CrossMark

Applied Psycholinguistics **36** (2015), 613–637
doi:10.1017/S0142716413000362

Facilitating the task for second language processing research: A comparison of two testing paradigms

A. KATE MILLER
Indiana University–Purdue University Indianapolis

Received: August 22, 2012 Accepted for publication: February 24, 2013

ADDRESS FOR CORRESPONDENCE
Kate Miller, Department of World Languages and Cultures, Indiana University–Purdue University Indianapolis, 539E Cavanaugh Hall, 425 University Boulevard, Indianapolis, IN 46202.
E-mail: am27@iupui.edu

ABSTRACT
This study considers the effects of experimental task demands in research on second language sentence processing. Advanced learners and native speakers of French were presented with the same experimental sentences in two different tasks designed to probe for evidence of trace reactivation during processing: cross-modal priming (Nicol & Swinney, 1989) and probe classification during reading (Dekydtspotter, Miller, Schaefer, Chang, & Kim, 2010). Although the second language learners produced nontargetlike results on the cross-modal priming task, the probe classification during reading task revealed results suggestive of trace reactivation, which point to detailed structural representations during online sentence processing. The implications for current theories of second language sentence processing and for future research in this domain are discussed.

A current and ongoing debate centered on the nature of second language (L2) processing has emerged in recent years because much research has investigated the nature of the representations computed during online sentence processing in a L2. As noted by Bley-Vroman (2009), investigating online sentence processing routines in a L2 can provide further insight into the L2 acquisition process and how or why it differs from first language (L1) acquisition. The possibility that learners might not be able to process input sentences in their L2 by using the same reflexes employed by native speakers (NSs) has emerged as a new possible explanation for the source of the apparent differences between L1 and L2 grammars: if L2 learners are processing the input in nonnativelike ways, this could prevent their interlanguage grammars from converging on those developed by NSs during L1 acquisition.

A large body of research has argued that the same domain-specific mechanisms that underlie L1 processing are also used in L2 processing (e.g., Frenck-Mestre, 2002, 2004; Frenck-Mestre & Pynte, 1997; Hoover & Dwivedi, 1998; Juffs, 1998,

© Cambridge University Press 2013 0142-7164/13 $15.00

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

2005; Juffs & Harrington, 1995; Williams, Möbius, & Kim, 2001). From this perspective, although there may be quantitative differences (see, e.g., Frenck-Mestre, 2002, for a discussion of processing speed in L2), there are essentially no qualitative differences between L1 and L2 sentence processing. Provided sufficient L2 proficiency, the same factors that have been shown to influence processing in a L1 similarly affect processing in a L2. For example, L2 learners have demonstrated use of the active filler strategy (Clifton & Frazier, 1989) to structure input quickly (e.g., Williams et al., 2001), sensitivity to structural ambiguities (e.g., Hoover & Dwivedi, 1998), and use of lexical subcategorization information to resolve such ambiguities (e.g., Frenck-Mestre & Pynte, 1997). Thus, proficient L2 learners have demonstrated processing reflexes similar to those that have been observed in NSs, and any perceived differences between NS and L2 results on certain online processing tasks may simply be due to other mitigating factors, such as variations in reading strategies or reduced working memory (WM) resources, without the need to assume a fundamental difference between the processing routines of NSs and L2 learners (see Dekydtspotter, Schwartz, & Sprouse, 2006).

However, the nontargetlike results of L2 learners on experimental sentence processing tasks have been interpreted by some researchers as evidence for a more fundamental difference in sentence processing in L1 versus L2 (e.g., Felser & Roberts, 2007; Felser, Roberts, Gross, & Marinis, 2003; Hahne, 2001; Hahne & Friederici, 2001; Love, Maas, & Swinney, 2003; Marinis, Roberts, Felser, & Clahsen, 2005; Weber-Fox & Neville, 1996). Despite similar performance of L2 learners and NSs in the processing of lexical semantics and local syntactic dependencies such as agreement relationships (i.e., gender concord or subject–verb agreement), it has been argued that L2 learners process complex nonlocal dependencies (e.g., relative clauses, in which a noun phrase appears nonadjacent to its licensing verb) differently than do NSs (e.g., Felser & Clahsen, 2009). Clahsen and Felser (2006b) pointed out that language learners, both L2 learners and children acquiring their L1, may have difficulty integrating several different information sources (i.e., lexical, contextual, prosodic, etc.) in real time owing to limited cognitive resources. However, Clahsen and Felser argued that there is evidence that these two types of learners differ in their strategies for resolving structural ambiguities and long-distance dependencies: whereas children tend to focus on structural information while ignoring lexical–semantic information (Felser, Marinis, & Clahsen, 2003), L2 learners seem to exhibit greater reliance on this nonstructural information (Felser, Clahsen, & Münte, 2003; Papadopoulou & Clahsen, 2003).

According to Clahsen and Felser's (2006a, 2006b) shallow structure hypothesis, the difference between L1 and L2 processing can be attributed to the difficulty that even highly proficient L2 learners have in computing complex syntactic representations in real time, such that they instead tend to construct shallow representations of the input. Clahsen and Felser (2006b) noted that the basic parsing mechanisms are universal (and therefore do not need to be learned for a L2), whereas grammatical rules are specific to each language. Thus, although a structural processing route should be available to L2 learners in principle, owing to a purported deficiency in the L2 grammar (even at advanced proficiency levels), detailed syntactic information is unavailable for the processing of complex structures in real time.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

As a result, L2 learners will overrely on nonstructural cues and use contextual and semantic information to assign meaning to the input. Note that L2 learners and NSs ultimately converge on their interpretations; the difference lies in the way that the input is structured and the processing route that leads to this interpretation.

TRACES AND CROSS-MODAL PRIMING

The processing limitation proposed by Clahsen and Felser (2006a, 2006b) is particularly relevant to the way in which L2 learners are predicted to process filler-gap dependencies (nonlocal dependency structures in which a verbal argument has moved from its canonical position to an earlier sentential position at some distance from the verb). In a sentence as in (1), the relativizer *who* (the filler) has been moved up in the structure but leaves behind a silent copy (i.e., a trace) in the canonical object position (the gap). The trace exhibits the same features as the moved expression, which is co-indexed with *the boy*, forming a referential chain.

(1) *The policeman saw* [*the* [*boy*$_i$ [*who*$_i$ *the crowd at the party accused* <who$_i$> *of the crime*.]]]

A detailed parse of the sentence in (1) would yield a fully specified hierarchical representation that includes the trace, which can be used to link the antecedent of the *wh*-filler to its thematic position. For the L2 learner parsing shallowly, however, the sentence would be segmented into chunks, and semantic knowledge of the argument structure of the transitive verb *accused* would be used to assign *the boy* its thematic role with respect to this verb. Because meaning is derived through segmentation and semantic association, the representation lacks hierarchical structure and syntactic gaps.

The presence of traces as structural reflexes of sentence processing may be revealed through studies that use cross-modal priming (Nicol & Swinney, 1989). In these experiments, participants seated at a computer are asked to make simple categorization decisions about (picture) probes presented visually on the monitor at specific moments, while at the same time listening to aurally presented sentences for comprehension (which is verified through questions related to the content of the sentence). Priming effects occur when the antecedent of the *wh*-filler, which has been temporarily stored in memory, is mentally reactivated at the hypothesized gap site for integration into the structure, thus facilitating responses to probes that are semantically related or identical to this antecedent. If encountering the trace triggers reactivation of the antecedent, response times (RTs) to matching or related probes at the gap position should be faster than those to unrelated probes appearing in this same position and faster than those to related probes presented at other moments during processing.

The cross-modal priming methodology has also been used to examine sentence processing in a L2. The shallow structure hypothesis is based largely on the results of L2 studies that have used tasks that were originally designed to test L1 sentence processing. The argument for shallow L2 processing based on different NS and learner results thus assumes that all else (phonological decoding, automaticity of lexical access, etc.) is equal, which is almost certainly not the case. As

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

Dekydtspotter et al. (2006) pointed out, direct comparison of NS and learner results on such tasks may lead to erroneous conclusions: given potential differences in (a) speed of processing, (b) prosodic contours imposed during reading, (c) speed and automaticity of lexical access, and (d) heteromorphy of L1–L2 semantic fields, there can be no guarantee that, at a planned target segment, the response patterns of NSs and L2 learners will reflect the same moments or stages in the parsing process. This argument applies to both reading and listening tasks.

