical words in our experiment involved Italian minimal pairs
that differed in their final vowel (and gender), but also in their lexical-
semantic meaning. These words, therefore, were orthographic and
phonological neighbors of the words that were intended to fit the
sentence context. We examined ERP effects in response to sentences
where the intended target word was switched with its minimal pair
(“Swap” condition), as well as to sentences where an orthographically and
semantically unrelated word was inserted into the sentence context
(“Mismatch” condition). Both of these semantic violations were
compared to semantically-correct sentences (see Table 1).

While the “Mismatch” sentences were similar to classical violations
tested in ERP paradigms of lexical-semantic processing (e.g. “He
spread the warm bread with socks”; Kutas and Hillyard, 1980, 1984), the “Swap” sentences – though equally semantically implausible – arguably consisted of a less salient type of lexical-semantic violation,
given the confusability of the target words.

In terms of condition effects, we expected non-attriters native
speakers (Controls) to elicit large N400 effects for both “Mismatch”
and “Swap” anomalies relative to “Correct” sentences. A graded N400
pattern (Mismatch > Swap > Correct) would be conceivable, as re-
duced N400s have been reported for conditions where the presented
target word overlaps with the contextually appropriate word in ortho-
graphic/phonological form (Laszlo and Federmeier, 2009) or in
semantic meaning (e.g., The pizza was too hot to eat <*drink
< **cry; Kutas and Hillyard, 1984; see also Federmeier and Kutas,
1999; Federmeier et al., 2007; Wlooto and Federmeier, 2016). Crucially,
however, we expected Controls to significantly distinguish between
“Swap” and “Correct” sentences in the N400 time-window. Further
downstream, in the P600 interval, we expected that the semantically anomalous sentences could potentially elicit P600 effects.

Although both Mismatch and Swap sentences are equally implausible,
one could argue that Swap violations are less severe (i.e., resulting in a
weaker conflict) because simply replacing the target-word with its
nearly identical orthographic neighbor would rescue the sentence. In
this view, we would expect Mismatch violations to elicit larger P600
responses than Swap sentences, as Mismatch errors would require a
more elaborate re-analysis to arrive at a plausible sentence interpreta-
ton. On the other hand, if the replacement of a target-noun with its
orthographic/phonological neighbor in Swap violations causes en-
hanced conflict and results in readers having to double-check the input
for the source of the problem, then Swap violations may elicit larger
P600 effects than Mismatch sentences.

A key question was whether processing patterns in L1 attriters and
late L2 learners would resemble those observed in native monolinguals.
If L1 attriters, having acquired their L1 since birth and having lived in
an exclusively monolingual environment until adulthood, remain
native-like in their Italian lexical-semantic processing despite their
limited L1 exposure and predominant L2 use, we would expect both
violation conditions violations to elicit large N400 effects. In contrast,
attriters may not show native automatic processing responses, espe-
cially for errors involving confusable words, and be more inclined to
process Swap sentences as correct; this would confirm their anecdotal
reports of “confusing words” and could point to weakening L1 form-
meaning links (Gollan et al., 2008; Jared and Kroll, 2001) that make
lexical access and integration less efficient in attriters. The same
possibilities hold for the L2 learners who, with their late age-of-
acquisition, may differ from L1 native speakers (monolinguals and
attriters) in their ERP responses.

However, we expected processing patterns to depend on Italian
proficiency level, where ERP responses in individuals with higher
Italian proficiency scores would converge upon those observed in
non-attriter native speakers. We expected individuals with lower
Italian proficiency to show a reduced N400 effect in response to
Swap violations compared to Mismatch violations, with – crucially –
little to no difference between Swap and Correct sentences. Our aim in
comparing L1 attriters and L2 learners was to explore whether the
native-like-ness of N400 responses was predicted by Italian profi-
ciency, irrespective of group membership, thus pointing to a con-
tinuum between L1 and L2 processing rather than a qualitative
difference, irrespective of group membership, thus pointing to a con-
tinuum between L1 and L2 processing rather than a qualitative
difference between the two. Such a finding would strongly advocate
favor of ongoing neuroplasticity for language in adulthood, and
proficiency-based brain responses, regardless of whether the target
language is one’s L1 or L2.

Finally, we also examined whether group and/or proficiency
differences were observed at later stages of lexical-semantic processing,
namely on the P600s. One possibility was that the groups (or
proficiency subgroups) would differ most from native controls in the
N400 time-window, failing to automatically detect the anomaly, but
would converge on their P600 responses, having detected the violation
in a more conscious “double-take” or “second-thought” realization of
the conflict further downstream (Kasparian et al., unpublished manu-
script, for a similar finding in L1 attriters’ and L2 learners’ processing
of Italian-English homographs/cognates). An alternate possibility was
that the groups would deviate most from native controls in the re-
analysis stage reflected by the late P600 rather than in the N400 time-
window, thus demonstrating processing differences at the level of re-
analyzing or repairing the input (Kasparian et al., unpublished manuscript).

To date, ours is the first ERP study to examine lexical-semantic processing in sentence contexts in L1 attritors and L2 learners, and one of the first steps in systematically exploring the impact of proficiency and changes of exposure/use in the target language.

2. Method

2.1. Participants

2.1.1. Attritors

Participants in the attrition group consisted of twenty-four Italian native speakers (14 female; M age: 36; Range: 25–50) who had immigrated to Canada in adulthood (M age at immigration (AoA) of English)=28.2 years; M length of residence=11 years). Participants unanimously reported having very limited use of their native Italian in contrast to their predominant use of English, and having noticed changes to their native Italian fluency since immigration. For many individuals (43%), these changes were described as difficulties in vocabulary (e.g., difficulties in accessing the intended word, semantic intrusions from English and the production of “false friends”), while other participants (9%) reported difficulties in grammar (e.g., influences of English grammar or awkward non-native Italian constructions). Approximately 48% reported difficulties in both grammar and vocabulary. Given the unanimous self-reports of L1 difficulties while living in an L2 environmental context, we refer to these individuals as “Attritors”. Although Attritors were living in the French-speaking province of Québec, in order to minimize the potential influence of cross-linguistic similarities between Italian and French, we recruited Attritors who either had minimal knowledge of French or whose use of English outweighed their use of French.

