THE ROLE OF MEMORY PROCESSES AND QUALITY OF LEXICAL REPRESENTATIONS IN NATIVE AND NON-NATIVE READING COMPREHENSION

BY

NAYOUNG KIM

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Educational Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Kiel Christianson, Chair
Professor Gary Dell
Professor Carolyn Anderson
Associate Professor Jill Jegerski

ABSTRACT

This dissertation investigated what memory mechanisms support parsing and how they constrain sentence comprehension during first-language (L1) and second-language (L2) sentence reading. Although two memory-based accounts in sentence processing research, the capacity-based model (Just & Carpenter, 1992; King & Just, 1991) and the cue-based retrieval model (McElree & Vasishth, 2005; McElree, Vasishth, & Van Dyke, 2006; Nicenboim & Vasishth, 2018), demonstrated memory mechanisms supporting sentence comprehension, how readers access linguistic representations outside focal attention during reading is a largely unresolved issue, especially in L2 processing. Thus, the current research compared the predictions of the cue-based retrieval model and the capacity-based model in sentence comprehension using eye-tracking. Based on previous evidence for the cue-based retrieval model, this dissertation also examined whether enhancing the quality of lexical representations through semantic elaboration influences retrieval efficiency, given the assumption that providing additional semantic information for the target and/or the distractor increases the uniqueness of the target representation in memory by reducing similarity-based retrieval interference. Importantly, in order to understand whether the ability to use an efficient, cue-driven operation determines skilled versus less-skilled reading, L1 and L2 speakers' reading patterns were compared. The findings that both L1 and L2 readers were sensitive to similarity-based retrieval interference during sentence comprehension suggest that sentence processing relies on a series of cue-based retrievals, but the ability to employ this operation itself may not distinguish skilled reading from less-skilled reading. In particular, the observed L1-L2 differences in reading patterns suggest that the most likely predictor of reading ability may be individuals' quality of lexical representation.

TABLE OF CONTENTS

Chapter 1. Introduction

A clear understanding of memory mechanisms that support sentence comprehension is important to understand reading ability. The notion that memory capacity is a central bottleneck to comprehension has impacted numerous studies on reading ability. These studies have demonstrated that comprehension difficulty tends to increase as working memory (WM) capacity decreases. WM capacity has generally been manipulated in this line of research either by the use of dual-tasks (Fedorenko, Gibson, & Rohde, 2006, 2007; Gordon, Hendrick, & Levine, 2002) or by participant selection, as in participants who scored high on WM capacity measures compared with those who scored low (Just & Carpenter, 1980, 1992; Just, Carpenter, & Keller, 1996; Just & Varma, 2002, 2007). Important assumptions of the WM capacity model are that WM – the ability to temporarily maintain and manipulate information simultaneously – is an independent system from long-term memory (LTM), and also that WM is fixed in its capacity (see Adams, Nguyen, & Cowan, 2018; Baddeley, 2012; Caplan & Waters, 2013; Pickering & van Gompel, 2006, for reviews). The capacity-based account thus predicts that sentence comprehension becomes more difficult as the storage and processing demands of a sentence increase, and comprehension fails when these demands exceed some fixed WM capacity (Just & Carpenter, 1992; Just, Carpenter, & Keller, 1996).

On the other hand, a cue-based retrieval account (Lewis & Vasishth, 2005), a recent alternative approach to sentence processing, claims that sentence comprehension relies on a cue-driven, direct-access operation, rather than some fixed capacity. This account assumes a unitary store model where information that the capacity account assigns in WM is in fact the activated portion of LTM (Anderson et al., 2004; Crowder, 1976, Cowan, 2001, 2006; Ericsson & Kintsch, 1995, McElree, 2001, 2006; Oberauer, 2002; Verhaeghen, Cerella, & Basak, 2004; Van Dyke &

Shankweiler, 2012). Thus, the cue-based retrieval model is based on the assumptions that memory representations are stored only in LTM, which is content-addressable and directly accessible via a cue-driven retrieval operation (Clark & Gronlund, 1996; Dosher & McElree, 2002; Kohonen, 1984), and that representations in LTM vary in activation strength, with the most activated representation available for retrieval when required (Lewis & Vasishth, 2005; Van Dyke & Shankweiler, 2012). Importantly, cue-based retrieval posits similarity-based interference as the source of comprehension difficulty. When retrieval cues are associated not only with target representations but also with other representations in memory (i.e. "cue-overload," Van Dyke & Johns, 2012), this creates retrieval interference because the amount of activation available for boosting matching cues is shared between the target and other items, increasing the likelihood that an incorrect item will be retrieved during sentence reading.

Although both the capacity-based and the cue-based retrieval accounts have demonstrated memory mechanisms that support sentence comprehension, how readers mentally encode and retrieve linguistic representations in memory during sentence processing still remains controversial. This dissertation, therefore, tested the predictions of the cue-based retrieval model in sentence comprehension using eye-tracking, in comparison to those of the capacity-based model. Based on the evidence for cue-based retrieval interference from previous research, this dissertation also investigated whether enhancing the quality of lexical representations of the target and/or other items through semantic elaboration increases retrieval efficiency, given the assumption that semantic elaboration for the target and/or other items increases the uniqueness of the target representation in memory, leading to reduced similarity-based retrieval interference during reading. Along with the experimental manipulations, the relationships between individual difference measures (here, WM and reading speed) and retrieval ability during sentence comprehension were also explored in order to understand whether these individual difference measures are determinant predictors of reading ability. Importantly, less-skilled second language (L2) readers were included in the current research in comparison to skilled native language (L1) readers to examine whether the ability to use an efficient cue-based retrieval operation during reading determines skilled versus less-skilled reading.

The findings of the current study are important in that they contribute to the understanding of what memory mechanisms support parsing, and how they constrain sentence comprehension. Supportive evidence for the cue-based retrieval account specifically suggests that language processing may be subject to domain-general principles and constraints (Lewis et al, Vasishth, & Van Dyke, 2006). Also, given that the cue-based retrieval account does not make predictions about the features that are not cued by the retrieval trigger, examining the effects of semantic elaboration on retrieval efficiency is import to further develop the cue-based retrieval model. Furthermore, less-skilled L2 readers' data shed light on how L2 parsing is accomplished, what types of cues L2 readers rely on, and where L2 comprehension difficulty arises during reading. These findings are expected to be informative for designing L2 classroom reading instruction approaches. In particular, comparisons of L1 and L2 readers' sensitivity to semantic and syntactic interference will help us better understand the source of L1-L2 differences in sentence processing and provide a more diagnostic test for the cue-based retrieval model.

Chapter 2. Literature Review

In this chapter, I first discuss two well-established memory-based theories in sentence processing research, the capacity-based account and the cue-based retrieval account. Then I introduce the importance of the quality of lexical representations in successful target retrieval, and review empirical studies demonstrating semantic elaboration effects on efficient retrieval processing in sentence comprehension. In order to understand how skilled L1 and less-skilled L2 readers are different during sentence comprehension, the next section discusses L2 theories demonstrating L2 readers' comprehension bottleneck, especially focusing on L2 readers' inability to compute fully specified syntactic representations or capacity differences during sentence processing. Given that reading comprehension depends not only on the characteristics of the sentence material being read, but also on the individual reading abilities of the reader, the last section focuses on two individual difference measures, WM capacity and reading speed, and their relationships with retrieval ability during sentence comprehension.

Capacity-based account

The capacity-based approach in sentence processing research suggests that any constraint that memory imposes on sentence processing stems from capacity limits on WM, which is the ability to maintain and manipulate information simultaneously (see Baddeley, 2012 for a recent review). The important assumptions of the capacity-based model are that WM and LTM are separate, independent systems of information processing, and that WM is a fixed capacity resource (see Caplan & Waters, 2013; Adams et al., 2018; Pickering & van Gompel, 2006, for reviews).

The findings of Carpenter, Miyake, and Just (1994) that readers with brain injury or disease showed intact lexicon and production rules, but impaired storage and processing of language, provide supportive evidence for the idea that WM is a separate system from LTM. The assumption of the limited capacity of WM in the capacity-based account was motivated by the multicomponent WM model (Baddeley, 2000; Baddeley & Hitch, 1974), where WM is divided into the phonological loop (holding verbal information), the visuospatial store (holding visual and spatial information), the episodic buffer (holding semantic information and associations between different types of information), and the central executive (responsible for attentional control and information manipulation).

Baddeley's WM model was originally developed to demonstrate the findings of memory recall studies where participants were asked to retain lists of items while carrying out other processes, and then recall the lists at the end of the task (Baddeley, 1966; Conrad, 1964; Murray, 1968; Wickelgren, 1965). Studies examining performance of these complex WM tasks, for example reading span (Daneman & Carpenter, 1980) or operation span (Turner & Engle, 1989), have shown that participants' performance declines rapidly with an increase in memory demand in various experimental tasks, and that WM capacity is strongly correlated with higher-order cognitive abilities, for example, reasoning ability, scholastic aptitude, fluid intelligence, and executive function (Cowan et al., 2005; Engle, Tuholski, Laughlin, & Conway, 1999; Fukuda, Vogel, Mayr, & Awh, 2010; Kyllonen & Christal, 1990; Miyake, Friedman, Rettinger, Priti, & Hegarty, 2001). Although some aspects of Baddeley's working memory model have recently been questioned, especially the relationship between WM and LTM, this model has also had a tremendous impact on models of sentence processing (Caplan & Waters, 1999; Just & Carpenter, 1992; King & Just, 1991).

Critical evidence for the capacity-based approach to sentence comprehension comes from studies demonstrating increased comprehension difficulty due to reduced WM capacity. These studies tested sentences with increased processing load or distance (e.g., sentences with high syntactic complexity, ambiguities, or long-distance dependencies), and used either a dual-task paradigm, where memory load was manipulated as an index of WM capacity (Gordon et al., 2002; Fedorenko et al., 2006, 2007) or an individual-differences paradigm, in which participants who performed well on a measure of WM capacity were compared with those who performed poorly (Just & Carpenter, 1980, 1992; Just et al., 1996; Just & Varma, 2002, 2007).

King and Just (1991) and Just and Carpenter (1992), for example, showed that low-span readers, which were classified using the reading span task (Daneman & Carpenter, 1980), showed lower comprehension accuracy and slower reading times (RTs) at the main verb of syntactically complex sentences (object- vs. subject-extracted relative clauses) compared to high-span readers. Similarly, MacDonald, Just, and Carpenter (1992) reported differential reading patterns of low-span and high-span participants, such that high-span participants showed longer RTs for temporarily ambiguous sentences compared to unambiguous controls. MacDonald et al. suggested that these patterns were attributed to high-span readers maintaining more alternative interpretations of the ambiguity than did low span participants.

Cue-based retrieval account

Despite the prevalence of the capacity-based account (Just & Carpenter, 1992; Just, Carpenter, & Keller, 1996; King & Just, 1991), this approach has been challenged by recent evidence that the capacity for temporarily maintaining information is extremely limited (Cowan, 2001; McElree, 2006; Lewis et al., 2006; Oberauer, 2002; Van Dyke & Johns, 2012; Van Dyke, Johns, Kukona, 2014). This evidence is critical, because it implies that *capacity* is extremely limited for everybody, and consequently that capacity differences may not explain the variability in reading performance. In other words, WM capacity, which has been suggested as the primary comprehension bottleneck for decades, may not be a significant predictor of reading ability.

The direct evidence of severely limited WM capacity comes from the studies that used the speed-accuracy tradeoff (SAT) procedure (Dosher, 1979; Reed, 1973, 1976; Wickelgren, 1977), which provides conjoint measures of accuracy (an index of quality of information processing) and speed (an index of speed that information is computed or accessed). Measures of processing speed are especially informative to understand whether or not information that is assumed to be stored in WM remains more accessible than information in LTM, and this ultimately provides information about parsing mechanisms underlying sentence processing.

McElree (2006), for example, used the SAT paradigm and demonstrated that items predicted to be within WM did not show privileged access, but rather were retrieved with the same speed as items assumed to be in LTM. This suggests that the same type of retrieval operation may be employed for accessing representations in WM and LTM (Oztekin & McElree, 2006), or that as mentioned earlier, representations may be stored only in LTM, and information assumed to be in WM may be in fact activated portion of LTM (Anderson et al., 2004; Crowder, 1976, Cowan, 2001, 2006; Ericsson & Kintsch, 1995, McElree, 2001, 2006; Oberauer, 2002; Verhaeghen et al., 2004; Van Dyke & Shankweiler, 2012). Either way, his findings provide evidence against the capacity-based account that holds WM capacity to be a comprehension bottleneck. It is important to note that the only speed difference observed in his study was between items in focal attention (the items currently being processed) and all other items outside focal attention.

McElree (2000) and McElree et al. (2003) extended the logic used in memory research to examine retrieval operations in sentence comprehension with long-distance dependencies. They manipulated the amount of interpolated linguistic material between the two dependent constituents (see sentences in 1a-1c), and sentences were visually presented one word at a time.

(1a) *This was the book that the editor admired/amused*.

(1b) T*his was the book that the editor* ***believed that the publisher*** *admired/amused.*

(1c) *This was the book that the editor* ***believed that the journalist reported that the publisher*** *admired/amused*.

Participants were asked to make binary acceptability decisions after the onset of the final word in the sentence, and their response latency was measured along with the acceptability accuracy. As expected, the results revealed that as the amount of interpolated material increased, the probability of computing an acceptable interpretation decreased. However, what was very interesting was that the manipulation of the distance between the two dependent constituents did not affect the speed of comprehension (i.e. response latency), which is comparable to what is found in basic memory research (no privileged access for the items assumed to be within WM). Similar patterns were observed for other various dependencies, such as verb phrase ellipsis (Martin & McElree, 2009, 2011) and pronoun resolution (Foraker & McElree, 2007).

Another reason why the capacity-based approach may not accurately predict the comprehension bottleneck is that, as Van Dyke and Johns (2012) pointed out, the capacity-based account emphasizes memory decay as a source of comprehension difficulty. That is, information that is not actively maintained in memory decays faster, resulting in comprehension difficulty. According to Van Dyke and Johns (2012), the decay hypothesis for forgetting seems to be problematic, because prior memory research has shown that the primary source of forgetting is interference, rather than decay (Berman, Jonides, & Lewis, 2009; Underwood & Keppel, 1962; Waugh & Norman, 1965). Interference in memory retrieval has been known to arise when retrieval cues match with not only the target but also partially with other items in the memory (i.e. "cue-overload," Van Dyke & Johns, 2012), and thus when there are other items that share similar features with the retrieval cues, the probability of target retrieval is likely to decrease (Nairne, 2002; Oztekin & McElree, 2007; Watkins & Watkins, 1975).

Given the recent evidence of extremely limited WM capacity and similarity-based retrieval interference in memory research, Lewis and Vasishth (2005) proposed an alternative approach to sentence processing, the cue-based retrieval hypothesis, which suggests that sentence comprehension relies on a cue-driven, direct-access operation, and that similarity-based retrieval interference is the primary source of comprehension difficulty. According to this model, parsing is accomplished through a series of rapid cue-based retrievals, and that each incoming word triggers memory retrievals to integrate that word with the previously constructed structure. The retrieval cues, grammatically and contextually derived from the incoming word, are a subset of features of the to-be-retrieved target, and what is being retrieved is target representations (feature bundles), not the target word. This cue-based parsing model assumes that comprehension difficulty increases when the similarity between the target and the distractors increases.

Computational evidence for the cue-based retrieval theory comes from a model, implemented in the Adaptive Control of Thought-Rational (ACT-R; Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006). In ACT-R, items are stored as chunks in content-addressable memory, and an item’s retrieval latency and probability are governed by the item’s level of activation at the time of retrieval.

$$(2) \quad A_i = \ln\left(\sum_r t_{ri}^{-d}\right) + \sum_j W_j\left(S - \ln(fan_j)\right) - \sum_k P_k M_{ki} + \epsilon$$

The mathematical expression in (2) (from Parker, Shvartsman, & Van Dyke, 2017) describes the model implemented in the ACT-R. In (2), A_i indicates the activation level of an item *i*, which is the sum of four terms. The first term is the item's base activation, with t_{ri} indicating the time since the *r*th retrieval of the item. The parameter *d* is a constant, usually estimated to be 0.5 in most ACT-R models (Anderson et al., 2005). The second term reflects similarity-based inhibitory interference (i.e. increased RTs due to retrieval cues matching both target and distracting items). W_j indicates weights associated with elements of the target chunks, and S indicates the maximum associative activation boost received from retrieval cues, which is reduced by $\ln(fan_j)$, the number of items associated with the retrieval cues. The third term reflects facilitatory interference (e.g. reduced RTs due to the retrieval cues matching distracting items), P_k the partial matching penalty, applied to each cue *k* in proportion (M_{ki}). The last term indicates a noise term from the logistic distribution at retrieval.

A number of empirical findings have demonstrated that retrieval latency and probability in long-distance dependencies are affected by the similarity of distractors to the retrieval cues that are used to access the target (Lewis et al., 2006; McElree, 2000; McElree, Foraker, & Dyer, 2003; Tan et al., 2017; Van Dyke, 2007; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006, 2011). Van Dyke (2007) examined the effects of syntactic and semantic interference during L1 sentence comprehension by manipulating the syntactic and semantic properties of the intervening noun (distractor) between the subject and the verb, creating high and low syntactic and semantic interference conditions (see 3a-3d sentences below). In the low syntactic interference conditions (3a and 3b), the syntactic feature of the distractor *seat/man* [+Object] does not match with the syntactic feature of the target *lady* [+Subject]. On the other hand, in the high syntactic interference conditions (3c and 3d), the syntactic feature of the distractor *seat/man* [+Subject] matches with the syntactic feature of the target *lady* [+Subject]. Similarly, in the low semantic interference conditions (3a and 3c), the semantic feature of the distractor *seat* [+Inanimate] does not overlap with the semantic feature of the target *lady* [+Animate], whereas in the high semantic interference conditions (3b and 3d), both the distractor and the target have the animate semantic feature.

(3a) Low-Syntactic, Low-Semantic

The pilot remembered that the ***lady*** who was sitting in the smelly **seat** yesterday afternoon ***moaned*** about a refund for the ticket.

(3b) Low-Syntactic, High-Semantic

The pilot remembered that the ***lady*** who was sitting in the smelly **man** yesterday afternoon ***moaned*** about a refund for the ticket.

(3c) High-Syntactic, Low-Semantic

The pilot remembered that the ***lady*** who said that the **seat** was smelly yesterday afternoon ***moaned*** about a refund for the ticket.

(3d) High-Syntactic, High-Semantic

The pilot remembered that the ***lady*** who said that the **man** was smelly yesterday afternoon ***moaned*** about a refund for the ticket.

In order to examine the effects of semantic and syntactic interference on retrieval latency and probability, RTs on the retrieval region (*moaned*) and the spillover region (*about a refund*), and comprehension question-response accuracy were analyzed. As the cue-based retrieval account predicts, the results showed that the high semantic and syntactic interference conditions elicited longer RTs in the critical regions and lower question-response accuracy. However, no reliable two-way interaction was observed between the two fixed effects, suggesting that all retrieval cues may not be multiplicatively combined into a single retrieval probe.

Tan, Martin, and Van Dyke (2017) also examined English L1 readers' sensitivity to semantic and syntactic interference, with the sentences adopted from Van Dyke (2007). Their results revealed similar reading patterns as those in Van Dyke (2007), such that readers showed lower comprehension accuracy and longer RTs at the retrieval cites for the high semantic and syntactic interference conditions. Their findings, therefore, provided supportive evidence for the direct-access, cue-driven parsing mechanism during sentence comprehension. What was different from Van Dyke (2007) was that they additionally investigated the influence of WM capacity on retrieval interference in sentence comprehension. Their analysis of the three-way interaction between WM capacity and two fixed effects (semantic and syntactic interference) revealed reliable interaction, showing that high WM capacity readers were less susceptible to interference effects compared to low-span readers.

Quality of Lexical Representations on Retrieval

Assuming that the direct-access, cue-driven operation is the underlying memory mechanism supporting language comprehension, the next important question is what creates comprehension bottleneck. As discussed above, when the retrieval cues are associated not only with the target but also with other items in memory (i.e. cue-overload), similarity-based retrieval interference occurs, resulting in comprehension difficulty. Retrieval efficiency could also be reduced when the target and other distractors in memory have low-quality lexical representations. According to Perfetti (2007), low-quality representations refer to those whose linguistic features are not fully specified, for example, words with a greater variability in grapheme-phoneme correspondence or fewer meaning dimensions. Thus, accessing items whose features are underspecified during reading may increase spurious activations of irrelevant linguistic information, adding noise to processing, and makes it difficult for readers to discriminate items with similar features, leading to increased retrieval interference.

Examining the role of lexical representation quality on retrieval efficiency is particularly important to further develop the cue-based retrieval model, because the cue-based retrieval model does not make predictions about the features that are not cued by the retrieval trigger, for example semantic or syntactic complexity. Thus, following up on an earlier study by Hofmeister and Vasishth (2014), the current study tested whether or not enhancing the quality of lexical representations of the target and the distractor through semantic elaboration facilitates retrieval processing during sentence comprehension. The underlying assumption was that semantic elaboration for the target and/or the distractor would increase uniqueness of the target representations, and consequently, reduce similarity-based interference, leading to successful target retrieval (Hofmeister 2011; Hofmeister & Vasishth, 2014).

Few studies have demonstrated that semantically complex expressions facilitated target retrieval during sentence processing (Hofmeister, 2011; Hofmeister & Vasishth, 2014) and discourse processing (Troyer, Hofmeister, & Kutas, 2016). Hofmeister and Vasishth (2014, Experiment 1) examined semantic elaboration of the target and the non-target on retrieval efficiency during sentence reading. In their study, critical sentences contained a transitive matrix clause, in which the object noun phrase was modified by an object-extracted relative clause (RC) (see below sentences 4a - 4d). Also, the number of modifying words for the matrix subject NP (distractor) and the matrix object NP (target) was manipulated, creating four conditions (Target Complexity: Simple vs. Complex, Distractor Complexity: Simple vs. Complex).

(4a) Simple, Simple

The **congressman** interrogated the **general** who a lawyer for the White House ***advised*** to not comment on the prisoners.

(4b) Simple, Complex

The **congressman** interrogated the **victorious four-star general** who a lawyer for the White House ***advised*** to not comment on the prisoners.

(4c) Complex, Simple

The **conservative U.S. congressman** interrogated the **general** who a lawyer for the White House ***advised*** to not comment on the prisoners.

(4d) Complex, Complex

The **conservative U.S. congressman** interrogated the **victorious four-star general** who a lawyer for the White House ***advised*** to not comment on the prisoners.

Their analysis of RTs at the retrieval site (*advise)* showed facilitated retrieval when the target (the matrix object NP) was semantically elaborated during the encoding phase, but elaboration of the distractor (the matrix subject NP) showed weaker effects. Important to note is that the critical sentences in their study contained two distractor nouns (*congressman, lawyer*) that have the same syntactic [+subject] and semantic [+animate] features before and after the target, respectively. The issue is that it is not clear whether the observed RTs at the retrieval sites reflect semantic elaboration effects or spillover effects of encoding interference of the most recent distractor (*lawyer*). It is also possible that RTs at the retrieval sites could have been influenced by both proactive interference (from the matrix subject NP distractor, *congressman*) and retroactive interference (from the subject of the object-extracted RC, *lawyer*), making the RT data less interpretable.

Following up on Hofmeister and Vasishth, the current study (Experiments 3 and 4) adopted their design, manipulating the number of modifying words for the target and the distractor to examine semantic elaboration effects on retrieval efficiency. However, instead of using the same syntactic construction as in Hofmeister and Vasishth, Experiments 3 and 4 used the sentences in the High semantic-High syntactic interference condition from Experiments 1 and 2 in the current study (e.g., *The* ***resident*** *who said that the* ***neighbor*** *was dangerous last month* ***had complained*** *about the investigation*), so that sentences include only one distractor that shares the same semantic and syntactic features with the target. This allows us to clearly test the effects of semantic elaboration on *retroactive* interference in sentence comprehension. Given that one of the assumptions of the cue-based retrieval model is that linguistic chunks in memory decay as a function of time and prior retrievals, comparing the findings of Hofmeister and Vasishth with those of the current study (Experiment 3) will be informative to understand how the position of the distractor in relation to the target (proactive vs. retroactive) influences retrieval efficiency.

Less Skilled L2 Readers' Reading Comprehension

Most studies demonstrating L2 readers' comprehension bottleneck have focused on their inability to compute fully specified syntactic representations or capacity differences during sentence processing. The shallow structure hypothesis (SSH) by Clahsen and Felser (2006), which is one of the most cited theories of L2 sentence processing, claims that comprehension difficulty in L2 sentence processing arises from L2 readers' inability to use syntactic information, and thus L2 readers construct shallow syntactic representations, relying more on semantic and pragmatic information during sentence processing. Supportive evidence for the SSH was presented in studies demonstrating L1 vs. L2 contrasts in resolving syntactic ambiguities and processing long-distance dependencies during online sentence reading (Felser, Roberts, Marinis, & Gross, 2003; Papadopoulou & Clahsen, 2003; Felser & Roberts, 2007; Marinis, Roberts, Felser, & Clahsen, 2005).

Felser et al. (2003), for example, examined L1 and high-proficiency L2 readers' attachment preferences of ambiguous RCs when the head nouns were complex noun phrases (NPs) including the genitive preposition *of* (e.g., "*The dean liked the secretary <u>of</u> the professors who <u>was/were</u> reading a letter.*"). Comparisons between L1 and L2 readers' RTs at the morphologically disambiguating verb revealed L1-L2 differences. Whereas L1 readers showed increased RTs when the RCs had to be attached to the NP1s (given the preference for low attachment in English), L2 readers did not show a strong preference for either attachment site. On the other hand, when the thematic preposition *with* was included in the complex NPs (e.g. *The dean liked the professors <u>with</u> the secretary who were/was reading a letter.*"), L2 readers demonstrated similar behavioral patterns as those of L1 readers (i.e. increased RTs in the attachment of RCs to NP1s). Based on these findings, Felser et al. claimed that L2 readers are unable to use syntactic information to the same extent that L1 readers do, and that they heavily rely on lexical-semantic information during L2 parsing.