Felser and Clahsen (2009) consider the use of cross-modal priming preferable to other possible methodologies such as self-paced reading, because it allows for direct comparison with young NS children (another type of language learner) who have not yet developed reading proficiency in their L1. However, the cross-modal priming paradigm may not be ideal for testing L2 learners. As Dekydtspotter et al. (2010) pointed out, listening to decontextualized speech without recourse to visual cues to aid in comprehension creates an additional computational burden on the L2 learner's processing resources. This same argument could also be applied to self-paced listening tasks, although perhaps to a lesser extent, given that the participant is able to control the speed of presentation, which is not possible in a cross-modal priming task. In addition, because classroom L2 learners generally receive much written input, reading proficiency usually develops simultaneously with, or even prior to, the emergence of oral proficiency. Thus, presenting experimental sentences visually might be more useful for the investigation of L2 processing of complex structures, assuming that the visual presentation will help to lessen the computational load. Both self-paced reading and eye-tracking measures have proven to be accurate tools for assessing online computations among NSs and, although perhaps to a lesser extent, in L2 learners (e.g., for a review, see Frenck-Mestre, 2005).

Furthermore, nonlocal dependencies of the type most often cited in discussions of shallow L2 processing appear to be especially taxing in that the referent of the antecedent of a displaced filler must be retained in memory until it can be integrated into the structure and must thus compete for resources as other referents are introduced (e.g., Gibson, 1998). This difficulty is not unique to processing in a L2; evidence compatible with trace reactivation as an indicator of structurally based processing has been found only for NS adults and children with high WM capacities (e.g., Roberts, Marinis, Felser, & Clahsen, 2007). Thus, as Indefrey (2006) points out, it is very likely that the experimental power of testing materials varies as a function of computational complexity for NSs and L2 learners alike.

PROCESSING FILLER-GAP DEPENDENCIES IN L1 AND L2

Reactivation effects, as evidenced by priming, have been argued to demonstrate that traces are part of the representations computed during online L1 sentence processing. For example, Love (2007) tested for trace reactivation during sentence processing by English NS children (ages 4–6) and college-aged young adults, using sentences as in (2).

(2) a. *The zebra that the* [#1] *hippo had kissed* [#2] *on the nose ran far away.*
b. *The camel that the* [#1] *rhino had kissed* [#2] *on the nose ran far away.*

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

The two experimental positions (gap, indicated by #2, or task, #1) were tested in subsequent experiments, which used a switched-targets design such that related probes (i.e., those that matched the filler) and unrelated probes (those matching the embedded subject) were counterbalanced and thus served as both experimental and task probes throughout the experiment. For example, a picture of a zebra served as a related probe in Sentence (2a) and as a control probe in (2b). Participants listened to the sentences for comprehension (which was periodically verified through simple questions) and had the additional task of indicating, by pressing a button as quickly as possible, whether images appearing on the computer screen depicted something that was edible or inedible. The main experimental items all featured pictures of animals, and the filler sentences involved images of common edible items as well as inedible inanimate objects. Participants were trained to a 100% accuracy level in the classification of images before completing the task. Both adult and child NS participants produced faster RTs for antecedent-matching probes that appeared concurrently with the gap position.

However, given that the trace immediately followed the verb in Love's (2007) sentences, the reactivation effects found at this test point could have been induced either by the syntactic gap or by the verb itself. In other words, it is not clear whether the presence of the trace triggered reactivation of the filler for integration into the structure or whether participants used their thematic knowledge of the verb to integrate the filler. These results are thus compatible with either full or shallow processing. Roberts et al. (2007) attempted to differentiate between verb- and trace-driven filler integration among English NS children (ages 5–7) and adults on a cross-modal priming task that used sentences involving indirect object relative clauses as in (3), where the structural gap is separated from the verb by the direct object.

(3) *Fred chased the squirrel to which the nice monkey explained the game's* [#1] *difficult rules* [#2] *in the class last Wednesday*.

While listening to these sentences, either at the gap position (#2) or in a control position (#1) that occurred approximately 500 ms earlier, participants pressed a button to classify picture probes as alive or not alive. Identical probes matched the antecedent (in this case, *squirrel*), whereas unrelated probes depicted common inanimate objects (for this example, the unrelated probe was a picture of a toothbrush). Periodic comprehension questions were used to verify that participants were actively paying attention to the content of the sentences. All participants completed a WM reading span test in addition to the main experimental task. The results of the cross-modal priming task revealed reactivation effects: RTs for the identical probes in the gap position were faster than those for unrelated probes at the same position as well as those to identical probes in the control position. However, only the children and adults with high WM capacity demonstrated these priming effects. The low WM adults produced relatively flat RTs, with no significant differences across the different conditions. The low WM children, in contrast, exhibited slower RTs to the identical probes.

Thus, experimental evidence from L1 studies is compatible with structural sentence processing routines that use traces of moved expressions to mediate

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

filler-gap dependencies. The picture for L2 sentence processing is a bit less clear. Recent studies conducted within the cross-modal priming paradigm have argued that L2 learners differ from both adult monolingual (Love et al., 2003) and child NSs (Felser & Roberts, 2007) in that NSs show significant priming effects suggestive of trace reactivation during sentence processing, whereas L2 learners do not.

Felser and Roberts (2007) tested adult advanced L1 Greek learners of L2 English on the same stimulus used by Roberts et al. (2007) and compared their learner results against those of each of the four subgroups of NSs from that study: low WM adults, high WM adults, low WM children, and high WM children. Note, as Felser and Roberts pointed out, that the structure of indirect object relative clauses such as those tested is very similar in English and Greek (the participants' L1). The L2 learners exhibited faster RTs for identical versus unrelated probes in both the gap and the control positions, which Felser and Roberts took to be evidence for maintained activation of the antecedent throughout sentence processing, as opposed to reactivation of the antecedent triggered by the syntactic gap. Furthermore, no effect of WM capacity was found among these L2 learners. Felser and Roberts pointed out that their L2 learners behaved differently from all four subgroups of NSs in Roberts et al.'s study: both the high WM NS adults and children showed a position-specific advantage for identical probes at the gap position only, whereas the L2 learners demonstrated this advantage in both the gap and the control positions. The low WM NS adults exhibited no significant differences in their RTs to identical versus unrelated probes in either position, while the low WM children actually produced longer RTs to identical targets in both positions (but much more so at the gap position). Felser and Roberts interpreted the finding that the L2 learners behaved differently from all NS subgroups and that they did not even seem to be able to process in nativelike ways a structure that could easily be transferred from the L1 as an indication of a more fundamental difference between L1 and L2 sentence processing. Specifically, because the L2 learners showed the same overall RT asymmetry regardless of position, which meant there was no effect of structure, Felser and Roberts (2007) concluded that these advanced learners were computing shallow representations of the L2 input.