2.1.2. L2 learners

Twenty English native speakers who were intermediate/advanced L2 learners of Italian were recruited (15 female; M age: 31.6; Range=21–51). Participants had learned Italian at an average age of 21 years (=AoA), mostly in a classroom-instruction context. Importantly, none of the learners were heritage speakers who had grown up hearing Italian, or a dialectal variety of it, spoken to (or around) them in childhood. Individuals were recruited from a number of organizations in the Montreal area: University-level Italian language programs and music programs, Italian student associations, language learning schools offering advanced Italian courses to adults, the Italian Cultural Institute, and an online “Meetup” social network group, where Italian speakers of varying levels and linguistic backgrounds sign up to attend events where they could practice with other Italian speakers in Montreal. We recruited L2 learners who were dominant in English though living in a French-speaking province. It can be argued that the two groups of bilinguals were more comparable in their linguistic background than if we had tested English-Italian L2 learners living in Italy.

2.1.3. Italian native speaker controls

Thirty Italian native speakers residing in Italy were recruited as our native speaker control group (17 female; M age: 31; Range=25–54). Participants had minimal exposure to second languages (including English and Italian dialects), which we operationally defined as less than five hours per week. All participants except one were right-handed and none had a known history of neurological disorders.

2.2. Proficiency measures

Participants in all groups completed a background questionnaire pertaining to their demographic information (age, gender, education level), and language background. Attritors answered additional questions pertaining to their immigration history, first language exposure and identity/attitudes felt towards each language and culture.

All participants (including Controls) completed four Italian proficiency measures: (1) 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; (2) A written C-test (Italian version: Krais, 2008), where they were asked to fill in the blanks in 5 short texts in which twenty words in each text had been partially deleted; (3) 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 (4) A timed reading fluency task where participants were asked to produce as many vocabulary items from two categories (“animals” and “fruits and vegetables”) as possible within one minute per category. Participants also completed (1) a timed reading fluency task where they had to silently read and answer as many true-false statements as possible in three minutes (Woodcock et al., 2003; adapted into Italian for this study); and (2) the letter-number-sequencing task from the Italian WAIS-IV as a measure of Working Memory (Orsini and 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.

Attritors scored numerically lower on all four proficiency measures, but did not differ significantly from Controls (p>0.1). Although our recruitment process targeted L2 learners in the advanced range of Italian proficiency, our group of L2 learners was not matched to Controls or Attritors on their proficiency level. Although, overall, our L2 learners were significantly lower in proficiency measures compared to the two groups of native-Italian speakers, there was also overlap between the groups, such that several (n=5) L2 learners were in the native speaker range on at least one measure (see Discussion on p. 41). This allowed us to study whether ERP correlates of lexical-semantic processing approach and converge upon native-like profiles with increasing proficiency. Recall that our main aim was to determine whether ERP responses fall along a continuum modulated by Italian proficiency level (but see Discussion for questions our data cannot immediately address and for required follow-up studies).

Based on their scores on the 3 individual proficiency tests (excluding self-report measures, as monolingual controls rated themselves at ceiling), subgroups of “high” and “low” proficiency were derived by median split. Participants’ were categorized by considering their ranking on all 3 tests as a composite measure; if participants were “high” on 2 out of the 3 proficiency tests, they were categorized as being “high” overall (and vice versa for low-proficiency). High and low proficiency subgroups are described in Table 3. High vs. low proficiency groups (including within native monolinguals and within L1 attritors) differed significantly on most proficiency tasks.

2.3. Stimuli

Sentence examples of the three experimental conditions are provided in Table 1. The target words in this experiment were each constituents of a minimal pair in Italian. We searched for pairs of nouns that differed only in their final vowel (e.g., cappello vs. cappella) and, consequently, in their semantic meaning (e.g., hat vs. chapel). We avoided nouns where the vowel change resulted in a difference in pronunciation or length, as well as nouns that overlapped in semantic meaning such that it would be difficult to create a sentence context that

---

1 As part of the larger study, the same Attritors also participated in English ERP studies and were given the English versions of these same proficiency measures (see Kasparian et al., unpublished manuscript).
highly constrained the meaning of only one word in the minimal pair (e.g., *cero* (candle) / *cera* (wax)). We also avoided vowel-initial words, so that the determiner preceding the target noun would be *determiner* (e.g., *highly* constrained the meaning of only one word in the minimal pair *mine* / *mio*). We decided to replace a masculine noun by a feminine one in the *Mismatch* condition as well. Quadruples (i.e., the two sets of minimal pairs) were matched as closely as possible on word length and frequency, but reassigned in cases where *Swap* and *Mismatch* sentences happened to be semantically plausible (even in a remote, figurative sense). The advantage of our design is that it is completely counterbalanced – every target noun contributed to each main condition, thus ruling out that any differences between conditions were due to target noun characteristics such as frequency or length. All sentences were eleven words long and were identical in grammatical structure. The target noun was in the 9th position in the sentence, preceded by its determiner, and followed by a prepositional phrase, such that, in all items, the target noun was the direct object of the verb. Sentences were checked by three Italian native speakers who did not take part in the study.

Although reporting effects of cloze probability is beyond the scope of the current paper, we obtained and coded cloze probability values for our experimental sentences for future examination, as sentence constraint is viewed as the most important factor influencing participants’ expectations/predictions of upcoming content words, thus modulating N400 amplitudes (see Kutas and Federmeier (2011) for a review). We assessed cloze probability with a sentence-completion task administered to 12 Italian native speakers who did not participate in the study. Cloze probability was operationally defined as the percentage of times a word was provided, and the stimuli were coded for high or low cloze probability, as determined by a median split (operationally defined as the percentage of times a word was provided, and the stimuli were coded for high or low cloze probability, as determined by a median split).
mean = 43%).

Three experimental lists were created such that, across lists, each sentence contributed equally to each condition, while no sentence context nor target noun was repeated within an experimental list. Each participant was presented with each noun within a minimal pair (e.g., cappello/cappella) only once, and in different conditions (i.e., different preceding contexts). Nouns belonging to the same quadruple (2 minimal pairs, e.g., cappello/cappella/cartello/cartella) were kept as separated as possible during the pseudo-randomization process, and never occurred within less than 8 intervening trials. Each participant also saw 228 filler sentences, which were part of the larger Italian study (testing morphosyntactic number agreement processing and relative clause word-orders) and which will be reported in separate papers (see Kasparian et al., 2013, 2014, 2016). Out of the total of 324 pseudorandomized stimuli (96 experimental and 228 fillers) per participant, 146 sentences (approx. 45%) were acceptable (grammatically and semantically), while 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 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: (1) 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; (2) to assess whether conditions showed graded response patterns (and whether these were related to graded ERP patterns); and (3) 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 permmissibility of these constructions.