However, the findings of Felser et al. have been criticized as their stimuli were not normed, and therefore could have been biased towards one of the attachment sites. Also, factors that may influence attachment preferences of ambiguous RCs were not controlled in their study, for example, prosodic structure (Fernandez, 2006), discourse context (Pan, Schimke, & Felser, 2015), L1 transfer (Kamide & Mitchell, 1997), and WM span (Kim & Christianson, 2013; Swets, Desmet, Hambrick, & Ferreira, 2007). In addition, given the findings that L1 readers may perform 'Good Enough' language processing under certain circumstances (i.e., even when confronted with a conflicting syntactic interpretation, L1 readers often fail to revise a strongly plausible initial misinterpretation; Christianson, 2016; Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Christianson, Williams, Zacks, & Ferreira, 2006; Christianson, Luke, & Ferreira, 2010, Ferreira, 2003; Ferreira, Christianson, & Hollingworth, 2001; Ferreira & Patson, 2007; Patson, Darowski, Moon, & Ferreira, 2009; Swets, Desmet, Clifton, & Ferreira, 2008), the SSH may not accurately capture L1-L2 processing differences. Importantly, recent studies have provided evidence for L2 readers' clear attachment preferences, although not always in the same direction as L1 readers (Dussias, 2003; Hopp, 2014; Witzel, Witzel, & Nicol, 2012). Hopp (2014), for example, showed that both L1 and high-proficiency L2 English readers preferred NP2 attachment to NP1 attachment in resolving ambiguous RC attachments. On the other hand, Witzel, Witzel and Nicol (2012) demonstrated Chinese L2 English readers' NP1 attachment preference, different from L1 English readers' NP2 preference.

While the SSH argues that L2 processing is qualitatively different from L1 processing, others have claimed that L1 and L2 parsing mechanisms are fundamentally similar, and that L1-

L2 differences can be attributed to computational resource limitations in L2 parsing (e.g., McDonald, 2006). This capacity-based approach suggests that L2 readers' less successful syntactic processing is at least partially due to limitations on their computational resources (e.g., WM), which may be heavily taxed by the activation and inhibition of L1 or reduced automaticity in L2 (Hopp, 2006; McDonald, 2006).

The observation that L2 readers' processing is generally slower than L1 readers (e.g., slower reaction times in grammatical judgment tasks and lexical decision tasks, Bialystock & Miller, 1999; McDonald, 2000; Scherag et al., 2004) may indicate high cognitive load that imposes on L2 readers' computational resources during online processing. Importantly, the findings that L1 readers' behavioral patterns resembled typical L2 behavioral patterns when their computational resources were taxed by cognitive load (i.e. simulating processing constraints applied to L2 processing) provides supportive evidence for the capacity-based account.

For example, some L1 studies that increased cognitive load by employing dual-task paradigms (Blackwell & Bates, 1995; McDonald, 2006), providing noisy input (Kilborn, 1991), or adding time constraints (Miyake, Carpenter, & Just, 1994) showed L1 speakers' impaired performance, resembling late L2 learners' performance. Further evidence in favor of the capacity-based theory is that both L1 and L2 processing occur in a 'good enough' manner, relying on plausibility and word order processing heuristics when resources are overloaded during online sentence processing (Christianson, 2016; Christianson et al., 2001; Christianson et al., 2006; Christianson et al., 2010, Ferreira, 2003; Ferreira et al., 2001; Ferreira & Patson, 2007; Lim & Christianson, 2013a, 2013b; Patson et al., 2009; Swets et al., 2008). Although incomplete processing in L2 may be more pronounced compared to L1 processing, the fact that 'good enough' processing in L1 and L2 partly occurs as a function of cognitive load supports the capacity-based account capturing L1-L2 differences (Hopp, 2006).

Recently, motivated by the cue-based retrieval account, Cunnings (2016) proposed that the L2 comprehension bottleneck can be attributed to L2 readers' susceptibility to similarity-based retrieval interference, and that L2 readers rely more on discourse cues than syntactic cues when resolving dependencies at retrieval. However, his proposal did not specifically demonstrate why L2 readers would suffer more from similarity-based interference than L1 readers, or why L2 readers assign more weights on discourse information than syntactic information at retrieval (Dillon, 2016; Kaan, 2016; Malko, Ehrenhofer, & Phillips, 2016; Omaki, 2016). In response to Cunnings (2016), Malko et al. (2016) commented that noisier L2 input may explain why L2 learners are more susceptible to similarity-based interference than L1 speakers, because noise may make it difficult to identify relevant retrieval cues in the input and distinguish target representations from the distractor representations. Although Malko et al. (2016) provided a possible explanation for L2 readers' reduced ability to retrieve relevant information from memory during sentence processing, this hypothesis has not been empirically tested in L2 studies. In addition, as Jacob, Lago, and Patterson (2016) pointed out, recent studies examining number interference effects have shown contradicting evidence to Cunnings's hypothesis, such that they demonstrated similar interference effects in L1 and L2 (Foote, 2010; Hoshino, Dussias, & Kroll, 2010; Nicol & Greth, 2003, Lim & Christianson, 2014; Tanner, Nicol, Herschensohn, & Osterhout, 2012). Thus, examining interference effects of less-skilled L2 readers in the current experiments provides insight into the L2 comprehension bottleneck during sentence comprehension, and furthermore helps us better understand L1-L2 processing differences.

Individual Differences Factors Moderating Retrieval Interference

Reading comprehension depends not only on the characteristics of the sentence material being read, but also on the individual reading abilities of the reader. Thus, the current study attempted to explore the relationship between participants' reading ability and their WM capacity and reading speed to specifically examine how these individual difference measures are associated with the ability to efficiently retrieve critical information from memory during sentence comprehension.

Working memory capacity. Although the nature of the WM system and the role of WM capacity in language processing are currently very controversial with recent evidence of extremely limited WM capacity, numerous studies have demonstrated a correlation between WM capacity and reading performance, and suggested the capacity of WM as the primary cognitive factor constraining reading performance (Caplan & Waters, 1999; Gordon, et al., 2002; Fedorenko, et al., 2006, 2007, Just & Carpenter, 1980, 1992; Just, et al., 1996; Just & Varma, 2002, 2007). However, as reviewed in the previous section of the cue-based retrieval approach, recent findings in sentence processing research seem to suggest that comprehension processes may rely on an efficient direct-access, cue-driven operation, rather than some fixed WM capacity (McElree & Vasishth, 2005; McElree et al., 2006; Nicenboim & Vasishth, 2018; Van Dyke, 2007; Van Dyke & Johns, 2012; Tan et al., 2017). This alternative framework proposes that comprehension is determined by whether or not information can be efficiently retrieved when it is needed. In other words, the ability to efficiently retrieve target information from memory is important for successful sentence comprehension.

Assuming that the cue-driven retrieval is the underlying mechanism supporting sentence processing, the next important question to explore is how individual differences in WM are associated with retrieval ability during sentence processing. As McElree et al. (2006) described, in the cue-based retrieval framework, what creates comprehension bottleneck is in fact limited focus of attention, rather than the capacity *per se*. If this is the case, no reliable interaction is predicted between interference effects and participants' WM capacity, because participants' WM capacity in the current study is measured by an operation span task (Turner & Engle, 1989), which is more likely to measure the storage component of WM, rather than the ability of attention control. If reliable interaction between interference effects and WM capacity is found, this result could be interpreted as indicating a correlation between retrieval ability and attentional control of WM within the cue-based retrieval framework.

Few studies have focused on the role of WM in the retrieval processing in sentence comprehension. Van Dyke, Johns, and Kukona (2014), for example, investigated what individual differences measures contribute to poor comprehension. They adopted the comprehension paradigm from Van Dyke and McElree (2006), in which a memory load (Load vs. No-load) and presence of interference (Interfering vs. Non-interfering) were manipulated. Memory load was manipulated, such that participants either did or did not have a list of three words to maintain in memory (e.g., *table*, *sink*, *truck*) while reading object-cleft sentences (e.g., *It was the boat that the guy who lived by the sea* <u>*fixed/sailed*</u> *in two sunny days*). Interference conditions were created with the verb manipulation, such that the words in the memory list either were or were not plausible direct objects for the manipulated verb (e.g., *fixed* for the Interfering condition and *sailed* for Non-interfering condition). Also, a battery of 24 different cognitive tasks (measuring print mapping, reading skill, oral language use, memory, and intelligence) were administered in order to examine factors that contribute to poor comprehension, particularly, the relationship between WM capacity and language comprehension. The critical result was that the interfering condition showed longer RTs in the critical verb region, compared to the Non-interfering condition, providing supportive evidence for the cue-based retrieval model. In addition, their analyses showed a high degree of multicollinearity between reading span scores and many of their skill measures. Importantly, reading span scores significantly correlated with RTs on the critical region, but after partialling out variability associated with intelligence, WM was no longer a strong predictor of comprehension. The only measure that remained significant after partialling out intelligence was receptive vocabulary. Based on these findings, Van Dyke et al. (2014) suggested that previous findings that emphasized the role of WM capacity in sentence processing may be due to its shared variance with many other cognitive and language-related abilities, and claimed that receptive vocabulary may be a key predictor of reading ability, rather than WM capacity.

Inconsistent with Van Dyke et al. (2014), Tan et al. (2017) showed interaction between readers' WM capacity and semantic and syntactic retrieval interference in online and offline sentence comprehension data. In their study, WM capacity modulated the magnitude of participants' syntactic interference effects in online RT data (even after partialling out variability associated with vocabulary knowledge), and semantic interference effects in offline question-response accuracy. The discrepancies between Tan et al. and Van Dyke et al. could be because the studies used different syntactic constructions and comprehension paradigms. Tan et al. tested sentences in which similarity-based retrieval interference was manipulated in the subject-verb long-distance dependency, whereas Van Dyke et al. used syntactically complex object-cleft sentences where not only retrieval interference but also memory load were manipulated. Another possibility is that the observed effects of WM capacity in Tan et al. may reflect the variance shared between WM capacity and intelligence, since intelligence was not partialled out from complex WM scores in their study. Further investigation is needed to determine the relationship between WM capacity and the ability to retrieve critical information from memory during reading in order to better understand factors that predict reading ability.

Reading speed. According to the Verbal Efficiency hypothesis (Perfetti, 1985), successful reading comprehension depends on the efficiency of word-level processing. The underlying idea is that rapid and automatic word-level processes (i.e., efficient retrieval of orthographic and lexical codes) preserve processing resources, which are then available for higher level processing (e.g., sentence and text level processing). Thus, when decoding is slow and effortful, resources are dedicated to word-level processing, and this leaves fewer resources for the higher level processing, resulting in poor comprehension.

In line with Perfetti's Verbal Efficiency theory, some empirical evidence has shown a correlation between word reading speed and reading comprehension in both adult and child populations (Hess & Radtke, 1981; Jackson & McClelland, 1979). Jackson and McClelland, for example, examined two groups of college students who differed in their verbal ability on a number of information processing tasks, and measured their listening comprehension, reading comprehension, and reading speed (as an index of the speed of accessing memory codes of visually presented letters). Their result showed that reading speed and listening comprehension explained nearly 75 percent of the variance in reading skill, and based on this finding, Jackson and McClelland suggested that reading speed is an important component of reading comprehension (reading speed and listening comprehension were not correlated). Perfetti and Hogaboam (1975) also reported that third- and fifth-grade poor readers were slower at naming high-frequency words compared to skilled readers. Along with the studies demonstrating the association between word reading speed and reading comprehension, numerous findings have also shown that less-skilled readers are typically slower to retrieval phonologically encoded information during reading compared to skilled readers (Perfetti, 1985; Swan & Goswami, 1997a,b; Wolf & Bowers, 1999; Goswami, 2011).

Motivated by Perfetti's (1985) Verbal Efficiency theory and evidence of a link between word reading and comprehension, the current study examined the relationship between word reading speed and retrieval ability in sentence comprehension. If reading speed is strongly associated with retrieval ability during reading, faster readers are predicted to be less susceptible to retrieval interference compared to slower readers, because faster readers with efficient decoding skill and extensive language experience may be more sensitive to semantic cues (Nicenboim et al., 2016; Traxler et al., 2012), and this may help them to efficiently discriminate lexical items in memory and retrieve the correct item at retrieval (Van Dyke & Johns, 2012; Van Dyke & Shankweiler, 2012).

Overview of Experiments

In the following chapters I present four experiments whose results demonstrate the role of memory retrieval and quality of lexical representations in reading comprehension. Experiments 1 and 2 in Chapters 3 and 4, respectively, evaluate the interference effects of syntactic and semantic properties of NPs that intervene between two syntactically dependent items. The first experiment is a replication of previous studies that provided supportive evidence of similarity-based interference in skilled L1 readers' sentence comprehension, whereas the second experiment provides an initial evaluation of less-skilled L2 readers' sensitivity to retrieval interference in L2 reading comprehension. Experiments 3 and 4 in Chapters 5 and 6, respectively, evaluate the effects of semantic elaboration on retrieval interference to examine whether enhanced quality of lexical representations through semantic elaboration facilitates target retrieval during sentence comprehension. Chapter 7 attempts to integrate the results from all the experiments in a direct-access, cue-based retrieval account, and demonstrate what individual difference factors (here, reading speed and working memory capacity) may influence retrieval efficiency during sentence comprehension.

Chapter 3: Retrieval Interference in L1 Reading Comprehension

The first experiment was designed to examine syntactic and semantic retrieval interference effects during L1 sentence reading. To create high and low interference conditions (Semantic Interference: High vs. Low, Syntactic Interference: High vs. Low), the current experiment manipulated the syntactic and semantic properties of the intervening noun between the subject and the main verb so that the syntactic and semantic features of the intervening noun either match or mismatch with the retrieval cues. The main verb region was the target retrieval site where the subject-verb, long-distance dependency was established, and RTs on this region and post-verb region were analyzed to examine syntactic and semantic interference effects during L1 sentence processing. In addition to the retrieval latency, retrieval probability of the target was examined by analyzing question-response accuracy.

The two competing memory-based accounts (capacity-based account vs. cue-based retrieval account) make differing predictions with regard to RTs on the critical regions and question-response accuracy. The cue-based retrieval account predicts increased processing difficulty at the point of establishing a long-distance dependency that requires retrieval when the similarity between the features of the distractor and the retrieval cues increases (Lewis et al., 2006). Thus, relative to the Low Interference conditions, the High Interference conditions are predicted to elicit longer RTs at the main verb region and result in lower question-response accuracy. In addition, the retrieval account also predicts the interaction between syntactic and semantic interference because global matching suggests that all cues are multiplicatively combined into a single retrieval probe (Van Dyke & McElree, 2006).

The capacity-based account, on the other hand, predicts no interference effects at the main verb region. According to the capacity-based account, comprehension difficulty increases as the distance between the head and the dependent increases. Given that the number of words between the subject and the main verb was held constant across all four conditions, the capacity-based account expects no significant RT differences between High and Low Interference conditions at the main verb region. Similarly, no significant differences in question-response accuracy are expected between the High and Low Interference conditions.

In addition to interference effects, the extent to which readers' WM capacity predicts retrieval ability during sentence processing was examined by testing the interaction between participants' operation span scores and interference effects. Because the cue-based retrieval account assumes that most linguistic processing employs an efficient direct-access, cue-based retrieval, instead of relying on limited active maintenance of information during sentence processing, this account predicts no significant interaction between operation span scores and interference effects.

Another individual differences measure examined as a predictor of retrieval ability was participants' reading speed. If reading speed is a strong predictor of the retrieval ability, faster readers are predicted to be less susceptible to similarity-based interference at retrieval because they may have greater experience with language and thus have a high proportion of high-quality lexical representations and be more sensitive to semantic cues (Nicenboim et al., 2016, Traxler et al., 2012), and this may help them to efficiently discriminate lexical items in memory and inhibit spurious activations of irrelevant linguistic information at retrieval (Van Dyke & Johns, 2012; Van Dyke & Shankweiler, 2012).

Method

Participants. Fifty-one students from University of Illinois at Urbana-Champaign participated in Experiment 1 in exchange for course credit or $7.00 compensation. All participants were native speakers of English and had normal vision and hearing, with no history of neurological impairment.

Apparatus. Participants' eye-movements (right eye monocular tracking) were recorded with an SR Research Ltd. Eyelink 1000, which records the position of a reader's eye at a sampling rate of 1000Hz. A chin-rest and a head-rest were used in order to minimize head movements. Stimuli were presented in a monospaced font (Courier New) with font size 18 on a 24-inch monitor with a 1920 x 1080 screen resolution, and the eye-to-screen distance measured approximately 60 cm. SR Research Experiment Builder was used to present the experimental stimuli.

Materials. The experiment used a repeated measures 2 x 2 design, manipulating Syntactic Interference (High, Low) and Semantic Interference (High, Low). The critical sentences consisted of thirty-two sets of sentences, which were simplified versions of those used in Van Dyke (2007). Each sentence contained a main clause and a subject-extracted RC that intervened between the subject and the verb of the main clause. The four types of sentences in each set had the same main clause but differed in the intervening region, in which the syntactic and semantic features of the intervening noun (distractor) were manipulated to be similar or dissimilar to the features of the subject noun of the main clause (target). See Table 3.1 for an example and Appendix A for a full list of stimuli.

Table 3.1. Example stimuli in Experiment 1.

	Low Semantic Interference	*High Sematic Interference*
Low Syntactic Interference	The ***resident*** who was living near the dangerous **<u>warehouse</u>** last month ***had complained*** about the investigation.	The ***resident*** who was living near the dangerous **<u>neighbor</u>** last month ***had complained*** about the investigation.
High Syntactic Interference	The ***resident*** who said that the **<u>warehouse</u>** was dangerous last month ***had complained*** about the investigation.	The ***resident*** who said that the **<u>neighbor</u>** was dangerous last month ***had complained*** about the investigation.

In the Low Syntactic Interference conditions, the distractor (*warehouse*, *neighbor*) was the object of the prepositional phrase, and thus its syntactic feature [+object] did not match with the syntactic feature of the target (*resident*) [+subject]. Whereas in the High Syntactic Interference conditions, both the target and the distractor had the [+subject] syntactic feature. In the Low Semantic Interference conditions, the distractor (*warehouse*) had an [+inanimate] semantic feature, which did not overlap with the [+animate] semantic feature of the target (*resident*). On the other hand, in the High Semantic Interference conditions, the distractor (*neighbor*) had an [+animate] semantic feature as those of the target (*resident*), and thus the distractor could also be a plausible agent for the main verb.

An adverbial phrase was inserted between the intervening region and the main verb region to avoid local coherence effects, which can arise in the Low Syntactic Interference condition where the main verb immediately follows the intervening noun without an adverbial clause (Glaset et al., 2013; Tabor, Galantucci, & Richardson, 2004; Tan et al., 2017; Van Dyke, 2007). The critical regions included the main verb region where the long-distance dependency was established and the post-verb region containing either an adverbial phrase or the patient of the main verb. The number of words prior to the main verb region was held constant in all four experimental conditions to control the positional confound. The 32 critical sentences were distributed across four lists using a Latin Square design, with conditions counterbalanced across lists. In addition to the critical sentences, 68 filler and 5 practice sentences were constructed with various syntactic structures, including object-extracted relative clauses, simple clauses with transitive and dative verbs, conjoined clauses, and subjunctive clauses.

Procedures. Participants were randomly assigned to one of four lists used for counterbalancing and were tested individually. During the session, participants completed an eye-tracking reading task and an operation span task at their own pace.

Eye-tracking reading. At the beginning of the eye-tracking experiment, participants performed a 9-point calibration procedure to make sure that the eye-tracker recordings were accurate. After the calibration phrase, participants were instructed to silently read each sentence for comprehension and provided a verbal response for each comprehension question during the eye-tracking. The comprehension questions for the critical sentences were identical for the four conditions and always focused on the subject of the main clause to examine whether participants correctly established the long-distance dependency in the sentences (for example, "*Who was complaining?*"). Comprehension questions for filler sentences focused on other regions of sentences to prevent participants from adopting a task relevant strategy, for example, intentionally avoiding the intervening region to provide correct comprehension question responses.

Operation span. Participants performed an operation span task, measuring the capacity of working memory (Turner & Engle, 1989). Although a modified version of the reading span task by Daneman and Carpenter (1980) has been extensively used in psycholinguistics studies to examine the role of working memory capacity in sentence processing, the present study chose to use the operation span, instead of the reading span, because the reading span measure may be predictive of not only working memory capacity but also verbal ability or reading experience (Conway et al., 2005; MacDonald & Christiansen, 2002; Nicenboim, Logacev, Gattei, & Vasishth, 2016). The operation span task was conducted using Paradigm Stimulus Presentation software (Perception Research Systems, 2007). During the task, participants first evaluated simple math equations and memorized letters that were shown between the equations. The letters were always consonants and appeared for 800ms. After a set of three to seven equations with letters, participants were instructed to type the letters in order of their presentation. To prevent participants from adopting a rehearsal strategy, they were asked to read the equations and letters out loud as they performed the task.

Analysis. Prior to the analyses of eye-movement measures and question-response latency and accuracy, two participants were excluded from the analyses due to low question-response accuracy for the fillers (72% accuracy) and low recall accuracy in the operation span task (7% accuracy), and this left forty-nine participants for the analysis.

To examine syntactic and semantic retrieval interference effects during sentence comprehension, various eye movement measures were analyzed in critical regions in sentences. The critical regions included the main verb where the distant subject must be retrieved in order to establish the long distance dependency and the post-verb spillover region including either an adverbial phrase or the patient of the main verb. The eye movement measures included gaze duration (GD: the sum of all fixations on the target word before leaving the target), regression-path duration (RPD: the sum of all fixations on the target word from first entering the target until leaving the target, including re-reading the earlier regions), total viewing time (TVT: the sum of all fixations on the target word including regressions), regression-in (RI: the probability of regressing back to the target region after leaving it), and regression-out (RO: the probability of regressing out of the target region).

Fixation duration and fixation probability measures in critical regions were analyzed with linear mixed-effects models (Baayen, Davidson, & Bates, 2008; Kliegl, Masson, & Richter, 2010) and logistic mixed-effects models (Jaeger, 2008), using the lmerTest package (version 3.0-

1) in R (version 3.4.4). Prior to the analysis of eye movement measures, fixations shorter than 80ms and longer than 1200ms were removed by using the EyeLink Data Viewer program. For each fixation duration measure, fixations above 2.5 standard deviations from the mean of each condition were excluded from data analyses. Raw fixations in all fixation duration measures were log-transformed to meet the error normality assumption. Given that words in critical regions differed in length, log-transformed, length-adjusted fixation durations were computed for each participant by regressing the log-transformed fixations against word length (measured in number of letters). After trimming the eye-movement data, fixation duration and fixation probability measures were modeled as a function of fixed effects of Syntactic Interference (High, Low), Semantic Interference (High, Low), Individual Difference measure (operation span score or reading speed), and their three-way interaction. Sum contrasts were used for the two fixed effects (High Interference condition as +1 and Low Interference condition as -1). All linear mixed-effects models included random effects of Participants and Items to account for variability in participants and sentence items, and the maximal random effect structure was justified by model comparison (Barr, Levy, Scheepers, & Tily, 2013). Responses to comprehension questions were also analyzed using the logistic linear mixed-effects model to examine the influence of syntactic and semantic interference on comprehension question error rates.

To investigate whether WM capacity and reading speed play an important role in resolving syntactic and semantic interference in long-distance dependencies, three-way interaction between the fixed effects (Syntactic Interference, Semantic Interference) and the individual difference measures (operation span score or reading speed) were included in the linear mixed-effects models. Reading speed for each participant was obtained by calculating an average RT per word from filler sentences, and both operation span scores and reading speed were mean-centered and scaled prior to the inference analysis.

Results

Eye-tracking. Descriptive statistics of eye-movements for the main verb and post-verb regions are reported in Table 3.2.

Table 3.2. Experiment 1: Mean performance for eye movement data (raw reading times in ms) in the main verb and post-verb spillover regions.

Sem	*Syn*	*GD*	*RPD*	*TVT*	*RI*	*RO*
Main verb region						
High	High	342 (157)	448 (347)	728 (431)	0.40 (0.49)	0.16 (0.37)
	Low	343 (146)	436 (290)	630 (366)	0.40 (0.49)	0.14 (0.34)
Low	High	367 (160)	478 (389)	705 (399)	0.40 (0.49)	0.14 (0.35)
	Low	349 (158)	401 (231)	628 (345)	0.38 (0.49)	0.11 (0.31)
Post-verb region						
High	High	369 (206)	2869 (2330)	582 (352)	-	0.91 (0.28)
	Low	385 (238)	2166 (1596)	561 (345)	-	0.91 (0.29)
Low	High	374 (209)	2335 (1747)	582 (344)	-	0.88 (0.32)
	Low	407 (236)	2019 (1551)	609 (368)	-	0.89 (0.32)

Note: Sem = Semantic Interference; Syn = Syntactic Interference; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out. Standard deviations in parentheses.

Main verb region. The main verb region was the retrieval site where the retrieval of the target (subject) was triggered, and retrieval cues were generated to establish a long-distance dependency. The words in this region differed in length, so log-transformed length-adjusted reading times for each eye movement measure were modeled as a function of fixed effects of Semantic Interference (High, Low), Syntactic Interference (High, Low), Individual Difference measure (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. Summaries of linear mixed-effects models for the main region are described in Table 3.3.

Table 3.3. Experiment 1: Mixed-effects modeling results for all dependent measures at the main verb region.

