CrossMark

Studies in Second Language Acquisition, 2015, **37**, 487–516.
doi:10.1017/S0272263114000588

INTERMEDIATE TRACES AND INTERMEDIATE LEARNERS

Evidence for the Use of Intermediate Structure During Sentence Processing in Second Language French

A. Kate Miller
Indiana University–Purdue University Indianapolis

This study reports on a sentence processing experiment in second language (L2) French that looks for evidence of trace reactivation at clause edge and in the canonical object position in indirect object cleft sentences with complex embedding and cyclic movement. Reaction time (RT) asymmetries were examined among low (n = 20) and high (n = 20) intermediate L2 learners and native speakers (n = 15) of French in a picture-classification-during-reading task. The results show that a subgroup of learners (13 from the low intermediate and 9 from the high intermediate group) as well as the native speakers produced response patterns consistent with reactivation—with the shortest RTs for antecedent-matching probes presented concurrently with the gap—at clause edge, followed by a second reactivation in the canonical object position. This finding suggests that L2 learners may be able to process real-time input in nativelike ways, despite arguments set forth in previous research of this kind.

I would like to thank Laurent Dekydtspotter and the attendees of GALANA 5 and SLRF 2012, as well as four anonymous *SSLA* reviewers, for their comments and feedback on this project.

Correspondence concerning this article should be addressed to Kate Miller, Department of World Languages, Cavanaugh Hall 539E, 425 University Boulevard, Indianapolis, IN 46202. E-mail: am27@iupui.edu

© Cambridge University Press 2014

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

Filler-gap dependencies—that is, structures in which a verbal argument (the *wh*-filler) has been moved from its canonical position (the gap) to an earlier sentential position at some distance from the verb—present a unique challenge for sentence processing in that the fronted constituent must be retained in memory until it can be integrated into the structure. Empirical research on the processing of such dependency structures in a native language (L1) has revealed response patterns that are highly suggestive of mental reactivation of the fronted *wh*-filler at the structurally defined gap site from which it originated (e.g., Love & Swinney, 1996; Roberts, Marinis, Felser, & Clahsen, 2007). These results suggest that a *wh*-filler is temporarily stored in working memory (WM) for subsequent retrieval and integration into the dependency structure; filler integration is facilitated by the presence of a movement trace (i.e., a silent copy of the displaced *wh*-filler). Thus, the activation and maintenance of the referent of the filler interacts with the computation of the movement chain: Encountering the trace of a moved *wh*-expression triggers reactivation of its referent, thereby modulating the processing load and leading to efficiencies in computations.

It has been argued that sentence processing routines are afforded by a specialized and (largely) universal computational system for human language (Dekydtspotter, 2009; Slabakova, 2008) and, consequently, that parsing does not need to be (re)learned concurrently with a language. Second language (L2) learners may thus benefit from the efficiencies made available through this universal parser in the real-time processing of target language input. Previous experimental results have shown that sentence processing in a L2 involves many of the same reflexes that have been observed in L1 sentence processing: rapid structuring of the input (Juffs & Harrington, 1995; Williams, Möbius, & Kim, 2001), knowledge of island constraints (Omaki & Schulz, 2011), sensitivity to prosodic cues (Dekydtspotter, Donaldson, Edmonds, Liljestrand-Fultz, & Petrush, 2008), and use of lexical subcategorization information (Frenck-Mestre & Pynte, 1997). In contrast, however, another body of research has argued that, despite universal processing routines, incomplete or impaired (i.e., nonnative) knowledge of the target language grammar can hinder L2 sentence processing (Clahsen & Felser, 2006a, 2006b), resulting in superficial representations that lack hierarchical structure, effectively neutralizing any benefit of the detailed chain computations that involve movement traces in resolving *wh*-dependencies.

Among native speakers (NSs), empirical evidence for trace-induced reactivation is observed in longer reading times on segments that include a syntactic gap—which suggests that additional computations are happening at this point (e.g., Gibson & Warren, 2004)—or in faster response times (RTs) in categorizing visual probes (images or words) presented at crucial moments during processing that are related to the

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

filler of an aurally presented dependency structure due to facilitative priming effects (e.g., Love, 2007). Similarly, robust evidence for traces in the online structuring of input has been lacking in much previous research on L2 sentence processing, thus leading to the conclusion that the representations constructed in real time are not the same for the L1 and the L2. However, the fact that some studies have failed to capture the predicted effects does not necessarily indicate that movement traces are entirely absent in L2 sentence processing. Nicol, Fodor, and Swinney (1994) noted that response patterns in sentence processing research could be affected by myriad factors, regardless of whether detailed syntactic representations are actually computed by participants. Two major assumptions underlie the interpretation of response patterns exhibited in studies that target trace reactivation. It is assumed, first, that it is the presence of a movement trace that mediates the integration of a fronted *wh*-filler into the sentence structure (for arguments that filler integration is mediated by the verb's thematic structure, see, e.g., Pickering & Barry, 1991; Sag & Fodor, 1994; Traxler & Pickering, 1996), and, second, that the motor response of pressing a button will reflect the computation of that trace, in either slowed reading times or shorter RTs. Furthermore, the crucial assumption is that there will be an exact overlap between the trace-induced reactivation of the filler and the participant's response. As Dekydtspotter, Schwartz, and Sprouse (2006) noted, however, there is no guarantee that L2 learners' response patterns for a planned target segment will reflect the relevant processing moments.

The current study looks for evidence of trace reactivation at clause edge due to cyclic movement and of the use of intermediate structure to facilitate the processing of sentences with complex embedding. The experimental materials have been designed to minimize the computational burden, thus allowing (L2 learner) participants to focus more resources on maintaining in memory the referents involved in the dependency structure. A picture-classification-during-reading task involving indirect object cleft sentences in L2 French was used to gather RT data from low (n = 20) and high (n = 20) intermediate L2 learners as well as from a group of French NSs (n = 15). The cleft structure is argued to be more easily processed due to the focalized nature of the *wh*-filler (Sanford, Price, & Sanford, 2009). Additionally, to facilitate lexical access—which is slower and less automatic in a L2 (Favreau & Segalowitz, 1983; Segalowitz & Segalowitz, 1993), potentially leading to processing lags—only English-French cognate vocabulary items were used as antecedents. Whereas, in previous L2 sentence processing research, reactivation effects could have been overshadowed or masked by other mitigating factors (having little or nothing to do with syntax), the current study endeavors to uncover such effects among L2 learners.

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

PROCESSING *WH*-DEPENDENCIES IN THE L1 AND L2

Trace Reactivation and Crossmodal Priming Studies

Evidence that the representations computed during online L1 sentence processing include movement traces comes in part from research done in the crossmodal priming paradigm (Nicol & Swinney, 1989). These experiments are administered via computer; participants listen to aurally presented sentences for comprehension (which is verified through questions related to the content of the sentence) and are asked to make simple classification decisions about (picture) probes presented visually on the monitor at specific moments. Priming effects occur when the referent 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 the filler. Response times to matching or related probes at the gap position are predicted to 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.

Roberts et al. (2007) used a crossmodal priming task targeting indirect object relative clause sentences such as in (1) to test English NS adults and children (ages 5 to 7).

(1) *Jo knew the ostrich to which the black spider explained the difficult* [#1] *problem* [#2] *at school last Monday.*

Participants listened to these sentences for comprehension and classified picture probes that appeared on a computer screen—at the gap position (#2) or in a control position (#1) that occurred approximately 500 ms earlier—as *alive* or *not alive* by pushing a button. Identical probes matched the filler (e.g., a picture of an ostrich), whereas control probes depicted some unrelated inanimate object (e.g., a toothbrush or a carrot). All participants also completed a WM reading span test in addition to the main experimental task. Adults and children with higher WM capacity produced faster RTs to identical probes in the gap position compared with those for unrelated probes at the same position; control position RTs did not reflect this facilitation for filler-matching probes. In contrast, the children and adults with lower WM capacity did not exhibit such priming effects.