Participants were seated in 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). Participants received a short demonstration to show how eye-movement, blinks and muscle movement create artifacts in the EEG signal, and 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 twenty 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 first character of the first word appeared in uppercase, and the rest of the words were in lowercase. 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 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 after a 500 ms ISI, indicating the onset of the response interval. The prompt remained on the screen until participants pressed a button from 1 to 5. Immediately after a response was provided, the image of the blue eye appeared at the center of the screen for a 2000 ms interval, encouraging participants to blink their eyes between trials. The next trial began after the

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 6 of which were external electrodes: 4 electro-oculogram (EOG) channels placed above and below the left eye (EOGV), and at the outer canthus of each eye (EOGH), as well as 2 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 kΩ for scalp and reference electrodes, and below 10 kΩ for EOG electrodes. Signals were amplified using NeuroScan (Canada) and BrainVision (Italy) and filtered online with a band-pass filter of 0.1–100 Hz sampling rate of 500 Hz. Data pre-processing and analyses were carried out using EEProbe (ANT, Enschede, Netherlands).

Offline, EEG recordings were 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 standard deviation of 30 microvolts. One participant in the attrition group was excluded due to baseline problems. On average, participants contributed 30 artifact-free trials per condition (range: 66–100%), with no differences across conditions (ps > 0.1). The number of artifact-free trials did not significantly differ across groups, for any of the conditions.

ERPs were time-locked to the onset of the target noun with a baseline correction of −500 to 0 ms. The original baseline of −200 to 0 ms was adjusted in order to compensate for early differences triggered by the preceding verb in some attriters and L2 learners. Two consecutive and non-overlapping time-windows of interest were examined, based on the ERP components we expected as well as visual inspection of the grand average data for each group: (1) 300–550 ms (N400) and (2) 650–1000 ms (P600).

Repeated-measures ANOVAs were performed separately for 4 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 Condition (correct, swap, mismatch) and Ant-Post (anterior, central, parietal, occipital). Lateral ANOVAs additionally included factors Hemisphere (left, right) and Laterality (lateral, medial). For all ANOVAs, Group (Controls, Attriters, L2 learners) and Proficiency (High, Low) were between-subjects factors (but see Section 3.4). All global ANOVAs which yielded significant interaction including the factor Condition were followed up with pair-wise ANOVAs to clarify the nature of the interaction. 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.

A subset of participants who exhibited severe drifts that affected random noise-levels in the grand-average data were filtered with a 0.5–30Hz band-pass filter, in order to avoid excluding Attriters and L2 learners who were very difficult to recruit. With the awareness that such a band-pass filter could affect slow-going waves such as the P600, single-subject data as well as grand-average data of filtered participants were compared to the original data with the standard filter, and the pattern of results was consistent.
3. Results

3.1. Acceptability judgments

Acceptability ratings (1-5) for each sentence condition are shown in Fig. 1a. A repeated-measures ANOVA with within-subjects factor Condition (Correct, Swap, Mismatch) and between-subjects factor Group (Controls, Attriters, L2 learners) revealed a significant main effect of Condition ($F(2,142)=307.914, p<0.0001$ after Greenhouse-Geisser correction), a significant main effect of Group ($F(2,71)=9.238, p<0.0001$), as well as a significant Condition x Group interaction ($F(4,142)=28.303, p<0.0001$). A one-way ANOVA by Condition revealed that Group differences existed for Correct ($F(2,73)=7.267, p<0.0001$) as well as for Swap ($F(2,73)=13.793, p<0.0001$) and Mismatch sentences ($F(2,73)=16.029, p<0.0001$). Post-hoc Bonferroni-corrected multiple comparisons between groups showed that both violation conditions were rated lower than the correct control condition (Correct vs. Swap: Condition: $F(1,71)=225.198, p<0.0001$; Correct vs. Mismatch: $F(1,71)=451.055, p<0.0001$). The significant group interaction was due to the Italian L2 learners differing from both native controls and Attriters on each of the three conditions ($p<0.01$). Controls and Attriters, however, did not differ significantly on any of the conditions ($p>0.1$). Differences in ratings between Mismatch vs. Swap sentences ($F(1,71)=135.086, p<0.0001$) were not significantly affected by Group ($p>0.1$). It remains to be seen whether online measures such as ERPs would reveal group differences in the real-time processing of Swap sentences relative to Mismatch violations, contrary to offline responses provided at the end of the sentence.

The L2 learners who seemed more willing to accept semantically-anomalous sentences as correct, either because they did not perceive the target-word errors, or because they were in a syntactic “processing mode”, giving higher acceptability ratings overall when errors involved vocabulary rather than grammar (i.e., in our other sub-experiments). Although the present study does not enable us to test this possibility, comparing L2 learners’ ratings across the different Italian sub-experiments that were part of the larger study showed that syntactic anomalies did receive lower ratings than lexical-semantic anomalies. While L2 learners’ average rating in response to lexical-semantic violations was 3.5 on 5 violations of number agreement received a significantly lower average rating of 2.88 ($p<0.0001$) and infrequent/dispreferred (but grammatical) relative-clause word orders received a lower rating of 3.23 ($p=0.09$). Thus, it could be conceivable that the L2 learners in our study were stricter with grammatical errors overall than with vocabulary errors.

3.2. Reaction times

Reaction times between the onset of the prompt and participants’ button-press are depicted in Fig. 1b. A repeated-measures ANOVA with within-subjects factor Condition (Correct, Swap, Mismatch) and between-subjects factor Group (Controls, Attriters, L2 learners) revealed a significant main effect of Condition ($F(2,142)=13.888, p<0.0001$ after Greenhouse-Geisser correction), as well as a significant main effect of Group ($F(2,71)=7.264, p<0.001$). However, the interaction between Condition x Group was only marginal ($F(4,142)=2.606, p=0.06$). Follow-up Bonferroni-corrected comparisons indicated that L2 learners took significantly longer to respond than native controls ($p<0.005$) but not compared to Attriters ($p>0.1$). Controls and Attriters only differed marginally in their reaction times ($p=0.09$). Follow-up comparisons by condition showed that Mismatch sentences were responded to significantly slower than those to Correct sentences ($F(1,71)=14.534, p<0.0001$) and to Swap sentences ($F(1,71)=32.045, p<0.0001$). Reaction times to Swap sentences were not significantly different from Correct sentences ($F(1,71)=0.181, p>0.1$).

3.3. ERPs elicited on the target noun

Grand average ERP waveforms elicited in native speaker Controls are presented in Fig. 2. Controls showed a broadly-distributed N400 in response to both Swap and Mismatch violations, followed by a posterior P600 effect that seemed more pronounced in the Swap than the Mismatch condition.

High-proficiency Attriters (Fig. 3a) showed a similar, broadly-distributed N400 in response to Mismatch and Swap violations, but the P600 effect that followed was larger and less focal than the posterior P600 observed in Controls. Low-proficiency Attriters (Fig. 3b) showed a broad N400 in response to Mismatch sentences but a visibly reduced N400 for the Swap condition, compared to High-Proficiency Attriters and Controls. A P600 effect was elicited in response to both kinds of lexical-semantic violations, but this P600 seemed smaller and more broadly distributed (i.e., less posterior) in Low- than in High-Proficiency Attriters.