	Fixation duration measure														
	GD			*RPD*			*TVT*			*RI*			*RO*		
Predictor	*β*	SE	*t*	*β*	SE	*t*	*β*	SE	*t*	*β*	SE	*z*	*β*	SE	*z*
OSpan															
Intercept	0.00	0.03	0.04	0.00	0.04	0.07	0.00	0.04	0.07	**-0.50**	**0.15**	**-3.43**	**-2.04**	**0.13**	**-15.49**
Sem	**-0.06**	**0.03**	**-2.02**	-0.01	0.04	-0.18	0.01	0.03	0.34	0.02	0.11	0.20	0.26	0.15	1.69
Syn	0.02	0.03	0.66	0.07	0.04	1.84	**0.16**	**0.03**	**4.54**	0.06	0.11	0.53	0.27	0.15	1.76
OSpan	0.00	0.01	-0.22	0.00	0.02	-0.10	0.00	0.01	-0.19	-0.14	0.13	-1.07	0.09	0.13	0.70
Sem x Syn	**-0.12**	**0.06**	**-2.16**	**-0.17**	**0.07**	**-2.44**	0.00	0.07	-0.02	-0.06	0.22	-0.26	-0.16	0.31	-0.52
Sem x OSpan	0.03	0.03	1.10	0.00	0.04	0.09	0.06	0.03	1.83	0.00	0.11	0.04	-0.23	0.16	-1.49
Syn x OSpan	0.02	0.03	0.54	0.07	0.04	1.83	0.04	0.04	1.21	0.07	0.11	0.63	-0.07	0.16	-0.45
Sem x Syn x OSpan	-0.04	0.06	-0.67	0.06	0.07	0.77	0.10	0.07	1.28	0.09	0.23	0.38	0.06	0.31	0.20
RS															
Intercept	0.00	0.03	0.04	0.00	0.04	0.07	0.00	0.04	0.07	**-0.50**	**0.15**	**-3.41**	**-2.03**	**0.13**	**-15.42**
Sem	**-0.06**	**0.03**	**-2.01**	-0.01	0.04	-0.18	0.01	0.03	0.33	0.02	0.11	0.17	0.24	0.15	1.58
Syn	0.02	0.03	0.67	0.07	0.04	1.85	**0.16**	**0.03**	**4.55**	0.06	0.11	0.53	0.26	0.15	1.72
RS	0.02	0.01	1.13	0.02	0.01	1.22	**0.03**	**0.02**	**2.04**	0.09	0.13	0.68	0.00	0.13	-0.01
Sem x Syn	**-0.12**	**0.06**	**-2.18**	**-0.17**	**0.07**	**-2.43**	0.00	0.07	-0.03	-0.06	0.22	-0.25	-0.14	0.30	-0.47
Sem x RS	-0.02	0.03	-0.70	-0.02	0.04	-0.46	0.01	0.03	0.25	0.16	0.11	1.44	0.03	0.16	0.20
Syn x RS	-0.02	0.03	-0.59	0.02	0.04	0.60	0.07	0.03	1.93	0.09	0.11	0.80	0.07	0.16	0.47
Sem x Syn x RS	0.05	0.06	0.92	0.12	0.07	1.73	0.06	0.07	0.92	-0.17	0.22	-0.77	0.18	0.32	0.57

Note: Sem = Semantic Interference; Syn = Syntactic Interference; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; OSpan = Operation span; RS = Reading speed; *β*=Estimate coefficients; SE=standard error. Bold indicates coefficients that are significant at the $p<0.05$ level.

Gaze duration. The model summaries of both three-way interaction models including operation span scores and reading speed, respectively, showed a significant two-way interaction between Semantic and Syntactic Interference (t=-2.16, p=0.03; t=-2.18, p=0.03), such that the semantic interference effects were greater in the High Syntactic Interference conditions compared to the Low Syntactic Interference conditions ($M_{SemHighSynHigh}$=342ms, $M_{SemLowSynHigh}$=367ms; $M_{SemHighSynLow}$=343ms, $M_{SemLowSynLow}$=349ms) (Figure 3.1). Table 3.4 summarizes the results from nested comparisons. Also, a significant main effect of Semantic Interference was found in both models (t=-2.02, p=0.04; t=-2.01, p=0.04), with longer gaze durations in the Low Semantic Interference conditions compared to the High Semantic Interference conditions ($M_{SemHigh}$=343ms, M_{SemLow}=358ms). The findings that the High Semantic Interference conditions resulted in shorter gaze durations relative to the Low Semantic Interference conditions could be due to the fact that readers may have spent less time on the main verb to regress back to the previous regions for re-reading the target and the distractor. On the other hand, in the case of the Low Semantic Interference conditions where the semantic features of the target and the distractor are not overlapping, readers stayed on the verb longer because they did not have to go back to the previous regions for re-check the target or the non-target. This explanation can be supported by the observed patterns of longer regression-path duration and total viewing time in the High Semantic Interference conditions compared to the Low Semantic Interference conditions. Other main effects and interactions in both models were not statistically significant (*t*s <1.10, *p*s>0.27; *t*s <1.13, *p*s>0.26).

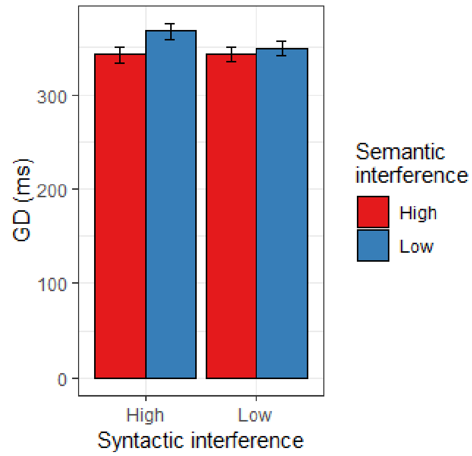

Figure 3.1. Experiment 1: Two-way interaction between Semantic and Syntactic Interference (GD) in the main verb region.

Table 3.4. Experiment 1: Pairwise differences of contrast for a two-way interaction between Semantic and Syntactic Interference (GD) in the main verb region.

Contrast	*β*	SE	*t*	*p* (>|t|)
SemHigh,SynHigh – SemLow,SynHigh	**-0.12**	**0.04**	**-2.96**	**0.02**
SemHigh,SynHigh – SemHigh,SynLow	-0.04	0.04	-1.06	0.71
SemHigh,SynHigh – SemLow,SynLow	-0.04	0.04	-0.96	0.77
SemLow,SynHigh – SemHigh,SynLow	0.08	0.04	1.89	0.23
SemLow,SynHigh – SemLow,SynLow	0.08	0.04	2.00	0.19
SemHigh,SynLow – SemLow,SynLow	0.00	0.04	0.11	1.00

Note: SemHigh = Semantic High Interference condition; SemLow = Semantic Low Interference condition; SynHigh: Syntactic High Interference condition; SynLow: Syntactic Low Interference condition; *β*=Estimate coefficients; SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

Regression-path duration. Both three-way interaction models including operation span scores and reading speed respectively revealed a significant two-way interaction between Semantic and Syntactic Interference (t=-2.44, p=0.01; t=-2.43, p=0.02). Syntactic interference effects were greater in the Low Semantic Interference conditions relative to the High Semantic Interference conditions ($M_{SemHighSynHigh}$=448ms, $M_{SemHighSynLow}$=436ms; $M_{SemLowSynHigh}$=478ms, $M_{SemLowSynLow}$=401ms) (Figure 3.2). Table 3.5 reports the results from pairwise comparisons for the observed interaction between Semantic and Syntactic Interference. Other main effects and interactions were not statistically significant (*t*s <1.84, *p*s>0.07; *t*s <1.85, *p*s>0.06).

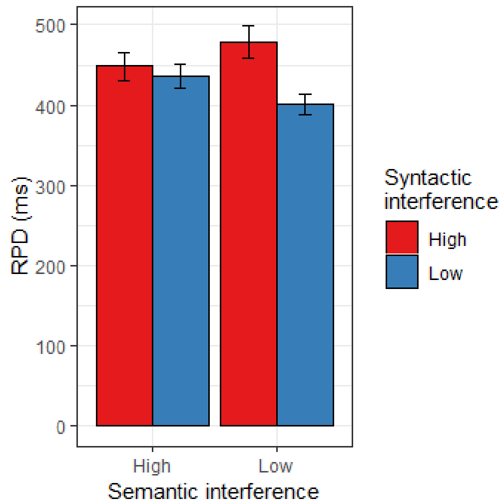

Figure 3.2. Experiment 1: Two-way interaction between Semantic and Syntactic Interference (RPD) in the main verb region.

Table 3.5. Experiment 1: Pairwise differences of contrast for a two-way interaction between Semantic and Syntactic Interference (RPD) in the main verb region.

Contrast	*β*	SE	*t*	*p* (>|t|)
SemHigh,SynHigh – SemLow,SynHigh	-0.09	0.05	-1.83	0.26
SemHigh,SynHigh – SemHigh,SynLow	-0.02	0.05	-0.40	0.98
SemHigh,SynHigh – SemLow,SynLow	0.06	0.05	1.19	0.63
SemLow,SynHigh – SemHigh,SynLow	0.07	0.06	1.43	0.48
SemLow,SynHigh – SemLow,SynLow	**0.15**	**0.05**	**3.01**	**0.01**
SemHigh,SynLow – SemLow,SynLow	0.08	0.05	1.60	0.38

Note: SemHigh = Semantic High Interference condition; SemLow = Semantic Low Interference condition; SynHigh: Syntactic High Interference condition; SynLow: Syntactic Low Interference condition; *β*=Estimate coefficients, SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

Total viewing time. The model summaries of both three-way interaction models with operation span scores and reading speed revealed a significant main effect of Syntactic Interference (t=4.54, p<0.001; t=4.55, p<0.001), with longer total reading times in the High Syntactic Interference conditions compared to the Low Syntactic Interference conditions ($M_{SynHigh}$=717ms, M_{SynLow}=629ms) (Figure 3.3). Also, a significant main effect of reading speed on total reading times was found from the model including reading speed as an individual difference measure (t=2.04, p=0.04), such that slower readers tended to be slower at reading the main verb region compared to the fast readers. Other main effects and interactions were not statistically significant (*t*s <1.83, *p*s>0.06; *t*s <1.93, *p*s>0.05)

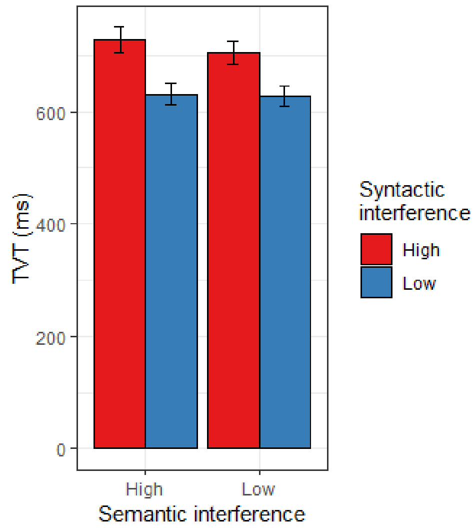

Figure 3.3. Experiment 1: Syntactic Interference main effects (TVT) in the main verb region.

Regression-in and Regression-out. No statistically significant effects were found in the regression-in and regression-out probability measures in both three-way interaction models with operation span scores and reading speed (regression-in: *z*s <-1.07, *p*s>0.28; *z*s <1.44, *p*s>0.15, regression-out: *z*s <1.76, *p*s>0.07; *z*s <1.72, *p*s>0.08).

Post-verb region. The post-verb region was a spillover region including either an adverbial phrase or the patient of the main verb. Log-transformed, length-adjusted reading times for each eye movement fixation measure were modeled as a function of Semantic Interference (High, Low), Syntactic Interference (High, Low), Individual Difference measure (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. Given that the post-verb region was the last region in the sentence, regression-in data were not available, and thus no analysis for the regression-in probability measure was conducted. Summaries of linear mixed-effects models for the post-verb region are described in Table 3.6.

Table 3.6. Experiment 1: Mixed-effects modeling results for all dependent measures at the post-verb spillover region.

	Fixation duration measure														
	GD			*RPD*			*TVT*			*RI*			*RO*		
Predictor	β	SE	*t*	β	SE	*t*	β	SE	*t*	β	SE	*z*	β	SE	*z*
OSpan															
Intercept	0.00	0.03	-0.06	-0.01	0.05	-0.12	0.00	0.04	-0.04	-	-	-	**2.26**	**0.11**	**20.49**
Sem	-0.04	0.03	-1.19	**0.16**	**0.05**	**3.33**	-0.07	0.04	-1.97	-	-	-	0.33	0.18	1.87
Syn	-0.06	0.03	-1.75	**0.21**	**0.05**	**4.22**	0.01	0.04	0.28	-	-	-	-0.04	0.18	-0.20
OSpan	-0.01	0.02	-0.56	-0.01	0.02	-0.38	-0.01	0.02	-0.78	-	-	-	-0.11	0.09	-1.20
Sem x Syn	0.09	0.07	1.28	**0.22**	**0.10**	**2.27**	**0.16**	**0.07**	**2.16**	-	-	-	0.01	0.36	0.04
Sem x OSpan	0.01	0.03	0.18	-0.06	0.05	-1.15	-0.01	0.04	-0.19	-	-	-	-0.34	0.18	-1.87
Syn x OSpan	-0.01	0.03	-0.37	0.06	0.05	1.21	0.06	0.04	1.77	-	-	-	0.24	0.18	1.32
Sem x Syn x OSpan	-0.01	0.07	-0.08	0.14	0.10	1.39	0.07	0.07	0.93	-	-	-	0.31	0.37	0.85
RS															
Intercept	0.00	0.03	-0.06	0.00	0.05	-0.13	0.00	0.04	-0.04	-	-	-	**2.87**	**0.21**	**13.91**
Sem	-0.04	0.03	-1.19	**0.16**	**0.05**	**3.35**	-0.07	0.04	-1.98	-	-	-	**0.72**	**0.25**	**2.84**
Syn	-0.06	0.03	-1.75	**0.21**	**0.05**	**4.23**	0.01	0.04	0.28	-	-	-	0.25	0.25	0.99
RS	0.02	0.02	0.90	0.04	0.02	1.47	**0.04**	**0.02**	**2.15**	-	-	-	**0.95**	**0.21**	**4.57**
Sem x Syn	0.09	0.07	1.30	**0.22**	**0.10**	**2.27**	**0.16**	**0.07**	**2.21**	-	-	-	0.60	0.50	1.19
Sem x RS	0.03	0.03	0.79	**0.12**	**0.05**	**2.52**	-0.01	0.04	-0.16	-	-	-	**0.68**	**0.27**	**2.51**
Syn x RS	0.01	0.03	0.32	0.08	0.05	1.70	-0.02	0.04	-0.56	-	-	-	0.32	0.26	1.20
Sem x Syn x RS	-0.05	0.07	-0.72	0.11	0.10	1.16	0.01	0.07	0.15	-	-	-	0.57	0.53	1.08

Note: Sem = Semantic Interference; Syn = Syntactic Interference; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; OSpan = Operation span; RS = Reading speed; β=Estimate coefficients, SE=standard error. Bold indicates coefficients that are significant at the $p<0.05$ level.

Gaze duration. Both three-way interaction models including operation span scores and reading speed respectively showed no significant effects in the post-verb spillover region (*t*s <-1.75, *p*s>0.08; *t*s <-1.75, *p*s>0.08).

Regression-path duration. The summaries of both models with operation span scores and reading speed revealed a reliable two-way interaction between Semantic and Syntactic Interference ($t=2.27$, $p=0.02$; $t=2.27$, $p=0.02$), a main effect of Semantic Interference ($t=3.33$, $p<0.001$; $t=3.35$, $p<0.001$), and a main effect of Syntactic Interference ($t=4.22$, $p<0.001$; $t=4.23$, $p<0.001$). The high semantic and syntactic interference conditions elicited longer regression-path durations at the post-verb region compared to the low semantic and syntactic interference conditions ($M_{SemHigh}$=2520ms, M_{SemLow} =2176ms; $M_{SynHigh}$ =2605ms, M_{SynLow} =2092ms). As for the interaction pattern, syntactic interference effects were greater in the High Semantic Interference conditions compared to the Semantic Low Interference conditions ($M_{SemHighSynHigh}$=2869ms, $M_{SemHighSynLow}$=2166ms; $M_{SemLowSynHigh}$=2335ms, $M_{SemLowSynLow}$=2019ms) (Figure 3.4). Table 3.7 reports the results of the nested comparisons for this observed two-way interaction. The three-way interaction model with reading speed also showed a reliable interaction between Semantic Interference and reading speed ($t=2.52$, $p=0.01$), such that semantic interference effects were reduced as reading speed measure (measured in mean reading times per word) decreased. This interaction pattern suggests that faster readers may be less susceptible to similarity-based interference and more efficient at target retrieval during sentence processing relative to slower readers (Figure 3.5)

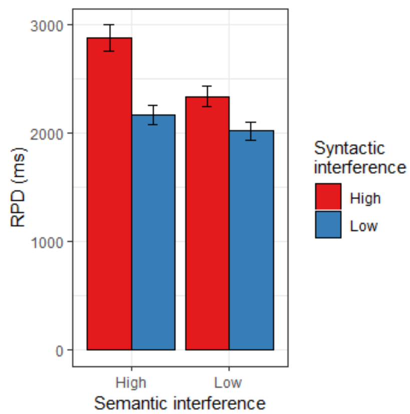

Figure 3.4. Experiment 1: Two-way interaction between Semantic and Syntactic Interference (RPD) in the post-verb region.

Table 3.7. Experiment 1: Pairwise differences of contrast for a two-way interaction between Semantic and Syntactic Interference (RPD) in the post-verb region.

Contrast	*β*	SE	*t*	*p* (>|t|)
SemHigh,SynHigh – SemLow,SynHigh	**0.28**	**0.07**	**3.98**	**<0.001**
SemHigh,SynHigh – SemHigh,SynLow	**0.32**	**0.07**	**4.60**	**<0.0001**
SemHigh,SynHigh – SemLow,SynLow	**0.37**	**0.07**	**5.36**	**<0.0001**
SemLow,SynHigh – SemHigh,SynLow	0.04	0.07	0.62	0.93
SemLow,SynHigh – SemLow,SynLow	0.10	0.07	1.37	0.52
SemHigh,SynLow – SemLow,SynLow	0.05	0.05	0.75	0.88

Note: SemHigh = Semantic High Interference condition; SemLow = Semantic Low Interference condition; SynHigh: Syntactic High Interference condition; SynLow: Syntactic Low Interference condition; *β*=Estimate coefficients, SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

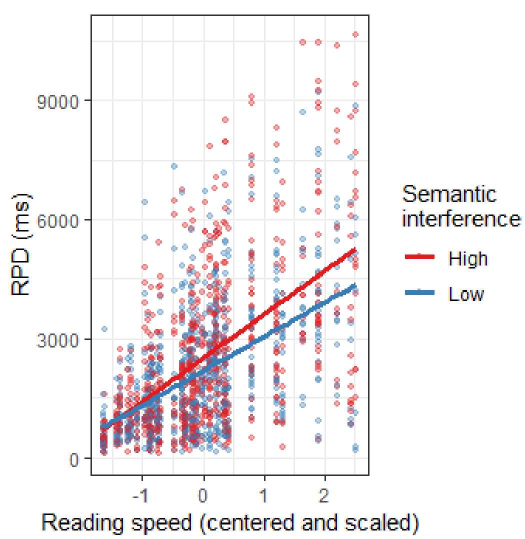

Figure 3.5. Experiment 1: Two-way interaction between Semantic Interference and reading speed (RPD) in the post-verb region.

Total viewing time. Similar to the patterns observed in the regression-path duration, the two three-way interaction models with operation span scores and reading speed revealed a reliable two-way interaction between Semantic and Syntactic Interference (t=2.16, p=0.03; t=2.21, p=0.03), such that semantic interference effects were greater in the Low Syntactic Interference conditions compared to the High Syntactic Interference conditions (Figure 3.6). Table 3.8 summarizes the results of pairwise comparisons for this observed interaction. Also, a significant effect of reading speed on total reading time was found from the model including reading speed (t=2.14, p=0.03), with longer total reading times elicited by slower readers.

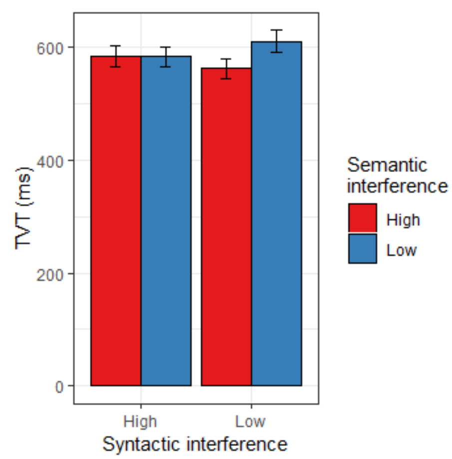

Figure 3.6. Experiment 1: Two-way interaction between Semantic and Syntactic Interference (TVT) in the post-verb region.

Table 3.8. Experiment 1: Pairwise differences of contrast for a two-way interaction between Semantic and Syntactic Interference (TVT) in the post-verb region.

Contrast	*β*	SE	*t*	*p* (>|t|)
SemHigh,SynHigh – SemLow,SynHigh	0.01	0.05	0.16	0.10
SemHigh,SynHigh – SemHigh,SynLow	0.09	0.05	1.73	0.31
SemHigh,SynHigh – SemLow,SynLow	-0.06	0.05	-1.20	0.63
SemLow,SynHigh – SemHigh,SynLow	0.08	0.05	1.57	0.40
SemLow,SynHigh – SemLow,SynLow	-0.07	0.05	-1.35	0.53
SemHigh,SynLow – SemLow,SynLow	**-0.15**	**0.05**	**-2.93**	**0.02**

Note: SemHigh = Semantic High Interference condition; SemLow = Semantic Low Interference condition; SynHigh: Syntactic High Interference condition; SynLow: Syntactic Low Interference condition; β = Estimate coefficients; SE = Standard error. Bold indicates coefficients that are significant at the p<0.05 level.

Regression-out. The three-way interaction model including operation span scores revealed no significant effects (*z*s <1.87, *p*s>0.06). On the other hand, the three-way interaction model with reading speed showed significant main effects of Semantic Interference (z=2.84, p<0.01) and reading speed (z=4.57, p<0.001), and their two-way interaction (z=2.51, p=0.01). As shown in Figure 3.7, the High Semantic Interference conditions elicited more regressions to the previous regions in the sentence relative to the Low Semantic Interference conditions ($M_{SemHigh}$=91%, M_{SemLow}=89%). Also, these semantic interference effects were reduced as reading speed measure (mean reading times per word) decreased (Figure 3.17). That is, faster readers tended to make fewer regressions in the High Semantic Interference conditions compared to the slower readers, and this suggests that fluent readers may be less susceptible to semantic interference at retrieval (Figure 3.8).

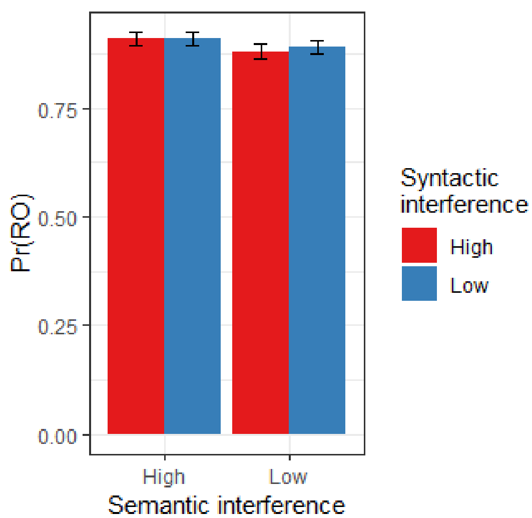

Figure 3.7. Experiment 1: Semantic Interference main effect (RO) in the post-verb region.

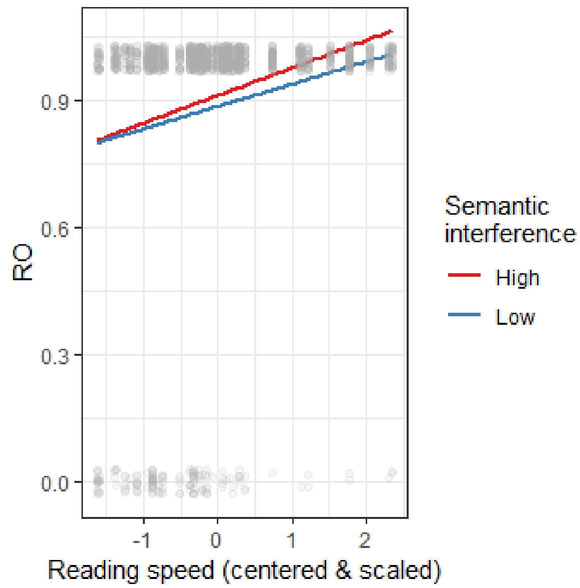

Figure 3.8. Experiment 1: Two-way interaction between Semantic Interference and reading speed (RO) in the post-verb region.

Question-response latency and accuracy. Question-response latency and accuracy were modeled as a function of fixed effects of Semantic Interference (High, Low), Syntactic Interference (High, Low), individual difference measures (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. The question-response latency data were analyzed using the linear mixed-effects models, and the binary accuracy data were analyzed using logit mixed models. Descriptive statistics and summaries of linear mixed-effects models for comprehension question response latency and accuracy are reported in Table 3.9 and 3.10 respectively.

Table 3.9. Experiment 1: Mean performance for comprehension question response latency (in ms) and accuracy.

Sem	*Syn*	*Response Latency*	*Response Accuracy*
High	High	1674 (807)	0.75 (0.44)
	Low	1617 (787)	0.84 (0.37)
Low	High	1533 (730)	0.88 (0.32)
	Low	1561 (768)	0.88 (0.33)

Note: Sem = Semantic Interference; Syn = Syntactic Interference. Standard deviations in parentheses.

Table 3.10. Experiment 1: Mixed-effects modeling results for comprehension question response latency and accuracy.

Predictor	*Response Latency*				*Response Accuracy*			
	β	*SE*	*t*	*p* (>|t|)	*β*	*SE*	*z*	*p* (>|t|)
OSpan								
Intercept	**10.51**	**0.06**	**185.74**	**< 0.0001**	**2.56**	**0.26**	**10.00**	**< 0.0001**
Sem	0.07	0.06	1.40	0.16	**-0.76**	**0.21**	**-3.65**	**< 0.001**
Syn	0.02	0.06	0.36	0.72	-0.42	0.22	-1.94	0.05
OSpan	**-0.10**	**0.04**	**-2.33**	**0.02**	0.17	0.19	0.85	0.39
Sem x Syn	0.04	0.09	0.49	0.62	**-0.79**	**0.16**	**-5.06**	**< 0.0001**
Sem x OSpan	0.04	0.03	1.17	0.24	**-0.79**	**0.21**	**-3.82**	**< 0.001**
Syn x OSpan	-0.01	0.02	-0.59	0.55	0.35	0.21	1.64	0.10
Sem x Syn x OSpan	-0.03	0.05	-0.63	0.52	**0.74**	**0.16**	**4.57**	**< 0.0001**
RS								
Intercept	**10.51**	**0.06**	**182.68**	**< 0.0001**	**2.58**	**0.26**	**9.85**	**< 0.0001**
Sem	0.07	0.05	1.40	0.16	**-0.81**	**0.23**	**-3.48**	**< 0.001**
Syn	0.02	0.06	0.37	0.71	-0.33	0.23	-1.45	0.15
RS	0.07	0.05	1.65	0.10	0.29	0.20	1.47	0.14
Sem x Syn	0.04	0.09	0.49	0.62	**-0.90**	**0.15**	**-5.83**	**< 0.0001**
Sem x RS	-0.02	0.03	-0.86	0.39	0.13	0.23	0.55	0.59
Syn x RS	-0.01	0.02	-0.58	0.56	0.01	0.23	0.05	0.96
Sem x Syn x RS	0.04	0.05	0.78	0.43	**-0.54**	**0.14**	**-3.74**	**< 0.001**

Note: Sem = Semantic Interference; Syn = Syntactic Interference; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; OSpan = Operation span; RS = Reading speed; *β*=Estimate coefficients, SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

Question-response latency. Both three-way interaction models with operation span scores and reading speed were analyzed, and the model including operation span scores revealed a significant effect of operation span scores on question-response latency (t=2.16, p=0.03), with high-capacity readers responding to comprehension questions faster relative to lower-capacity readers. No other reliable main effects and interactions were found (ts <1.40, ps>0.17; ts <1.65, ps>0.11).