Felser and Roberts (2007) tested advanced adult L1 Greek learners of L2 English on Roberts et al.'s (2007) crossmodal priming task. These L2 participants exhibited faster RTs for identical versus unrelated probes in both the gap and control position. Their performance on a WM reading span test seemed to have no effect on their RT patterns, in contrast with Roberts et al.'s NS results. Additionally, Felser and Roberts pointed out that their L2 participants behaved differently from all subgroups of NSs

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

from that study. First, whereas the high-WM NS adults and children showed a position-specific (i.e., at the gap position only) advantage for identical probes, the L2 learners demonstrated this advantage in both the gap and control positions. Second, the low-WM NS adults exhibited no significant differences in their RTs to identical versus unrelated probes in either position. Finally, the low-WM NS children actually produced longer RTs to identical targets in both positions. Felser and Roberts also noted that the learners did not even seem to be able to process in a supposedly nativelike way a structure that could easily be transferred from their L1 (it is important to note that the structure of indirect object relative clauses such as those tested is very similar in English and Greek). Given that the same RT asymmetry was produced regardless of position—indicating a lack of detailed structure—Felser and Roberts concluded that these advanced learners were maintaining the activation of the antecedent throughout sentence processing rather than temporarily storing it and using the trace to trigger reactivation.

Sag and Fodor (1994) noted that although the predicted response pattern for crossmodal priming experiments is certainly consistent with a trace-mediated sentence processing routine, such effects could be expected under almost any theory of sentence processing, even a purely linear strategy based on word order expectations in which a ditransitive verb is expected to be followed first by its direct and then by its indirect object. Sag and Fodor thus argued that priming effects really only indicate that a gap is detected at that specific point during the processing of a sentence but not whether the gap contains a movement trace. Thus, it may seem odd that the L2 learners of Felser and Roberts (2007) would not show any effect of structure whatsoever; the absence of any such effects may in fact indicate that something else—unrelated to the ability to use syntactic information during L2 sentence processing—may have affected the learners' RTs.

Intermediate Traces and Self-Paced Reading Experiments

Perhaps a more promising area of exploration with respect to the processing routines available to L2 learners involves looking at long-distance dependencies that contain intermediate structure. Purely structural traces are hypothesized for sentences such as in (2) to mediate the long-distance dependency across clausal boundaries.

(2) *Who*$_i$ *did the consultant claim* [t_i *that the proposal had pleased* t_i]*?*

Given that a *wh*-phrase can only cross one bounding node at a time (e.g., Chomsky, 1986, 2005), sentences such as in (2) require a

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

two-step movement operation. Direct movement from the base-generated position is prohibited; the direct object noun phrase (NP) first moves into the specifier position of the embedded complementizer phrase (CP) before moving on to the specifier position of the matrix CP. The fronted constituent leaves an intermediate trace at the landing site of its first movement (at clause edge). Gibson (1998) noted that, from a processing standpoint, this intermediate structure can serve to mediate the long-distance dependency by breaking it up into smaller, more manageable chunks: The intermediate trace at clause edge reactivates the *wh*-filler in memory, thus refreshing its referent and facilitating its maintenance until the dependency can be resolved.

Gibson and Warren (2004) used a self-paced reading task to look for evidence of the use of intermediate structure during the online processing of sentences such as in (3) by English NSs. The segmentation of the sentences is indicated by the slash marks.

(3) a. *The manager who*$_i$ / *the consultant claimed* / t_i *that* / *the new proposal* / *had pleased* t_i / *will hire* / *five workers tomorrow.*
b. *The manager who*$_i$ / *the consultant's claim* / *about* / *the new proposal* / *had pleased* t_i / *will hire* / *five workers tomorrow.*

The sentence in (3a) involves extraction across a verb phrase (VP) and includes the intermediate trace that results from cyclic movement. In contrast, the sentence in (3b) involves extraction across a NP, which can be performed in one step and thus involves no cyclic movement and includes no intermediate trace. The experiment also included non-extraction sentences of comparable length as in (4).

(4) *The consultant claimed* / *that* / *the new proposal* / *had pleased* / *the manager* / *who will hire five workers tomorrow.*

The participants exhibited longer reading times on the complementizer *that* in sentences of the type in (3a) as compared with sentences with no extraction, as in (4), a finding that suggests that the processing of the intermediate trace in (3a) slowed reading at this point. Additionally, reading times on the verbal segment *had pleased* were faster in sentences such as in (3a) as compared with that in (3b), which confirmed Gibson and Warren's prediction that reactivation of the filler at clause edge (triggered by encountering the intermediate trace) would mediate the dependency, allowing the filler to be integrated more quickly at its canonical object position. These results seem to point to the computation of intermediate traces in online L1 processing and the use thereof in helping to maintain a dependency structure across longer distances.

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

Marinis, Roberts, Felser, and Clahsen (2005) adapted Gibson and Warren's (2004) materials to test English NSs and L2 learners from various L1 backgrounds (including German and Greek, the grammars of which both license *wh*-movement, as well as Chinese and Japanese, which leave *wh*-elements in situ). An example quadruple is given in (5).

(5) a. *The manager who*$_i$ / *the secretary claimed* / t_i *that* / *the new salesman* / *had pleased* t_i / *will raise company salaries.*
b. *The manager who*$_i$ / *the secretary's claim* / *about* / *the new salesman* / *had pleased* t_i / *will raise company salaries.*
c. *The manager thought* / *the secretary claimed* / *that* / *the new salesman* / *had pleased* / *the boss in the meeting.*
d. *The manager thought* / *the secretary's claim* / *about* / *the new salesman* / *had pleased* / *the boss in the meeting.*

Both the NSs and L2 learners showed an effect of extraction, as evidenced by longer reading times on the fifth segment *had pleased* in (5a) and (5b) versus (5c) and (5d), which reflects the cost of filler integration at that point during processing. The NSs also produced increased reading times on the third segment (the complementizer *that*) for sentences such as in (5a) as compared to (5c).

Marinis et al. (2005) argued that it was the interaction of effects observed in the third segment and the fifth segment that was truly suggestive of the use of intermediate structure to mediate the dependency across clausal boundaries. This interaction was only observed in the NS data; these effects were not detected in any of the learner groups. Given that the L2 learners did not show the expected reading time asymmetries at the intermediate gap site nor the crucial interaction between extraction and sentence type, Marinis et al. concluded that L2 learners are not able to make use of the same resources available to NSs during the online processing of complex nonlocal dependencies. However, as noted by Dekydtspotter et al. (2006), the direct comparison of NSs and L2 learners on this stimulus assumes that the reading times on a given segment will reflect the same computational moments in L1 and L2 reading, which may not necessarily be the case. Indeed, Nicol et al. (1994) remarked, "in . . . any . . . paradigm in which the experimenter decides when and where to look for subjects' response to a stimulus, a null result may only indicate that the test point happens not to have been optimally chosen" (p. 1234). Dekydtspotter et al. thus reexamined Marinis et al.'s L2 learner data, focusing specifically on the reading times for the segment following the complementizer *that* (the fourth segment). One-tailed *t* tests comparing extraction (5a, 5b) versus nonextraction (5c, 5d) sentences revealed (spillover) asymmetries on this segment among the L1 German and L1 Japanese participants for VP extraction only (5a vs. 5c). Thus, a closer look at the data

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

demonstrated that the analytical procedures adopted by Marinis et al.—namely, their decision to examine only the reading times on the third and fifth segments—allowed them to overlook in the L2 learner data the RT asymmetries suggestive of a sensitivity to the intermediate trace.

Pliatsikas and Marinis (2013) followed up on Marinis et al. (2005), using the same self-paced reading task to compare a group of L2 English classroom learners with another group of learners that had an average of 9 years of naturalistic exposure to L2 English. To be sure not to miss any relevant asymmetries, Pliatsikas and Marinis analyzed the reading times for the third, fourth, and fifth segments of sentences such as those in (5). A NS control group predictably produced longer reading times on the segment *that* in extraction contexts, followed by shorter reading times on the verbal segment *had pleased* for sentences that contained intermediate gaps, echoing Marinis et al.'s results. Neither learner group showed an effect of extraction on the third segment, but such an effect was found for the following segment among both groups, suggesting a possible spillover effect. On the fifth segment, however, the results for the two learner groups differed: Whereas the classroom exposure group did not show any evidence of facilitation for *wh*-filler integration in sentences involving intermediate structure, the naturalistic exposure group converged with the NSs in producing shorter reading times on this segment—thus indicating speedier filler integration—when the dependency relation was modulated by an intermediate trace (5a).