However, it may seem a bit odd that these L2 learners would not show any such effect of structure whatsoever. Sag and Fodor (1994) argued that although reactivation effects are certainly consistent with a trace-mediated sentence processing routine, such effects can really only indicate that the human sentence processing mechanism has detected a gap (i.e., recognized that a constituent is missing) at that point during processing but not that the gap contains any kind of trace. Thus, even learners processing shallowly could still expect to encounter the indirect object of a ditransitive verb such as *give* in its canonical position following the direct object and would recognize the gap that the moved filler had left behind, even if their L2 representation did not contain a trace. Reactivation effects would still be expected to occur. The absence of any such effects in Felser and Roberts (2007) thus seems to suggest that there may have been some other factor (not related to the ability to access syntactic information during L2 sentence processing) that affected the learners' RTs.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

As noted by Nicol, Fodor, and Swinney (1994), there may be any number of factors that can affect response patterns in such studies, regardless of whether syntactic representations are computed. The predicted response pattern of the cross-modal priming methodology relies on two major assumptions: first, that the integration of a fronted *wh*-filler into the sentence structure is mediated by encountering the trace (see, e.g., Pickering & Barry, 1991; Sag & Fodor, 1994; Traxler & Pickering, 1996, for arguments that filler integration is mediated by the verb's thematic structure); and second, that RTs to matching or related probes will reflect facilitation priming effects (see Dekydtspotter & Miller, 2013, for a discussion of inhibition antipriming effects). Perhaps more important, however, is the crucial assumption that there will be an exact overlap between the trace-induced reactivation of the filler and the facilitation exhibited in the shorter RTs. Nicol et al. (1994) discussed the limited window of opportunity that this methodology affords for capturing the relevant effects, even in L1 processing experiments. Omaki and Schulz (2011) cited a study by McKoon, Ratcliff, and Ward (1994) as evidence that cross-modal priming effects are not consistently replicable even across L1 populations and pointed out that it may be difficult to rely on this methodology to postulate a model of real-time sentence processing. Furthermore, as noted by Dekydtspotter et al. (2006), there is no guarantee that L2 learners' response patterns for a planned target segment will reflect the relevant processing moments.

Returning to Felser and Roberts (2007), it is possible that their L2 learner results were simply due to a task effect: in the sentences used by both Roberts et al. (2007) and Felser and Roberts (2007), all of the identical/alive probes matched referents in the sentences and the unrelated/not alive targets came out of the blue, so to speak, having little or nothing to do with the auditory stimuli that participants were listening to. The shorter RTs to identical probes found across the board in the L2 learner data could thus be because the animals depicted by the identical targets had just been encountered in the sentence (and the semantic features activated in conceptual structure), whereas the unrelated targets had not.

TESTING L2 SENTENCE PROCESSING

The question thus remains as to how to best test sentence processing in a L2. Although it has been argued that cross-modal priming allows for direct comparison of adult L2 learners with young NS children (Felser & Clahsen, 2009), given the high cognitive demands of processing filler-gap dependencies, the cross-modal priming task may not be ideal for testing L2 learners. Classroom L2 learners generally receive much written input, and thus reading proficiency usually develops simultaneously with or even prior to oral proficiency. Dekydtspotter et al. (2010) designed a probe classification during reading task that combined the classification of picture probes with segmented forced-paced reading aloud. Participants are asked to read sentences aloud to themselves in a low voice, segment by segment, and to classify intervening probes into two categories. This methodology is argued to be preferable to self-paced reading in that participants do not have the opportunity to pause or to reflect on the structure as they read but instead must try to keep up as each segment appears on the screen. Whereas cross-modal priming

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

studies rely on comprehension questions to ensure that participants are paying attention (essentially making it a listening comprehension exercise and adding to the complexity of the task), having participants read the sentences aloud offers a simple alternative means of focusing attention on the experimental stimulus. Thus, some of the processing burden created by cross-modal priming tasks may be alleviated, which allows for testing at lower proficiency levels.

A similar unimodal priming task has been explored in the context of L1 processing (e.g., McKoon, Albritton, & Ratcliff, 1996; McKoon & Ratcliff, 1994), although these studies paired forced-paced silent reading with a lexical decision task: the visually presented sentences were interrupted by a string of letters to be classified as a word or nonword. The probe words were slightly offset from the center of the screen and were followed by two asterisks to differentiate them from the rest of the sentence. However, using this methodology, Nicol, Swinney, Love, and Hald (2006) showed that participants exhibited faster RTs when the probe word was compatible with the immediate context of the sentence and could be easily integrated into the structure (e.g., following the word *folded*, the probe word *newspaper* elicited faster RTs as compared with *technique*). These results suggest that this unimodal presentation may be more susceptible to congruency effects than in the cross-modal paradigm: the interrupting probe may be integrated as part of the sentence structure. This finding is potentially problematic because it could conceivably create an illusion of priming or mask any priming effects that may emerge. In cross-modal priming, in contrast, there is no interruption to the auditory stimulus (the sentence continues as the probes are presented visually) and thus less chance of probe integration into the sentence. However, the methodology developed by Dekydtspotter et al. (2010) uses pictures instead of words as probes to minimize such an effect, because visual processing is separate from language processing (e.g., Jackendoff, 1987; Marr, 1982). In addition, if congruency of the probes is controlled for such that related and unrelated probes are equally compatible with the sentence structure, this potential confound is not expected to affect the results.

Dekydtspotter et al. (2010) used probe classification during reading to test for evidence of trace reactivation at the clause edge (i.e., evidence of intermediate traces owing to cyclic movement required by extraction out of an embedded clause) in English NSs and high intermediate L2 learners from various L1 backgrounds (mostly Korean and Chinese) on sentences as in (4).

(4) *Harry* / *is* / *who* / *Mary* / *said* / *on* / [#1] / *Monday* / *that* / [#2] / *the headmaster* / *congratulated* / *at the assembly*.

The picture probe appearing either at the clause edge immediately following the complementizer *that* (#2) or in an earlier control position inside a prepositional phrase (#1) depicted either a boy (matching the antecedent *Harry* in gender) or a girl (matching the gender of the embedded subject *Mary*). Participants read the sentences aloud to themselves, as each segment (indicated by the slash marks in (4)) appeared on the screen, and classified intervening probes as human or nonhuman. The distracter items were sentences of the same structure and included pictures of animals. Participants also completed a WM reading span test in English. The

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

results grouped according to WM revealed only partial priming effects (shorter RTs to both boy and girl probes in the clause edge vs. the control position) among the NSs and low WM L2 learners, as well as asymmetries in the opposite direction among the high WM L2 learners (faster RTs for nonmatching probes at the clause edge). However, a closer examination of the individual data revealed that about half of both the NS and the L2 learner participant groups did show robust facilitation priming effects as expected, with faster RTs for matching versus nonmatching probes at the clause edge, whereas the other half, again of both participant groups, exhibited inhibition effects, producing shorter RTs for nonmatching probes in the clause-edge position.

Dekydtspotter et al. (2010) argued that these inhibitions are compatible with structurally based processing routines and the computation of intermediate traces during sentence processing. Inhibitions in lieu of the expected facilitations reflect weak activation of the antecedent of the filler (Carr & Dagenbach, 1990; Dagenbach, Carr, & Barnhardt, 1990), presumably owing to slowed or nonautomatic lexical access in the L2 (see, e.g., Segalowitz, 2003). In addition, recall that a similar inhibitory pattern has been found for L1 acquisition: the low WM children of Roberts et al.'s (2007) study also produced longer RTs to identical versus unrelated picture probes at the syntactic gap position. Thus, the probe classification during reading methodology is argued to tap into the same mental processes as the cross-modal priming methodology and to reflect sentence processing routines.

THE CURRENT STUDY

The current study directly compares the use of cross-modal picture priming and probe classification during reading for testing sentence processing in L2 French. Thus, the same experimental materials were presented in these two tasks.

Materials

The test sentences ($n = 20$) as in (5) involved indirect object relative clauses and were structurally very similar to those used by Felser and Roberts (2007), as in the example in (6).[1]

(5) *Georges déteste le zèbre à qui le jeune kangourou a offert le dernier gâteau après la fête mercredi soir.* (the current study)
"Georges hates the zebra to whom the young kangaroo offered the last cake after the party Wednesday evening."

(6) *Jane loved the tiger to which the black beetle offered the sweet strawberry cake at the party last week.* (Felser & Roberts, 2007)

Although direct translation from English to French of the sentences used in Felser & Roberts's study would have been possible in most cases, the vocabulary has been simplified: all of the animal names used as antecedents are either French–English homographic cognates (as in *zèbre*, "zebra," and *kangourou*, "kangaroo," in this example) or common vocabulary that learners would have seen from the beginning stages of L2 learning (e.g., *chien*, "dog").