Question-response accuracy. The summary of the three-way interaction model with operation span scores revealed a reliable effect of Semantic Interference (z=-3.65, p<0.001), and a two-way interaction between Semantic and Syntactic Interference (z=-5.06, p<0.0001). The High Semantic Interference conditions resulted in lower question-response accuracy relative to the Low Semantic Interference conditions ($M_{SemHigh}$=79%, M_{SemLow}=88%), and as shown in

Table 3.11, the observed two-way interaction between Semantic and Syntactic Interference was driven by semantic interference effects in the High Syntactic Interference condition (Figure 3. 9). There was also a two-way interaction between Semantic Interference and operation span scores (z=-3.82, p<0.001), such that overall, participants with higher operation span scores showed an advantage in comparison with those with lower operation span scores in responding to comprehension questions for the Low Semantic Interference conditions (Figure 3.10). Also, a three-way interaction between the two fixed effects and operation span scores was found (z=4.57, p<0.0001), such that for the Low Semantic Interference conditions, as operation span scores increased, syntactic interference effects decreased, showing higher WM capacity readers' advantage for the Low Syntactic Interference conditions. Whereas, for the High Semantic Interference conditions, this pattern was reversed given that participants with lower operation span scores (lower WM capacity readers) showed a greater advantage for the Low Syntactic Interference conditions compared to those with higher operation span scores (Figure 3.11). The three-way interaction model with reading speed also showed a significant effect of Semantic Interference (z=3.48, p<0.001) and interaction between Semantic and Syntactic Interference (z=-5.83, p<0.0001). Also, there was a reliable three-way interaction between the two fixed effects and reading speed (z=-3.74, p<0.001). The interaction had the following pattern: For the Low Semantic Interference conditions, syntactic interference effects were greater for the slower readers (indicated by increased reading speed measure) relative to faster readers. These effects were driven by slower readers' higher comprehension accuracy for the High Syntactic Interference conditions. Whereas, for the High Semantic Interference conditions, slower readers' greater syntactic interference effects were driven by their higher accuracy for the Low Syntactic Interference conditions (Figure 3.12).

Table 3.11. Experiment 1: Pairwise differences of contrast for a two-way interaction between Semantic and Syntactic Interference (ACC).

Contrast	*β*	*SE*	*z*	*p* (>|z|)
SemHigh,SynHigh – SemLow,SynHigh	**-1.17**	**0.23**	**-5.20**	**<0.0001**
SemHigh,SynHigh – SemHigh,SynLow	**-0.72**	**0.08**	**-8.49**	**<0.0001**
SemHigh,SynHigh – SemLow,SynLow	**-1.16**	**0.23**	**-5.16**	**<0.0001**
SemLow,SynHigh – SemHigh,SynLow	0.45	0.23	1.99	0.19
SemLow,SynHigh – SemLow,SynLow	0.01	0.10	0.09	0.10
SemHigh,SynLow – SemLow,SynLow	-0.45	0.23	-1.95	0.21

Note: SemHigh = Semantic High Interference condition; SemLow = Semantic Low Interference condition; SynHigh: Syntactic High Interference condition; SynLow: Syntactic Low Interference condition; *β*=Estimate coefficients, SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

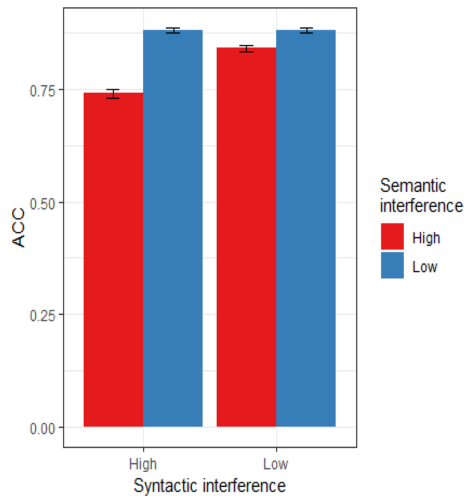

Figure 3.9. Experiment 1:Two-way interaction between Semantic and Syntactic Interference (ACC).

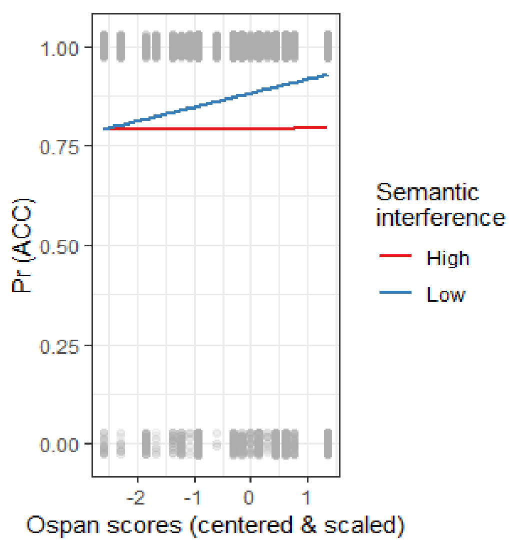

Figure 3.10. Experiment 1: Two-way interaction between Semantic Interference and operation span scores (ACC).

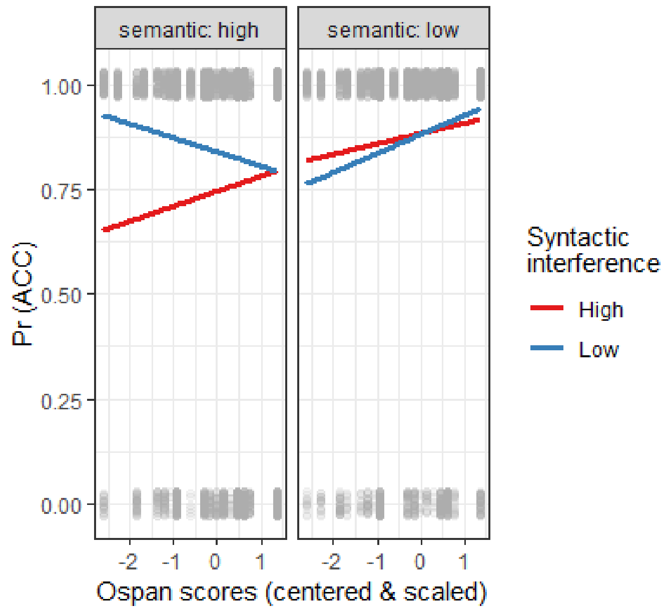

Figure 3.11. Experiment 1: Three-way interaction between Semantic Interference, Syntactic Interference, and operation span scores (ACC).

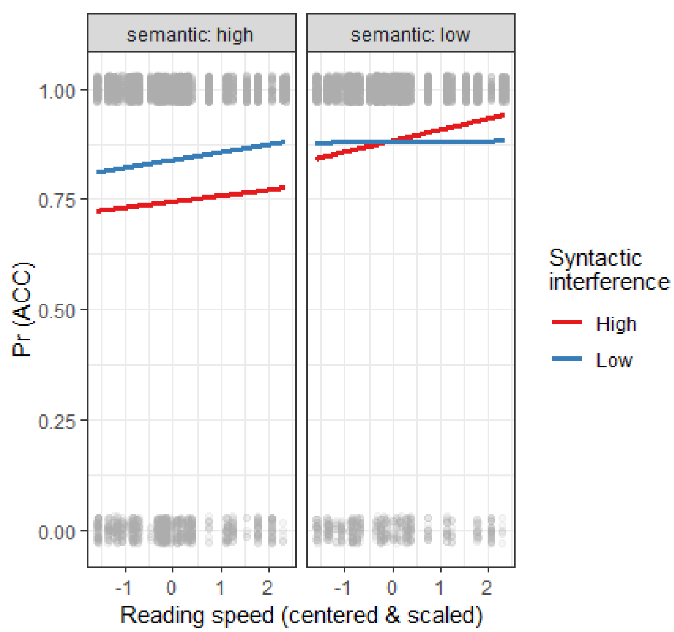

Figure 3.12. Experiment1: Three-way interaction between Semantic Interference, Syntactic Interference, and reading speed (ACC).

Discussion

The skilled L1 readers' sensitivity to semantic and syntactic interference effects provided supportive evidence for the cue-based retrieval account that sentence parsing is accomplished through a series of cue-based retrievals. While previous studies that examined skilled readers' use of cue-based retrieval during sentence reading failed to find interactions between semantic and syntactic interference (Tan et al., 2017; van Dyke et al., 2014), the current experiment found interactions between semantic and syntactic cues in early eye-movement measures, which suggests that all cues are multiplicatively combined into a single retrieval probe (Van Dyke & McElree, 2006).

Also, the finding that only syntactic interference effects were observed in total viewing time indicates that syntactic cues could have been weighted more than semantic cues at later stages of parsing, especially during reading sentences of the High Semantic, High Syntactic condition where semantic interference resolution may depend on using discriminative syntactic cues. This finding is important in that it suggests that cue weighting might need to be incorporated in the cue-based retrieval theory to better understand memory mechanisms underlying sentence processing.

In terms of predictors of the ability to effectively retrieve critical information from memory, only reading speed seemed to be associated with retrieval ability during online reading. Specifically, the findings showed that fast skilled readers tended to be less susceptible to semantic interference effects compared to slower skilled readers. This could be because fast readers with greater language experience may have a high proportion of high-quality lexical representations and are more sensitive to semantic cues (Nicenboim et al., 2016, Traxler et al., 2012), and this may help them to efficiently discriminate lexical items in memory and inhibit spurious activations of irrelevant linguistic information at retrieval (Van Dyke & Johns, 2012; Van Dyke & Shankweiler, 2012). Working memory capacity, on the other hand, was not a strong predictor of the retrieval ability during online sentence processing, which is consistent with the findings of van Dyke et al. (2014).

The models for offline question-response accuracy revealed interference effects, and overall the patterns were similar to those of online reading data, with high interference conditions eliciting lower comprehension accuracy compared to the low interference conditions. These observed interference effects indicate that cue-based retrieval is also employed in an offline task, responding to comprehension questions. What was interesting was that different from online reading data, offline accuracy data showed three-way interactions between interference effects and both individual difference measures, reading speed and working memory capacity. The fact that working memory capacity appeared as a significant predictor of retrieval ability only in the offline data suggests that different mechanisms may be recruited during online reading and answering comprehension questions offline. The implications of this result for education and testing will be discussed in the General Discussion.

Chapter 4: Retrieval Interference in L2 Reading Comprehension

Experiment 2 was designed to examine less-skilled L2 readers' sensitivity to semantic and syntactic retrieval interference during L2 sentence comprehension in order to understand their use of cue-based retrieval during reading. As in Experiment 1, L2 readers' working memory capacity and reading speed were measured to investigate whether these individual difference measures predict less-skilled readers' ability to efficiently retrieve a target in the face of retrieval interference during reading. Experiment 2 was identical to Experiment 1, except that L2 readers additionally took part in a cloze task (indicative of L2 proficiency) and a language background survey.

There has not been strong empirical evidence of L2 readers' sensitivity to similarity-based retrieval interference in L2 sentence comprehension. However, as the retrieval account claims for skilled reading, if we assume that L2 parsing is also accomplished through a series of cue-based memory retrievals, and that similarity-based retrieval interference is the primary source of comprehension difficulty during L2 sentence reading, we should observe increased processing difficulty at the retrieval site and lower question-response accuracy in the high semantic and syntactic interference conditions. Specifically, the High Semantic and Syntactic Interference conditions, in which semantic and syntactic features of the distractor match with those of retrieval cues are predicted to elicit longer reading times at the main verb and post-verb spillover regions and lower question-response accuracy, compared to the Low Semantic and Syntactic Interference conditions.

In particular, based on Cunnings's (2016) claim that L2 readers are more susceptible to interference effects than L1 readers, their syntactic and semantic interference effects are expected to be more pronounced than those of L1 readers. Although Cunnings did not demonstrate why L2 readers should experience more difficulty in resolving retrieval interference compared to L1 speakers, one possibility is that L2 readers may not be able to fully specify the retrieval cues as efficiently as L1 readers do, due to weaker L2 lexical, grammatical, and pragmatic knowledge. These underspecified retrieval cues may not unambiguously distinguish the target representation from the distractor representation, and thus L2 readers are less likely to retrieve the correct target in long-distance dependencies compared to L1 readers. Another possibility is that even if L2 readers fully specify the retrieval cues, they may fail to fully specify features of the target and the distractors at encoding, which may decrease the distinctiveness of the target representation in memory, leading to increased similarity-based interference in long-distance dependencies. If we observe only semantic retrieval interference effects, this may provide supportive evidence for the SSH, which claims that L2 readers are only sensitive to lexical-semantic information, not syntactic information during L2 parsing. Alternatively, if we observe both semantic and syntactic interference effects, this would provide supportive evidence for Good-enough processing of L2 readers because this indicates that L2 sentence processing does indeed proceed along both semantic and syntactic routes (Lim & Christianson, 2013).

No significant retrieval interference effects, on the other hand, may provide evidence in favor of the capacity-based account because while the type of information is varied by manipulating syntactic and semantic properties of the distractor noun, the memory load (the number of words between the subject and the verb) is held constant across the conditions.

As for the predictors of retrieval ability during reading, if working memory capacity is a primary cognitive factor limiting reading ability, L2 readers with lower operation span scores should have more difficulty in resolving semantic and syntactic interference during reading compared to those with higher span scores. That is, interference effects should be larger for readers with lower working memory than higher memory capacity L2 readers. If reading speed is a predictor of L2 readers' ability to resolve similarity-based interference during L2 reading, faster L2 readers are predicted to show smaller interference effects than slower L2 readers, because faster L2 readers with greater L2 experience may be more sensitive to relevant retrieval cues (Nicenboim et al., 2016, Traxler et al., 2012) and thus better at discriminating lexical items in memory, and this ultimately helps them efficiently overcome retrieval interference in long-distance dependencies.

Method

Participants. Forty students from University of Illinois at Urbana-Champaign participated in Experiment 2 in exchange for course credit or $7.00 compensation. All participants were non-native speakers of English, and they have used English for about 8 years on average. The mean scores of participants' Test of English as a Foreign Language (TOEFL) and the cloze task were 102 (out of 120) and 30 (out of 40) respectively, indicating that the recruited participants were highly proficient in English. All participants had normal vision and hearing, with no history of neurological impairment.

Apparatus, materials, procedures, Analysis. Apparatus, materials, and procedures were identical to those in Experiment 1, except that the L2 participants additionally participated in a cloze task and completed a language background survey. During the session, eye-tracking reading, an operation span task, a cloze task, and a short language background survey were conducted in order. The language background survey was adopted and revised from Li, Zhang, Tsai, and Puls (2014).

Results

Eye-tracking. Descriptive statistics of eye-movements for the main verb and post-verb spillover regions are reported in Table 4.1.

Table 4.1. Experiment 2: Mean performance for eye movement data (raw reading times in ms) in the main verb and post-verb spillover regions.

Sem	*Syn*	*GD*	*RPD*	*TVT*	*RI*	*RO*
Main verb region						
High	High	495 (231)	573 (314)	1157 (665)	0.48 (0.50)	0.12 (0.32)
	Low	472 (202)	554 (292)	975 (558)	0.41 (0.49)	0.10 (0.30)
Low	High	477 (245)	631 (420)	1158 (727)	0.49 (0.50)	0.16 (0.37)
	Low	461 (206)	582 (387)	1005 (562)	0.51 (0.50)	0.13 (0.34)
Post-verb region						
High	High	523 (302)	5244 (4023)	980 (605)	-	0.97 (0.17)
	Low	529 (277)	4658 (3628)	945 (562)	-	0.97 (0.17)
Low	High	549 (324)	4847 (3669)	1049 (658)	-	0.97 (0.18)
	Low	522 (304)	4247 (3245)	953 (597)	-	0.95 (0.21)

Note: Sem = Semantic Interference; Syn = Syntactic Interference; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out. Standard deviations in parentheses.

Main verb region. The main verb region was the retrieval site where the retrieval of the subject was triggered, and retrieval cues were generated to establish the subject-verb long-distance dependency. Log-transformed length-adjusted reading times for each eye movement measure were modeled as a function of fixed effects of Semantic Interference (High, Low), Syntactic Interference (High, Low), Individual Difference measure (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. Summaries of linear mixed-effects models for the main verb region is described in Table 4.2.

Gaze duration. Both three-way interaction models including operation span scores and reading speed respectively revealed a significant main effect of Semantic Interference ($t=2.02$, $p=0.04$; $t=2.00$, $p=0.05$), with longer gaze durations in the High Semantic Interference conditions compared to the Low Semantic Interference conditions ($M_{SemHigh}$=484ms, M_{SemLow}=469ms) (Figure 4.1). Also, the summary of the model with operation span scores additionally showed a reliable three-way interaction (t=2.06, p=0.04). For the Low Syntactic Interference conditions, as operation span scores increased semantic interference effects were reduced, suggesting that high memory span L2 readers may be less susceptible to semantic interference compared to low span L2 readers when the syntactic interference is low. This interaction pattern was reversed in the High Syntactic interference conditions (Figure 4.2).

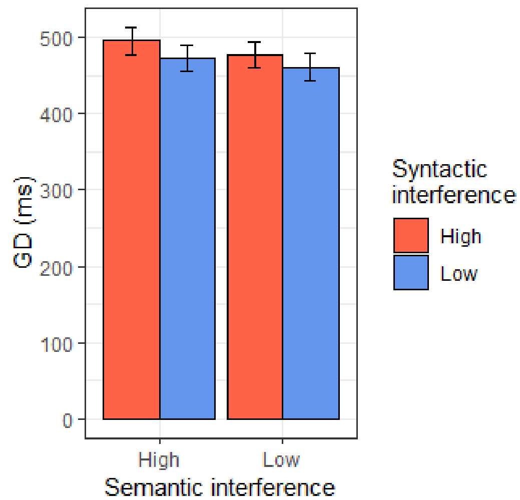

Figure 4.1. Experiment 2: Semantic Interference main effect (GD) in the main verb region.

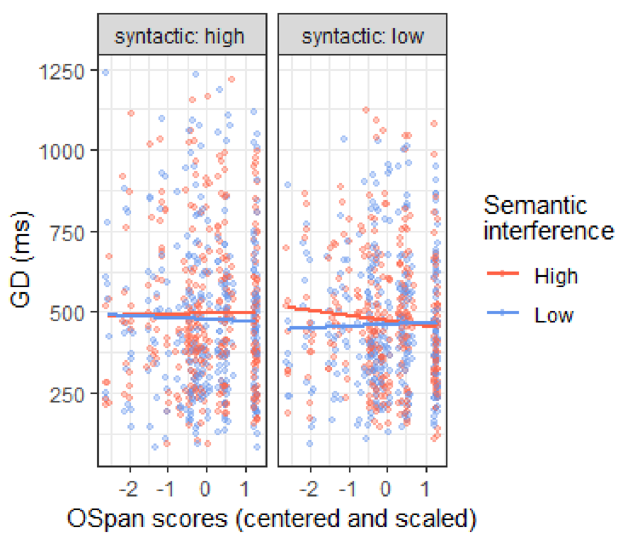

Figure 4.2. Experiment 2: Three-way interaction between Semantic Interference, Syntactic Interference, and operation span scores (GD) in the main verb region.

Table 4.2. Experiment 2: Mixed-effects modeling results for all dependent measures at the main verb region.

	Fixation duration measure														
	GD			*RPD*			*TVT*			*RI*			*RO*		
Predictor	β	SE	t	β	SE	t	β	SE	t	β	SE	z	β	SE	z
OSpan															
Intercept	0.00	0.03	0.01	0.00	0.04	0.02	0.00	0.05	0.02	-0.12	0.13	-0.92	**-2.12**	**0.15**	**-14.30**
Sem	**0.06**	**0.03**	**2.02**	-0.05	0.04	-1.43	-0.02	0.04	-0.51	**-0.25**	**0.12**	**-2.07**	-0.32	0.17	-1.85
Syn	0.02	0.03	0.86	0.05	0.04	1.41	**0.20**	**0.04**	**5.10**	0.12	0.12	1.03	0.22	0.17	1.25
OSpan	0.00	0.02	0.01	-0.01	0.02	-0.33	-0.01	0.02	-0.34	-0.14	0.10	-1.33	-0.12	0.10	-1.17
Sem x Syn	0.05	0.07	0.77	-0.02	0.07	-0.28	0.11	0.08	1.40	0.41	0.24	1.71	-0.12	0.35	-0.36
Sem x OSpan	-0.02	0.03	-0.78	0.04	0.04	1.11	-0.04	0.04	-1.08	-0.15	0.12	-1.21	0.19	0.17	1.06
Syn x OSpan	0.00	0.03	0.04	-0.04	0.04	-1.05	-0.03	0.04	-0.67	0.00	0.12	-0.01	-0.10	0.17	-0.56
Sem x Syn x OSpan	**0.13**	**0.07**	**2.06**	0.14	0.07	1.88	0.06	0.08	0.80	-0.24	0.24	-0.99	-0.61	0.35	-1.75
RS															
Intercept	0.00	0.03	0.00	0.00	0.04	0.03	0.00	0.05	0.02	-0.12	0.13	-0.93	**-2.12**	**0.15**	**-14.19**
Sem	**0.07**	**0.03**	**2.00**	-0.05	0.04	-1.45	-0.02	0.04	-0.50	**-0.25**	**0.12**	**-2.10**	-0.33	0.17	-1.89
Syn	0.03	0.03	0.86	0.05	0.04	1.41	**0.20**	**0.04**	**5.12**	0.12	0.12	1.01	0.20	0.17	1.15
RS	0.02	0.02	0.95	0.02	0.02	0.96	0.04	0.02	1.90	**0.24**	**0.10**	**2.40**	-0.13	0.11	-1.13
Sem x Syn	0.05	0.07	0.77	-0.02	0.07	-0.30	0.11	0.08	1.41	0.41	0.24	1.72	-0.13	0.35	-0.37
Sem x RS	-0.04	0.03	-1.10	-0.04	0.04	-0.96	-0.08	0.04	-1.90	-0.11	0.12	-0.86	-0.17	0.21	-0.80
Syn x RS	-0.01	0.03	-0.23	0.01	0.04	0.32	-0.03	0.04	-0.83	-0.17	0.12	-1.36	0.10	0.21	0.47
Sem x Syn x RS	0.13	0.07	1.90	0.00	0.07	0.05	-0.11	0.08	-1.42	-0.20	0.24	-0.79	-0.55	0.41	-1.34

Note: Sem = Semantic Interference; Syn = Syntactic Interference; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; OSpan = Operation span; RS = Reading speed; β=Estimate coefficients; SE=standard error. Bold indicates coefficients that are significant at the $p<0.05$ level.

Regression-path duration. No significant effects were found in both three-way interaction models with operation span scores and reading speed (*t*s <1.88, *p*s>0.06; *t*s <1.45, *p*s>0.15).

Total viewing time. Summaries of both three-way interaction models including operation span scores and reading speed revealed a significant effect of Syntactic Interference (t=5.10, p<0.0001; t=5.11, p<0.0001), with longer total reading times in the High Syntactic Interference conditions compared to the Low Syntactic Interference conditions ($M_{SynHigh}$=1158ms, M_{SynLow}=990ms) (Figure 4.3).

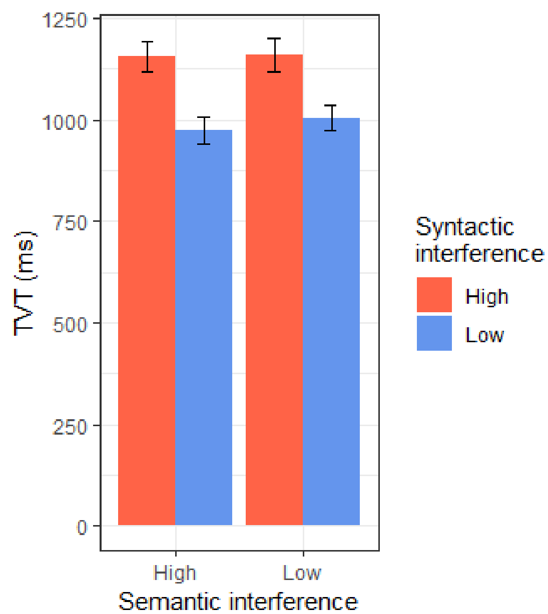

Figure 4.3. Experiment 2: Syntactic Interference main effects (TVT) in the main verb region.

Regression-in. Summaries of both three-way interaction models with operation span scores and reading speed showed a significant effect of Semantic Interference effect (z=-2.07, p=0.04; z=-2.10, p=0.04), such that the probability of regression made back into the main verb region was smaller in the High Semantic Interference conditions compared to the Low Semantic Interference conditions ($M_{SemHigh}$=45%, M_{SemLow}=50%) (Figure 4.4). Also, the model with reading speed showed an average reading speed effect across Semantic and Syntactic Interference fixed factors (z= 2.40, p=0.02); slower readers tended to regress back into the main verb region more often than the fast readers on average.

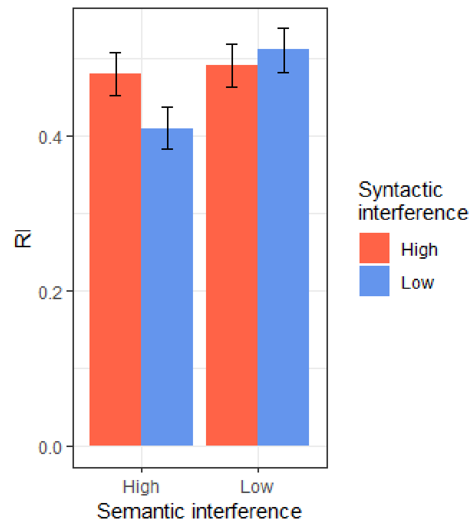

Figure 4.4. Experiment 2: Semantic Interference main effects (RI) in the main verb region.