Pliatsikas and Marinis (2013) thus concluded, following Frenck-Mestre (2002) and Dussias (2003), that, provided sufficient naturalistic exposure, L2 learners appear to be capable of developing nativelike online sentence processing routines and are able to benefit from computational efficiencies, such as clause-edge reactivation to refresh fillers, while processing real-time input in their L2. As indicated by the average length of L2 immersion for the naturalistic exposure group of Pliatsikas and Marinis's study (9 years), it would seem that the development of such nativelike processing routines is a difficult goal that very few L2 learners will ever achieve. However, the notion that so-called nativelike parsing is something to be developed or acquired is inconsistent with theories of universal parsing routines that are independent of any specific language (Dekydtspotter, 2009; Slabakova, 2008). Assuming that these universal parsing routines are available to all types of language comprehenders, and, crucially, assuming that all else is equal, there should be no difference in the performance of classroom learners and naturalistic exposure learners on sentence processing tasks. Once again, it could be argued that the L1-L2 differences that have arisen in previous experimental results may reflect other issues not related to the nature of the

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

representations computed by nonnative speakers while processing their L2.

A Comparative Fallacy? Other Possible Factors

One potential difficulty with previous studies might be the type of sentence structure used. Indeed, even Pliatsikas and Marinis (2013) noted that the sentence structure under investigation, with its successive cyclic movement and complex embedding, was relatively rare in naturalistic speech. Additionally, Rodríguez (2008) questioned Marinis et al.'s (2005) interpretation of their L2 results, noting that significant differences in reading times on segments such as *the secretary claimed* and *the secretary's claim* may in fact reflect an initial processing difficulty encountered as a result of the genitive structure. The complexity of this structure, Rodríguez argued, may have induced a processing delay that lingered, rendering the integration of the *wh*-filler at the original gap site more difficult. Rodríguez also pointed out that the sentences used by Marinis et al. (and, subsequently, Pliatsikas & Marinis) involved three human referents (*the nurse*, *the doctor*, and *the patient*)—namely, three of the same kind of entity—which could create processing difficulties even for NSs (Gordon, Hendrick, & Johnson, 2004).

Furthermore, with respect to the referents themselves, previous research has shown that lexical access is slower and less automatic in a L2 (Favreau & Segalowitz, 1983; Segalowitz & Segalowitz, 1993); this must surely have consequences for sentence processing. In the context of L1 processing among older adults, Angwin et al. (2006) pointed out that "since trace reactivation effects are dependent on rapid information processing, delays in semantic activation would be expected to interfere with both the initial activation of the antecedent and its subsequent reactivation" (p. 112). Many of the antecedents in the filler-gap dependency structures that were targeted in Felser and Roberts's (2007) study referred to some fairly exotic animals (e.g., *peacock* and *ostrich*). Although classroom L2 learners may explicitly learn such vocabulary, it seems that most (even highly advanced) learners would neither encounter nor produce such words in their normal daily interactions. Thus, it is plausible that the lexical access routines to such L2 vocabulary items would be underroutinized, requiring effortful retrieval from the lexicon and creating processing lags, and thus masking any reactivation effects.

Dekydtspotter and Miller (2013) reported on a sentence processing experiment with intermediate learners of L2 English. The sentences were presented in a computer-paced reading task that involved concurrent

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

picture classification. This methodology is similar to crossmodal priming but with the experimental items presented visually rather than aurally. The experimental stimulus included sentences such as in (6), with slash marks indicating segmentation for reading.

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

Participants read each sentence segment aloud as it appeared on the computer screen and classified interrupting picture probes as *human* or *(nonhuman) animal*. Reading speed was controlled by the computer software. Picture probes appeared either at clause edge (test position #2), immediately following the complementizer *that*, or in an earlier control position inside of the prepositional phrase *on Monday* (#1). The human probes depicted either a boy or a girl, thus matching either the filler (e.g., *Harry*) or the embedded subject (e.g., *Mary*) in gender. The animal control probes appeared in the distracter items. As in crossmodal priming experiments, the classification decision was predicted to be facilitated when a picture probe matching the antecedent appeared concurrently with the movement trace—namely, at clause edge, where the intermediate trace had just induced reactivation of its antecedent.

A group of higher WM L1 Chinese–L2 English intermediate learners exhibited an antipriming pattern, with longer RTs to matching probes appearing at clause edge. This response pattern was the opposite of that revealed in the data of a NS control group, who produced the shortest RTs in this same condition. Dekydtspotter and Miller (2013) argued that although the learner and NS results differed, this finding was not evidence for any important differences in the (syntactic) representations computed by learners and by NSs. Rather, the direction of the asymmetry reflected differences in lexical access routines: The NSs were able to quickly access and then reactivate the referents of the fillers in these sentences, due to their well-established, overlearned lexical access to the familiar first names that were used in the experimental sentences. The learners, in contrast, seemed to have more difficulty activating these referents in real time. Dekydtspotter and Miller noted that this difficulty may have induced a center-surround mechanism (Carr & Dagenbach, 1990; Dagenbach, Carr, & Barnhardt, 1990), which acts as a spotlight in the semantic network to focus all resources on the weakly activated referent, thereby blocking competition from surrounding (i.e., closely related) concepts and leading to inhibition effects for the matching probes. The authors argued that such a pattern suggested that the learners did indeed show a sensitivity to intermediate structure, given that the RT asymmetry was found only at clause edge (and not in the earlier control position).

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

THE CURRENT STUDY

The current study reports on a sentence processing experiment in L2 French designed to investigate patterns of filler reactivation during online sentence processing among classroom L2 learners of French with minimal study abroad experience. In an effort to avoid (at least some of) the potential confounds that may have skewed the results of previous studies, the experimental materials were designed to minimize the computational burden.

Materials and Methodology

The main experimental task targeted indirect object cleft sentences with complex embedding such as in (7).

(7) *C'est à l'éléphant à qui*$_i$ *Sarah a découvert vendredi soir* t$_i$ *que l'on avait expliqué le nouveau jeu* t$_i$ *chez lui.*
"It's to the elephant that Sarah discovered Friday evening that someone had explained the new game at his house."

As noted by Gibson (1998) and Gibson and Warren (2004), during the real-time processing of filler-gap dependencies, the cost of maintaining a referent in memory is proportional to the number of referents introduced into the discourse context. Furthermore, as Rodríguez (2008) pointed out, having to maintain multiple referents of the same type can add to the processing load. Whereas the experimental sentences used in Marinis et al. (2005) contained three human referents (e.g., *manager*, *secretary*, and *salesman*, which were also related semantically), the sentence in (7) involves three animate referents: one animal, one human, and one indefinite (*on* "someone"). Furthermore, the cleft structure, which is very frequently used in spoken French, is predicted to be more easily processed than its relative clause counterpart; in the cleft structure, the *wh*-filler receives sentential focus, thus rendering it more salient in memory (e.g., Sanford et al., 2009). Finally, to foster speedy lexical access, all of the antecedents used in these sentences were English-French cognate animal names—that is, words that overlap in both form and meaning (e.g., *zèbre* "zebra," *crocodile* "crocodile"). The similarities between the two languages with respect to the cognate vocabulary are predicted to mitigate L2 lexical access via the L1 lexicon (e.g., Costa, Caramazza, & Sebastian-Galles, 2000), thus allowing for processing resources to be focused elsewhere.

The sentences were presented in a probe-classification-during-reading task, which pairs the classification of picture probes (as in the crossmodal

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

priming methodology) with computer-paced segmented reading aloud. Participants read aloud and must try to keep up as each segment appears on the screen, thus eliminating any opportunity to pause or to reflect on the structure, as might be possible with self-paced reading. Having participants read aloud to themselves, quietly but not silently—rather than having them listen and answer comprehension questions, as in the crossmodal priming paradigm—offers a simple alternative means of focusing attention on the experimental stimuli. Participants are monitored to ensure that they are reading aloud, although they are explicitly told that the researcher will not be listening to evaluate their L2 pronunciation (this is meant to make the learners feel more relaxed during testing). An anonymous *SSLA* reviewer noted that, without any comprehension checks in the main experimental task, there is no way to know that the L2 participants even understood the sentences that they read; they may have even chosen to focus solely on the picture-classification aspect of the task. The omission of follow-up comprehension questions was intended to make the task a little easier for these learner participants, so that they did not feel any stress that might be associated with a L2 reading comprehension exercise. The experimental design assumes that access to phonetic representations while reading aloud will induce mandatory firing of the processing system (Fodor, 1983). Additionally, it should be noted that, although they were a bit long and complex, the French sentences used in the task can be translated almost word for word into English.