High-proficiency L2 learners (Fig. 4a) showed an N400 effect in response to both semantically anomalous sentences. On the other hand, low-proficiency L2 learners (Fig. 4b) showed a somewhat posterior N400 only in the Mismatch condition, whereas the Swap condition overlapped with the Correct condition in the N400 time-window. Overall, neither subgroup of L2 learners elicited a P600 effect.
in either violation condition. Instead, the High Proficiency L2 learners elicited a negativity in the Mismatch condition.

Fig. 5 illustrates the continuum-like pattern of ERP responses to Swap (Fig. 5a) and Mismatch (Fig. 5b) sentences by proficiency subgroup. In sum, visual inspection of ERP patterns suggests that proficiency level modulated the amplitude and scalp distribution of the N400, and more so for Swap than for Mismatch violations. The P600 effect appeared to be affected by Condition (Swap > Mismatch), by Proficiency (High > Low) but also by Group (Attriters > Controls > L2 learners), in terms of its amplitude and scalp distribution.

3.3.1. N400 between 300–550 ms

The global ANOVA in the 300–550 ms time-window for midline electrodes revealed a significant main effect of Condition (F(2,134)=25.41, p < 0.0001) and a significant Condition×Ant-Post interaction (F(6,402)=3.98, p < 0.001). Interactions with Group did not reach significance (p > 0.1), indicating that the N400 was shared between Controls, Attriters and L2 learners. However, a significant interaction between Condition×Proficiency was found (F(2,134)=3.81, p < 0.05).

No significant interactions involved both between-subject factors Proficiency×Group, suggesting that the impact of both factors was independent.

Follow-up analyses by Condition pairs were then performed to elucidate the nature of the interactions. The comparison between Mismatch vs. Correct conditions (i.e., difference wave computed for Mismatch minus Correct) yielded a significant main effect of Condition (F(1,67)=58.02, p < 0.0001) as well as a significant Condition×Ant-Post interaction (F(3,201)=5.00, p < 0.05). The interaction between Condition×Proficiency also reached significance (F(1,67)=6.49, p < 0.05), as did the three-way interaction between Condition×Ant-Post×Proficiency (F(3,201)=3.21, p < 0.05). When followed-up by proficiency subgroup, results indicated that high proficiency individuals showed a main effect of Condition (F(1,36)=49.71, p < 0.0001) and a significant Condition×Ant-Post interaction (F(3,108)=5.46, p < 0.05), reflecting the maximal amplitude of the N400 among midline electrodes at Cz and Pz in high proficiency individuals. Conversely, low proficiency individuals only showed a significant main effect of Condition (F(1,35)=17.05, p < 0.05), suggest-

---

\(^{3}\) Low Proficiency L2 learners showed an apparent positivity in the Swap condition, but the latter turned out to be non-significant and was primarily driven by two participants.

\(^{4}\) When two ERP conditions are compared, subtracting the control condition from the violation condition is a standard in the ERP literature. The rationale is that this difference would reflect the ‘additional’ cognitive processing required by the more difficult violation condition.

---

Fig. 2. ERPs elicited by the target noun in response to Mismatch violations and Swap violations compared to Correct sentences in native monolingual Controls. Time ranges (in milliseconds) on the x-axis are relative to the onset of the target noun (0 ms). Negative values are plotted up. Controls elicited a biphasic N400+P600 pattern in response to both violation conditions.
ing that the N400 effect for Mismatch sentences was more broadly distributed in low proficiency compared to high proficiency speakers.

The comparison between Swap vs. Correct conditions also revealed a main effect of Condition \((F(1,67)=30.41, p<0.0001)\) and a significant interaction between Condition\(\times\)Proficiency \((F(1,67)=7.60, p<0.01)\), which when followed-up by proficiency subgroup indicated that only high proficiency individuals showed a significant N400 effect in the Swap condition \((F(1,36)=32.86, p<0.0001)\), whereas for low proficiency individuals, Swap sentences were not significantly different from Correct control sentences in the N400 time-window \((p>0.1)\).

The comparison between Mismatch vs. Swap violation conditions yielded a marginal main effect of Condition \((F(1,67)=2.92, p=0.09)\) and a significant interaction between Condition\(\times\)Ant-Post \((F(3,201)=6.19, p<0.0005)\), reflecting a graded pattern in N400 responses at Pz \((F(1,67)=13.15, p<0.05)\) and Oz \((F(1,67)=4.69, p<0.05)\) but not Fz or Cz \((p>0.1)\).

Crucially, we compared the difference waves between Mismatch vs. Swap and Swap vs. Correct as two levels of Condition in an additional ANOVA. We expected to find a significant Condition\(\times\)Proficiency interaction, indicating that, for high-proficiency individuals, the major N400 difference lies in the Swap-Correct condition whereas Mismatch-Swap difference does not add much more (as both violations elicited large N400 effects in high-proficiency individuals). Conversely, in low-proficiency subgroups, we would expect an N400 difference only for the Mismatch vs. Swap condition and not for Swap vs. Correct. Our prediction was confirmed, although the effect was marginal \((F(1,71)=3.47, p=0.06)\).

3.3.2. P600 between 650–1000 ms

On the midline in the P600 window, a significant main effect of Condition \((F(2,134)=4.92, p<0.01)\) was qualified by a significant interaction between Condition\(\times\)Ant-Post \((F(6,402)=4.67, p<0.005)\), reflecting the largely posterior prominence of the positivity. Contrary to the preceding N400, the presence and distribution of the P600 was significantly modulated both by Proficiency \((\text{Condition}\times\text{Proficiency}: F(6,402)=2.48, p<0.05)\) and by Group \((\text{Condition}\times\text{Group}: F(2,134)=3.14, p<0.05)\).