Regression-out. No significant effects were found in both three-way interaction models with operation span scores and reading speed (zs <1.85, ps>0.06; zs <1.89, ps>0.05).

Post-verb region. The post-verb region was the spillover region following the retrieval site (the main verb) and included either an adverbial phrase or the patient of the main verb. Log-transformed, length-adjusted reading times for each eye movement measure were modeled as a function of Semantic Interference (High, Low), Syntactic Interference (High, Low), Individual Difference measure (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. Summaries of linear mixed-effects models for the post-verb region is described in Table 4.3. Given that the post-verb region was the last region in the sentence, regression-in data were not available, and thus no analysis for the regression-in measure was conducted.

Table 4.3. Experiment 2: Mixed-effects modeling results for all dependent measures at the post-verb spillover region.

	Fixation duration measure														
	GD			*RPD*			*TVT*			*RI*			*RO*		
Predictor	*β*	SE	*t*	*β*	SE	*t*	*β*	SE	*t*	*β*	SE	*z*	*β*	SE	*z*
OSpan															
Intercept	0.00	0.04	0.01	0.00	0.05	0.02	0.00	0.05	0.01	-	-	-	**3.38**	**0.16**	**21.18**
Sem	0.00	0.04	-0.04	0.10	0.05	1.85	-0.02	0.04	-0.41	-	-	-	0.32	0.32	1.01
Syn	-0.01	0.04	-0.21	**0.17**	**0.05**	**3.27**	0.06	0.05	1.05	-	-	-	0.21	0.32	0.65
OSpan	0.00	0.02	0.05	0.00	0.03	-0.11	-0.01	0.02	-0.29	-	-	-	-0.05	0.16	-0.33
Sem x Syn	-0.07	0.07	-0.96	0.05	0.11	0.43	-0.04	0.08	-0.54	-	-	-	-0.40	0.64	-0.63
Sem x OSpan	-0.05	0.04	-1.30	-0.09	0.05	-1.60	**-0.09**	**0.04**	**-2.30**	-	-	-	-0.13	0.32	-0.42
Syn x OSpan	-0.07	0.04	-1.95	0.02	0.05	0.30	-0.03	0.04	-0.85	-	-	-	-0.14	0.32	-0.43
Sem x Syn x OSpan	-0.08	0.07	-1.08	0.02	0.11	0.19	-0.05	0.08	-0.60	-	-	-	0.21	0.65	0.32
RS															
Intercept	0.00	0.04	0.02	0.00	0.05	0.05	0.00	0.05	0.01	-	-	-	**5.13**	**0.58**	**8.92**
Sem	0.00	0.04	-0.01	0.10	0.07	1.48	-0.02	0.04	-0.40	-	-	-	0.05	0.68	0.07
Syn	-0.01	0.04	-0.19	**0.17**	**0.07**	**2.62**	0.06	0.05	1.03	-	-	-	-0.02	0.68	-0.02
RS	0.02	0.02	0.83	0.05	0.03	1.74	0.03	0.02	1.65	-	-	-	**2.43**	**0.82**	**2.95**
Sem x Syn	-0.07	0.07	-0.96	0.04	0.10	0.42	-0.04	0.08	-0.51	-	-	-	-1.29	1.37	-0.94
Sem x RS	0.01	0.04	0.34	-0.06	0.05	-1.17	**-0.08**	**0.04**	**-2.04**	-	-	-	-0.59	1.02	-0.58
Syn x RS	0.04	0.04	0.96	-0.04	0.05	-0.76	0.01	0.04	0.18	-	-	-	-0.49	1.02	-0.48
Sem x Syn x RS	**0.15**	**0.08**	**2.01**	0.02	0.11	0.15	0.06	0.08	0.80	-	-	-	-1.45	2.06	-0.70

Gaze duration. The only reliable effect found in gaze duration was a three-way interaction between Semantic Interference, Syntactic Interference, and reading speed (t=2.01, p=0.05). As shown in Figure 4.5, overall slower L2 readers were more susceptible to syntactic interference effects given that as reading speed increased, syntactic interference effects increased. The influence of reading speed on syntactic interference effects, however, was greater in the Low Semantic Interference conditions compared to those in the High Semantic Interference conditions.

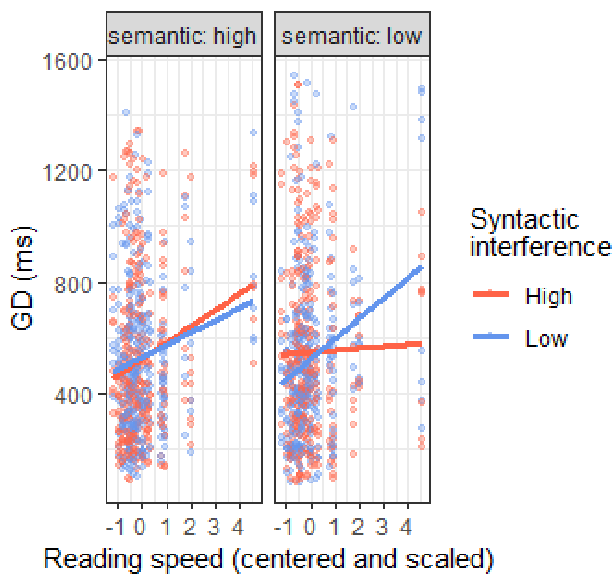

Figure 4.5. Experiment 2: Three-way interaction between Semantic Interference, Syntactic Interference, and operation span scores (GD) in the post-verb region.

Regression-path duration. Summaries of both three-way interaction models including operation span scores and reading speed revealed a significant main effect of Syntactic Interference (t=3.27, p<0.01; t=2.62, p=0.01), with longer regression-path duration for the High Syntactic Interference conditions compared to the Low Syntactic Interference conditions ($M_{SynHigh}$ =5047ms, M_{SynLow} =4452ms) (Figure 4.6).

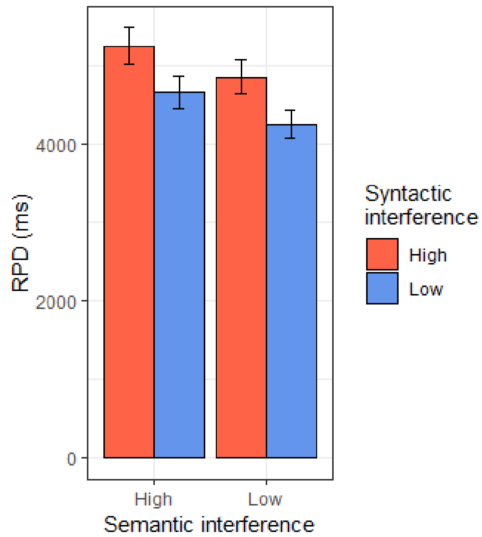

Figure 4.6. Experiment 2: Syntactic Interference main effects (RPD) in the post-verb region.

Total viewing time. The three-way interaction model including operation span scores showed a reliable two-way interaction between Semantic Interference and operation span scores (t=-2.30, p=0.02). As shown in Figure 4.7, L2 readers with higher capacity of working memory tended to show greater advantage in processing sentences with high semantic interference compared to low memory span L2 readers. Also, a significant interaction between Semantic Interference and reading speed was found in the three-way interaction model including reading speed (t=-2.04, p=0.04). As reading speed increased, semantic interference effects increased, indicating that slower L2 readers tended to be more susceptible to semantic interference during sentence reading (Figure 4.8).

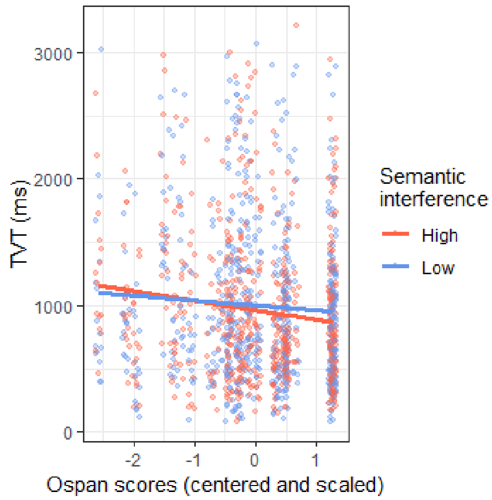

Figure 4.7. Experiment 2: Two-way interaction between Semantic Interference and operation span scores (TVT) in the post-verb region.

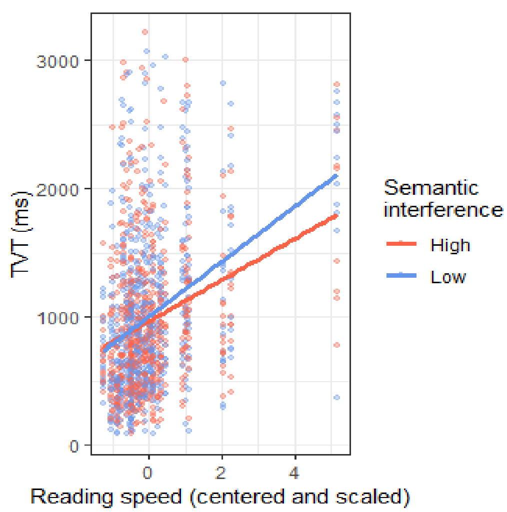

Figure 4.8. Experiment 2: Two-way interaction between Semantic Interference and Reading speed (TVT) in the post-verb region.

Regression-out. The only significant effect found in the regression-out measure was a main effect of reading speed (z=2.95, p<0.01); Slower L2 readers seemed to regress back into the previous regions more often than faster L2 readers.

Question-response latency and accuracy. As in Experiment 1, question-response latency and accuracy were modeled as a function of fixed effects of Semantic Interference (High, Low), Syntactic Interference (High, Low), individual difference measures (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. Question-response latency and accuracy were analyzed using linear mixed-effects and logit mixed-effects models, respectively. Table 4.4 and 4.5 show their descriptive statistics and summaries of linear mixed-effects models.

Table 4.4. Experiment 2: Mean performance for comprehension question response latency (in ms) and accuracy.

Sem	*Syn*	*Response Latency*	*Response Accuracy*
High	High	2551 (1579)	0.72 (0.45)
	Low	2213 (1529)	0.88 (0.32)
Low	High	2279 (1390)	0.85 (0.36)
	Low	2035 (1026)	0.92 (0.27)

Note: Sem = Semantic Interference; Syn = Syntactic Interference. Standard deviations in parentheses.

Table 4.5. Experiment 2: Mixed-effects modeling results for comprehension question response latency and accuracy.

	Response Latency				*Response Accuracy*			
Predictor	*β*	*SE*	*t*	*p* (>|t|)	*β*	*SE*	*z*	*p* (>|t|)
OSpan								
Intercept	**10.95**	**0.08**	**135.37**	**< 0.0001**	**2.07**	**0.16**	**12.56**	**< 0.0001**
Sem	0.09	0.06	1.54	0.12	**-0.68**	**0.07**	**-9.49**	**< 0.0001**
Syn	**0.14**	**0.07**	**2.16**	**0.03**	**-0.98**	**0.07**	**-13.69**	**< 0.0001**
OSpan	-0.02	0.07	-0.23	0.81	0.02	0.16	0.14	0.80
Sem x Syn	0.06	0.12	0.49	0.63	**-0.42**	**0.14**	**-2.96**	**< 0.001**
Sem x OSpan	-0.06	0.03	-1.63	0.11	**0.36**	**0.07**	**5.41**	**< 0.0001**
Syn x OSpan	0.00	0.03	-0.01	0.99	0.10	0.07	1.51	0.10
Sem x Syn x OSpan	-0.01	0.07	-0.19	0.85	-0.24	0.13	-1.79	0.07
RS								
Intercept	**10.93**	**0.07**	**161.03**	**< 0.0001**	**2.08**	**0.17**	**12.51**	**< 0.0001**
Sem	0.09	0.06	1.52	0.13	**-0.70**	**0.07**	**-9.72**	**< 0.0001**
Syn	**0.15**	**0.07**	**2.16**	**0.03**	**-1.00**	**0.07**	**-13.70**	**< 0.0001**
RS	**0.24**	**0.05**	**5.04**	**< 0.0001**	0.14	0.17	0.83	0.40
Sem x Syn	0.06	0.12	0.51	0.61	**-0.40**	**0.14**	**-2.74**	**<0.01**
Sem x RS	-0.01	0.03	-0.15	0.88	-0.15	0.09	-1.68	0.09
Syn x RS	-0.01	0.03	-0.24	0.81	**-0.36**	**0.09**	**-4.04**	**< 0.0001**
Sem x Syn x RS	-0.04	0.07	-0.60	0.55	**0.41**	**0.17**	**2.40**	**0.01**

Note: Sem = Semantic Interference; Syn = Syntactic Interference; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; OSpan = Operation span; RS = Reading speed; *β*=Estimate coefficients, SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

Question-response latency. Both three-way interaction models including operation span scores and reading speed showed a significant main effect of Syntactic Interference (*t*=2.16, *p*=0.03; *t*=2.16, *p*=0.03), with longer question-response times for the High Syntactic Interference conditions relative to the Low syntactic Interference conditions ($M_{SynHigh}$ =2415ms, M_{SynLow}=2123ms) (Figure 4.9). Also, there was a significant effect of reading speed on question-response latency (*t*=5.04, *p*<0.0001), such that slower L2 readers showed a longer question-response latency compared to faster L2 readers.

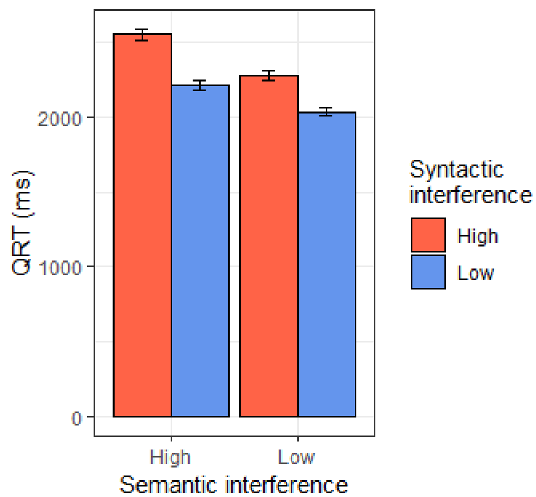

Figure 4.9. Experiment 2: Syntactic Interference main effects (QRT).

Question-response accuracy. Summaries of both three-way interaction models including operation span scores and reading speed revealed significant main effects of Semantic Interference (*z*=-9.49, *p*<0.0001; *z*=-9.72, *p*<0.0001), Syntactic Interference (*z*=-13.69, *p*<0.0001; *z*=-13.70, *p*<0.0001), and their two-way interaction (*z*=-2.96, *p*<0.01; *z*=-2.74, *p*<0.01). As shown in Figure 4.10 and Table 4.6, the High Semantic Interference conditions resulted in a lower question-response accuracy compared to the Low Semantic Interference conditions ($M_{SemHigh}$=80%, M_{SemLow}=88%). The same pattern was also observed for syntactic interference effects, with the High Syntactic Interference conditions eliciting lower question-response accuracy relative to the Low Syntactic interference conditions ($M_{SynHigh}$=78%, M_{SynLow}=90 %). As for the two-way interaction, syntactic interference effects were greater in the High Semantic Interference conditions than those in the Low Semantic Interference conditions. Also, the models showed two significant two-way interaction effects between Semantic Interference and operation span scores (z=5.41, p<0.0001), as well as between Syntactic Interference and reading speed (z=-4.04, p<0.001). Lower span L2 readers showed greater semantic interference effects than higher span L2 readers (Figure 4.11), and slower L2 readers showed greater syntactic interference effects than fast L2 readers (Figure 4.12). Additionally, a significant three-way interaction between the two fixed effects and reading speed was found (z=2.40, p=0.01). As shown in Figure 4.13, slower L2 readers were more susceptible to syntactic interference effects, and overall the influence of reading speed on syntactic interference effects was larger for the High Semantic Interference conditions than the Low Semantic Interference conditions.

Table 4.6. Experiment 2: Pairwise differences of contrast for a two-way interaction between Semantic and Syntactic Interference (ACC).

Contrast	*β*	*SE*	*z*	*p* (>|z|)
SemHigh,SynHigh – SemLow,SynHigh	**-1.17**	**0.26**	**-4.57**	**<0.0001**
SemHigh,SynHigh – SemHigh,SynLow	**-1.30**	**0.10**	**-13.63**	**<0.0001**
SemHigh,SynHigh – SemLow,SynLow	**-1.97**	**0.26**	**-7.47**	**<0.0001**
SemLow,SynHigh – SemHigh,SynLow	-0.13	0.26	-0.48	0.96
SemLow,SynHigh – SemLow,SynLow	**-0.80**	**0.11**	**-7.20**	**<0.0001**
SemHigh,SynLow – SemLow,SynLow	-0.67	0.27	-2.49	0.06

Note: SemHigh = Semantic High Interference condition; SemLow = Semantic Low Interference condition; SynHigh: Syntactic High Interference condition; SynLow: Syntactic Low Interference condition; *β*=Estimate coefficients, SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

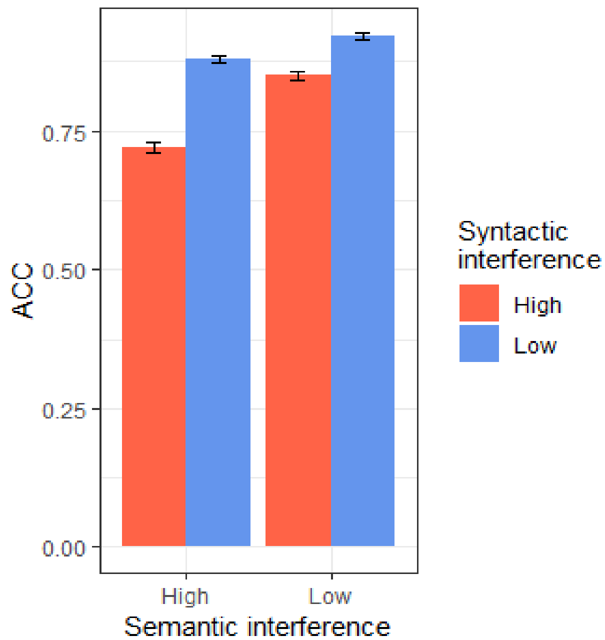

Figure 4.10. Experiment 2: Main effects of Semantic and Syntactic Interference, and their two-way interaction (ACC).

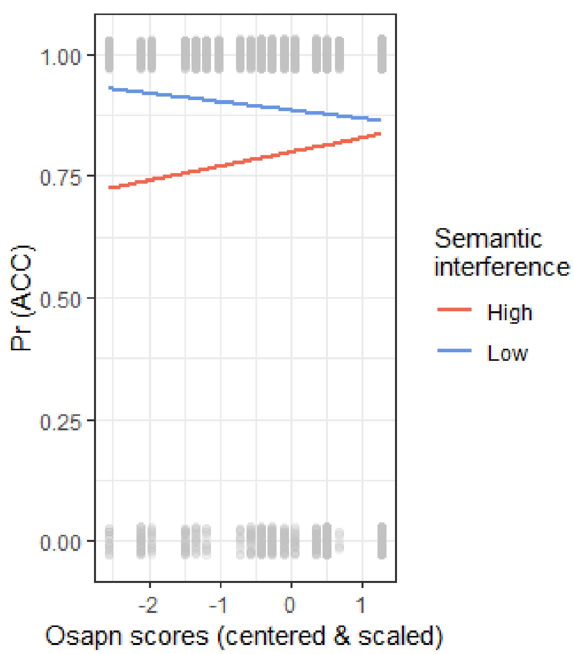

Figure 4.11. Experiment 2: Two-way interaction between Semantic Interference and operation span scores (ACC).

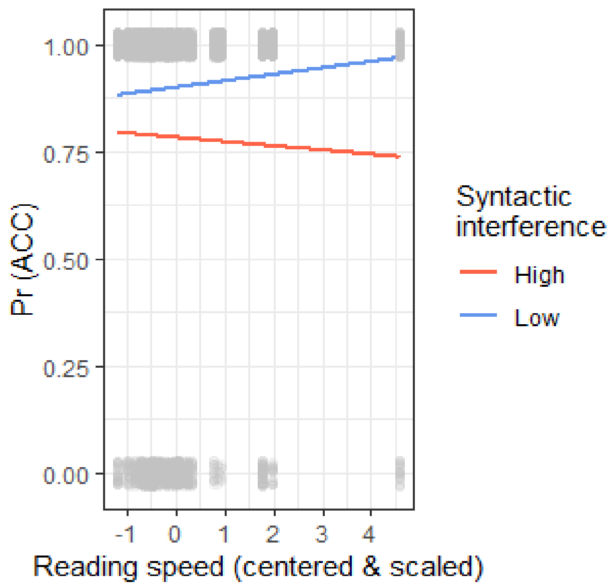

Figure 4.12. Experiment 2: Two-way interaction between Syntactic Interference and reading speed (ACC).

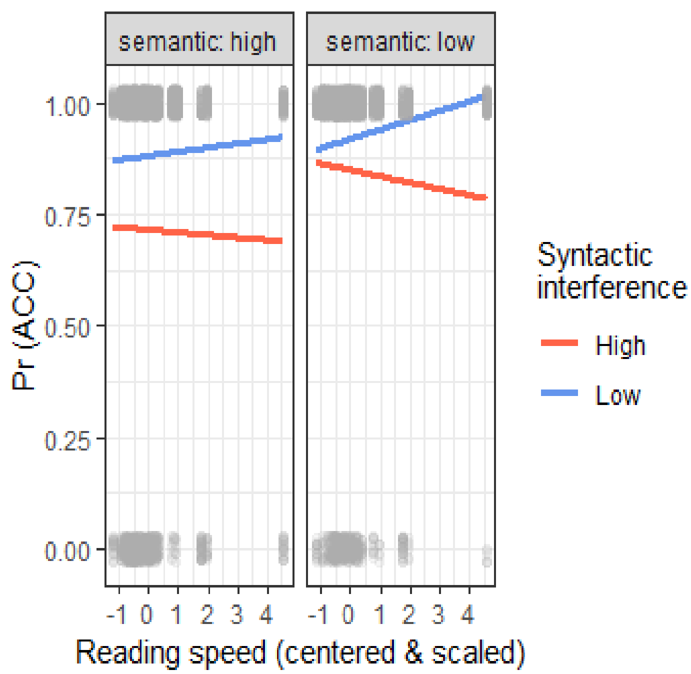

Figure 4.13. Experiment 2: Three-way interaction between Semantic Interference, Syntactic Interference, and reading speed (ACC).

Discussion

The observed interference effects in the L2 data show that less-skilled L2 readers also employ efficient cue-based retrieval during L2 sentence comprehension, as skilled L1 readers do. However, different from the L1 readers' data, which showed a two-way interaction between semantic and syntactic interference in all fixation duration measures, the L2 data failed to reveal reliable two-way interactions. This suggests that L2 readers may have difficulty in combining all relevant cues into a single retrieval probe or simultaneously using multiple cues at retrieval, possibly due to limited computational resources, taxed by the activation and inhibition of L1 or reduced automaticity in L2.

The findings that both semantic and syntactic interference effects were found in online reading data provides evidence against the Shallow Structure Hypothesis, claiming that L2 learners are only sensitive to lexical-semantic information during L2 parsing. In particular, the L2 readers and L1 readers alike relied more on syntactic cues at the later stage of parsing to resolve interference. This finding supports the Good-enough processing theory that L2 (and L1) sentence processing proceeds along both semantic and syntactic routes. Furthermore, it appears that L2 readers might be less efficient at integrating semantic and syntactic cues, as suggested by Lim and Christianson (2013b).

What was especially interesting in the L2 data was that both reading speed and working memory capacity moderated interference effects in fixation duration and probability measures. Although only reading speed was a predictor of the retrieval ability of skilled L1 readers, in the case of the L2 data, not only reading speed but also working memory capacity influenced L2 readers' performance in resolving interference during sentence processing. Overall L2 readers who read faster and had higher working memory capacity were less susceptible to interference effects compared to slower and lower span L2 readers.

Similar to the L1 data, the models for question-response accuracy revealed interference effects and moderating effects of reading speed and working memory capacity on interference effects. These findings demonstrate that, like skilled readers, less-skilled readers also employ a cue-based retrieval mechanism in offline comprehension tasks, and show that faster and higher-span L2 readers tend to be more successful in retrieving the correct target in responding comprehension questions.

Chapter 5: Quality of Lexical Representations in L1 Reading Comprehension

Whereas Experiment 1 and 2 examined similarity-based retrieval interference during L1 and L2 sentence comprehension, Experiment 3 was designed to investigate whether enhancing the quality of lexical representations through semantic elaboration leads to facilitated retrieval in the subject-verb, long-distance dependency. Given that the cue-based retrieval account does not make predictions about the features that are not cued by the retrieval trigger, investigating semantic elaboration effects on retrieval efficiency is import to further develop the cue-based retrieval theory.

To vary the quality of lexical representations of the target and the distractor, the number of modifying words for the target and the distractor was manipulated, creating simple and complex semantic elaboration conditions (Target Complexity: Complex vs. Simple, Distractor Complexity: Complex vs. Simple). The critical sentences were adopted from those of the High Semantic, High Syntactic Interference condition in Experiment 1 and 2, and were revised so that they had either complex or simple targets and distractors. RTs at the main verb and the spillover regions, and question-response latency and accuracy were analyzed to understand whether the quality of lexical representations influences skilled L1 readers' retrieval efficiency during sentence comprehension. In addition, three-way interaction between the two fixed effects and individual difference measures, which are reading speed and working memory capacity were examined to understand their moderating effects on semantic elaboration effects.

Based on prior research that showed semantic elaboration effects on memory retrieval (Gallo et al., 2008; Hofmeister, 2011; Hofmeister & Vasishth, 2014; Troyer et al., 2016), the Complex semantic elaboration conditions were predicted to elicit shorter RTs at the retrieval sites and higher question-response accuracies, compared to the Simple semantic elaboration conditions, because semantic elaboration for either the target or the distractor may increase the uniqueness of the target representation, leading to efficient retrieval. Along with effects of semantic complexity, an interaction of the target and the distractor semantic elaboration effects was expected to be significant, because the target semantic complexity may lead to faster RTs at the retrieval sites only when the distractor is simple. The rationale underlying this predicted interaction was that adding a lot of unique information will be less likely to contribute to efficient retrieval given the law of diminishing returns (Hofmeister & Vasishth, 2014). Thus, the Complex Target, Complex Distractor condition may not facilitate the target retrieval to the same extent as the other two conditions do (the Complex Target, Simple Distractor, and the Simple Target, Complex Distractor conditions).