The experimental stimuli consisted of 20 critical items (a complete list is included in Appendix A). A 2 × 2 × 2 design crossed test position (gap or control), probe type (antecedent matching or nonmatching), and gap type (clause edge or canonical) to create eight versions of each experimental item, as in (8a–h). The slashes represent the segmentation of each item.

(8) a. Filler-matching probe at clause edge
C'est / au zèbre / à qui / Marc / a compris / lundi / soir / que / [picture probe: ZEBRA] */ l'on / avait montré / la jolie / photo / dans sa chambre.*
b. Nonmatching probe at clause edge
C'est / au zèbre / à qui / Marc / a compris / lundi / soir / que / [picture probe: HAMMER] */ l'on / avait montré / la jolie / photo / dans sa chambre.*
c. Filler-matching probe in the matrix control position
C'est / au zèbre / à qui / Marc / a compris / lundi / [picture probe: ZEBRA] */ soir / que / l'on / avait montré / la jolie / photo / dans sa chambre.*
d. Nonmatching probe in the matrix control position
C'est / au zèbre / à qui / Marc / a compris / lundi / [picture probe: HAMMER] */ soir / que / l'on / avait montré / la jolie / photo / dans sa chambre.*
e. Filler-matching probe in the canonical gap position
C'est / au zèbre / à qui / Marc / a compris / lundi / soir / que / l'on / avait montré / la jolie / photo / [picture probe: ZEBRA] */ dans sa chambre.*

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

f. Nonmatching probe in the canonical gap position
C'est / au zèbre / à qui / Marc / a compris / lundi / soir / que / l'on / avait montré / la jolie / photo / [picture probe: HAMMER] */ dans sa chambre.*
g. Filler-matching probe in the embedded control position
C'est / au zèbre / à qui / Marc / a compris / lundi / soir / que / l'on / avait montré / la jolie / [picture probe: ZEBRA] */ photo / dans sa chambre.*
h. Nonmatching probe in the embedded control position
C'est / au zèbre / à qui / Marc / a compris / lundi / soir / que / l'on / avait montré / la jolie / [picture probe: HAMMER] */ photo / dans sa chambre.*
"It's to the zebra that Marc understood Monday evening that someone had shown the pretty picture in his room."

The picture probes were cartoon images that depicted animals that matched the referent of the antecedent or unrelated inanimate objects. The pictures were taken from digital clip art software packages and were adjusted to a uniform size. Examples of the picture probes are shown in Figure 1.

Although sentence processing research usually endeavors to control for frequency effects by matching related and unrelated probes, the design of the current study did not consider probe frequency. However, this does not seem to pose a problem for this specific task. The probes used were pictures rather than actual words, and the classification decision was not lexically based (i.e., word or nonword) but rather related to the animacy of what was pictured. The French words for the inanimate (nonmatching) probes did not appear anywhere in the experimental sentences, and participants were never required to name the animal or thing depicted. For classroom L2 learners, most of whom had little or no immersion experience, it is not clear what lexical frequency would even mean. Another potential concern with respect to the picture probes is that the matching animate picture probes depicted cognate vocabulary, whereas the nonmatching inanimate probes depicted mostly noncognate vocabulary. Cognate facilitation effects have been found among bilinguals for a variety of tasks—some of which do not



Figure 1. Examples of the alive and not alive picture probes used in probe classification during the reading task.

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

even require activation of the phonological form of the word—including picture naming (Costa et al., 2000; Poarch & van Hell, 2012), lexical decision (Dijkstra, Grainger, & Van Heuven, 1999), and word recognition (Lemhöfer et al., 2008). This could introduce a potential confound that may yield shorter RTs to matching versus nonmatching probes, thus creating an illusion of priming or masking any relevant patterns that may otherwise emerge in the data. However, Gollan, Montoya, Fennema-Notestine, and Morris (2005) compared the response patterns of monolingual native English speakers and English-Spanish bilinguals across picture-naming and picture-classification tasks and reported that the picture-classification task (in which participants indicated whether an image depicted something man-made or natural) did not reveal any differences between the participant groups, which may indicate that cognate facilitation effects are much less robust for tasks that do not require lexical retrieval.

Participants saw two versions of each critical item: one testing reactivation at the clause edge region (8a–d) and one testing reactivation at the canonical indirect object position (8e–h). The sentences were separated into two presentation blocks such that participants only saw one version of each sentence per block. Across the two blocks, participants saw a total of five items in each condition.[1] An additional 60 sentences were created as distracters. The distracter sentences included 50 direct and 10 indirect object cleft structures and involved probes that depicted animals and inanimate objects. The picture probes appeared in various positions throughout the distracter sentences, thus giving participants the impression that either type of probe could appear at any point during reading. Additionally, the number of animate and inanimate probes was balanced across the distracter items—as in the critical sentences—so that participants would make an equal number of *alive* versus *not alive* decisions over the course of the experiment. Participants completed both parts of the main experimental task in a single testing session; the two presentation blocks were separated by two additional tasks—a WM reading span test and a C-test, both described in the next section—and a background questionnaire.

Participants and Procedures

Two groups of (low and high) intermediate L1 English-L2 French learners and a group of French NS controls participated in the current study. The learner participants were undergraduate students at a large Midwestern university, enrolled in a 200-level (low intermediate) or 300-400-level (high intermediate) course at the time of testing. All of the L2 learners had begun their study of French at or after the age of 12,

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

with a mean age of acquisition of 15.1 (*SD* = 2.04 years) for the low intermediate learners and of 14.8 (*SD* = 1.75 years) for the high intermediate learners. The NSs were mostly graduate students or postdoctoral researchers at the same university, with a few additional participants from the community. All participants completed a WM reading span test adapted for L2 French from Harrington and Sawyer (1992). In this test, participants read six-to-eight-word sentences in French (a NS version of the task used sentences that had been lengthened by five or six words), which were grouped into increasingly large sets of two to five sentences. The test was administered via PowerPoint, with each slide displaying a single sentence and automatically advancing to the next slide after 5 s. As they read, participants indicated whether each sentence was grammatically correct by marking their responses on an answer sheet. At the end of each set of sentences, participants turned the answer sheet over and wrote down the last word of each sentence from that set, in order. All sentence-final words were monosyllabic nouns (debriefing revealed that these words were familiar to the learner participants). Participants were given 14–30 s, depending on set size, to write down their answers. The maximum score was 42 words remembered. Slight misspellings that did not alter the meaning of the word were accepted as correct (e.g., *math* instead of *maths* "mathematics" but not *lis* "lily" instead of *lit* "bed"); words given out of order were not. The choice to use L2 reading span as a measure of WM is consistent with previous research (i.e., Felser & Roberts, 2007). However, using this kind of test, especially in L2 research, introduces a potential confound whereby WM capacity becomes entangled with reading skill. A one-way ANOVA revealed significant differences across group means, $F(2, 52) = 7.06$, $p = .002$. Post hoc Tukey HSD tests revealed that although the NS group's scores differed from those of both the low intermediate, $p = .003$, and high intermediate, $p = .008$, learner groups, the two L2 groups did not differ from each other in terms of WM score, $p = .925$.

Participants also completed a pen-and-paper C-test consisting of two short independent paragraph-length texts of 74 and 97 words each. The first sentence of each of these paragraphs was left intact; in subsequent sentences, the second half of every other word was deleted (in cases of an odd number of letters, the larger half of the word was removed). This test, borrowed from Renaud (2010), is given in Appendix B. The maximum score was 49. A one-way ANOVA confirmed significant differences in C-test scores across participant groups, $F(2, 52) = 118.18$, $p < .001$, with post hoc comparisons using the Tukey HSD test revealing that the mean C-test scores of all three groups were statistically distinct, $p < .001$. Participant characteristics are reported in Table 1.