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

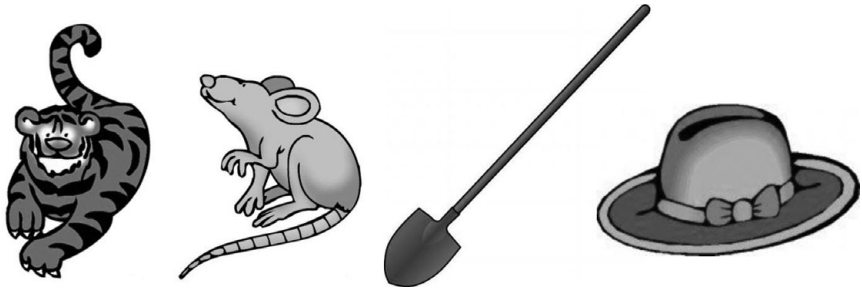


Figure 1. Examples of alive and not alive picture probes.

A context was created in which a French schoolteacher was looking through storybooks that she had read to her class and remembering what the children had said about the various characters in the books. Thus, the experimental sentences presented a human character's opinion about an animal who had had been given, shown, sent, sold, explained, or told something by another animal on a specific day of the week. The direct object of the embedded verb was always a noun modified by a prenominal adjective (e.g., *le dernier gateau*, "the last cake"); this created an earlier control position for the probe to appear between the adjective and the noun, where no trace should be posited. The picture probes were meant to be illustrations from the books the class had read. They were cartoon-style images taken from computer software clip art packages and modified to a uniform size. All the experimental probes in this task were animal pictures; pictures of inanimate objects were used in the distracter items. Examples of picture probes are given in Figure 1. To avoid a possible task effect that may have shaped the advanced L2 learner results of Felser and Roberts (2007), where RTs were faster for identical pictures that matched referents in the sentence than for unrelated pictures, in this experiment, nonmatching control probes matched the embedded subject of the sentence.

Participants read or listened to the teacher's comments and pressed a button to indicate whether picture probes, which matched either the antecedent or the embedded subject and appeared either in the gap or in the control position, depicted something alive or not alive. This yielded four experimental conditions, as illustrated in (7). The slashes represent the segmentation for the probe classification during reading task. Each participant encountered only one version of each sentence but equal numbers of sentences in all four conditions.

(7) a. CONDITION 1 (C1): Matching probe in the gap position
Georges / déteste / le zèbre / à qui / le jeune kangourou / a offert / le dernier / gâteau / [picture probe: ZEBRA] */ après la fête / mercredi soir.*

b. CONDITION 2 (C2): Nonmatching probe in the gap position
Georges / déteste / le zèbre / à qui / le jeune kangourou / a offert / le dernier / gâteau / [picture probe: KANGAROO] */ après la fête / mercredi soir.*

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

c. CONDITION 3 (C3): Matching probe in the control position
Georges / déteste / le zèbre / à qui / le jeune kangourou / a offert / le dernier / [picture probe: ZEBRA] */ gâteau / après la fête / mercredi soir*.
d. CONDITION 4 (C4): Nonmatching probe in the control position
Georges / déteste / le zèbre / à qui / le jeune kangourou / a offert / le dernier / [picture probe: KANGAROO] */ gâteau / après la fête / mercredi soir*.
"Georges hates the zebra to whom the young kangaroo gave the last cake after the party Wednesday evening."

In addition to the 20 experimental items, the stimulus included an additional 30 distracter items. Twenty of the distracter sentences contained direct object relative clauses and pictures of inanimate objects, with 5 more direct object relative sentences with animal probes and 5 more indirect object relative clauses with inanimate object probes. The probes appeared at various points throughout the distracter items, so as to give participants the impression that a picture probe could appear at any moment during reading.

Procedure

The experimental sentences were presented in two distinct tasks: cross-modal priming and probe classification during reading. Both tasks were administered on a computer via DMDX software (Forster & Forster, 2003), which controlled probe presentation and reading speed (in the case of the probe classification during reading task), and recorded RTs to picture probes. This program also randomized the presentation order of the sentences. For the cross-modal priming task, the sentences were recorded by a female NS of French. A single red star in the center of the computer screen served as a fixation point while participants listened to each sentence. The probes were timed to appear on the screen with the onset of the direct object noun *gâteau* (control position) or the prepositional phrase *après la fête* (gap position). Probes were presented for 650 ms, and participants were asked to decide as quickly as possible whether the probe depicted something that was alive or not alive and to press the left or right arrow keys on the computer keyboard to indicate their classification decision. To ensure that participants were actively listening to the sentences for comprehension, the cross-modal priming task also included simple true or false follow-up statements. These questions focused on the content of various parts of the sentences: for example, for the item in (7), the true–false statement was *Georges adore le zèbre*, "Georges loves the zebra" (false). Other possible statements might have been *le kangourou était jeune*, "the kangaroo was young" (true), or *c'était lundi soir*, "it was Monday evening" (false), and so on. The statements appeared on the computer screen after each aurally presented sentence, and participants indicated whether the statement was accurate by pushing the Y or N key.

The probe classification during reading methodology (Dekydtspotter et al., 2010) involves segmented forced-paced reading aloud while simultaneously categorizing intervening picture probes as alive or not alive. Participants read sentences aloud as each segment appeared on a computer screen and classifed picture probes as alive or not alive as quickly as possible by pressing the left or right arrow key.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

Picture probes appeared for 650 ms. Each sentence was preceded by a red star in the center of the screen that signaled that the sentence was about to begin and that the participant needed to be ready to start reading aloud. The end of each sentence was indicated by a period following the last word. A blank white screen appeared for 1 s between the end of an item and the red star that signaled the beginning of the next sentence. It is important to recall that this task did not include any comprehension questions, because it is assumed that having participants read aloud will serve to focus their attention on the experimental stimuli (a point that will be taken up again in the Discussion Section). Sentence segments were displayed for 500–700 ms, depending on length. Reading speed, which was controlled by the experimental software, was tested and calibrated through trial and error and piloting with NS and L2 learner volunteers, and was established to be a fluent and appropriate speed of presentation. Participants were monitored to ensure that they were reading aloud and were reminded to do so if the researcher noticed that they had stopped (although for the most part, participants did not need to be reminded to read aloud).

The same participants completed both the cross-modal priming and the probe classification during reading tasks, in two separate testing sessions that occurred approximately 1 week apart. Participants were tested individually. Most testing sessions took place in a quiet laboratory designed for this purpose; however, a few participants were tested in other quiet locations (e.g., campus offices or the library). Participants were asked to fill out a background questionnaire that gathered basic demographic information (e.g., age, gender, and L1) as well as information concerning their experience with learning French or any other languages (e.g., length of L2 study or time spent abroad). In addition to both versions of the main experimental task, participants also completed a WM reading span test adapted for L2 French from Harrington and Sawyer (1992), in which they read sets of two to five sentences and tried to remember the last word of each sentence within each set while also judging the grammaticality of each sentence that they read. There were a couple of other preliminary tasks, as described in the following sections. Instructions for all tasks were provided both orally and in writing in English (the learner participants' L1), and participants were encouraged to ask for clarification if they did not understand.

Participants

I report here on two participant groups: advanced learners ($n = 15$; 8 males, 7 females) and NSs ($n = 12$; 2 males, 10 females). Portions of this data (i.e., the results of the probe classification during reading task) have been reported and discussed in Miller (2011), which also included an intermediate learner group whose (flat) results have been omitted here, and in Dekydtspotter and Miller (2013). The L2 learners were graduate students at a large Midwestern university pursuing degrees in French literature or linguistics at the time of testing. The L1 of all learner participants was American English. The NSs were graduate students or postdoctoral researchers at the same university. Descriptive information concerning the participants' age, length of L2 French study (in years), time spent abroad (in months), and WM scores is provided in Table 1.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

Table 1. *Participant characteristics*

	Advanced ($n = 15$)			Native ($n = 12$)		
Characteristic	Mean	*SD*	Range	Mean	*SD*	Range
Age	27.0	3.5	23–36	30.3	6.4	25–43
Working memory	31.4	6.5	18–41	36.6	4.3	28–42
Years of study	10.3	3.4	5–16		NA	
Time abroad (months)	15.4	11.7	1.5–36		NA	

A one-way analysis of variance (ANOVA) revealed a significant difference between the WM scores of the two groups, $F(1, 25) = 5.622, p < .05$. This finding may seem problematic, given that previous studies have revealed a significant role for WM, at least in research on L1 sentence processing, and these results seem to suggest a WM advantage for the NSs. However, this difference might simply be because the test was administered in French; the mean WM scores for the omitted group of intermediate learners was significantly lower than that of the advanced learner group (26.9 vs. 31.4, respectively), $F(1, 32) = 5.440, p < .05$, suggesting that the scores on this test improved with increased proficiency. WM measures based on reading span in a L2 may be inherently problematic, given that WM capacity becomes confounded with reading abilities. Nonetheless, for the purposes of the current study, a conscious decision was made to test reading span in the target language to maintain consistency with previous research (e.g., Felser & Roberts, 2007, measured WM through L2 reading span).