Follow-up analyses within Groups demonstrated a main effect of Condition in Attriters \((F(1,22)=5.30, p<0.05)\), reflecting the broad positivity elicited by Mismatch sentences, and a main effect of Condition in L2 learners \((F(1,19)=5.11, p<0.05)\), reflecting the unexpected negativity observed in this time-window. Controls showed only a significant Condition\(\times\)Ant-Post interaction \((F(3,87)=10.18, p<0.0005)\), confirming the more posterior focus of the P600 in Controls \((Fz: p>0.1; Cz: p>0.1; Pz: p<0.05; Oz: p<0.005)\) compared to Attriters \((Fz: p>0.1; Cz: p<0.05; Pz: p<0.05; Oz: p<0.05)\). In addition to Group differences, the P600 effect for Mismatch vs. Correct sentences was also significantly modulated by Proficiency \((\text{Condition}\times\text{Ant-Post}\times\text{Proficiency}: F(3,201)=3.82, p<0.05)\), with high-proficiency individuals eliciting a robust and more posterior P600 \((\text{Condition}\times\text{Ant-Post}: F(3,108)=6.96, p<0.005; \text{Pz and Oz}: p<0.05)\), compared to low proficiency individuals \((\text{Condition}\times\text{Ant-Post}: F(3,105)=2.94, p=0.06)\).
The comparison between Swap vs. Correct sentences revealed a Condition main effect ($F(1,67)=9.10$, $p<0.005$) qualified by a Condition × Ant-Post interaction ($F(3,201)=8.97$, $p<0.0001$), which in turn was significantly modulated both by Group (Condition × Ant-Post × Group: $F(6,201)=2.97$, $p<0.05$) and by Proficiency (Condition × Ant-Post × Proficiency: $F(3,201)=3.11$, $p<0.03$). Follow-up analyses by Group demonstrated that Swap violations elicited a P600 effect in Controls (Condition: $F(1,29)=3.28$, $p=0.08$; Condition × Ant-Post: $F(3,87)=9.98$, $p<0.0001$) and Attriters.
Given that Controls and Attributers were the only two subgroups to show a clear P600 effect in response to lexical-semantic anomalies in our study, we conducted a lateral ANOVA with only those two groups (keeping Proiciency as a second between-subjects factor), to assess whether both groups of native Italian speakers differed significantly in the amplitude and/or scalp distribution of the P600. A Condition×Group interaction (F(4,196)=13.64, p < 0.0001) revealed that Controls and Attributers differed significantly in their P600 responses, both for Mismatch vs. Correct (Condition×Group: F(1,49)=5.31, p < 0.05) and Swap vs. Correct comparisons (Condition×Group: F(1,49)=4.38, p < 0.05). A significant five-way interaction between Condition×Laterality×Ant-Post×Proiciency×Group reflected the pattern that Attributers elicited more broadly-distributed P600 effects than Controls, and that amplitudes were larger in high vs. low proiciency individuals. Thus, the significant Group differences found in our ANOVA with all 3 groups were not simply due to L2 learners’ lack of a P600 effect, but also to Attributers eliciting larger and less focal/posterior P600 effects (for both violations) compared to non-attribting Italian native speakers.

3.4. Further investigations of proiciency subgroups

In additional analyses, we compared each proiciency subgroup of bilingual speakers to the Controls to determine how “native-like” (i.e., how similar to native monolinguial non-attribting Controls) their ERP responses were in each time-window.

Two supplementary ANOVAs were run with “Group” (3 levels) as the only between-subjects factor: (1) all Controls compared to High Proiciency Attriters and to High Proiciency L2 learners; (2) all Controls compared to Low Proiciency Attriters and to Low Proiciency L2 learners. Our aim was to confirm the findings from the global ANOVAs above (Section 3.3) that native-like-ness of N400 responses depended on proiciency level regardless of group membership (i.e., L1/L2 status), and to rule out that proiciency effects observed in our initial ANOVAs were due to the L2 learners alone.

We expected that the two High groups would be indistinguishable from Controls in the N400 time-window (and would be indistinguishable from one another), whereas the two Low groups would differ significantly from Controls, at least in the Swap condition, though not differing significantly from one another. On the P600, we expected Group differences for Controls vs. Highs as well as Controls vs. Lows comparisons, given that the L2 learners stood out from the other two groups and did not elicid P600 effects, regardless of their proiciency level. The results (below) confirmed our predictions.

3.4.1. Controls vs. High Proiciency Attriters vs. High Proiciency L2 learners

In the N400 time-window, a significant Condition main effect (F(2,100)=22.50, p < 0.0001) and Condition×Ant-Post interaction (F(6,300)=3.39, p < 0.05) reached significance in the absence of any interactions with Group (p > 0.5). The correct condition significantly differed from both Mismatch (F(1,50)=41.36, p < 0.0001; Condition×Ant-Post: F(3,150)=4.50, p < 0.01) and Swap (F(1,50)=38.67, p < 0.0001; Condition×Ant-Post: F(3,150)=3.01, p < 0.05) violations, which in turn were only marginally different from each other (Condition×Ant-Post: F(3,150)=2.68, p = 0.07). Crucially, planned comparisons revealed that the two High proiciency bilingual groups did not differ from each other in any of the N400 effects they elicited (all ps > 0.2).

In the P600 window, Group did exert an effect (F(4,10)=2.70, p < 0.05), confirming that, for both violations, the P600 effect was present only in the two L1-Italian groups (Condition×Ant-Post effects in Controls: F(6,174)=7.68, p < 0.001; Attriters: F(6,72)=2.26, p < 0.05), but not in High-Proiciency L2 learners (p > 0.1).

3.4.2. Controls vs. low proiciency attriters vs. low proiciency L2 learners

The global ANOVA revealed a marginal interaction between Condition×Group (F(4,94)=2.01, p = 0.09) which, when followed-up by Condition pairs supported our predictions that the groups differed most on the N400 effect in Swap conditions rather than Mismatch conditions. For Mismatch vs. Correct sentences, a significant Condition×Group interaction (F(2,47)=2.98, p < 0.05) reflected the large N400 response elicited in Controls (F(1,29)=39.45, p < 0.0001) relative to the weaker N400 effect in Low Proiciency Attriters (F(1,9)=6.34, p < 0.05) and Low Proiciency L2 learners (F(1,9)=4.13, p < 0.05). For Swap vs. Correct sentences, the significant Condition×Group interaction (F(2,47)=4.26, p < 0.05) confirmed thatControls were the only group to show a significant main effect of Condition (F(1,29)=23.16, p < 0.0001), while neither one of the Low Proiciency bilingual groups showed a significant N400 for the Swap condition (ps > 0.1). Importantly, planned comparisons confirmed that the two Low Proiciency bilingual groups did not differ from each other on any N400 effects (ps > 0.2).

As predicted, Group also exerted an effect in the P600 window Condition×Ant-Post×Group (F(12,282)=3.09, p < 0.005), given Low Proiciency L2 learners did not elicid P600 responses in either violation condition relative to correct sentences (ps > 0.1).