Assuming that adding modifiers for either the target or the distractor facilitates retrieval latency and leads to successful retrieval probability as previous research demonstrated (Hofmeister & Vasishth, 2014), and that both reading speed and working memory capacity are individual difference measures that moderate semantic elaboration effects on retrieval efficiency, faster and higher working memory span L1 readers were predicted to show a greater advantage for the complex semantic elaboration conditions compared to slower and lower span L1 readers (shorter RTs at the retrieval sites and a higher question-response accuracy for the complex semantic elaboration conditions). Faster readers, assumed to have greater language experience than slower readers, may be more sensitive to semantically elaborated information (Traxler et al., 2012), and thus this would help them efficiently distinguish the target representation from that of the distractor. Although the nature of working memory still remains controversial, based on the assumption that working memory is involved in either storage or retrieval processes during sentence reading, high span readers were expected to be more efficient at encoding additional semantic information or holding/retrieving semantically elaborated information at retrieval.

Method

Participants. Fifty-seven students from University of Illinois at Urbana-Champaign participated in Experiment 3 in exchange for course credit or $7.00 compensation. All participants were native speakers of English and had normal vision and hearing, with no history of neurological impairment.

Apparatus and procedures. The apparatus and procedures were identical to those in Experiment 1, except that during the eye-tracking reading task, participants were instructed to respond to a comprehension question by pressing the button on the controller that corresponded to the correct answer, instead of verbally providing their responses to comprehension questions as in Experiment 1.

Materials. Experiment 3 had a repeated measures 2 x 2 design (Target Complexity: Complex vs. Simple, Distractor Complexity: Complex vs. Simple), manipulating the number of modifying words for the subject noun (target) and the intervening noun (distractor). To vary the quality of lexical representations of the target and the distractor through semantic elaboration, 32 critical sentences of the High Semantic, High Syntactic Interference condition from Experiment 1 were adapted so that they had either simple or complex modifiers for the target and the distractor. As in Table 5.1, each sentence contained a main clause and a subject-extracted RC intervening the subject and the main verb. In the Simple Target conditions, the target (*resident*) appeared without any modifying words, whereas in the Complex Target conditions, the target appeared with two modifying words. Similarly, in the Simple Distractor conditions, the distractor (*neighbor*) appeared without any modifying words, whereas in the Complex Distractor condition the distractor appeared with two modifying words.

Table 5.1. Example stimuli in Experiment 3.

		Target Complexity	
		Complex	*Simple*
Distractor Complexity	*Complex*	The **cautious senior resident** who said that the **tall, young neighbor** was dangerous last month ***had complained*** about the investigation.	The **resident** who said that the **tall, young neighbor** was dangerous last month ***had complained*** about the investigation.
	Simple	The **cautious senior resident** who said that the **neighbor** was dangerous last month ***had complained*** about the investigation.	The **resident** who said that the **neighbor** was dangerous last month ***had complained*** about the investigation.

Hofmeister (2011) demonstrated that modifying words for the target and the distractor that were unlikely based on real-world knowledge did not facilitate target retrieval, relative to the baseline condition, so Experiment 3 only included modifying words that were likely to describe attributes of the target and the distractor in real-world. As descried in Experiment 1, an adverbial phrase was inserted between the intervening region and the main verb region to avoid local coherence effects (Glaser et al., 2013; Tabor et al., 2004; Tan et al., 2017, Van Dyke, 2007). The critical regions were the main verb region where the target retrieval was triggered and the post-verb spillover region containing either an adverbial phrase or the patient of the main verb. RTs at the critical regions and question-response latency and accuracy were obtained. Operation span scores and average reading times per word were computed for each participant to measure their working memory capacity and reading speed. In addition to 32 sets of critical sentences, 68 fillers and 5 practice sentences from Experiment 1 were included in all four lists. Critical sentences were distributed across four lists using a Latin Square design, with conditions counterbalanced across lists.

Analysis. Prior to the analyses of eye-movement measures and question-response latency and accuracy, one participant was excluded from the analyses due low recall accuracy in the operation span task (20% accuracy), which left fifty-six participants for the analysis.

As in Experiments 1 and 2, fixation duration and probability measures in critical regions were analyzed with linear mixed-effects and logistic mixed-effects models, using the lmerTest package (version 3.0-1) in R (version 3.4.4). Prior to the statistical analyses of eye movement measures, fixations shorter than 80ms and longer than 1200ms were removed by using the EyeLink Data Viewer program. For each fixation duration measure, fixations above 2.5 standard deviations from the mean of each condition were excluded from data analyses. To meet the error normality assumption and residualize word length effects, raw fixations in all fixation duration measures were log-transformed and length-adjusted. After trimming the eye-movement data, fixation duration and probability measures were modeled as a function of Target Complexity (Complex, Simple), Distractor Complexity (Complex, Simple), individual Difference measure (operation span score or reading speed), and their three-way interaction to examine semantic elaboration effects and moderating effects of individual difference measures on semantic elaboration effects. Sum contrasts were used for the two categorical fixed effects (Complex as +1 and Simple as -1). All linear mixed-effects models included random effects of Participants and Items to account for variability in participants and sentence items, and the maximal random effect structure was justified by model comparison. Question-response latency and accuracy were also analyzed using linear mixed-effects and logistic mixed-effects models, respectively, and the same fixed effect parameters as those for eye-movement data were contained in the models. Reading speed and operation span scores were obtained in the same way as in Experiments 1 and 2, and they were all centered and scaled prior to statistical analyses.

Results

Eye-tracking. Descriptive statistics of eye-movements for the main verb and post-verb regions are reported in Table 5.2.

Table 5.2. Experiment 3: Mean performance for eye movement data (raw reading times in ms) in the main verb and post-verb spillover regions.

TargComp	*DistComp*	*GD*	*RPD*	*TVT*	*RI*	*RO*
Main verb region						
Comp	Comp	346 (150)	548 (540)	639 (371)	0.34 (0.48)	0.20 (0.40)
	Sim	340 (153)	521 (484)	646 (356)	0.36 (0.48)	0.19 (0.39)
Sim	Comp	343 (153)	472 (383)	678 (393)	0.36 (0.48)	0.16 (0.37)
	Sim	347 (145)	483 (428)	674 (381)	0.31 (0.46)	0.15 (0.36)
Post-verb region						
Comp	Comp	371 (207)	2467 (2352)	493 (348)	-	0.84 (0.37)
	Sim	368 (205)	2259 (1971)	515 (366)	-	0.85 (0.36)
Sim	Comp	348 (191)	2504 (2038)	518 (361)	-	0.88 (0.32)
	Sim	377 (234)	2234 (1846)	522 (387)	-	0.84 (0.37)

Note: TargComp = Target Complexity; DistComp = Distractor Complexity; Comp = Complex; Sim = Simple; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out. Standard deviations in parentheses.

Main verb region. The main verb region was the retrieval site where target retrieval was triggered, and retrieval cues were generated to establish a long-distance dependency. Log-transformed, length-adjusted fixation duration and fixation probability measures were modeled as a function of fixed effects of Target Complexity (Complex, Simple), Distractor Complexity (Complex, Simple), individual difference measure (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. Summaries of linear mixed-effects models for the main region are described in Table 5.3.

Table 5.3. Experiment 3: Mixed-effects modeling results for all dependent measures at the main verb region.

	Fixation duration measure														
	GD			*RPD*			*TVT*			*RI*			*RO*		
Predictor	*β*	SE	*t*	*β*	SE	*t*	*β*	SE	*t*	*β*	SE	*z*	*β*	SE	*z*
OSpan															
Intercept	0.00	0.03	0.01	0.00	0.05	0.08	0.00	0.04	0.08	**-0.76**	**0.13**	**-5.98**	**-1.83**	**0.18**	**-10.31**
TargComp	-0.01	0.03	-0.48	**0.08**	**0.04**	**2.01**	-0.07	0.03	-1.93	0.08	0.11	0.80	**0.30**	**0.13**	**2.26**
DistComp	0.00	0.03	-0.10	0.00	0.04	0.08	-0.01	0.03	-0.33	0.09	0.11	0.82	0.10	0.13	0.77
OSpan	0.01	0.01	0.75	0.02	0.02	0.82	0.01	0.02	0.52	0.00	0.11	-0.01	0.00	0.11	-0.04
TargComp x DistComp	0.06	0.06	1.11	0.05	0.08	0.67	-0.01	0.04	-0.14	-0.35	0.21	-1.65	-0.01	0.26	-0.03
TargComp x OSpan	-0.01	0.03	-0.42	-0.01	0.04	-0.19	-0.05	0.03	-1.38	-0.11	0.11	-1.07	-0.11	0.13	-0.87
DistComp x OSpan	0.26	0.03	0.94	0.03	0.04	0.74	0.04	0.03	1.29	0.12	0.11	1.15	-0.03	0.13	-0.21
TargComp x DistComp x OSpan	-0.05	0.06	-0.82	0.00	0.08	0.02	-0.04	0.07	-0.65	0.24	0.21	1.19	0.02	0.26	0.07
RS															
Intercept	0.00	0.03	0.01	0.00	0.05	0.07	0.00	0.04	0.08	**-0.76**	**0.11**	**-6.87**	**-1.83**	**0.18**	**-10.39**
TargComp	-0.01	0.03	-0.48	**0.08**	**0.04**	**2.03**	-0.07	0.03	-1.95	0.11	0.11	1.03	**0.30**	**0.13**	**2.29**
DistComp	0.00	0.03	-0.10	0.00	0.04	0.07	-0.01	0.03	-0.33	0.10	0.11	0.94	0.09	0.13	0.68
RS	0.01	0.01	1.05	0.02	0.02	1.17	**0.04**	**0.02**	**2.26**	**0.46**	**0.10**	**4.74**	0.07	0.11	0.64
TargComp x DistComp	0.06	0.06	1.07	0.05	0.08	0.67	-0.01	0.07	-0.15	-0.39	0.21	-1.81	0.02	0.26	0.07
TargComp x RS	0.05	0.03	1.73	-0.03	0.04	-0.73	0.00	0.03	-0.08	-0.16	0.11	-1.51	-0.13	0.14	-0.94
DistComp x RS	0.04	0.03	1.32	0.01	0.04	-0.19	-0.03	0.04	-0.88	-0.07	0.11	-0.63	0.13	0.14	0.96
TargComp x DistComp x RS	-0.04	0.06	-0.82	-0.13	0.08	-1.51	-0.13	0.07	-1.87	0.20	0.22	0.91	-0.10	0.27	-0.37

Note: TargComp = Target Complexity; DistComp = Distractor Complexity; OSpan = Operation span; RS = Reading speed; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; *β*=Estimate coefficients; SE=standard error. Bold indicates coefficients that are significant at the $p<0.05$ level.

Gaze duration. No significant effects were found in either three-way interaction model for gaze durations (*t*s <1.11, *p*s>0.27; *t*s <1.73, *p*s>0.08).

Regression-path duration. Results of both three-way interaction models revealed only a significant main effect of Target Complexity (t=2.01, p=0.04; t=2.03, p=0.04), such that the Complex Target conditions elicited longer RTs at the main verb compared to the Simple Target conditions ($M_{CompTarg}$=534ms, $M_{SimTarg}$=477ms) (Figure 5.1).

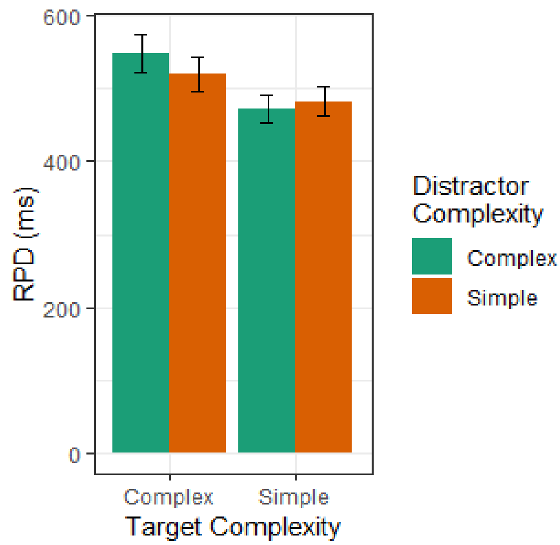

Figure 5.1. Experiment 3: Target Complexity main effect (RPD) in the main verb region.

Total viewing time and Regression-in. The only significant effect found in both total viewing time and regression-in measures was reading speed (t=2.26, p=0.02; *z*=4.74, *p*<0.0001). Slower L1 readers in comparison with faster readers showed longer total viewing time and made more backward regressions into the main verb region from the post-verb region.

Regression-out. Results of both three-way interaction models for regression-out revealed a main effect of Target Complexity (*z*=2.26, *p*=0.02; *z*=2.29, *p*=0.02); the probability of regression made back into the previous regions from the main verb region was higher in the

Complex Target conditions compared to the Simple Target conditions ($M_{CompTarg}$=19%, $M_{SimTarg}$=16%) (Figure 5.2).

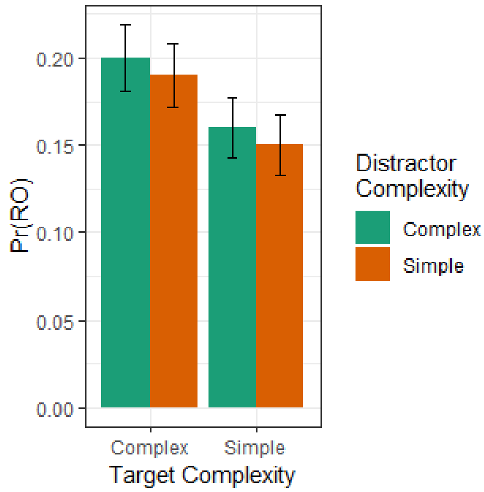

Figure 5.2. Experiment 3: Target Complexity main effect (RO) in the main verb region.

Post-verb region. The post-verb spillover region contained either an adverbial phrase or the patient of the main verb. Log-transformed, length-adjusted fixation duration and probability measures were modeled as a function of Target Complexity (Complex, Simple), Distractor Complexity (Complex, Simple), Individual Difference measure (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. Again, because the post-verb region was the last region of the sentence, regression-in data were not available, and thus no analysis was conducted. Summaries of linear mixed-effects models for the post-verb region are described in Table 5.4.

Gaze duration and total viewing time. No significant effects were found in either three-way interaction model for gaze durations and total viewing time (*t*s <1.55, *p*s>0.12; *t*s <1.58, *p*s>0.11).

Table 5.4. Experiment 3: Mixed-effects modeling results for all dependent measures at the post-verb region.

	Fixation duration measure														
	GD			*RPD*			*TVT*			*RI*			*RO*		
Predictor	β	SE	*t*	β	SE	*t*	β	SE	*t*	β	SE	*z*	β	SE	*z*
OSpan															
Intercept	0.00	0.03	0.01	0.00	0.05	0.01	0.00	0.04	-0.05	-	-	-	**2.27**	**0.21**	**11.06**
TargComp	0.04	0.03	1.19	-0.10	0.06	-1.69	-0.03	0.03	-0.79	-	-	-	-0.23	0.15	-1.53
DistComp	0.00	0.03	-1.30	**0.13**	**0.06**	**2.30**	0.01	0.03	0.27	-	-	-	0.20	0.15	1.36
OSpan	0.01	0.02	0.63	0.01	0.03	0.36	0.01	0.02	0.39	-	-	-	0.01	0.20	0.03
TargComp x DistComp	0.04	0.07	0.65	-0.16	0.11	-1.40	-0.04	0.07	-0.64	-	-	-	-0.60	0.30	-1.98
TargComp x OSpan	0.02	0.03	0.66	0.02	0.06	0.36	-0.01	0.04	-0.22	-	-	-	0.05	0.16	0.34
DistComp x OSpan	-0.05	0.03	-1.55	-0.08	0.06	-1.48	-0.02	0.03	-0.56	-	-	-	-0.01	0.16	-0.04
TargComp x DistComp x OSpan	-0.06	0.07	-0.87	-0.06	0.11	-0.50	-0.09	0.07	-1.28	-	-	-	-0.06	0.31	-0.19
RS															
Intercept	0.00	0.03	0.01	0.00	0.05	-0.01	0.00	0.04	-0.06	-	-	-	**2.32**	**0.19**	**12.27**
TargComp	0.04	0.03	1.19	-0.09	0.06	-1.68	-0.03	0.03	-0.81	-	-	-	-0.24	0.17	-1.39
DistComp	-0.01	0.03	-0.26	**0.13**	**0.06**	**2.30**	0.01	0.03	0.29	-	-	-	0.08	0.17	0.45
RS	0.02	0.02	1.09	0.05	0.03	1.63	0.03	0.02	1.58	-	-	-	**0.79**	**0.19**	**4.23**
TargComp x DistComp	0.04	0.07	0.64	-0.16	0.11	-1.42	-0.05	0.07	-0.66	-	-	-	-0.59	0.31	-1.91
TargComp x RS	-0.02	0.03	-0.57	0.05	0.06	0.90	-0.01	0.04	-0.19	-	-	-	-0.04	0.17	-0.21
DistComp x RS	0.04	0.03	1.20	-0.08	0.06	-1.36	0.04	0.04	1.19	-	-	-	-0.34	0.17	-1.99
TargComp x DistComp x RS	0.03	0.07	0.38	-0.02	0.11	-0.17	0.04	0.07	0.49	-	-	-	-	-	-

Note: TargComp = Target Complexity; DistComp = Distractor Complexity; OSpan = Operation span; RS = Reading speed; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; β=Estimate coefficients; SE=standard error. Bold indicates coefficients that are significant at the $p<0.05$ level.

Regression-path duration. Both three-way interaction models showed a significant main effect of Distractor Complexity (t=2.30, p=0.02; t=2.30, p=0.02), with the Complex Distractor conditions eliciting longer regression-path durations compared to the Simple Distractor conditions ($M_{CompDist}$=2486ms, $M_{SimDist}$=2246ms) (Figure 5.3).

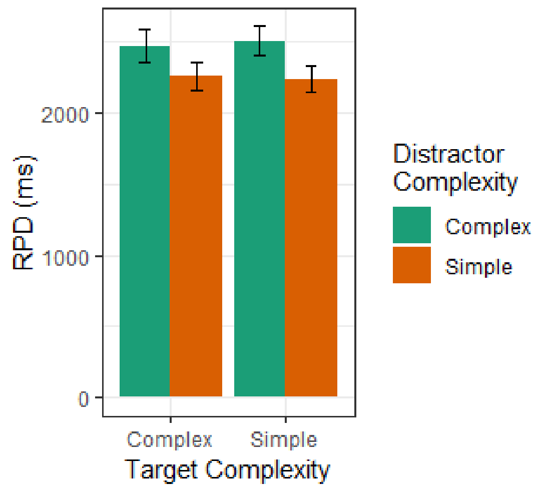

Figure 5.3. Experiment 3: Distractor Complexity main effect (RPD) in the post-verb region.

Regression-out. The three-way interaction model including reading speed revealed a reading speed effect on regression-out (z=4.22, p<0.0001). Similar to the pattern observed previously, slower L1 readers tended to make more regressions back into the main verb from the post-verb spillover region.

Question-response latency and accuracy. Mixed effects models for question-response latency and accuracy contained fixed effect parameters for Target Complexity (Complex, Simple), Distractor Complexity (Complex, Simple), individual difference measures (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. Log-transformed question-response latency data were analyzed using the linear mixed-effects models, and the binary accuracy data were analyzed using logit mixed models.

Descriptive statistics and summaries of linear mixed-effects models for question-response latency and accuracy are reported in Table 5.5 and 5.6 respectively.

Table 5.5. Experiment 3: Mean performance for comprehension question response latency (in ms) and accuracy.

TargComp	*DistComp*	*Response Latency*	*Response Accuracy*
Comp	Comp	2088 (992)	0.77 (0.42)
	Sim	1992 (890)	0.83 (0.38)
Sim	Comp	1780 (857)	0.82 (0.38)
	Sim	1705 (771)	0.84 (0.37)

Note: Note: TargComp = Target Complexity; DistComp = Distractor Complexity; Comp = Complex; Sim = Simple; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out. Standard deviations in parentheses.

Table 5.6. Experiment 3: Mixed-effects modeling results for comprehension question response latency and accuracy.

	Response Latency				*Response Accuracy*			
Predictor	*β*	*SE*	*t*	*p* (>|t|)	*β*	*SE*	*z*	*p* (>|t|)
OSpan								
Intercept	**10.75**	**0.05**	**220.56**	**< 0.0001**	**2.03**	**0.20**	**9.99**	**< 0.0001**
TargComp	**0.23**	**0.05**	**4.91**	**< 0.0001**	-0.30	0.16	-1.92	0.55
DistComp	0.05	0.05	0.99	0.32	**-0.29**	**0.06**	**-5.10**	**< 0.0001**
OSpan	0.01	0.04	0.20	0.85	0.21	0.15	1.38	0.16
TargComp x DistComp	0.01	0.04	0.27	0.79	**-0.29**	**0.11**	**-2.57**	**0.01**
TargComp x OSpan	0.02	0.03	0.64	0.53	0.30	0.15	1.98	0.04
DistComp x OSpan	-0.02	0.03	-0.70	0.49	**-0.15**	**0.05**	**-2.74**	**< 0.01**
TargComp x DistComp x OSpan	0.01	0.05	0.20	0.84	0.18	0.11	1.65	0.09
RS								
Intercept	**10.75**	**0.05**	**228.20**	**< 0.0001**	**2.04**	**0.21**	**9.93**	**< 0.0001**
TargComp	**0.23**	**0.05**	**4.94**	**< 0.0001**	-0.28	0.16	-1.72	0.08
DistComp	0.04	0.05	0.97	0.33	**-0.30**	**0.06**	**-5.34**	**< 0.0001**
RS	**0.09**	**0.03**	**2.58**	**0.01**	0.18	0.16	1.12	0.26
TargComp x DistComp	0.01	0.07	0.20	0.84	**-0.30**	**0.11**	**-2.64**	**<0.01**
TargComp x RS	0.04	0.03	1.45	0.15	-0.14	0.16	-0.88	0.37
DistComp x RS	0.03	0.03	0.97	0.34	**-0.31**	**0.06**	**-5.53**	**< 0.0001**
TargComp x DistComp x RS	-0.07	0.05	-1.49	0.14	-0.20	0.11	-1.78	0.07

Note: TargComp = Target Complexity; DistComp = Distractor Complexity; OSpan = Operation span; RS = Reading speed; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; *β*=Estimate coefficients; SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

Question-response latency. Both three-way interaction models including operation span scores and reading speed respectively revealed a significant main effect of Target Complexity

(t=4.91, p<0.0001; t=4.94, p<0.0001), with the Complex Target conditions eliciting longer response latencies relative to the Simple Target conditions ($M_{CompTarg}$=2040ms, $M_{SimTarg}$=1743ms) (Figure 5.4). Also, there was an effect of reading speed on question-response latency (t=2.58, p=0.01), such that slower L1 readers took more time to respond to comprehension questions compared to faster readers.

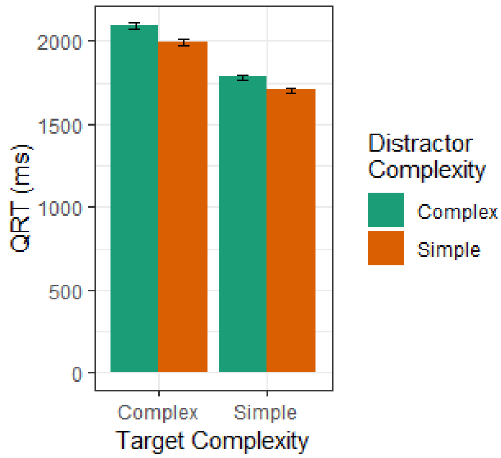

Figure 5.4. Experiment 3: Target Complexity main effect (QRT).

Question-response accuracy. Summaries of both three-way interaction models showed a main effect of Distractor Complexity (z=-5.10, p<0.0001; z=-5.34, p<0.0001) and a reliable interaction between Target Complexity and Distractor Complexity (z=-2.57, p=0.01; z=-2.64, p<0.01). The Complex Distractor conditions resulted in lower question-response accuracies relative to the Simple Distractor conditions ($M_{CompDist}$=80%, $M_{SimDist}$=83%), and the distractor complexity effects were greater in the Complex Target conditions compared to the Simple Target conditions (Figure 5.5 and Table 5.7). There was also a two-way interaction between Distractor Complexity and operation span scores (z=-2.74, p<0.01), such that overall, higher span L1 readers showed an advantage in comparison with those with lower span readers in responding to comprehension questions for the Simple Distractor conditions (Figure 5.6). Also, there was a reliable two-way interaction between Distractor Complexity and reading speed (z=-5.53, *p*<0.0001). Interestingly and unexpectedly, distractor semantic complexity effects were greater for the slower readers than the faster readers (Figure 5.7). In particular, slower readers seem to show a greater advantage in responding to the comprehension questions for the Simple Distractor conditions.

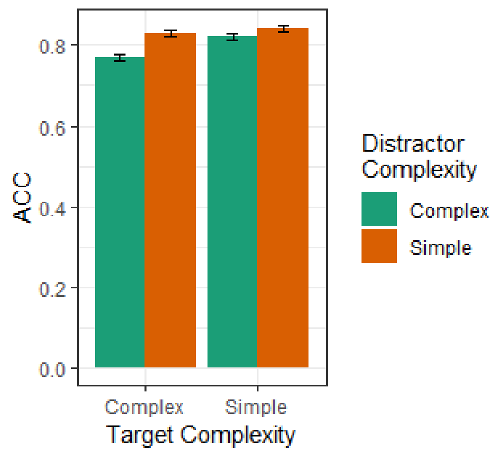

Figure 5.5. Experiment 3: Distractor Complexity main effect and two-way interaction between Target Complexity and Distractor Complexity (ACC).

Table 5.7. Experiment 3: Pairwise differences of contrast for a two-way interaction between Target Complexity and Distractor Complexity (ACC).