Picture classification pretest

To establish a baseline categorization RT for the images used as picture probes, and to ensure that any RT asymmetries in the main experimental tasks were not induced by the pictures themselves, a picture classification pretest was administered to all participants before they completed the probe classification during reading and the cross-modal priming tasks. This test also introduced participants to the task of categorizing each image as alive or not alive as quickly as possible by pressing the left or right arrow key on the computer. Given the experimental design, whereby the same pictures were used as both matching and nonmatching probes appearing in both positions, thus cycling through all four conditions, significant RT asymmetries to the experimental probes are not expected (it should be recalled that all of the pictures used in the main experimental items depicted animals; the inanimate object pictures were used in the distracter items). The results of this task, as illustrated in Figure 2, confirm that there was no difference in participants' responses to the pictures used in different conditions on each version of the main experimental task. An ANOVA revealed no significant effect of picture set, $F(1, 25) = 0.528, p = .474$, and no interaction between picture set and participant group, $F(1, 25) = 0.547, p = .467$. This finding indicates that any RT asymmetries

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

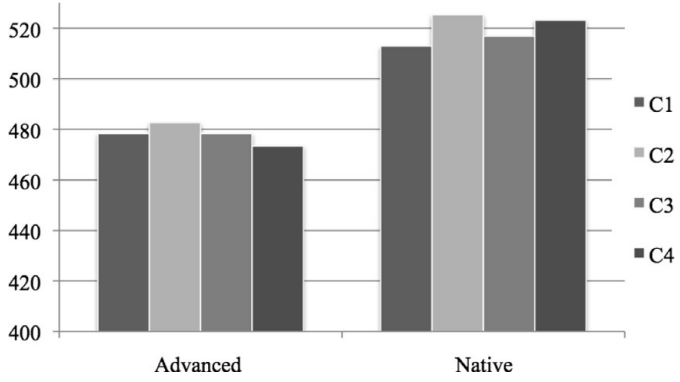


Figure 2. Results of the picture classification pretest: response times (ms) to pictures used in each experimental condition on the main experimental tasks.

revealed in the probe classification during reading and cross-modal priming tasks will not be due to the pictures themselves.

Masked priming task

Because the results of such sentence processing tasks crucially rely on priming effects, a masked priming pretest was administered to gather information about participants' priming behavior separate from sentence processing. Participants were told that, as a continuation of the training phase, they were to again classify images, each of which would be introduced by a series of hatch marks (#####). What participants did not realize was that masked prime words appeared between the hatch marks and picture probes for only 40 ms, rendering these words imperceptible to participants. The task incorporated the same vocabulary and images that appeared in the main experimental tasks, in three experimental conditions that differed in terms of the relationship between the masked prime animal word and the picture: in the match condition, the picture and the prime were an exact match (e.g., a picture of a zebra following the masked prime *zèbre*); in the mismatch condition, a picture of a different animal followed the prime (e.g., a picture of a kangaroo after the masked prime *zèbre*); and in the object condition, a picture of an inanimate object appeared after the prime. RTs for the masked priming task are given in Table 2 and illustrated in Figure 3.

An ANOVA revealed a main effect of condition, F (1, 25) = 7.351, $p < .05$, $\eta_p^2 = 0.227$, and a significant effect of participant group, $F(1, 25) = 8.677$, $p < .01$, $\eta_p^2 = 0.258$. Paired-samples t tests revealed significantly longer RTs in the object condition in both participant groups: advanced learners, match versus object, t (14) = 4.236, $p < .01$, mismatch versus object, t (14) = 6.034, $p < .001$; NSs, match versus object, t (11) = 5.176, $p < .001$, mismatch versus object, t (11) = 6.905, $p < .001$. However, whereas the advanced learners exhibited no difference in their RTs to matching and mismatching animal pictures following an animal

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

Table 2. *Response times (ms) to pictures in masked priming task*

Condition	Advanced ($n = 15$)	Native ($n = 12$)
Match	441 (23)	477 (45)
Mismatch	444 (32)	491 (52)
Object	477 (26)	522 (59)

Note: Standard deviations are in parentheses.

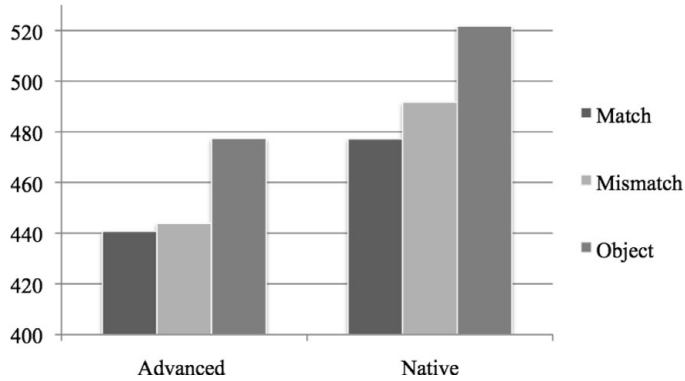


Figure 3. Response times (ms) to pictures in the masked priming task.

prime word, $t\ (14) = 0.526$, $p = .607$, the NS results revealed a more gradient asymmetry, with shorter RTs in the match versus the mismatch conditions, $t\ (11) = 2.218$, $p < .05$ (one tailed).

Analysis and predictions

Only RTs for which an image was identified correctly as alive or not alive were included for analysis. All participants were highly accurate in their classification of images, with less than 2% of the data excluded owing to incorrect classifications. Data pruning affected an additional 4% of the data: within each participant group, any RT that fell outside two standard deviations from the mean was removed and replaced by the new group mean. A $2 \times 2 \times 2$ ANOVA investigated the effects and interactions of probe position, probe type, and task as within-subject factors and participant group and WM score as between-subject factors. Planned paired-samples *t* tests, with one-tailed α (given expectations of facilitation due to matching probes) set at 0.05, were used to determine whether any asymmetries in the mean RTs in each condition were statistically significant. If processing is structurally based, and if learners use syntactic gaps to resolve filler-gap dependencies, RTs

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

are expected to be shorter when a picture probe appearing at the moment the gap is encountered matches the antecedent than when the image does not match (i.e., faster RTs for C1 vs. C2). In addition, RTs for matching probes in the gap position are also predicted to be faster than those to matching pictures appearing in the control position (C1 vs. C3). No other significant RT differences are expected.

It is important to note that this predicted priming pattern is the hallmark of a theory of sentence processing in which the referent of the antecedent of a moved constituent is reactivated at the extraction site to mediate the filler-gap dependency. Such a response pattern cannot be attributed to domain-general factors such as linear distance or greater salience of the relativized referent. The condition of interest, where antecedent matching probes appear in the gap position, actually represents the greatest linear distance between the antecedent and the picture probe; intuitively, shorter RTs would not be expected in this condition compared with the control position, which occurs earlier in the sentence. In addition, salience of the referent of the relative clause head noun would lead to shorter RTs to matching probes in both the gap and the control positions, similar to the advanced learner results of Felser and Roberts (2007). Thus, a priming pattern that is characterized by the shortest RTs in the classification of antecedent-matching probes in the gap position, when RTs in the other three conditions are nondistinct, is indicative of a sensitivity to syntactic structure during sentence processing and is suggestive of movement dependencies mediated through traces.