3.5. Proiciency as a continuous variable (across all groups)

3.5.1. Proiciency measures and behavioral performance

Correlational analyses indicated that Italian proiciency level was positively correlated with participants’ ratings, such that individuals with higher proiciency scores were better able to discriminate between correct sentences (more favorable ratings) and anomalous ones (less favorable ratings) than individuals with lower proiciency scores. Specifically, proiciency was positively correlated with ratings of correct control sentences (C-test: r=0.482, p < 0.0001; Error-detection test: r=0.557, p < 0.0001; Semantic fluency: r=0.44, p < 0.0001), and negatively-correlated with ratings of incorrect sentences, both for Swap (C-test: r=-0.357, p < 0.001; Error-detection test: r=-0.408, p < 0.0001) and Mismatch conditions (C-test: r=-0.363, p < 0.001; Error-detection test: r=-0.509, p < 0.0001; Semantic fluency: r=-0.450, p < 0.0001).
3.5.2. Proficiency measures and the N400

Correlational analyses revealed that the amplitude of the N400 effect elicited in response to Swap vs. Correct (difference wave) was negatively correlated with individual Italian proficiency measures, such that individuals with higher proficiency scores elicited a larger N400 difference (i.e., less positive amplitude) than individuals with lower proficiency scores (C-test: \( r = -0.306, p < 0.0001 \); Error-detection test: \( r = -0.387, p < 0.0001 \); Semantic fluency: \( r = -0.272, p < 0.01 \)). The same pattern was found for the N400 amplitude for Mismatch vs. Correct conditions (C-test: \( r = -0.332, p < 0.01 \); Error-detection test: \( r = -0.293, p < 0.01 \); Semantic fluency: \( r = -0.302, p < 0.01 \)). Thus, the higher the Italian proficiency scores, the stronger the violation effect in response to lexical-grammatically anomalous sentences.

It is important to note that since the groups were not completely matched on proficiency, the factor Group is partly confounded with proficiency. However, the factor Proficiency is not confounded with factor Group. An illustration that proficiency effects are not simply due to group effects on the N400 is that the correlations reported above remain significant with group partialed out (except for the semantic fluency measure). With group controlled for, individuals with higher proficiency scores elicited larger N400 amplitudes for each violation condition relative to the correct condition (i.e., Swap minus Correct and Mismatch minus Correct).

3.5.3. Proficiency measures and the P600

The posterior P600 elicited by Swap violations had a larger amplitude in participants with higher scores on the Error-detection (\( r = 0.259, p < 0.05 \)) and on the Semantic fluency task (\( r = 0.285, p < 0.05 \)), whereas the C-test was not a significant predictor of the P600 following Swap errors (\( r > 0.1 \)). In contrast, all three individual tests were positively correlated with P600 amplitudes elicited in response to Mismatch errors (C-test: \( r = 0.321, p < 0.01 \); Error-detection test: \( r = 0.390, p < 0.01 \); Semantic fluency: \( r = 0.281, p < 0.01 \)), such that individuals with higher proficiency scores elicited a larger P600 effect following Mismatch violations.

3.6. The effect of background factors in L1 attrition and L2 processing

Although, on average, age at testing was superior in Attriters, there were no modulations of ERP responses by age of testing or level of education (all \( p > 0.1 \)).

In Attriters, length of residence (LoR) was significantly correlated with the N400 amplitude in Mismatch violations relative to Correct sentences (Fz: \( r = 0.356, p < 0.05 \); Cz: \( r = 0.369, p < 0.05 \)), indicating that individuals who spent a greater number of years outside of Italy elicited weaker N400 effects in response to Mismatch violations, and therefore were less native-like in Italian. Hours of L1-Italian exposure/use were also negatively correlated with N400 amplitudes in response to Mismatch violations (Fz: \( r = -0.397, p < 0.05 \), such that Attriters with more exposure to Italian elicited a larger N400 effect than Attriters with less exposure to Italian.

In L2 learners, correlations between ERP responses and age-of-acquisition of Italian (AoA) or amount of L2 exposure did not reach significance (\( p > 0.1 \)).

4. Discussion

The present study examined the neurocognitive correlates of real-time lexical-semantic processing in adult L1 attriters and late L2 learners of Italian compared to non-attriting Italian monolingual speakers. We investigated whether L1 attriters and/or late L2 learners showed evidence of non-native-like patterns of sentence processing, particularly for pairs of words that are notoriously “confusable” in Italian. A second aim was to examine whether ERP responses were predicted by proficiency in the target language (Italian), with processing profiles falling along a continuum of native-like-ness (from high to low proficiency levels), rather than resulting in qualitative processing differences depending on whether Italian was the speakers’ L1 or L2. Finally, we explored whether group and/or proficiency differences were also observed at later stages of lexical-semantic processing, as indexed by the P600.

4.1. N400 differences predicted by proficiency level

Our findings revealed that proficiency level but not group membership (i.e., L1/L2 status) modulated the amplitude and scalp distribution of the N400 (350–500 ms) in response to lexical-semantic violations relative to correct control sentences. When comparing Mismatch vs. Correct sentences, high-proficiency individuals (from all three groups) elicited a large N400 effect which was maximal at Cz and Pz, whereas low-proficiency individuals elicited a weaker and more broadly-distributed N400 effect. Proficiency effects were even more robust for Swap sentences (i.e., the condition involving the confusable words) – only high-proficiency individuals showed an N400 effect for Swap errors, whereas in lower-proficiency subgroups, these sentences were statistically indistinguishable from Correct sentences. Proficiency differences therefore modulated the graded response pattern (N400 for Mismatch > Swap > Correct) between conditions – in high-proficiency subgroups, the major N400 difference was between Swap vs. Correct conditions, whereas the two violation conditions did not significantly differ from one another. This pattern differed in low-proficiency subgroups, however, who showed an N400 difference for the Mismatch vs. Swap comparison but not for Swap vs. Correct.

Proficiency effects in the N400 time-window in the absence of group effects were further supported by analyses in which we contrasted Controls with each of the two high-proficiency or low-proficiency subgroups. Controls, high-proficiency Attriters as well as high-proficiency L2 learners elicited large N400 effects in response to both kinds of violation conditions, which in turn were not statistically different from one another. Conversely, group differences existed between Controls, low-proficiency Attriters and lower-proficiency L2 learners, with the low-proficiency subgroups failing to show an N400 effect in response to Swap violations. Crucially, we showed that L1 attriters and L2 learners were indistinguishable in their N400 responses when they belonged to the same Italian proficiency cluster.