For the probe-classification-during-reading task, participants sat at a computer and were asked to read aloud to themselves in a low voice, paying careful attention to the sentences, and to indicate whether an interrupting image depicted something alive or not alive by pressing a

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

Table 1. Participant characteristics

Characteristic	Low intermediate (n = 20)	High intermediate (n = 20)	Native speakers (n = 15)
Age	20.50 (2.35)	20.55 (1.39)	29.80 (3.36)
Length of study (in years)	5.38 (1.60)	6.75 (1.43)	NA
Time spent abroad (in months)	0.08 (0.24)	2.51 (3.22)	NA
C-test score[a]	24.75 (3.73)	38.45 (6.63)	48.87 (1.46)
WM reading span score[b]	26.15 (6.12)	27.00 (9.36)	34.67 (4.48)

Note. Standard deviations appear in parentheses. NA = not applicable; WM = working memory.
[a] The maximum possible C-test score was 49.
[b] The maximum possible WM score was 42.

button (the left or right arrow key, respectively) as quickly as possible. Each sentence segment appeared in the center of the screen in black letters against a white background for 500 ms plus 20 ms per letter. The first segment of each sentence was preceded by a red asterisk in the center of the screen, and the end of each sentence was indicated by a period following the final word of the last segment. Picture probes were displayed for 650 ms. DMDX software (Forster & Forster, 2003) was used to measure RTs and to control reading speed. Because the sentences involved animals interacting with people in very humanlike ways, a context was created to render the sentences more felicitous: A group of French schoolchildren had spent a week in an enchanted zoo, where they befriended the animal inhabitants. Back in class the following week, they were talking with their teacher about the events of their magical holiday. This introductory context—and indeed all instructions for all of the tasks completed—was given in written and spoken English. Participants were encouraged to ask questions if they did not understand.

Analysis and Predictions

Only RTs for which an image was identified correctly as alive or not alive were included for analysis (overall, participants were highly accurate in their classifications); RTs potentially affected by display errors within the computer software were also discarded. Within each participant group, any RT that fell outside two standard deviations from the mean was removed and replaced by the new group mean.[2] These data-trimming procedures affected less than 8% of the data overall. A repeated measures ANOVA investigated the effects and interactions of

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

probe position (gap or control), probe type (matching or nonmatching), and gap type (intermediate or canonical). Planned paired-samples *t* tests, with one-tailed α set at .05 (given expectations of facilitation due to matching probes), were used to determine whether RT asymmetries were statistically significant.

If L2 learners are indeed able to compute movement traces as they read in their L2, their RTs are expected to be shorter when a picture probe appearing concurrently with the syntactic gap matches the antecedent than when the image does not match. Additionally, RTs for matching probes are predicted to be faster in the gap position than in the control position, with no RT asymmetry predicted for matching and nonmatching picture probes at the earlier control position. Such a response pattern suggests temporary storage of the *wh*-filler followed by reactivation of its referent when the gap is encountered. It is important to note that this response pattern cannot be attributed to the salience of the referent (recall that the filler receives sentential focus in the cleft structure) or to the cognate status of the word for the animal depicted in the matching probe, both of which could yield shorter RTs to matching probes regardless of test position (as in the L2 results of Felser & Roberts, 2007). Nor does this pattern reflect the linear distance between the referent and the test position, which would lead to shorter RTs for matching probes in the control position (which occurs earlier in the sentence, before the gap). Finally, for the clause-edge testing positions, it cannot be argued that this pattern may arise from simple word order expectations—that is, that the parser will detect a "missing" object following a ditransitive verb—because intermediate traces at clause edge are purely structural empty categories that would never be phonologically realized.

Additionally, if L2 learners are able to use intermediate structure to mediate long-distance dependencies—that is, to use the intermediate trace left at clause edge to reactivate the filler midway through the dependency chain—then reactivation is predicted to be similarly facilitated both at clause edge and in the canonical indirect object position (i.e., the original gap site). Based on this prediction that the effective use of the intermediate trace at clause edge in managing referents (as evidenced by priming effects) will yield a benefit further downstream for filler integration in its thematic position (with priming effects appearing at this point during processing as well), a post hoc analysis was conducted to identify those participants who showed evidence of priming at the clause edge and to examine their response patterns at the canonical gap site.

RESULTS

A 2 × 2 × 2 repeated-measures ANOVA, with test position (gap or control), probe (matching or nonmatching), and gap type (intermediate or

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

canonical) as within-subjects factors and participant group and WM score as between-subject factors,[3] revealed a main effect of position, $F_1(1, 49) = 17.40$, $p < .001$, $\eta_p^2 = .26$, $F_2(1, 57) = 4.94$, $p = .030$, $\eta_p^2 = .08$, and a main effect of probe, $F_1(1, 49) = 17.74$, $p < .001$, $\eta_p^2 = .27$, $F_2(1, 57) = 14.21$, $p < .001$, $\eta_p^2 = .20$. Additionally, participant group was a significant factor, $F_1(2, 49) = 3.22$, $p = .048$, $\eta_p^2 = .12$; however, WM score was not significant, $F_1(1, 49) = 0.86$, $p = .358$, $\eta_p^2 = .02$. Crucially, there was a significant interaction between position, probe, and gap type, $F_1(1, 49) = 4.91$, $p = .031$, $\eta_p^2 = .09$. The effects of test position and probe were maintained among the more proficient participant groups: high intermediate learners, position, $F_1(1, 19) = 5.71$, $p = .048$, $\eta_p^2 = .23$, and probe, $F_1(1, 19) = 7.08$, $p = .015$, $\eta_p^2 = .27$; NSs, position, $F_1(1, 14) = 11.83$, $p = .004$, $\eta_p^2 = .46$, $F_2(1, 19) = 6.82$, $p = .017$, $\eta_p^2 = .26$, and probe, $F_1(1, 14) = 6.48$, $p = .023$, $\eta_p^2 = .32$, $F_2(1, 19) = 9.78$, $p = .006$, $\eta_p^2 = .34$. However, these effects only approached significance among the lowest proficiency participants: low intermediate learners, position, $F_1(1, 19) = 3.24$, $p = .088$, $\eta_p^2 = .15$, and probe, $F_1(1, 19) = 3.90$, $p = .063$, $\eta_p^2 = .17$, $F_2(1, 19) = 18.13$, $p < .001$, $\eta_p^2 = .49$.

Mean RTs for each group are reported in Table 2 and depicted in Figure 2.

Filler Reactivation at Clause Edge

At clause edge, paired-samples *t* tests revealed that the NS control group responded more quickly to filler-matching probes than to nonmatching probes, $t(14) = 3.33$, $p = .005$. Similarly, matching probes were classified more quickly in the gap versus control position, $t(14) = 3.11$, $p = .008$. Crucially, however, in the earlier control position, there was no significant difference in RTs to matching and nonmatching probes, $t(14) =$

Table 2. RTs (in ms) in probe-classification-during-reading task

	Clause edge			Canonical object position		
Condition	LI ($n = 20$)	HI ($n = 20$)	NSs ($n = 15$)	LI ($n = 20$)	HI ($n = 20$)	NSs ($n = 15$)
Matching, gap	495 (38)	527 (53)	495 (59)	503 (33)	536 (51)	515 (64)
Nonmatching, gap	527 (44)	542 (68)	538 (59)	512 (34)	554 (66)	535 (59)
Matching, control	515 (38)	545 (70)	534 (58)	518 (45)	538 (55)	535 (62)
Nonmatching, control	522 (44)	559 (68)	553 (59)	526 (47)	567 (74)	561 (90)

Note. Standard deviations are given in parentheses. LI = low intermediate; HI = high intermediate; NSs = native speakers.

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

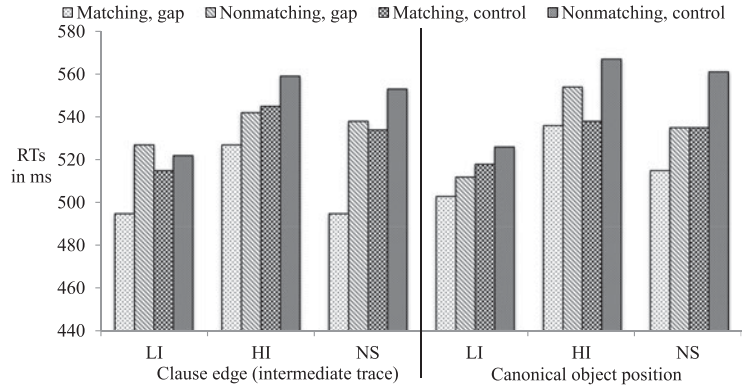


Figure 2. RTs in probe-classification-during-reading task at clause edge (left) and in canonical object position (right) for low intermediate (LI) learners (n = 20), high intermediate (HI) learners (n = 20), and native speakers (NSs; n = 15).