Contrast	*β*	*SE*	*z*	*p* (>|z|)
CompTarg,CompDist – SimTarg,CompDist	-0.43	0.17	-2.50	0.06
CompTarg,CompDist – CompTarg,SimDist	**-0.42**	**0.08**	**-5.48**	**<0.0001**
CompTarg,CompDist – SimTarg,SimDist	**-0.57**	**0.17**	**-3.29**	**0.01**
SimTarg,CompDist – CompTarg,SimDist	0.01	0.17	0.07	1.00
SimTarg,CompDist – SimTarg,SimDist	-0.14	0.08	-1.72	0.31
CompTarg,SimDist – SimTarg,SimDist	-0.15	0.18	-0.86	0.82

Note: CompTarg = Complex Target condition; CompDist = Complex Distractor condition; SimTarg = Simple Target condition; SimDist = Simple Distractor; *β*=Estimate coefficients, SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

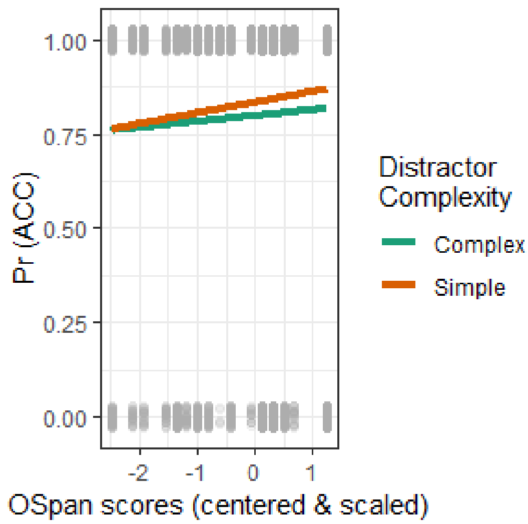

Figure 5.6. Experiment 3: Two-way interaction between Distractor Complexity and operation span scores (ACC).

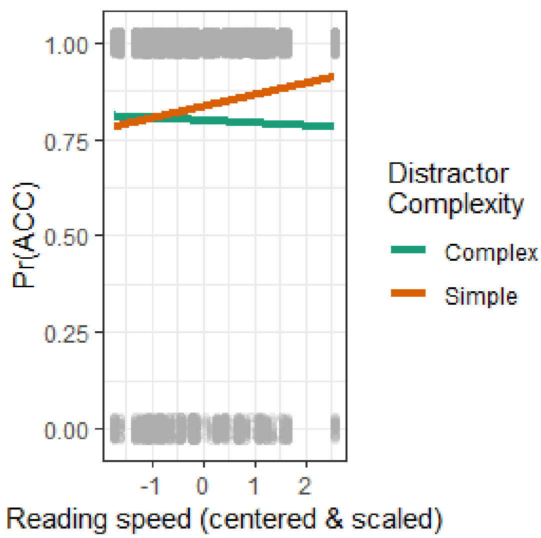

Figure 5.7. Experiment 3: Two-way interaction between Distractor Complexity and reading speed (ACC).

Discussion

The initial prediction was that semantic elaboration for the target or the distractor would facilitate target retrieval efficiency because providing additional semantic information for the target or the distractor would enhance their quality of lexical representations and consequently increase the uniqueness of the target representation, leading to efficient retrieval. Hofmeister and Vasishth (2014) provided reading time evidence of a facilitative effect of semantic elaboration on target retrieval (reduced RTs at the retrieval site when the target was semantically elaborated). However, the results of the current experiment show a contrasting pattern, such that adding additional semantic information for the target and/or the distractor led to inflated fixation durations and lower question-response accuracies.

The discrepancy between the current study and Hofmeister and Vasishth could be due to the fact that different sentence structures were tested in the two studies. In Hofmeister and Vasishth, the critical sentences contained a transitive matrix clause, in which the object noun phrase (target) was modified by an object-extracted relative clause, and the number of modifying words for the matrix subject NP (distractor) and the matrix object NP (target) was manipulated (e.g., The ***congressman / conservative U.S. congressman*** interrogated the ***general / victorious four-star general*** who a lawyer for the White House ***advised*** to not comment on the prisoners). Important to note is that the structure of their sentences was more complex compared to those used in the current study. Also, whereas their sentences included both proactive and retroactive interference (proactive interference from the matrix subject NP with semantic elaboration, and retroactive interference from the RC subject NP without semantic elaboration), the sentences in the current experiment only included retroactive interference, providing more accurate and interpretable data to examine semantic elaboration effects on retrieval efficiency.

Although the observed patterns of semantic elaboration effects on retrieval efficiency were opposite to what has been found in previous studies (Hofmeister, 2011; Hofmeister & Vasishth, 2014), the fact that the quality of lexical representations of the target and the distractor, which were manipulated by semantic elaboration, negatively influenced target retrieval latency and probability (i.e., longer RTs and lower question-response accuracy for the complex targets and distractors) suggests that the cue-based retrieval model needs be further developed to incorporate factors that are not cued by the retrieval trigger.

Three-way interactions between semantic elaboration effects and individual difference measures (operation span scores and reading speed) was explored, and the finding that no reliable interaction was observed in any online reading measures suggests that skilled L1 readers' working memory capacity and reading speed may not be associated with the ability to use additional semantic information for efficient retrieval latency.

Different from the online reading data, the offline question-response accuracy data revealed a reliable two-way interaction between Target and Distractor Complexity. This interaction, as demonstrated in the results section, was driven by the Complex Target, Complex Distractor condition. The lowest accuracy for the Complex Target, Complex Distractor condition could be from adding too much semantic information, which may result in spurious activations of irrelevant linguistic information at retrieval, leading to lower comprehension accuracy. Also, both reading speed and operation span scores moderated distractor semantic elaboration effects in question-response accuracy. Interestingly, high span and *slower* L1 readers showed a greater advantage in responding to comprehension questions for the Simple Distractor conditions, relative to lower span and faster readers. The finding that slower readers showed higher question-response accuracy than faster readers is opposite to the initial prediction that faster readers with greater language experience would be more efficient in processing semantic information and resolving semantic interference. A possibility for this unexpected pattern is that slower readers could have spent more time at the target and/or the distractor during reading, and that these longer encoding times for the target and/or the distractor helped them to retrieve the correct target more accurately after reading.

Chapter 6: Quality of Lexical Representations in L2 Reading Comprehension

Experiment 4 was designed to examine semantic elaboration effects on retrieval efficiency in less-skilled L2 readers. If increasing the quality of lexical representations of the target or the distractor through semantic elaboration enhances L2 readers' retrieval latency and probability, this finding will be informative to design L2 classroom reading instructions and L2 reading assessment, and also help further develop the cue-based retrieval theory that does not make predictions about the features that are not cued by the retrieval trigger. The experiment design, materials, measures, and analysis were identical to those of Experiment 3, except that the L2 participants additionally completed a cloze task and a language background survey.

Assuming that less-skilled L2 readers successfully encode the target and the distractor with additional semantic information, and that they use this elaborated semantic information to distinguish the target representation from that of the distractor, reduced RTs at the retrieval site and higher question-response accuracies were predicted for the complex semantic elaboration conditions compared to the simple semantic elaboration conditions, because as evidenced in previous L1 research, increasing the quality of lexical representations of the target or the distractor through semantic elaboration may increase uniqueness of the target representation, leading to efficient target retrieval. On the other hand, if less-skilled L2 readers fail to successfully encode the target and/or the distractor with additional modifiers, due to their limited cognitive resources or poor quality of L2 lexical representations, or if L2 readers adopt a reading strategy that does not rely on subordinate information like adjunctive phrases to reserve resources for other required computations during sentence processing (cf. Frazier & Clifton, 1996), the complex semantic elaboration conditions were expected to elicit inflated RTs at the retrieval site and lower question-response accuracies relative to the simple semantic elaboration conditions, or no semantic elaboration main effects for RTs at the retrieval site and question-response accuracies. In addition, if target semantic elaboration effects are influenced by the quality of lexical representation of the distractor that comes after the target, there should be interactions between Target Complexity and Distractor Complexity. Specifically, the target complexity was predicted to reduce RTs at the retrieval site only when the distractor was simple, given the law of diminishing returns (Hofmeister & Vasishth, 2014). Thus, the Complex Target-Complex Distractor condition may not facilitate the target retrieval and elicit longer RTs at the retrieval sites compared to those of the Complex Target-Simple Distractor and the Simple Target-Complex Distractor conditions.

Similar to the previous experiments, Experiment 4 also explored whether L2 readers' individual difference measures moderate semantic elaboration effects on target retrieval in long-distance dependencies. Assuming that working memory capacity and reading speed are associated with the ability to use additional semantic information for the target or the distractor and efficiently retrieve a correct target during reading, the higher span, faster L2 readers were predicted to show shorter RTs at the retrieval site and higher question-response accuracies for the complex semantic elaboration effects compared to the lower span, slower L2 readers.

Method

Participants. Forty-two students from University of Illinois at Urbana-Champaign took part in Experiment 4 in exchange for course credit or $7.00 compensation. All participants were non-native speakers of English, and they had used English for about 9 years on average. The recruited participants were highly proficient L2 English speakers, given that their mean scores of TOEFL and the cloze task were 103 (out of 120) and 29 (out of 40), respectively. All participants had normal vision and hearing, with no history of neurological impairment.

Apparatus, Materials, Procedures, Analysis. Apparatus, materials, and procedures were identical to those in Experiment 3, except that the L2 participants additionally completed a cloze task and a language background survey after eye-tracking reading and an operation span task.

Results

Eye-tracking. Descriptive statistics of eye-movements for the main verb and post-verb regions are reported in Table 6.1.

Table 6.1. Experiment 4. Mean performance for eye movement data (raw reading times in ms) in the main verb and post-verb spillover regions.

TargComp	*DistComp*	*GD*	*RPD*	*TVT*	*RI*	*RO*
Main verb region						
Comp	Comp	491 (257)	726 (522)	1056 (628)	0.43 (0.50)	0.20 (0.40)
	Sim	491 (247)	707 (523)	1026 (642)	0.46 (0.50)	0.20 (0.40)
Sim	Comp	479 (237)	653 (475)	1088 (617)	0.40 (0.49)	0.16 (0.37)
	Sim	472 (241)	630 (435)	1037 (598)	0.42 (0.50)	0.16 (0.37)
Post-verb region						
Comp	Comp	479 (260)	4690 (4005)	814 (485)	-	0.93 (0.25)
	Sim	478 (249)	4346 (3078)	820 (486)	-	0.92 (0.28)
Sim	Comp	483 (275)	4969 (3439)	901 (558)	-	0.95 (0.22)
	Sim	463 (276)	3662 (2795)	766 (434)	-	0.94 (0.24)

Note: TargComp = Target Complexity; DistComp = Distractor Complexity; Comp = Complex; Sim = Simple; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out. Standard deviations in parentheses.

Main verb region. The main verb region was the retrieval site where the target retrieval was triggered, and retrieval cues were generated to establish a subject-verb, long-distance dependency. Log-transformed, length-adjusted fixation duration and fixation probability measures were modeled as a function of fixed effects of Target Complexity (Complex, Simple), Distractor Complexity (Complex, Simple), individual difference measure (Ospan scores or reading speed), and their interaction, with random effects of Participants and Items. Summaries of linear mixed-effects models for the main region are described in Table 6.2.

Gaze duration. No significant effects were found in either three-way interaction model for gaze durations (*t*s <1.55, *p*s>0.12; *t*s <1.31, *p*s>0.19).

Regression-path duration. Summaries of both three-way interaction models revealed a significant main effect of Target Complexity (t=2.71, p=0.01; t=2.71, p=0.01), with the Complex Target conditions eliciting longer regression-path durations compared to the Simple Target conditions ($M_{CompTarg}$=713ms, $M_{SimTarg}$=629ms) (Figure 6.1). Also, both reading speed and Ospan scores moderated target semantic elaboration effects (t=-2.07, p=0.04; t=-2.07, p=0.04). Target semantic elaboration effects were greater for faster L2 readers compared to the slower L2 readers, and this interaction was driven by faster L2 readers' reduced regression-path duration for the Simple Target conditions (Figure 6.2). On the other hand, L2 readers with smaller working memory showed greater semantic elaboration effects than those with larger spans, and this was due to lower span L2 readers' increased regression-path duration for the Complex Target conditions (Figure 6.3).

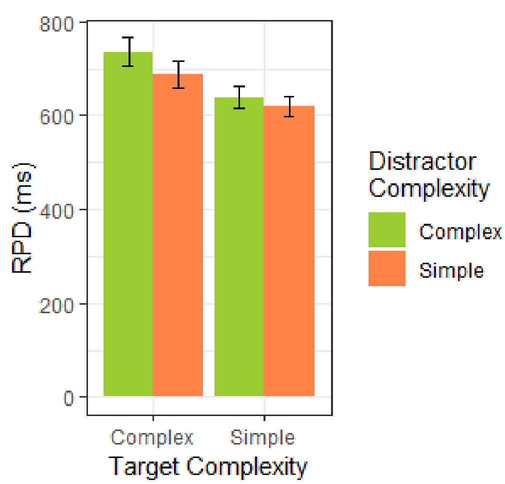

Figure 6.1. Experiment 4: Target Complexity main effect (RPD) in the main verb region.

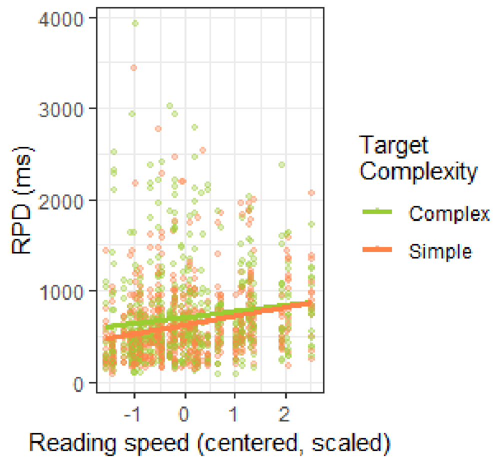

Figure 6.2. Experiment 4: Two-way interaction between Target Complexity and reading speed (RPD) in the main verb region.

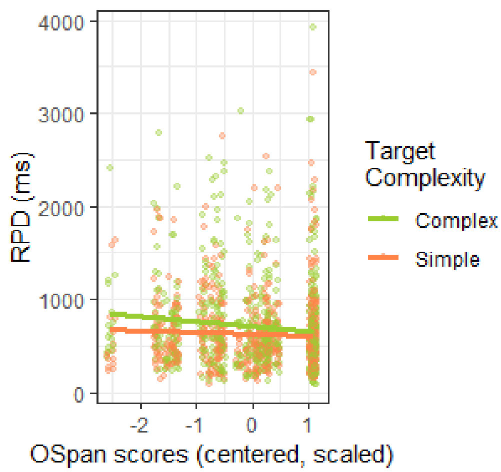

Figure 6.3. Experiment 4: Two-way interaction between Target Complexity and operation span scores (RPD) in the main verb region.

Total viewing time. Both three-way interaction models revealed a significant effect of Distractor Complexity (t=2.56, p=0.01, t=2.57, p=0.01), with the Complex Distractor conditions showing longer total viewing times at the main verb region compared to the Simple Distractor conditions ($M_{CompDist}$=1081ms, $M_{SimDist}$=1023ms) (Figure 6.4). The model including reading speed additionally showed an effect of reading speed on total viewing time (t=2.07, p=0.04), such that slower L2 readers showed longer total viewing times than faster L2 readers.

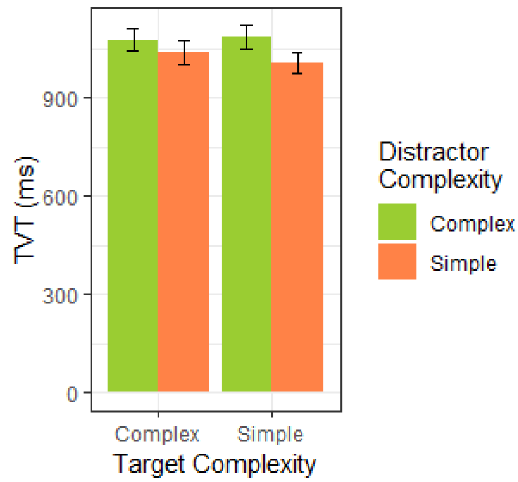

Figure 6.4. Experiment 4: Distractor Complexity main effect (TVT) in the main verb region.

Regression-in. The only effect found in regression-in was an effect of reading speed (z=4.74, p<0.0001). Slower L2 readers made more regressions into the main verb region from the post-verb region compared to faster L2 readers.

Regression-out. Summaries of both three-way interaction models revealed a main effect of Target Complexity (z=2.54, p=0.01; z=2.42, p=0.01); The probability of regression into the previous regions from the main verb region was higher in the Complex Target conditions relative to the Simple Target conditions ($M_{CompTarg}$=20%, $M_{SimTarg}$=15%) (Figure 6.5).

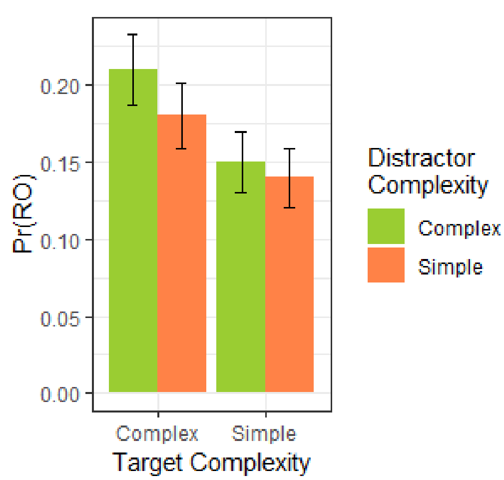

Figure 6.5. Experiment 4: Target Complexity main effect (RO) in the main verb region.

Table 6.2. Experiment 4: Mixed-effects modeling results for all dependent measures at the main verb region.

	Fixation duration measure														
	GD			*RPD*			*TVT*			*RI*			*RO*		
Predictor	β	SE	*t*	β	SE	*t*	β	SE	*t*	β	SE	*z*	β	SE	*z*
OSpan															
Intercept	0.00	0.03	0.04	0.00	0.04	0.04	0.00	0.04	0.09	**-0.36**	**0.14**	**-2.60**	**-1.77**	**0.15**	**-11.47**
TargComp	0.01	0.04	0.21	**0.12**	**0.04**	**2.71**	-0.01	0.04	-0.17	0.16	0.12	1.39	**0.38**	**0.15**	**2.54**
DistComp	0.01	0.04	0.15	0.04	0.04	1.50	**0.11**	**0.04**	**2.56**	-0.03	0.12	-0.27	0.09	0.15	0.61
OSpan	-0.01	0.02	-0.43	-0.02	0.02	-0.78	0.00	0.02	-0.03	0.16	0.11	1.45	-0.02	0.10	-0.24
TargComp x DistComp	-0.03	0.07	-0.42	0.05	0.09	0.54	-0.01	0.08	-0.10	-0.24	0.24	-0.99	0.12	0.40	0.39
TargComp x OSpan	-0.01	0.04	-0.30	**-0.09**	**0.04**	**-2.07**	0.00	0.04	-0.07	-0.06	0.12	-0.48	-0.08	0.15	-0.51
DistComp x OSpan	-0.06	0.04	-1.55	0.04	0.04	0.83	0.02	0.04	0.54	0.20	0.12	1.67	0.23	0.15	1.48
TargComp x DistComp x OSpan	0.07	0.07	0.98	0.11	0.09	1.29	0.10	0.08	1.13	0.10	0.24	0.41	-0.15	0.30	-0.50
RS															
Intercept	0.00	0.03	0.05	0.00	0.04	0.04	0.00	0.04	0.08	**-0.36**	**0.14**	**-2.73**	**-1.77**	**0.15**	**-11.49**
TargComp	0.01	0.04	0.18	**0.12**	**0.04**	**2.71**	-0.01	0.04	-0.18	0.17	0.12	1.46	**0.39**	**0.16**	**2.42**
DistComp	0.01	0.04	0.16	0.06	0.04	1.50	**0.11**	**0.04**	**2.57**	-0.03	0.12	-0.24	0.09	0.15	0.61
RS	0.02	0.02	1.31	-0.02	0.02	-0.78	**0.04**	**0.02**	**2.07**	**0.28**	**0.10**	**2.76**	0.05	0.10	0.49
TargComp x DistComp	-0.03	0.07	-0.43	0.05	0.09	0.55	-0.01	0.08	-0.07	-0.23	0.24	-0.98	0.11	0.30	0.36
TargComp x RS	-0.04	0.04	-1.12	**-0.09**	**0.04**	**-2.07**	0.03	0.04	0.84	-0.23	0.12	-1.92	-0.03	0.15	-0.20
DistComp x RS	-0.01	0.04	-0.23	0.04	0.04	0.83	-0.04	0.04	-1.09	-0.04	0.12	-0.36	-0.01	0.15	-0.05
TargComp x DistComp x RS	-0.05	0.07	-0.68	0.11	0.09	1.29	-0.11	0.08	-1.39	0.02	0.24	0.07	0.20	0.30	0.67

Note: TargComp = Target Complexity; DistComp = Distractor Complexity; OSpan = Operation span; RS = Reading speed; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; β=Estimate coefficients; SE=standard error. Bold indicates coefficients that are significant at the $p<0.05$ level.

Post-verb region. The post-verb spillover region contained either an adverbial phrase or the patient of the main verb. Log-transformed, length-adjusted fixation duration and probability measures were modeled as a function of Target Complexity (Complex, Simple), Distractor Complexity (Complex, Simple), Individual Difference measure (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. The post-verb region was the last region of the sentence, thus no analysis was conducted for regression-in. Summaries of linear mixed-effects models for the post-verb region are described in Table 6.4.

Gaze duration. No significant effects were found in either three-way interaction model (ts <1.01, ps>0.31).

Regression-path duration. Both three-way interaction models showed a significant main effect of Distractor Complexity ($t=4.81$, $p<0.001$; $t=4.82$, $p<0.001$), with the Complex Distractor conditions eliciting longer regression-path durations compared to the Simple Distractor conditions ($M_{CompDist}$=5083ms, $M_{SimDist}$=4202ms) (Figure 6.6). Also, there was a reliable two-way interaction between Target Complexity and Distractor Complexity, such that target semantic elaboration effects were greater in the Simple Distractor conditions than in the Complex Distractor conditions (Figure 6.6) (Table 6.3). This pattern indicates that the distractor's quality of lexical representation did interfere the retrieval of the target during reading. Important to note is that the distractor came after the target in the sentence (retroactive interference), whereas in Hofmeister and Vasishth (2014), which demonstrated weaker effects of semantic elaboration for the distractor on target retrieval, the distractor came before the target (proactive interference). The contrasting results between their study and the current experiment suggests a further examination of the influence of the distractor's position in the sentence (proactive vs. retroactive) on target retrieval, because as the cue-based retrieval model suggests, chunks in memory during reading decay as a function of time and prior retrievals, and thus the extent to which the distractor interferes the target retrieval may vary depending on the position of the distractor.

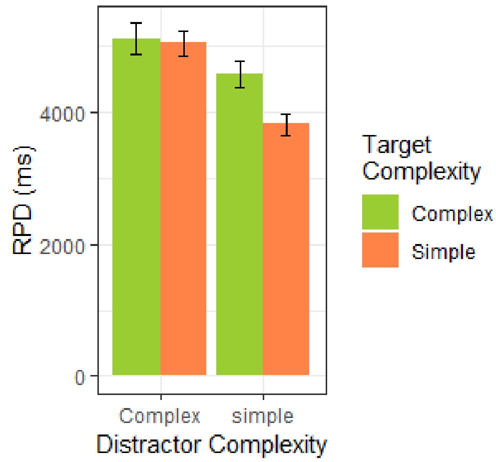

Figure 6.6. Experiment 4: Distractor Complexity main effect and two-way interaction between Target Complexity and Distractor Complexity (RPD) in the post-verb region.

Table 6.3. Experiment 4: Pairwise differences of contrast for a two-way interaction between Target Complexity and Distractor Complexity (RPD).

Contrast	*β*	*SE*	*t*	*p* (>|z|)
CompTarg,CompDist – SimTarg,CompDist	-0.08	0.09	-0.87	0.82
CompTarg,CompDist – CompTarg,SimDist	0.14	0.10	1.38	0.52
CompTarg,CompDist – SimTarg,SimDist	**0.35**	**0.11**	**3.14**	**0.02**
SimTarg,CompDist – CompTarg,SimDist	**0.22**	**0.10**	**2.25**	**0.14**
SimTarg,CompDist – SimTarg,SimDist	**0.43**	**0.09**	**4.74**	**0.00**
CompTarg,SimDist – SimTarg,SimDist	**0.21**	**0.09**	**2.24**	**0.12**

Note: CompTarg = Complex Target condition; CompDist = Complex Distractor condition; SimTarg = Simple Target condition; SimDist = Simple Distractor; *β*=Estimate coefficients, SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

Total viewing time. Both three-way interaction models revealed a significant effect of Distractor Complexity (t=2.77, p=0.01, t=2.77, p=0.01), such that the Complex Distractor conditions showed longer total viewing times relative to the Simple Distractor conditions ($M_{CompDist}$=903ms, $M_{SimDist}$=831ms) (Figure 6.7).

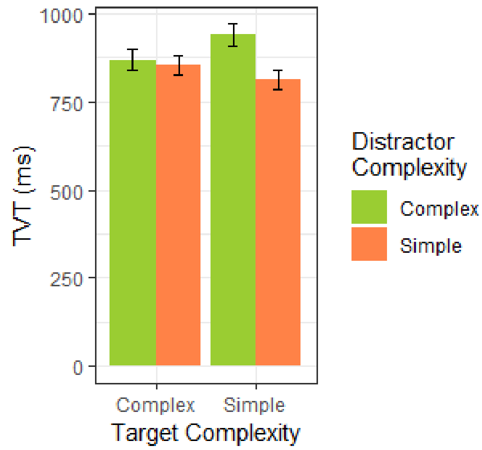

Figure 6.7. Experiment 4: Distractor Complexity main effect (TVT) in the post-verb region.

Regression-out. The three-way interaction model including reading speed revealed a reading speed effect on regression-out (z=2.26, p=0.02). Similar to the pattern observed previously, the probability of regressing into the main verb region from the post-verb spillover region was higher for slower L2 readers compared to faster L2 readers.

Table 6.4. Experiment 4: Mixed-effects modeling results for all dependent measures at the post-verb region.