Some previous studies had shown that sentences where the target noun had been replaced by an orthographic or phonological neighbor elicited reduced N400 effects compared to non-neighbor sentence conditions, even though both types of violation sentences were semantically implausible to the same degree (Laszlo and Federmeier, 2009). In line with such findings, we may have expected a graded N400 response pattern even in native controls, where the N400 to Swap violations would have been significantly reduced relative to Mismatch sentences, given that the presented target word in Swap sentences only differed from the expected word by its final vowel. Such graded N400 responses have been argued to reflect co-activation during lexical access, where the presented target word and related word-forms are activated in parallel (Holcomb et al., 2002; Laszlo and Federmeier, 2011). However, in our study, the N400 difference between Mismatch and Swap violations was only marginal in native controls and high-proficiency subgroups of L1 attriters and L2 learners. A possible reason for this differential finding is that our study included a range of high and low cloze-probability items. In more highly-constraining contexts, for example, readers may make stronger predictions for the target-word, therefore activating the target-word before its actual presentation (Federmeier et al., 2007; Foucart et al., 2014; Martin et al., 2013; Wicha et al., 2004). It would be interesting to investigate the possible impact of cloze probability on the N400 differences in our study.\footnote{Cloze probability differences between conditions (and potentially between groups and/or proficiency levels) on N400 and P600 effects will be addressed in a separate paper, so as not to detract from the main goals and findings of the present study.}
elicited a large N400 in response to both violations relative to the support our view of a continuum of native-like-ness. These L2 learners speakers, there was also overlap between the groups, such that several processing. Our semantics do not pose a challenge for late L2 learners. However, consistent with the posterior late-positive-complex reported by Laszlo 2004; Osterhout et al., 2006; White et al., 2012) and in L2 attrition (Pitkänen, 2010). Our data extend this finding to lexical-semantic processing in a group of L1 attriters, and reveal proficiency effects on the N400 across groups, rather than differences in processing determined by L1-L2 status (i.e., AoA).

4.2. A conflict-monitoring account of the P600 in lexical-semantic processing

Our finding of P600 effects following N400 responses to lexical-semantic violations was in line with a number of previous studies, suggesting that neurocognitively demanding revision of the input triggers a P600 effect even in lexical-semantic violations as opposed to strictly morphosyntactic ones (e.g., Diaz and Swaab, 2007; Kuperberg, 2007; Laszlo and Federman, 2009; van den Brink et al., 2001; Van de Meerdendonk et al., 2009; see Van Petten and Luka, 2012 for a review). The presence, amplitude and scalp distribution of the P600 was influenced by the type of violation, proficiency level and by group membership. These effects will be discussed in term, and interpreted within the conflict monitoring framework of the P600 (Van de Meerdendonk et al., 2009; Vissers et al., 2006; Vissers et al., 2008).

A P600 effect for Mismatch sentences relative to Correct sentences was found in both native speaker groups, but the P600 response in Attriters was larger in amplitude and less focal/posterior than the P600 in non-attribting Controls. Our group of L2 learners, however, did not show any indication of a P600 effect. In addition to group differences, both the amplitude and the scalp distribution of the P600 was additionally modulated by proficiency level, such that individuals with higher Italian proficiency scores showed more robust and more posterior P600 responses for Mismatch violations compared to lower-proficiency individuals. The P600 effect elicited in response to Swap violations was similarly influenced by Group and Proficiency, but its amplitude was significantly larger than that of the P600 effect for Mismatch sentences. Only the two L1-groups showed P600 effects for Swap vs. Correct sentences, whereas L2 learners elicited no significant effects in the P600 time-window. When L2 learners were removed from the analysis, Attriters were found to differ significantly from non-attribting native speakers by eliciting larger and less focal (more broadly distributed) P600 responses to Swap sentences, thus echoing the pattern observed for Mismatch violations.

The condition effect on P600 amplitude (Swap > Mismatch) was consistent with the posterior late-positive-complex reported by Laszlo and Federman (2009) in response to sentences where the final-noun was a semantically incongruent orthographic neighbor (relative to violations with orthographically-unrelated nouns). The authors discussed the effect as reflecting the explicit realization that the target-word was orthographically close to the expected word. One could argue that if the larger P600 merely reflects the “realization” of the orthographic similarity between the intended word and the presented word, late L2 learners in our study do not elicited a large P600 effect in response to Swap sentences because they did not recognize that the incongruent target-noun was only one vowel away from the noun that was intended to fit the sentence context. However, this would seem at odds with our finding that at least high-proficiency L2 learners were able to detect the erroneous insertion of an orthographic neighbor in the Swap condition early on, eliciting a similar N400 effect as high-proficiency native speakers.

Posterior P600 effects in response to both lexical-semantic violations have been discussed as reflecting a general conflict monitoring process. When a strong conflict is detected, the reader must engage in a
general (as opposed to purely syntactic) re-analysis of the input in order to determine where processing errors may have occurred (e.g., Van de Meerendonk et al., 2009; Vissers et al., 2006; Vissers et al., 2008). The more severe the conflict, the larger the P600 response. In our study, Mismatch and Swap sentences were equally semantically implausible, but only the Swap condition involved an additional conflict at the word-form level, as the target-noun differed in its (gender-marking) final vowel. The enhanced conflict caused by semantic anomaly+orthographic ambiguity may have triggered more elaborated re-analysis processes to double-check the preceding input relative to Mismatch violations where the only conflict was on a semantic level.

The proficiency and group differences on P600 amplitudes would also seem to fit this theoretical framework. A number of studies have shown that highly-proficient speakers engage in re-analysis/repair processes as indexed by the P600, whereas low-proficiency learners may not elicit P600 responses during early stages of learning (e.g., Osterhout et al., 2006; White et al., 2012). It is conceivable that P600 effects elicited in response to costly lexical-semantic anomalies requiring re-analysis may also be modulated by proficiency, insofar as there may be some degree of overlap between a general conflict monitoring process at play in both lexical-semantics and morphosyntax. Our correlational analyses revealed that the amplitude of the posterior P600 elicited in response to both Swap and Mismatch violations was larger in individuals with higher scores on the various proficiency measures. Given that, as a group, the L2 learners had the lowest averages on proficiency tasks, it remains to be seen in a follow-up study whether their proficiency level was still not high enough for native-like P600 responses to emerge, or whether this remains a qualitative difference from speakers processing their L1.

However, that the largest P600 responses were elicited in L1 attriters – even compared to monolingual Controls – suggests there is more to the story than proficiency variation. It is conceivable that the larger P600 effects for L1 attriters indicate increased conflict-monitoring and a more explicit “second-thought” process, especially for Swap violations, where a re-analysis of the input is necessary to determine “It should have been X”. This hypothesis of a more conscious “double-take” during real-time processing for Attriters is worth investigating further. Given that Attriters are a special group of bilinguals who are aware of their special circumstances and of the change they have been experiencing in their L1 since immigration, it is likely that increased attention and a more explicit metalinguistic analysis of the sentences during the experiment might play a role in this group. It would be interesting to examine whether their motivation to perform well in their L1 may be comparable or stronger than L2 learners’ motivation to perform well in a second language.