1.41, p = .180. The low intermediate learner group exhibited a similar response pattern, with the fastest RTs to matching probes at the gap site, when compared with RTs to nonmatching probes in the same position, $t(19)$ = 3.67, p = .002, and RTs to matching probes in the earlier control position, $t(19)$ = 2.25, p = .037. Reaction times to matching and nonmatching probes in the earlier control position did not differ, $t(19)$ = 0.73, p = .473. The high intermediate learners, however, failed to produce a statistically distinct response pattern. Although the shape of this group's results as shown in Figure 2 appears to reflect the predicted pattern, the RT asymmetries did not reach statistical significance for matching and nonmatching probes in the gap position, $t(19)$ = 1.23, p = .235, or for matching probes in the gap versus control position, $t(19)$ = 1.67, p = .112.

Filler Integration in the Canonical Indirect Object Position

As can be observed in Figure 2, in the canonical indirect object position, the results are not as robust. The shape of the NS group results is suggestive of trace reactivation consistent with filler integration at this point, with marginally faster RTs for matching versus nonmatching probes in the gap position, $t(14)$ = 1.82, p = .090, and for matching probes in the gap versus control position, $t(14)$ = 1.71, p = .109. Indeed, given one-tailed expectations of facilitation for matching probes, these results approach statistical significance. The two learner groups, however, did not produce any statistically significant asymmetries at this point during processing.

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

Interactions between Clause-Edge and Canonical Reactivations

Where the group results as a whole indicate that the filler was reactivated at clause edge, this does not necessarily mean that each and every one of that group's participants experienced this reactivation. Thus, within each participant group, all individuals who did show evidence of clause-edge reactivation were identified: Each participant's mean RTs to nonmatching probes in the gap position and to matching probes in the control position were averaged, and then their mean RTs to matching probes in the gap position were subtracted from these averages. Participants who exhibited more than a 20 ms advantage for the matching probes in the gap position were separated from those who did not (who had flat or slower RTs).[4] Twelve NSs, 13 low intermediate learners, and nine high intermediate learners were included in the *clause-edge reactivation* group (leaving three NSs, seven low intermediate learners, and 11 high intermediate learners in the *no reactivation* group). The RTs in the canonical indirect object position for these two groups are given in Table 3 and shown in Figure 3.

The visual comparison of the two groups' results in Figure 3 is striking. Overall, the participants who failed to experience facilitation priming effects at clause edge produced largely flat RTs further downstream at the canonical indirect object position, as shown in the graph on the right of Figure 3. It is important to note that based on a purely thematic strategy, reactivation would be predicted at this point during processing (recall that all groups of L2 participants in Marinis et al., 2005, exhibited longer reading times at the verbal segment). In contrast, the graph on the left of Figure 3 shows that those participants who did exhibit the predicted priming pattern at clause edge also appear to have reactivated

Table 3. RTs (in ms) in the canonical indirect object position among participants who experienced clause-edge reactivation and those who did not

	Clause-edge reactivation group			No reactivation group		
Condition	LI (n = 13)	HI (n = 9)	NSs (n = 12)	LI (n = 7)	HI (n = 11)	NSs (n = 3)
Matching, gap	496 (39)	528 (52)	498 (56)	516 (13)	543 (52)	585 (44)
Nonmatching, gap	511 (41)	569 (63)	520 (57)	513 (19)	540 (68)	593 (19)
Matching, control	513 (40)	558 (67)	522 (63)	527 (56)	521 (37)	586 (14)
Nonmatching, control	531 (54)	588 (90)	546 (92)	517 (31)	550 (57)	618 (62)

Note. Standard deviations are given in parentheses. LI = low intermediate; HI = high intermediate; NSs = native speakers.

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

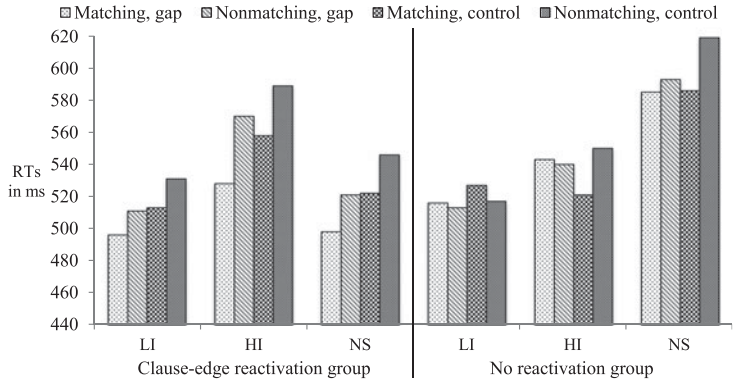


Figure 3. RTs in the canonical object position for two groups of participants: those who had reactivated the *wh*-filler at clause edge (left) and those who had not (right).

the filler again in the canonical indirect object position. Pairwise comparisons revealed marginally significant RT asymmetries at the canonical thematic position in all proficiency groups for these clause-edge reactivation participants: matching versus nonmatching probes in the gap position, NSs, $t(11) = 1.93$, $p = .079$, low intermediate learners, $t(12) = 1.87$, $p = .086$, and high intermediate learners, $t(8) = 3.69$, $p = .006$; matching probes in the gap versus control position, NSs, $t(11) = 1.78$, $p = .102$, low intermediate learners, $t(12) = 1.77$, $p = .102$, and high intermediate learners, $t(8) = 2.22$, $p = .057$. In the control position—which occurred earlier in the sentence—RTs to matching and nonmatching probes were not statistically distinct among NSs, $t(11) = 1.16$, $p = .271$, or high intermediate learners, $t(8) = 1.15$, $p = .282$; among low intermediate learners, however, RT asymmetries for matching and nonmatching probes were closer to approaching marginal significance, $t(12) = 1.66$, $p = .123$.

DISCUSSION

The results of the current study suggest that (at least some) classroom learners of French are sensitive to intermediate structure in sentences derived through cyclic movement of a *wh*-filler across a clausal boundary and that they are able to use a movement trace at clause edge to reactivate the filler, which facilitates a second reactivation in the canonical object position. Overall, the low intermediate group exhibited the predicted priming pattern at clause edge (with the shortest RTs to matching probes at the gap site) but failed to produce a discernible RT pattern further downstream in the canonical indirect object position. However, a closer examination of the data through the post hoc analysis of the

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

data of only those participants who experienced trace reactivation at clause edge revealed a more targetlike response pattern at the original gap site from which the filler was extracted. Similarly, although the high intermediate learners produced nonsignificant results for clause-edge reactivation (the shape of the pattern was nevertheless promising) as well as in the canonical position, the post hoc analysis revealed that a subset of participants from this group did appear to use the intermediate structure while processing in their L2. The same trends were found among the NSs. Although the results of the post hoc analysis should be interpreted with caution (given the reduced sample sizes), the shape of the results and the statistical trends are certainly promising. This finding seems to indicate that these L2 learners were indeed able to exploit the processing efficiencies made available to them via a universal parser: Those who reactivated the *wh*-filler at clause edge were then able to reactivate the referent again in the canonical object position to integrate the filler into the structure, thus resolving the dependency.

Assuming that parsing routines are indeed universal and therefore do not need to be learned for each (additional) language that a person acquires, and assuming L1 transfer of the relevant grammatical structures (recall that the French test sentences used in the current study can be translated almost word for word into English), then one thing that might potentially hinder processing in a L2 would be lexical access. The current study has endeavored to facilitate lexical access through the use of French-English cognate vocabulary and has also striven to simplify the testing stimuli as much as possible by targeting the indirect object cleft structure. These conscious choices were intended to facilitate processing such that evidence of structurally based sentence processing among L2 learners that may have been masked by potentially confounding factors (such as slowed lexical access or difficulties in maintaining referents in memory) would emerge, especially among intermediate classroom learners. It may seem somewhat counterintuitive, then, that a less experienced group of L2 learners (the low-proficiency group) would produce more robust results overall than the higher proficiency group, rather than the reverse. This discrepancy between the two groups and the nonsignificant results of the high intermediate learners are most likely due to this group's comparatively longer RTs and greater variability in performance. As can be observed in Figure 2, the high intermediate group was somewhat slower overall in their RTs: At clause edge and in the canonical object position, this group's fastest mean RT was the same or slower than the low intermediate group's slowest RT. Additionally, the standard deviations reported in Table 2 for the high intermediate group's RTs were slightly higher than those of the low intermediate group. It is unclear as to why this may be the case; however, given the relatively small sample sizes, these differences may simply be attributed to individual variation within the group.