RESULTS

For the cross-modal priming and probe classification during reading tasks, a $2 \times 2 \times 2$ ANOVA considered the effects of position (gap vs. control), probe (matching vs. nonmatching), and task (reading vs. listening) as within-subjects factors, with participant group and WM as between-subjects factors. This revealed a main effect of position, $F_1\ (1, 25) = 4.772$, $p < .05$, $F_2\ (1, 38) = 3.754$, $p = .060$, $\eta_p^2 = 0.160$; a main effect of probe, $F_1\ (1, 25) = 4.153$, $p = .052$, $F_2\ (1, 38) = 5.927$, $p < .05$, $\eta_p^2 = 0.142$; and a main effect of task, $F_1\ (1, 25) = 30.70$, $p < .001$, $F_2\ (1, 38) = 109.128$, $p < .001$, $\eta_p^2 = 0.552$. Furthermore, significant interactions were revealed between probe and participant group, $F_1\ (1, 25) = 6.708$, $p < .05$, $F_2\ (1, 38) = 6.708$, $p < .05$, $\eta_p^2 = 0.212$, and probe, modality, and participant group, $F_1\ (1, 25) = 12.288$, $p < .01$, $F_2\ (1, 38) = 5.370$, $p < .05$, $\eta_p^2 = 0.330$. A significant interaction between position and probe, $F_1\ (1, 25) = 6.126$, $p < .05$, $\eta_p^2 = .0197$, was further qualified by participant group, $F_1\ (1, 25) = 7.203$, $p < .05$, $\eta_p^2 = 0.211$.

Separate ANOVAs were thus conducted for each participant group. In the advanced learner data, there was a main effect of position, $F_1\ (1, 14) = 60.141$, $p < .05$, $F_2\ (1, 19) = 6.623$, $p < .05$, $\eta_p^2 = 0.305$; a main effect of probe, $F_1\ (1, 14) = 19.573$, $p < .01$, $F_2\ (1, 19) = 13.152$, $p < .01$, $\eta_p^2 = 0.583$; and a main effect of task, $F_1\ (1, 14) = 40.825$, $p < .001$, $F_2\ (1, 19) = 81.768$, $p < .001$, $\eta_p^2 = 0.745$, as well as a significant Probe × Task interaction, $F_1\ (1, 14) = 12.090$, $p < .01$, $\eta_p^2 = 0.463$. The NS data showed a significant main effect of task, $F_1\ (1, 11) = 6.836$, $p < .05$, $F_2\ (1, 19) = 42.070$, $p < .001$, $\eta_p^2 = 0.383$, as well as a significant Position

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

Table 3. *Response times (ms) for probe classification during reading and cross-modal picture priming tasks*

Condition	Probe Classification During Reading		Cross-Modal Picture Priming	
	Advanced ($n = 15$)	Native ($n = 12$)	Advanced ($n = 15$)	Native ($n = 12$)
C1: Matching, gap	498 (38)	524 (72)	571 (55)	612 (127)
C2: Nonmatching, gap	501 (42)	555 (80)	605 (75)	626 (98)
C3: Matching, control	519 (38)	562 (79)	579 (74)	649 (123)
C4: Nonmatching, control	511 (44)	560 (75)	627 (74)	597 (112)

Note: Standard deviations are in parentheses.

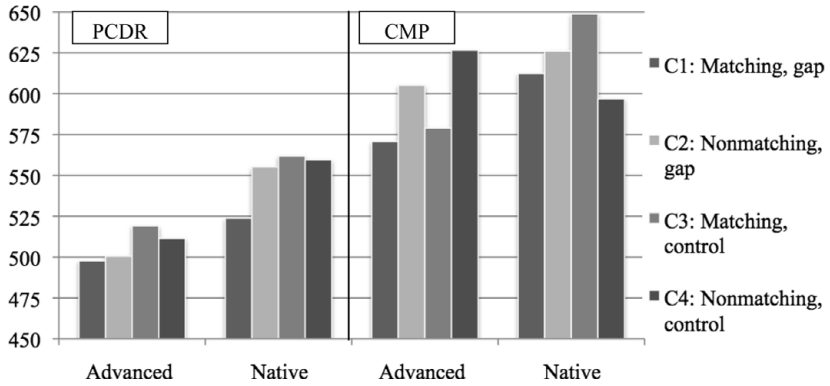


Figure 4. Response times (ms) for (left) the probe classification during reading task and (right) the cross-modal priming task.

× Probe interaction, F_1 (1, 11) = 10.670, $p < .01$, $\eta_p^2 = 0.492$. The RTs for the two tasks are given in Table 3 and illustrated in Figure 4.[2]

As can be seen in Figure 4, on the probe classification during reading task, the advanced learners responded more quickly to both matching and nonmatching probes in the gap position (C1, C2) compared to the control position (C3, C4), suggesting a partial priming effect. An ANOVA revealed an effect of position that approached significance in this data, F (1, 15) = 4.107, $p = .062$, $\eta_p^2 = 0.227$, and paired-samples t tests confirmed that although RTs to matching probes in the gap position (C1) were significantly shorter than those to matching probes in the control position (C3), t (14) = 2.970, $p < .05$, RTs to matching and nonmatching probes in the gap position (C1 vs. C2) were not statistically distinct, t (14) = 0.356, $p = .727$. The NS participants, in contrast, produced RT asymmetries suggestive of the expected priming effects, with significantly faster RTs to matching versus nonmatching probes in the gap position (C1 vs. C2), t (11) = 2.504, $p < .05$ (one

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

tailed), and to matching probes in the gap versus the control positions (C1 vs. C3), $t\ (11) = 2.481$, $p < .05$ (one tailed). Crucially, the matching–nonmatching asymmetry occurred only in the gap position, with no such differences exhibited in the control position (C3 versus C4), $t\ (11) = 0.164$, $p = .872$.

On the cross-modal priming task, among the advanced learners, an ANOVA revealed a main effect of probe, $F\ (1, 14) = 16.325$, $p < .01$, $\eta_p^2 = 0.538$, and paired-samples *t* tests revealed RT asymmetries for matching versus nonmatching probes in both the gap, $t\ (14) = 2.093$, $p = .055$, and the control positions, $t\ (14) = 3.045$, $p < .01$. This pattern thus replicates the advanced learner results of Felser and Roberts's (2007) cross-modal priming task. The NSs results, however, depict no discernible RT pattern, although an ANOVA revealed a significant Position × Probe interaction in this data, $F\ (1, 11) = 7.356$, $p < .05$, $\eta_p^2 = 0.401$. Such a null result for NS participants may seem surprising, but it should be recalled that the sensitivity and accuracy of the cross-modal priming task, even as a measure of L1 processing routines, has been questioned and discussed in previous research. Nicol et al. (1994) noted that there may be a number of factors, independent of language processing, that can affect the results of such studies. One possible explanation for the NS results will be considered in the Discussion section.

DISCUSSION

The advanced learners of Felser and Roberts's (2007) study experienced an across-the-board facilitation priming effect in the classification of picture probes that were identical to the antecedent of the *wh*-filler in both the gap and the control positions. In the current study, this same pattern was produced by advanced learners on the cross-modal priming task. Felser and Roberts argued that this response pattern constituted evidence of maintained activation, whereby the learners "retained fronted *wh*-phrases in WM during the processing of the experimental sentences, [but] they did not retrieve them from WM (i.e. reactivate them) at structurally defined gap sites" (p. 26). At first, it might seem that the fact that the advanced learner pattern of Felser and Roberts (2007) was replicated in the current study, with a different target language and slightly different materials, only serves to confirm the previous conclusions that L2 learners process input shallowly in real time. However, a comparison of the results of the cross-modal priming task with those of the probe classification during reading task, which tested the same learners on the same experimental sentences, casts serious doubt on this type of argument.