The finding that high-proficiency Attriters and high-proficiency L2 learners were indistinguishable from each other – and from native monolinguals – in the N400 time-window but diverged in their processing mechanisms further downstream emphasizes that simply focusing on the N400 effect in lexical-semantic processing may only provide part of the picture.

4.3. Conclusions and future directions

The present study is the first to examine L1 attrition in lexical-semantic processing in sentence comprehension, and the first to reveal non-native ERP responses in adult attriters compared to non-attriting native speakers. We found that lower-proficiency Attriters differed from non-attriting Controls on their N400 responses, and that Attriters as a group differed from Controls on their P600 responses. These differences were not simply due to comparing bilingual Attriters to monolingual native speakers of Italian, given that, even within the bilingual group of Attriters, factors such as length of residence and amount of L1 exposure (relative to L2 exposure) modulated how native-like their ERP response patterns were.

Interestingly, despite these processing differences, Attriters and Controls did not differ significantly on their acceptability judgment ratings or other behavioral tasks, although Attriters unanimously reported having noticed changes in some aspects of their behavior. It has been shown that attrition effects can be found in some tasks but not in others (Schmid, 2011). Longitudinal with L2 learners have also revealed differences at the neurocognitive level before they appear in behavior (e.g., McLaughlin et al., 2004; Pitkanen, 2010). The relationship between attrition effects in online processing vs. linguistic behavior therefore remains an open question.

The finding of processing differences in attriters who grew up in an exclusively monolingual L1 context until adulthood suggests that L1 response patterns are amenable to change, even later in life, and that early exposure to an L1 is not sufficient to guarantee native-like ERP responses to a language. This speaks directly to claims that L1 attrition – like L2 acquisition – is subject to maturational constraints due to an “entrenched” or stabilized L1 around puberty (Bylund, 2008; see reviews by Köpke and Schmid (2004) and Pallier (2007)). Changes in L1 processing patterns (that were not found to be modulated by age-of-immigration) therefore lend support to views of ongoing neuroplasticity for language in adulthood and would suggest that an adult speaker’s native language is not as stable and as “entrenched” as is widely assumed. Moreover, we found that Attriters’ ERP responses were modulated by their length of residence (LoR) in the L2 environment, as well as amount of current L1 exposure; Attriters who spent a greater number of years outside of Italy elicited weaker N400 responses to Mismatch violations, as did Attriters who were less exposed to their L1-Italian since immigration. Thus, the longer Attriters spent immersed in the L2 environment and the less L1 exposure they maintained, the less native-like they were in L1 lexical-semantic processing.

With respect to L2 processing in late learners, we showed that proficiency was the critical determinant of N400 responses in Italian L2 learners, and that our L2 learners’ ERP responses to these experimental stimuli were not modulated by AoA. It remains to be seen whether processing similarities between L1 attriters and L2 learners may also be found in complex areas of morphosyntax (see Kasparian, 2015), where it has been argued that neuroplasticity in adulthood is more limited than in the lexical-semantic domain (Newport et al., 2001; Wartenburger et al., 2003; Weber-Fox and Neville, 1996). Similarly, it remains an open question whether P600 responses indexing stages of input-checking and re-analysis in lexical-semantic processing may become native-like at even higher proficiency levels in late L2 learners, or whether these responses depend on having acquired the target language implicitly (L1 groups only) or on more general characteristics such as attention and motivation (largest in Attriters). It would also be of interest to test L2 learners who are immersed in Italian and who acquired their L2 in the same implicit way as the Attriters, to clarify the impact of immersion on native-like processing.

The direct comparison of L1 attriters and L2 learners in our study was a necessary first step towards exploring the impact of proficiency, exposure and dominance on both L1 and L2 processing. Further work is also warranted to determine how factors such as sentence constraint and word-frequency may interact with proficiency and exposure in determining the degree of native-like-ness of N400 responses, and whether these factors may modulate processing in a similar or different way in L1 attriters vs. L2 learners. Studies have shown that the selectivity of lexical access is influenced by the preceding sentence context (Duyck et al., 2007; Van Hell and De Groot, 2008; Eleton-Göttler, 2000) and that individual differences exist in how efficiently contextual cues can be used in real-time lexical-semantic processing (e.g., Foucart et al., 2014; Kaan, 2014; Martín et al., 2013). It would be of interest to further investigate whether L1 attriters and L2 learners are similarly efficient in predicting upcoming words in a sentence based on contextual cues, and whether such predictive processes are modulated by proficiency or amount of exposure, regardless of whether the target language is the L1 or L2.

Analyzing the impact of sentence constraint would also allow us to
examine whether condition differences exist between Swap and Mismatch for high-constraint sentences where the target word is highly predictable but substituted by its neighbor. Such sentences may show a reduction in the N400 window but a processing cost further downstream (larger P600?) and these effects may in turn be modulated by group membership and/or proficiency. Similarly, frequency modulates competition during lexical access in bilinguals (Dijkstra et al., 1998; Kerkhofs et al., 2006), and such frequency effects have been accounted for by language proficiency (Diestepaal et al., 2013) or language exposure/dominance (Whitford and Titone, 2012). Given the attention that frequency has received in behavioral investigations and theoretical models of L1 attrition (e.g., Paradis, 2007; see review by Schmid and Kopke (2004)), examining frequency effects during online lexical access and integration in sentence contexts would be an important extension of this work.

Exploring these avenues and recruiting more highly-proficient late L2 learners, as well as a group of Italian-English bilinguals who continue to be dominant in their L1-Italian, are important next steps to pursue to move towards a unified view of multilingual language processing and neuroplasticity.

Acknowledgements

K.K. was supported by a Vanier Canada Graduate Scholarship from the Canadian Institutes of Health Research, a Richard Tomlinson Doctoral Fellowship awarded by the Faculty of Medicine of McGill University, and 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 # 201876), the Natural Sciences and Engineering Research Council of Canada (NSERC; RGPIN 312835-05 and RPIN 402678-11), the Social Sciences and Humanities Research Council of Canada (SSHR; # 435-2013-2052), and the Fonds de Recherche Société et Culture, Québec (FQRSC; # 2010-SE-107372). The authors sincerely thank Francesco Vespignani for his collaboration and support during K.K.’s research stay in Italy, as well as John Drury, Eleonora Rossi and Francesca Postiglione for valuable feedback on our experimental stimuli. We are also indebted to research assistants Linna Jin, Kristina Maiorino, Filippo Vicari and Paolo Zandomeneghi for their invaluable help with data collection.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neuropsychologia.2016.10.007.

References

and L2-acquisition


Kotz, S.A., Elston-Güttler, K., 2004. The role of pro...