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

Indeed, with this type of experiment, it can sometimes be difficult to obtain robust results, due to individual variation. In a somewhat exploratory move, a post hoc analysis was conducted that used a specific criterion—namely, a RT differential of at least 20 ms—to identify those participants who had exhibited facilitation priming effects at clause edge, suggesting reactivation of the *wh*-filler at this point during processing. This uncovered more targetlike results, especially among certain high intermediate learner participants, even though the high intermediate learners had produced nonsignificant results as a group. It is interesting to note that the post hoc analysis revealed that more of the participants in the low intermediate group (13 out of 20) experienced reactivation at clause edge as compared to the high intermediate learners (only 9 out of 20 were included in the clause-edge reactivation group). However, the high intermediate clause-reactivation group's response patterns were actually more robust than those of the low intermediate learners. Again, this seems to point to interparticipant variation as the cause of the lack of statistical significance of the high intermediate group's aggregate results.

Furthermore, the nearly even split (9 and 11 participants) of this group also underlines the importance of considering individual differences in investigations of (L2) sentence processing, which have received more focus in recent research (e.g., Hopp, 2013; Roberts, 2012; Tanner, Inoue, & Ousterhout, 2014). However, in the case of the current study, the predictor for the difference between those participants who experienced reactivation at clause edge and those who did not is unclear at this point. One-way ANOVAs within each proficiency group revealed no significant differences in the WM scores of the clause-edge reactivation subgroup and the no reactivation subgroup: NSs, $F(1, 13) = 0.02$, $p = .892$; low intermediate learners, $F(1, 18) = 1.00$, $p = .330$; high intermediate learners, $F(1, 18) = 0.51$, $p = .486$. Similarly, the C-test scores were not significantly different between the two subgroups within each learner proficiency level: low intermediate learners, $F(1, 18) = 0.95$, $p = .343$, and high intermediate learners, $F(1, 18) = 0.22$, $p = .645$. Thus, WM and proficiency do not seem to be able to predict whether a learner will be able to use intermediate structure while processing in his or her L2. Another possible explanation for this difference is the automaticity of lexical access (which has also been considered by Hopp, 2013), although the experimental design in which only French-English cognate vocabulary was used as antecedents should have minimized this type of effect. Whatever the explanation as to why only certain individuals benefited from the use of intermediate structure, the findings of the current study demonstrate that targetlike response patterns may indeed be obscured in the results of L2 processing research. Previous arguments that L2 learners are unable to process target language input in the same detailed ways as NSs may have been put forth a bit too early. It will be

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

important in future research to consider the whole range of possible factors that may affect sentence processing routines, in both the L1 and the L2.

CONCLUSION

The current study looked for evidence that classroom learners would be able to use intermediate structure to mediate long-distance dependencies that cross clausal boundaries. Intermediate traces at clause edge can serve to refresh the referent of the filler in memory, thus facilitating integration into the structure at the canonical position and making it easier to maintain referents in memory across long distances during sentence processing. Most of the NSs (12 out of 15) produced a response pattern consistent with antecedent reactivation at clause edge, followed by a second reactivation in the canonical position. Although in the aggregate data the high intermediate learners failed to produce statistically significant results and the low intermediate learners appeared to reactivate the *wh*-filler at clause edge only (with no subsequent reactivation in thematic position), further inspection of the data found that a subgroup of each of these proficiency groups did produce the expected response pattern. The results of the current study suggest that (at least some) L2 learners are able to make use of chain computations (involving intermediate traces) in managing referential loads, as predicted by syntactic theory and universal parsing routines.

Received 27 June 2013
Accepted 10 April 2014
Final Version Received 11 June 2014

NOTES

1. As an anonymous *SSLA* reviewer noted, the number of items encountered in each condition by each participant may seem a bit low, given that some RTs will inevitably be discarded due to inaccurate classification decisions, display errors, and outlier status. However, with eight different experimental conditions (four test positions and two probe types), having participants read even just a few more items in each condition would make for a much longer task. Furthermore, it seems that this distribution is consistent with previous research of this type (e.g., Felser & Roberts, 2007; Marinis et al., 2005).

2. Missing values were replaced by the group mean rather than that individual participant's mean due to the small number of items per condition.

3. F_1 values refer to the by-subject analysis, whereas F_2 values refer to the by-item analysis. These values are reported where significant.

4. Recall that Dekydtspotter and Miller (2013) argued that inhibition patterns (i.e., longer rather than shorter RTs for matching probes in the gap position) were consistent with filler reactivation interacting with underroutinized lexical access. However, given that the current study used only cognates as antecedents in the dependency

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

structure, with the specific aim of facilitating lexical access, those (learner) participants who showed inhibition effects were not included in the clause-edge reactivation group with those who showed priming effects.

REFERENCES

Angwin, A. J., Chenery, H. J., Copland, D. A., Cardell, E. A., Murdoch, B. E., & Ingram, J. C. L. (2006). Searching for the trace: The influence of age, lexical activation and working memory on sentence processing. *Journal of Psycholinguistic Research*, *35*, 101–117.

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

Chomsky, N. (1986). *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. (2005). *Rules and representations*. New York, NY: Columbia University Press.

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.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1283–1296.

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: Learning, Memory, and Cognition*, *16*, 328–340.

Dekydtspotter, L. (2009). Second language epistemology: Take two. *Studies in Second Language Acquisition*, *31*, 291–321.

Dekydtspotter, L., Donaldson, B., Edmonds, A. C., Liljestrand-Fultz, A., & Petrush, R. A. (2008). Syntactic and prosodic computations in the resolution of relative clause attachment ambiguity by English-French learners. *Studies in Second Language Acquisition*, *30*, 453–480.

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., 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.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*, 496–518.

Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learners: Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition*, *25*, 529–557.

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first and second language reading of fluent bilinguals. *Memory & Cognition*, *11*, 565–574.

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

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

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, the Netherlands: Benjamins.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology*, *50A*, 119–148.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Gibson, E., & Warren, T. (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax*, *7*, 55–78.

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

Gollan, T., Montoya, R. L., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*, 1220–1234.
Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*, 97–114.
Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*, 25–38.
Hopp, H. (2013). Individual differences in the second language processing of object-subject ambiguities. *Applied Psycholinguistics*. Advance online publication. doi:10.1017/S0142716413000180
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.
Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 12–31.
Love, T. (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., & Swinney, D. (1996). Coreference processing and levels of analysis in object-relative constructions: Demonstration of antecedent reactivation in the cross-modal priming paradigm. *Journal of Psycholinguistic Research*, *25*, 5–24.
Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, *27*, 53–78.
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.
Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constraints in second language sentence processing. *Studies in Second Language Acquisition*, *33*, 563–588.
Pickering, M., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes*, *6*, 229–259.
Pliatsikas, C., & Marinis, T. (2013). Processing empty categories in a second language: When naturalistic exposure fills the (intermediate) gap. *Bilingualism: Language and Cognition*, *16*, 167–182.
Poarch, G. J., & van Hell, J. G. (2012). Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *111*, 419–438.
Renaud, C. (2010). *On the nature of agreement in English-French acquisition: A processing investigation in the verbal and nominal domains* (Unpublished doctoral dissertation). Indiana University, Bloomington.
Roberts, L. (2012). Individual differences in second language sentence processing. *Language Learning*, *62*, 172–188.
Roberts, L., Marinis, T., Felser, C., & Clahsen, H. (2007). Antecedent priming at trace positions in children's sentence processing. *Journal of Psycholinguistic Research*, *36*, 175–188.
Rodríguez, G. A. (2008). *Second language sentence processing: Is it fundamentally different?* (Unpublished doctoral dissertation). University of Pittsburgh, Pennsylvania.
Sag, I., & Fodor, J. D. (1994). Extraction without traces. In R. Aranovich, W. Byren, S. Preuss, & M. Senturia (Eds.), *Proceedings of the 13th West Coast Conference on Formal Linguistics* (pp. 365–384). Stanford, CA: CSLI Publications.
Sanford, A. J. S., Price, J., & Sanford, A. J. (2009). Enhancement and suppression effects resulting from information structuring in sentences. *Memory & Cognition*, *37*, 880–888.
Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, *14*, 369–385.
Slabakova, R. (2008). *Meaning in the second language*. Berlin, Germany: Mouton de Gruyter.