Figure 5 presents a comparison of the learner results of the current study on the cross-modal priming and the probe classification during reading tasks with those from Felser and Roberts (2007). Figure 5 illustrates that although the results of the cross-modal priming task are the same in both the current study and in Felser and Roberts (2007), a very different pattern was produced in response to the same experimental sentences when presented in the probe classification during reading task. On the probe classification during reading task, the advanced learners produced a partial priming pattern, with faster RTs to matching probes in the gap versus the control position but no difference between RTs to matching and nonmatching probes in the gap position.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

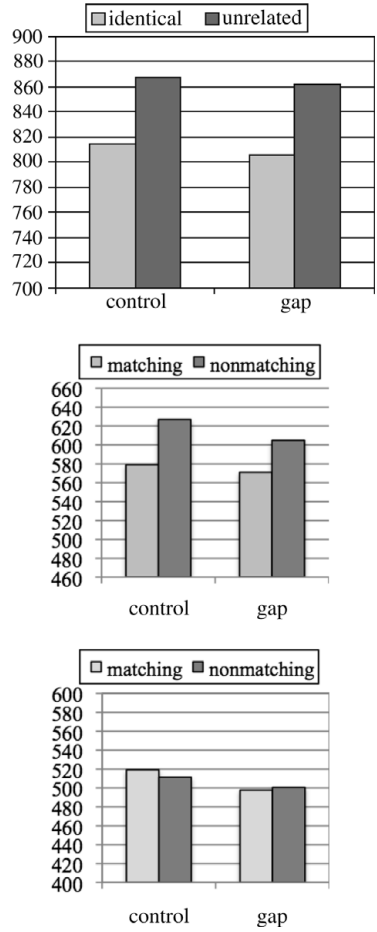


Figure 5. (Top) Felser and Roberts' (2007) advanced learner results on a cross-modal priming task, (middle) the current study's advanced learner results on the cross-modal priming task, and (bottom) the probe classification during the reading task. The top image is reprinted from "Grammatical Processing of Spoken Language in Child and Adult Language Learners," by C. Felser and H. Clahsen, 2009, *Journal of Psycholinguistic Research, 38,* p. 315. Copyright 2009 by Springer. Reprinted with permission.

This partial priming pattern produced by the advanced learners could reflect the fact that both probe types matched referents in the sentence (recall that in this experiment the nonmatching probes actually matched the embedded subject), which were thus both active in the mental representation of the discourse model. Alternatively, such a result could be because both the matching and nonmatching probes depicted animals: perhaps there was a large degree of semantic overlap in conceptual structure between the two probe types. Although there may not

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

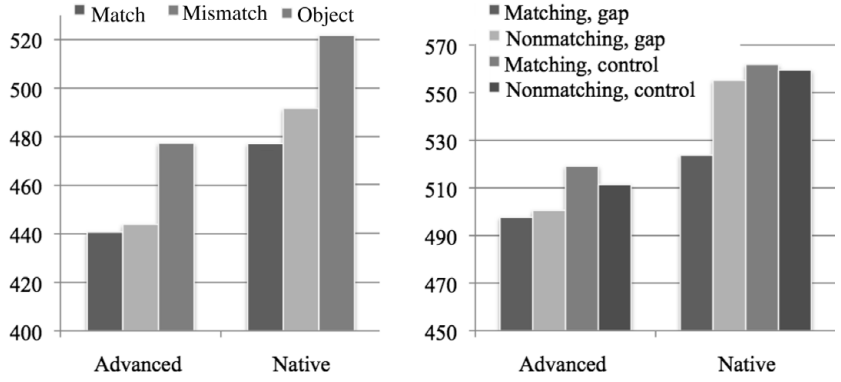


Figure 6. Response times (ms) for the masked priming task (left) and the probe classification during reading task (right).

seem to be much similarity between, for example, a giraffe and a turtle, both are animals, with four legs, that might be seen at a zoo, and so on. Thus, it is possible that such an association influenced learners' RTs. Note that no attempt was made to control for degree of similarity between the matching and the nonmatching probes, and this type of partial priming effect was unanticipated. Love's (2007) related and control probes both depicted animals (that seemed very similar in some pairs, such as *zebra–camel* and *hippo–rhino*), similarly to the current study, where the matching and the nonmatching probes were all images of animals.

The results of the masked priming task seem to confirm that the lack of asymmetry in RTs to matching and nonmatching probes in the gap position was due to the semantic overlap between the two probe types, although the interference of the embedded subject may have also played a role and cannot be ruled out at this point. The results of the masked priming task demonstrated that the advanced learners as well as the NS participants experienced robust priming effects following the covert presentation of a masked prime word; however, the response patterns also seemed to point to differences between learners and NSs in the degree of activation spreading in the semantic network (Collins & Loftus, 1975): whereas the NSs exhibited differing levels of facilitation for pictures that matched the primes exactly compared with probes that were merely related to the prime (i.e., a different animal), the L2 learner participants showed the same level of facilitation for all animal pictures, regardless of whether they matched the prime word exactly. This finding had important implications for the results of the probe classification during reading task. Two distinct priming profiles were revealed in the masked priming and the probe classification during reading tasks: in both tasks, whereas advanced learners exhibited no difference in their RTs to matching and nonmatching pictures, NS RTs did reflect significant differences in the classification of matching and nonmatching picture probes. These two profiles are illustrated in Figure 6.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

The results of the current study thus indicate that the task used can affect the results (and interpretation thereof) in (L2) sentence processing research. The task demands of the cross-modal priming task were much greater, given that this task included a listening comprehension component, with follow-up questions, whereas the probe classification during reading task did not require the participants to respond to any comprehension checks. This alone may account for the differing results on the two tasks: it is possible that the greater demands of the cross-modal priming task may have induced a processing breakdown among the learner participants, leading to the fastest RTs for pictures that matched the most salient referent. The probe classification during reading task was argued to remove some of the complexity of the cross-modal priming task. However, with no questions to verify participants' comprehension of the sentential content and to thus hold them accountable for reading, how can we really be sure that the design of the probe classification during reading task truly ensures that participants are processing and understanding the sentences that they read aloud?[3] The results of the task seem to indicate that sentence processing does take place during reading aloud. Recall that a response strategy based on linear distance between referent and picture probe would yield shorter RTs in the earlier control position relative to the gap position and that the salience of the relative clause head should induce faster RTs for matching probes relative to nonmatching probes; both of these possible response patterns should be available to (both native and nonnative) participants regardless of their processing routines. However, the NSs produced a priming pattern consistent with a sentence processing theory involving reactivation induced by movement traces, indicating detailed processing (and thus, comprehension). In addition, the partial priming pattern produced by the learners, when considered alongside the results of the masked priming task, is similarly more compatible with such processing.

Even the NS participants exhibited differing RT patterns on the two tasks: on the probe classification during reading, native French speakers produced the predicted response pattern, with the shortest RTs to matching probes appearing in the gap position, compared with the other three conditions. Again, these results confirm that the reading task is able to tap into the same mental processes as the cross-modal priming task. However, on the cross-modal priming task, the NSs did not exhibit the predicted RT pattern. It is not entirely clear what may have shaped the NSs' response pattern, but it may seem problematic for the current study that this task did not yield the expected results among NS participants. However, it could be that the increased task demands of the cross-modal priming task affected the NS results as well. As noted earlier, there may be any number of (nonlinguistic) factors that can affect the results of such studies. For example, recall that in Roberts et al.'s (2007) study, it was only the NS with high WM capacity who showed reactivation effects on the cross-modal priming task. In the current study, WM score as a continuous variable did not have a statistically significant effect on the overall results; however, it may have been a contributing factor in the NS results on this specific task. Although the sample size is too small to draw any real conclusions, the NS results split by WM group, which are depicted in Figure 7, seem to suggest that only the higher WM NSs produce a targetlike pattern, whereas the lower WM NSs exhibit an indiscernible pattern.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

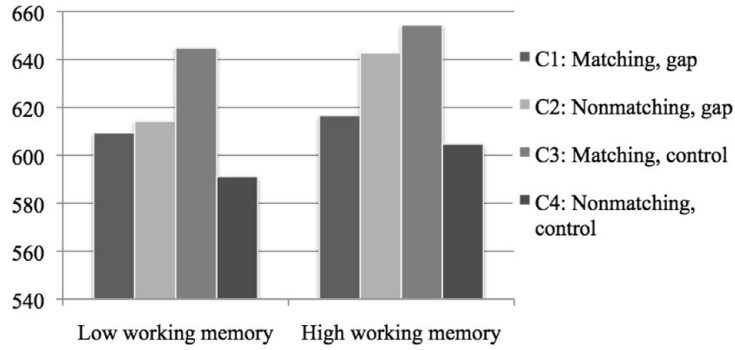


Figure 7. Results on the cross-modal priming task for low ($n = 5$) and high ($n = 7$) working memory native speakers.