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

Tanner, D., Inoue, K., & Ousterhout, L. (2014). Brain-based individual differences in on-line L2 grammatical comprehension. *Bilingualism: Language and Cognition*, *17*, 277–293.

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.

Williams, J. N., Möbius, P., & Kim, C. (2001). Native and non-native processing of English *wh*-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics*, *22*, 509–540.

APPENDIX A

EXPERIMENTAL SENTENCES (AND PICTURE PROBES IN ALL CAPS)

1. *C'est au lion à qui Christine a insisté lundi matin que l'on avait donné le beau chapeau dans le jardin.* LION WAGON
   "It's to the lion that Christine insisted Monday morning that someone gave the nice hat in the garden."
2. *C'est au panda à qui Marie a supposé mardi matin que l'on avait montré le beau dessin avant l'école.* PANDA COAT
   "It's to the panda that Marie supposed Tuesday morning that someone showed the nice drawing before school."
3. *C'est au rat à qui Sophie a imaginé mercredi soir que l'on avait offert le dernier gâteau après la fête.* RAT DART
   "It's to the rat that Sophie imagined Wednesday evening that someone gave the last cake after the party."
4. *C'est au gorille à qui Cécile a observé jeudi soir que l'on avait envoyé le petit cadeau chez lui.* GORILLA VIOLIN
   "It's to the gorilla that Cecile observed Thursday evening that someone sent the little gift at his house."
5. *C'est à l'éléphant à qui Sarah a découvert vendredi soir que l'on avait expliqué le nouveau jeu chez lui.* ELEPHANT BIKE
   "It's to the elephant that Sarah discovered Friday evening that someone explained the new game at his house."
6. *C'est à l'hippopotame à qui Nicole a compris lundi soir que l'on avait prêté le vieux vélo dans le parc.* HIPPO COMPUTER
   "It's to the hippopotamus that Nicole understood Monday evening that someone lent the old bike in the park."
7. *C'est au crocodile à qui Charlotte a su mardi matin que l'on avait emprunté le mauvais livre avant l'école.* CROCODILE PAPER CLIP
   "It's from the crocodile that Charlotte knew Tuesday morning that someone borrowed the wrong book before school."
8. *C'est à la panthère à qui Julie a insisté mercredi soir que l'on avait raconté la bonne histoire après la fête.* PANTHER BOX
   "It's to the panther that Julie insisted Wednesday evening that someone told the good story after the party."

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

9. *C'est au hamster à qui Sylvie a supposé que jeudi soir l'on avait prêté le beau pullover pendant le film.* HAMSTER GIFT
"It's to the hamster that Sylvie supposed Thursday evening that someone lent the nice sweater during the movie."
10. *C'est au léopard à qui Claire a imaginé vendredi matin que l'on avait donné le nouveau stylo en classe.* LEOPARD CLOCK
"It's to the leopard that Claire imagined Friday evening that someone gave the new pen in class."
11. *C'est au tigre à qui Vincent a observé lundi matin que l'on avait envoyé la petite carte pour dire merci.* TIGER CRAYON
"It's to the tiger that Vincent observed Monday morning that someone sent the little card to say thank you."
12. *C'est au zèbre à qui Marc a découvert mardi soir que l'on avait montré la jolie photo dans sa chambre.* ZEBRA HAMMER
"It's to the zebra that Marc discovered Tuesday evening that someone showed the pretty picture in his room."
13. *C'est au lézard à qui Luc a compris mercredi soir que l'on avait offert le grand cadeau pendant le dîner.* LIZARD LAMP
"It's to the lizard that Luc understood Wednesday evening that someone gave the big gift during dinner."
14. *C'est à la girafe à qui François a su jeudi soir que l'on avait expliqué le long devoir après l'école.* GIRAFFE MAILBOX
"It's to the giraffe that François learned Thursday evening that someone explained the long homework after school."
15. *C'est au rhinocéros à qui Antoine a insisté vendredi soir que l'on avait envoyé les belles fleurs avant la fête.* RHINO TELEPHONE
"It's to the rhinoceros that Antoine insisted Friday evening that someone sent the pretty flowers before the party."
16. *C'est au serpent à qui David a supposé lundi matin que l'on avait expliqué le gros problème dans son bureau.* SNAKE NOTEBOOK
"It's the snake to whom David supposed Monday morning that someone explained the big problem in his office."
17. *C'est au kangourou à qui Thomas a imaginé mardi soir que l'on avait donné la bonne orange au pique-nique.* KANGAROO FLOWERPOT
"It's to the kangaroo that Thomas imagined Tuesday evening that someone gave the good orange at the picnic."
18. *C'est à l'insecte à qui Nicolas a observé mercredi matin que l'on avait prêté le grand sac dans le magasin.* INSECT JEWEL
"It's to the insect that Nicolas observed Wednesday morning that someone lent the big bag in the store."
19. *C'est au pélican à qui Robert a découvert jeudi matin que l'on avait offert le nouveau ballon dans le jardin.* PELICAN BALLOON
"It's to the pelican that Robert discovered Thursday morning that someone gave the new ball in the garden."
20. *C'est à la gazelle à qui Philippe a compris vendredi soir que l'on avait envoyé la longue lettre après l'école.* GAZELLE SHOVEL
"It's to the gazelle that Philippe understood Friday evening that someone sent the long letter after school."

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

APPENDIX B

C-TEST

The Task

Un livre qui prétend introduire des aspects de la culture française ne serait pas complet sans un chapitre sur les beaux-arts. En fa____, de nomb_____ touristes vo____ en Fra____ dans l'inte_____ d'admirer s____ chefs-d'oeu____ de pein_____, d'archit_____ et d____ sculpture. Q____ n'a p____ entendu par____ du Louvre? d____ la cath_____ Notre-Dame d____ Paris? des scul_____ de Rodin? No____ ne pou_____ pas vo____ présenter u____ étude e____ profondeur d____ beaux-ar____ en Fra_____.

Quand on revient d'un voyage dans un pays étranger, la première chose dont on se souvient est presque toujours la cuisine: non seul_____ la nourr_____ mais au____ la fa____ de l____ préparer, d____ la man____, les heu____ des re____, tous l____ rites q____ les accomp_____ et q____ caractérisent l____ gens d____ pays mi____ que n'imp_____ quel au____ aspect d____ la v____. En Fra____, la gastr_____ est particul_____ importante, c____ c'est u____ véritable art; et il ne s'agit pas d'un art pratiqué par un petit nombre de spécialistes, mais d'un art auquel participe toute la population.

Answer Key and Translation

Un livre qui prétend introduire des aspects de la culture française ne serait pas complet sans un chapitre sur les beaux-arts. En fait, de nombreux touristes vont en France dans l'intention d'admirer ses chefs-d'oeuvre de peinture, d'architecture et de sculpture. Qui n'a pas entendu parler du Louvre? de la cathédrale Notre-Dame de Paris? des sculptures de Rodin? Nous ne pouvons pas vous présenter une étude en profondeur des beaux-arts en France.

"A book that claims to introduce aspects of French culture would not be complete without a chapter on fine arts. In fact, many tourists go to France with the intention of admiring its masterpieces of painting, architecture, and sculpture. Who hasn't heard of the Louvre? of the Notre-Dame cathedral in Paris? of Rodin's sculptures? We are not able to give you an in-depth study of France's fine arts."

Quand on revient d'un voyage dans un pays étranger, la première chose dont on se souvient est presque toujours la cuisine: non seulement la

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press

nourriture mais aussi la façon de la préparer, de la manger, les heures des repas, tous les rites qui les accompagnent et qui caractérisent les gens du pays mieux que n'importe quel autre aspect de la vie. En France, la gastronomie est particulièrement importante, car c'est un véritable art; et il ne s'agit pas d'un art pratiqué par un petit nombre de spécialistes, mais d'un art auquel participe toute la population.

"When someone returns from a trip to a foreign country, the first thing they remember is almost always the food: not just the food itself but also the way that it's prepared and eaten, meal times, all of the rituals that come with it and that characterize the country's people better than any other part of life. In France, fine food is especially important because it's a true art form; and it's not an art that only a small number of specialists practice, but one in which the entire population participates."

https://doi.org/10.1017/S0272263114000588 Published online by Cambridge University Press