CONCLUSION

The current study explored whether L2 learners are able to compute structural representations of the input in real time and whether evidence of such computations would be found when the experimental task was simplified. The results indicate that experimental task demands can determine the results of sentence processing studies, in both L1 and L2, and that L2 learners and NSs alike seem to be better able to demonstrate use of structural processing routines under certain conditions, for example, when the computational burden is alleviated through the use of a simplified task (that is more tailored to examining L2 processing). It therefore seems that previous conclusions concerning the nature of L2 sentence processing may merit further consideration. In future research, it will be important to consider the range of factors that may affect experimental results and to design sentence processing tasks that are appropriate for L2 learners. The current study constitutes an initial step toward this goal, by showing that the type of task used can affect experimental results.

NOTES

1. Although the (human) relative pronoun *to which* was used in Felser and Roberts's sentences, the current study used *à qui* ("to whom") instead, given that the animals were engaged in very humanlike activities such as attending parties and sharing cake.
2. It is unexpected that overall, the NSs' RTs would be slower than those of the learners. Given the greater cognitive demands of sentence processing in a L2, faster RTs among the NSs may intuitively be expected, and the adult NSs of Roberts et al. (2007) exhibited faster RTs than the advanced learners of Felser and Roberts (2007). In addition, the NS children tested in Roberts et al. exhibited much longer RTs than did both the NS adults and the L2 learners, presumably owing to slower auditory processing or impaired lexical access (Felser & Clahsen, 2009). However, this RT latency observed among the NSs of the current study with respect to the learner participants was also present in both the pretest and the masked priming data; thus, it seems to be a characteristic of this particular group of participants.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

3. I thank an anonymous reviewer for raising this question and for motivating me to better justify my methodological choices.

REFERENCES

Bley-Vroman, R. (2009). The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition, 31*, 175–198.

Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology, 16*, 341–350.

Clahsen, H., & Felser, C. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics, 27*, 107–126.

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

Clifton, C., & Frazier, L. (1989). Comprehending sentences with long distance dependencies. In G. M. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 273–317). Dordrecht: Kluwer.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Dagenbach, D., Carr, T. H., & Barnhardt, T. M. (1990). Inhibitory semantic priming of lexical decisions due to failure to retrieve weakly activated codes. *Journal of Experimental Psychology, 16*, 328–340.

Dekydtspotter, L., & Miller, A. K. (2013). Inhibitive and facilitative priming induced by traces in the processing of *wh*-dependencies in a second language. *Second Language Research, 29*, 345–372.

Dekydtspotter, L., Miller, A. K., Schaefer, V., Chang, Y., & Kim, O.-H. (2010). Clause-edge reactivations of fillers in processing English as a second language. In M. T. Prior, Y. Watanabe, & S.-K. Lee (Eds.), *Selected proceedings of the 2008 Second Language Research Forum* (pp. 108–122). Somerville, MA: Cascadilla Press.

Dekydtspotter, L., Schwartz, B. D., & Sprouse, R. A. (2006). The comparative fallacy in L2 processing research. In M. Grantham O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006): The Banff Conference* (pp. 33–40). Somerville, MA: Cascadilla Press.

Felser, C., & Clahsen, H. (2009). Grammatical processing of spoken language in child and adult language learners. *Journal of Psycholinguistic Research, 38*, 305–319.

Felser, C., Clahsen, H., & Münte, T. F. (2003). Storage and integration in the processing of filler-gap dependencies: An ERP study of topicalization and *wh*-movement in German. *Brain and Language, 87*, 345–354.

Felser, C., Marinis, T., & Clahsen, H. (2003). Children's processing of ambiguous sentences: A study of relative clause attachment. *Language Acquisition, 11*, 127–163.

Felser, C., & Roberts, L. (2007). Processing *wh*-dependencies in a second language: A cross-modal priming study. *Second Language Research, 31*, 9–36.

Felser, C., Roberts, L., Gross, R., & Marinis, T. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics, 24*, 453–489.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers, 35*, 116–124.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217–235). Amsterdam: John Benjamins.

Frenck-Mestre, C. (2004). Ambiguities and anomalies: What can eye-movements and event-related potentials reveal about second language sentence processing? In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism* (pp. 268–284). New York: Oxford University Press.

Frenck-Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research, 21*, 175–198.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology, 50*, 119–148.
Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 78*, 1–76.
Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research, 30*, 251–266.
Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4*, 123–141.
Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14*, 25–38.
Hoover, M. L., & Dwivedi, V. D. (1998). Syntactic processing by skilled bilinguals. *Language Learning, 48*, 1–29.
Indefrey, P. (2006). It is time to work toward explicit processing models for native and second language speakers. *Applied Psycholinguistics, 27*, 66–69.
Jackendoff, R. (1987). *Consciousness and the computational mind*. Cambridge, MA: MIT Press.
Juffs, A. (1998). Main verb vs. reduced relative clause ambiguity resolution in second language sentence processing. *Language Learning, 48*, 107–147.
Juffs, A. (2005). The influence of first language on the processing of *wh*-movement in English as a second language. *Second Language Research, 21*, 121–151.
Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in *wh*-extraction. *Studies in Second Language Acquisition, 17*, 483–516.
Love, T. E. (2007). The processing of non-canonically ordered constituents in long distance dependencies by pre-school children: A real-time investigation. *Journal of Psycholinguistic Research, 36*, 191–206.
Love, T., Maas, E., & Swinney, D. (2003). The influence of language exposure on lexical and syntactic language processing. *Experimental Psychology, 50*, 204–216.
Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27*, 53–78.
Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.
McKoon, G., Albritton, D., & Ratcliff, R. (1996). Sentential context effects on lexical decisions with a cross-modal instead of an all-visual procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1494–1497.
McKoon, G., & Ratcliff, R. (1994). Sentential context and on-line lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1239–1243.
McKoon, G., Ratcliff, R., & Ward, G. (1994). Testing theories of language processing: An empirical investigation of the on-line lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1219–1228.
Miller, A. K. (2011). Rethinking nonnative processing constraints: Evidence from L2 French. In L. Plonsky & M. Schierloh (Eds.), *Selected proceedings of the 2009 Second Language Research Forum* (pp. 109–129). Somerville, MA: Cascadilla Proceedings Project.
Nicol, J. L., Fodor, J. D., & Swinney, D. (1994). Using cross-modal lexical decision tasks to investigate sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1229–1238.
Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18*, 5–19.
Nicol, J., Swinney, D., Love, T., & Hald, L. (2006). The on-line study of sentence comprehension: An examination of dual task paradigms. *Journal of Psycholinguistic Research, 35*, 215–231.
Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constraints in second language sentence processing. *Studies in Second Language Acquisition, 33*, 563–588.
Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L2 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition, 25*, 501–528.
Pickering, M., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes, 6*, 229–259.
Roberts, L., Marinis, T., Felser, C., & Clahsen, H. (2007). Antecedent priming at gap positions in children's sentence processing. *Journal of Psycholinguistic Research, 36*, 175–188.
Sag, I. A., & Fodor, J. D. (1994). Extraction without traces. In R. Aranovich, W. Byrne, S. Preuss, & M. Senturia (Eds.), *Proceedings of the 13th West Coast Conference on Formal Linguistics* (pp. 365–384). Stanford, CA: CSLI Publications.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press

Segalowitz, N. (2003). Automaticity and second languages. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Oxford: Blackwell.
Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language, 35*, 454–475.
Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience, 8*, 231–256.
Williams, J., Möbius, P., & Kim, C. (2001). Native and non-native processing of English *wh*-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics, 22*, 509–540.

https://doi.org/10.1017/S0142716413000362 Published online by Cambridge University Press