	Fixation duration measure														
	GD			*RPD*			*TVT*			*RI*			*RO*		
Predictor	*β*	SE	*t*	*β*	SE	*t*	*β*	SE	*t*	*β*	SE	*z*	*β*	SE	*z*
OSpan															
Intercept	0.00	0.04	-0.05	0.00	0.05	0.04	0.00	0.05	0.01	-	-	-	**3.50**	**0.40**	**8.84**
TargComp	0.00	0.04	0.05	0.06	0.06	1.04	0.00	0.04	-0.08	-	-	-	0.47	0.38	1.19
DistComp	0.04	0.04	1.01	**0.28**	**0.06**	**4.81**	0.10	0.04	0.77	-	-	-	-0.38	0.32	-1.18
OSpan	-0.01	0.02	-0.45	0.00	0.03	0.15	0.00	0.02	-0.11	-	-	-	0.02	0.37	0.06
TargComp x DistComp	-0.03	0.07	-0.45	**-0.29**	**0.12**	**-2.48**	-0.12	0.07	-1.62	-	-	-	0.01	0.49	0.01
TargComp x OSpan	-0.03	0.04	-0.93	0.06	0.06	1.09	-0.02	0.04	-0.46	-	-	-	0.55	0.38	1.47
DistComp x OSpan	0.03	0.04	0.91	0.03	0.06	0.54	0.04	0.04	1.07	-	-	-	0.12	0.34	0.34
TargComp x DistComp x OSpan	0.07	0.08	0.85	0.08	0.12	0.67	0.00	0.08	-0.01	-	-	-	-0.79	0.51	-1.55
RS															
Intercept	0.00	0.04	-0.05	0.00	0.05	0.03	0.00	0.05	0.01	-	-	-	**3.44**	**0.40**	**8.72**
TargComp	0.00	0.04	0.07	0.06	0.06	1.04	0.00	0.04	-0.08	-	-	-	0.58	0.52	1.12
DistComp	0.04	0.04	0.99	**0.28**	**0.06**	**4.82**	**0.10**	**0.04**	**2.77**	-	-	-	-0.15	0.44	-0.34
RS	0.02	0.02	0.96	0.04	0.03	1.28	0.03	0.02	1.88	-	-	-	**1.00**	**0.44**	**2.26**
TargComp x DistComp	-0.03	0.07	-0.43	**-0.29**	**0.12**	**-2.46**	-0.12	0.07	-1.59	-	-	-	0.31	0.76	0.40
TargComp x RS	-0.06	0.04	-1.50	0.04	0.06	0.69	-0.01	0.04	-0.34	-	-	-	0.45	0.61	0.74
DistComp x RS	0.03	0.04	0.76	-0.03	0.06	-0.46	0.03	0.04	0.81	-	-	-	0.41	0.52	0.79
TargComp x DistComp x RS	-0.04	0.07	-0.47	-0.09	0.12	-0.73	0.01	0.07	0.07	-	-	-	0.13	0.87	0.15

Note: TargComp = Target Complexity; DistComp = Distractor Complexity; OSpan = Operation span; RS = Reading speed; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; *β*=Estimate coefficients; SE=standard error. Bold indicates coefficients that are significant at the $p<0.05$ level.

Question-response latency and accuracy. Mixed-effects models for question-response latency and accuracy contained fixed effect parameters for Target Complexity (Complex, Simple), Distractor Complexity (Complex, Simple), individual difference measures (operation span scores or reading speed), and their interaction, with random effects of Participants and Items. Log-transformed question-response latency and binary accuracy data were analyzed using the linear mixed-effects models and logit mixed models, respectively. Descriptive statistics and summaries of linear mixed-effects models for question-response latency and accuracy are reported in Table 6.5 and 6.6 respectively.

Table 6.5. Experiment 4: Mean performance for comprehension question response latency (in ms) and accuracy.

TargComp	*DistComp*	*Response Latency*	*Response Accuracy*
Comp	Comp	2537 (1060)	0.72 (0.45)
	Sim	2469 (1172)	0.80 (0.40)
Sim	Comp	2121 (947)	0.73 (0.44)
	Sim	2305 (1163c)	0.71 (0.46)

Note: Note: TargComp = Target Complexity; DistComp = Distractor Complexity; Comp = Complex; Sim = Simple; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out. Standard deviations in parentheses.

T*able 6.6*. Experiment 4: Mixed-effects modeling results for comprehension question response latency and accuracy.

	Response Latency				*Response Accuracy*			
Predictor	*β*	*SE*	*t*	*p* (>|t|)	*β*	*SE*	*z*	*p* (>|t|)
OSpan								
Intercept	**11.05**	**0.06**	**189.94**	**< 0.0001**	**2.01**	**0.29**	**6.94**	**< 0.0001**
TargComp	**0.19**	**0.05**	**3.49**	**< 0.001**	0.04	0.24	0.15	0.88
DistComp	-0.02	0.05	-0.33	0.74	**-0.21**	**0.07**	**-3.19**	**< 0.01**
OSpan	-0.04	0.04	-0.82	0.41	0.44	0.25	1.75	0.80
TargComp x DistComp	0.14	0.10	1.39	0.17	**-0.68**	**0.13**	**-5.21**	**< 0.0001**
TargComp x OSpan	0.04	0.04	1.02	0.31	0.07	0.23	0.30	0.77
DistComp x OSpan	**0.07**	**0.03**	**2.47**	**0.18**	0.03	0.06	0.48	0.63
TargComp x DistComp x OSpan	0.00	0.07	0.07	0.94	**-0.34**	**0.13**	**-2.68**	**< 0.01**
RS								
Intercept	**11.05**	**0.05**	**203.67**	**< 0.0001**	**2.01**	**0.30**	**6.80**	**< 0.0001**
TargComp	**0.19**	**0.05**	**3.49**	**< 0.001**	0.02	0.23	0.10	0.93
DistComp	-0.01	0.05	-0.29	0.77	**-0.22**	**0.07**	**-3.36**	**< 0.001**
RS	**0.14**	**0.04**	**3.74**	**< 0.001**	0.19	0.26	0.75	0.46
TargComp x DistComp	0.14	0.10	1.40	0.17	**-0.66**	**0.13**	**-5.07**	**< 0.0001**

(Table 6.6 continue)

TargComp x RS	-0.02	0.04	-0.43	0.67	0.34	0.23	1.49	0.14
DistComp x RS	-0.03	0.03	-1.12	0.27	**-0.14**	**0.07**	**-2.02**	**0.04**
TargComp x DistComp x RS	-0.01	0.06	-0.13	0.90	-0.01	0.14	-0.06	0.95

Note: TargComp = Target Complexity; DistComp = Distractor Complexity; OSpan = Operation span; RS = Reading speed; GD = gaze duration; RPD = regression-path duration; TVT = total viewing time; RI = regression-in; RO = regression-out; β=Estimate coefficients; SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

Question-response latency. Both three-way interaction models revealed a significant main effect of Target Complexity (t=3.49, p<0.001; t=3.49, p<0.001), with the Complex Target conditions eliciting longer response latencies relative to the Simple Target conditions ($M_{CompTarg}$=2490ms, $M_{SimTarg}$=2191ms) (Figure 6.8). Also, there was a reliable interaction between Distractor Complexity and Ospan scores (t=2.47, p=0.18), as well as an effect of reading speed on question-response latency (t=3.74, p<0.001). As shown in Figure 6.9, the effects of semantic elaboration for the distractor were greater for L2 readers with small spans, compared to those with larger spans. Also, slower L1 readers took more time to respond to comprehension questions than did faster readers.

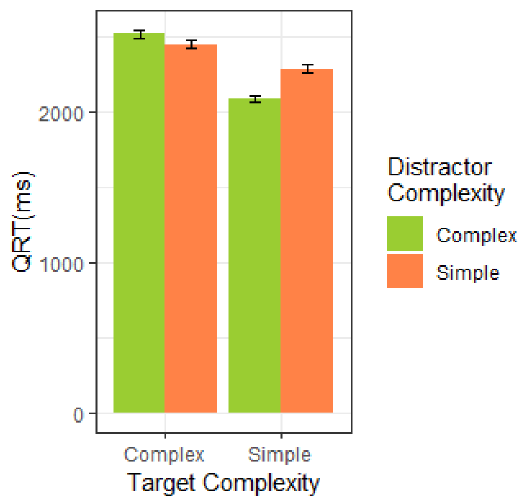

Figure 6.8. Experiment 4: Target Complexity main effect (QRT).

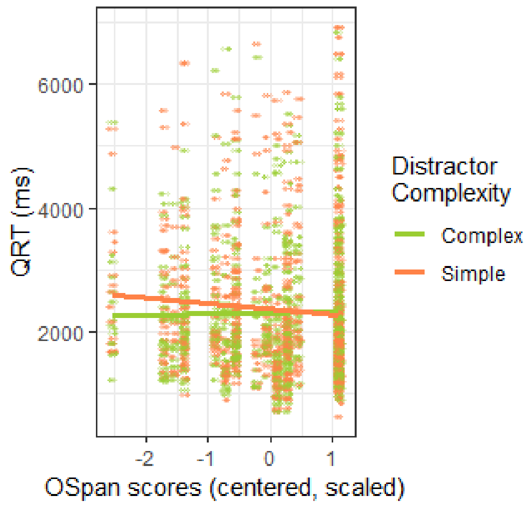

Figure 6.9. Experiment 4: Two-way interaction between Distractor Complexity and operation span scores (QRT).

Question-response accuracy. Summaries of both three-way interaction models showed a main effect of Distractor Complexity (z=-3.19, p<0.01; z=-3.36, p<0.001) and a significant interaction between Target Complexity and Distractor Complexity (z=-5.21, p<0.0001 z=-5.07, p<0.0001). The Complex Distractor conditions resulted in lower question-response accuracies relative to the Simple Distractor conditions ($M_{CompDist}$=76%, $M_{SimDist}$=79%), and the distractor complexity effects were greater in the Complex Target conditions compared to the Simple Target conditions (Figure 6.10 and Table 6.7). There was also a three-way interaction between Target Complexity, Distractor Complexity, and operation span scores (z=-2.68, p<0.01), as well as two-way interaction between Distractor Complexity and reading speed (z=-2.02, p=0.04). As shown in Figure 6.11, when there was a simple distractor, both high and low span L2 readers showed a higher question-response accuracy for the Complex Target conditions than for the Simple Target conditions. On the other hand, when there was a complex distractor, higher L2 span readers were more accurate in responding to questions for the Simple Target conditions than the Complex Target conditions, whereas the pattern was reversed for the lower span L2 readers, such that they were more accurate in responding to questions for the Complex Target conditions. Also, as shown in Figure 6.12, slower L2 readers showed a greater Distractor Complexity effect in comparison to faster L2 readers. Specifically, slower L2 readers were more accurate in responding to questions for the Simple Distractor conditions than for the Complex Distractor conditions.

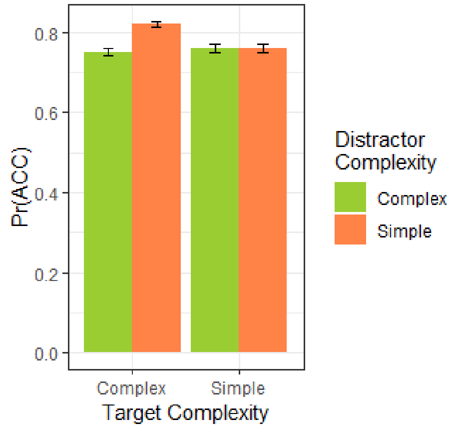

Figure 6.10. Experiment 4: Distractor Complexity main effect and two-way interaction between Target Complexity and Distractor Complexity (ACC).

Table 6.7. Experiment 4: Pairwise differences of contrast for a two-way interaction between Target Complexity and Distractor Complexity (ACC).

Contrast	*β*	*SE*	*z*	*p* (>|z|)
CompTarg,CompDist – SimTarg,CompDist	-0.12	0.09	-1.32	0.55
CompTarg,CompDist – CompTarg,SimDist	**-0.51**	**0.09**	**-5.59**	**<0.0001**
CompTarg,CompDist – SimTarg,SimDist	-0.03	0.09	-0.35	0.99
SimTarg,CompDist – CompTarg,SimDist	**-0.39**	**0.09**	**-4.32**	**0.00**
SimTarg,CompDist – SimTarg,SimDist	0.09	0.09	0.98	0.76
CompTarg,SimDist – SimTarg,SimDist	**0.48**	**0.09**	**5.31**	**<0.00-01**

Note: CompTarg = Complex Target condition; CompDist = Complex Distractor condition; SimTarg = Simple Target condition; SimDist = Simple Distractor; *β*=Estimate coefficients, SE=standard error. Bold indicates coefficients that are significant at the p<0.05 level.

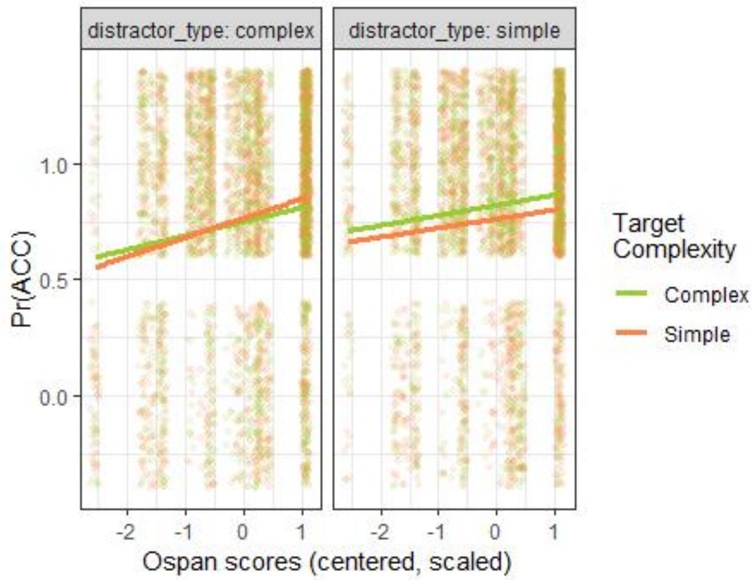

Figure 6.11. Experiment 4: Three-way interaction between Target Complexity, Distractor Complexity and operation span scores (ACC).

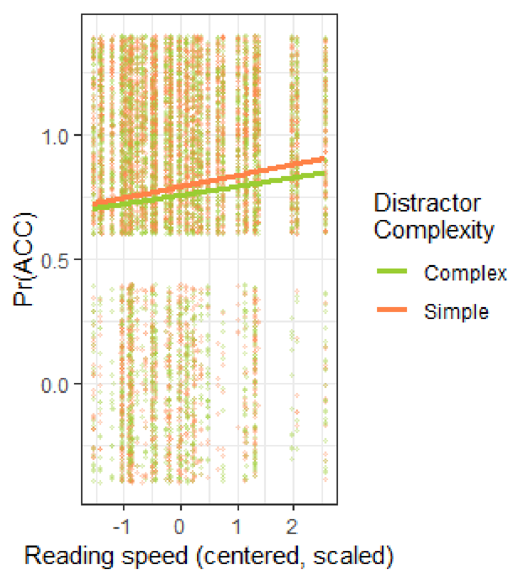

Figure 6.12. Experiment 4: Two-way interaction between Distractor Complexity and reading speed (ACC).

Discussion

Overall, the patterns of the observed semantic elaboration effects on retrieval latency and probability in less-skilled L2 readers were similar to those of the skilled L1 readers in

Experiment 3. Providing additional semantic information for the target and/or the distractor led to increased retrieval latencies at the retrieval sites and lower probabilities of successful target retrieval. These patterns contrast with previous research demonstrating facilitative effects of semantic elaboration on retrieval efficiency (Hofmeister, 2011; Hofmeister & Vasishth, 2014). As mentioned in Experiment 3, one of the assumptions of the cue-based retrieval model is that linguistic chunks in memory decay as a function of time and prior retrievals, and thus it is important to further examine how the position of the distractor in relation to the target (proactive vs. retroactive) influences retrieval efficiency, and this will be informative to resolve the discrepancy between the findings of the current experiments and previous research.

While the patterns of semantic elaboration effects on retrieval latency and probability were similar between skilled L1 and less-skilled L2 readers, some differences did arise between the L1 and L2 data. One difference was that there was a reliable two-way interaction between Target Complexity and Distractor Complexity in regression-path duration in the L2 data, whereas no interaction was observed in the L1 data. This observed interaction indicates that L2 readers' target semantic elaboration effects were modulated by the quality of the lexical representation of the distractor that came after the target. The finding that the interaction was observed only in L2 data suggests that L2 readers were more susceptible to interference from the distractor with semantic elaboration than L1 readers. Given that L2 readers may have a higher proportion of low-quality lexical representations compared to those of L1 readers, it is likely that adding additional semantic elaboration for the distractor could have increased L2 readers' spurious activations of irrelevant linguistic information at target retrieval, leading to increased retrieval latency.

Another interesting difference was that L2 readers' reading speed and working memory capacity moderated the effects of target semantic elaboration on retrieval latency, whereas L1 data revealed no effects of individual difference measures on semantic elaboration. Specifically, the interaction between operation span scores and target semantic elaboration was driven by lower span L2 readers' inflated regression-path durations in the Complex Target conditions. In the case of reading speed, the slower L2 readers were predicted to show greater elaboration effects on retrieval latency compared to faster L2 readers. However, the results showed the opposite pattern: greater semantic elaboration effects were found for the faster L2 readers, and this interaction was driven by faster L2 readers' reduced regression-path durations for the Simple Target conditions. The observed L1-L2 difference that only L2 data showed moderating effects of individual differences on semantic elaboration could be due to the fact that sub-processes of L2 sentence processing that are still developing with respect to their efficiency (e.g., lexical access) may be less automatic compared to those of skilled L1 readers, and thus L2 readers may need to recruit other cognitive resources like working memory capacity to complement their less-efficient or still-developing sentence processing in the L2.

Similar to the L1 data, question-response accuracy data in L2 also revealed a reliable two-way interaction between Target Complexity and Distractor Complexity, but the interaction pattern was different from the L1 data, such that the interaction was driven by the Complex Target, Simple Distractor condition showing the highest accuracy. This pattern is interesting because it demonstrates that providing additional semantic information for the target did lead to a higher probability of successful target retrieval only when the distractor was simple. Important to note is that the L2 retrieval latency data showed slow-downs at the retrieval sites when the target or the distractor was semantically elaborated, but the question-response accuracy data revealed the highest accuracy for the Complex Target, Simple Distractor condition. This suggests that L2 readers' increased reading times at the retrieval sites may reflect their processing of additional semantic information for the target and the distractor, leading to successful target retrieval, rather than reflecting processing difficulty at target retrieval. In addition, as in the L1 data, L2 readers' reading speed and working memory capacity moderated the effects of semantic elaboration on retrieval probability. The patterns of two-way interactions between reading speed and distractor semantic elaboration effects were similar, such that greater semantic elaboration effects for the distractor were found for the slower L2 readers compared to faster L2 readers. What was different between L1 and L2 was that there was a three-way interaction between working memory capacity and the two fixed effects in L2, whereas only L1 readers' working memory capacity moderated distractor elaboration effects. Again, this L1-L2 difference could be due to L2 readers' less efficient lower-level reading processes (e.g. lexical access), possibly increasing their reliance on other cognitive resources.

Chapter 7. General Discussion

In natural language, the meanings of expressions are encoded linearly. However, linguistic relations underlying these sequences are not linear in that they often extend over multiple words, phrases, or clauses, creating non-adjacent dependencies. For example, readers must establish a syntactic dependency between a subject and a verb in order to understand a sentence in English. However, the subject is often separated from the verb by intervening linguistic items, as in *The editor who was recently hired suddenly quit*. Non-adjacent dependencies are commonly found in other grammatical relations, such as clefting (*This was the book that the editor admired*), embedding (*The dog that my neighbor recently adopted was sweet*), gapping (*The estate auctioned off the books, but not the furniture*), or verb ellipsis (*Sam ate an apple*, *but Eddie did not.*), as well as in discourse processing when building a coherent discourse representation (e.g., anaphoric interpretation).

All of these non-adjacent dependencies require readers to access previously processed items to establish dependencies between distant items. In other words, the ability to retrieve previously encountered elements is critical in resolving non-adjacent dependencies and successfully comprehending sentences. A wealth of research has examined the role of memory in relation with this ability. In particular, motivated by the multicomponent WM model (Baddeley, 2000; Baddeley & Hitch, 1974), the capacity-based approach has emphasized limited WM capacity (see Baddeley, 2012, for a recent review) as primary comprehension bottleneck, and numerous studies have demonstrated increased comprehension difficulty due to readers' decreased WM capacity (Fedorenko et al., 2006, 2007; Gordon et al., 2002; Just & Carpenter, 1980, 1992; Just et al., 1996; Just & Varma, 2002, 2007).

However, evidence of extremely limited WM capacity from recent memory and reading research (Cowan, 2001; Dosher, 1979; Lewis et al., 2006; McElree, 2006; Oberauer, 2002; Reed, 1973, 1976; Van Dyke & Johns, 2012; Van Dyke et al., 2014; Wickelgren, 1977) has suggested that because WM capacity is extremely limited for everybody, capacity differences may not explain variability in reading performance. Rather, sentence processing relies on a direct-access, cue-drive retrieval operation, instead of fixed WM capacity (McElree & Vasishth, 2005; McElree et al., 2006; Nicenboim & Vasishth, 2018).

Both of these memory-based accounts are theoretically well motivated and supported by empirical evidence from numerous studies. However, what memory mechanisms support parsing and how they constrain sentence comprehension still remain controversial. Furthermore, the nature of the WM system and the role of WM in sentence comprehension are unclear, so it is challenging to understand how readers access linguistic representations outside focal attention during reading.

Accordingly, this dissertation tested predictions of the cue-based retrieval model in comparison with those of the capacity-based model in order to address memory mechanisms supporting sentence comprehension. Assuming that sentence processing relies on an efficient direct-access, cue-driven retrieval operation, this dissertation further explored the effects of semantic elaboration on retrieval efficiency to investigate whether enhancing the quality of lexical representations through semantic elaboration influences retrieval latency and accuracy. In particular, L2 readers, who are assumed to be on average less skilled readers than L1 readers due to relatively less experience with the language, were included in the current research to better understand whether the ability to employ an efficient cue-driven retrieval operation determines skilled versus less-skilled reading. Lastly, given that reading performance depends on the interaction between the material being read and individual differences in abilities of the reader, the current research examined the relationships between retrieval ability and two individual difference measures, working memory capacity and reading speed.

In the remainder of this chapter, I will address the goals discussed above in light of the findings from the experiments reported here, and demonstrate the implication of these findings for theories of sentence comprehension, individual differences in reading ability, and L2 reading.

Memory mechanisms supporting sentence comprehension

As discussed earlier, although both the capacity-based model and the cue-based model propose that memory mechanisms enable sentence comprehension, how readers access information outside focal attention during reading is a largely unresolved issue. Thus, in order to contribute to the understanding of what memory mechanisms guide parsing, and how they constrain sentence processing, Experiments 1 and 2 specifically tested the predictions of the cue-based retrieval model in comparison with those of the capacity-based model by examining syntactic and semantic retrieval interference during L1 and L2 sentence processing. In these experiments, the semantic and syntactic properties of the intervening noun between the subject and the verb were manipulated so that these features of the intervening noun either matched or mismatched the retrieval cues, creating high and low semantic and syntactic interference conditions.

Both L1 and L2 data revealed semantic and syntactic interference effects on retrieval latency and probability, such that high semantic and syntactic interference conditions elicited longer RTs on the retrieval sites (the main verb and the post-verb spillover regions) and lower comprehension question-response accuracy. These results provide supportive evidence for the cue-based retrieval model because the observed interference arose due to the similarity of the features of the distractor to the retrieval cues (i.e. cue-overload).

Given that the number of words was held constant between the subject and the verb (i.e. memory load was held constant across all four conditions), the observed retrieval interference effects provide robust evidence *against* the capacity-based model, which claims that comprehension difficulty increases as the distance between the head and the dependent increases. However, it is important to note that the number of words intervening in dependencies may not be an ideal metric for quantifying the storage and processing demands of the sentence because each word varies in terms of its syntactic function, length, or frequency, and all these lexical factors greatly influence reading times during online sentence processing.

Other metrics have been used to quantify the storage and processing demands of different types of sentences, for example, the number of embeddings in a sentence (Miller & Chomsky, 1963) or the number of incomplete dependencies at any given point in a sentence (Abney & Johnson, 1991; Gibson, 1998; Kimball, 1973). In fact, the number of embeddings between the head and the dependent could be a possible confound in the current experiments because in the case of the sentences in the high syntactic interference conditions, not only did the syntactic feature overlap between the target and the distractor, creating high syntactic interference, but also sentences included two embedded clauses (a subject-extracted RC and a *that*-clause), creating high processing/memory load. In other words, the observed syntactic interference in the current experiments could have resulted from high syntactic interference as well as high memory load. Nonetheless, the fact that semantic retrieval interference was also observed along with syntactic interference in target retrieval latency and probability suggests that sentence comprehension relies on a cue-driven retrieval operation.

Along with the main effects of semantic and syntactic interference, their interaction was also found in the L1 online data, which suggests that all cues may be multiplicatively combined into a single retrieval probe (Van Dyke & McElree, 2006). The pattern of the two-way interaction found in regression-path duration in the verb region showed greater syntactic interference effects for the Low Semantic Interference conditions than for the High Semantic Interference conditions. This interaction pattern is interesting because the High Semantic, High Syntactic Interference condition was predicted to elicit the longest RTs at the retrieval site, leading to greater syntactic interference effects for the High Semantic Interference conditions than for the Low Semantic Interference conditions. However, the result unexpectedly showed that the syntactic interference effects, in fact, were greatly reduced when there was high semantic interference in the sentence. This may suggest that semantic cues could be weighted more heavily than syntactic cues at the initial retrieval stage (cf. semantic heuristics within the Good Enough Theory [Christianson, 2016; Ferreira & Patson, 2007]). In comparison, the L2 online data failed to show this two-way interaction between the two fixed effects. This lack of interaction suggests that L2 readers may have difficulty in combining all relevant cues into a single retrieval probe or simultaneously employing multiple cues at retrieval, possibly due to their limited resources, taxed by the activation and inhibition of L1 or reduced automaticity in L2 (cf. Lim & Christianson, 2013b). Differences between L1 and L2 reading will be further discussed below.

Another interesting result was that both L1 and L2 data revealed only syntactic interference effects in total viewing time in the main verb. This may indicate that both L1 and L2 readers relied more on syntactic cues than semantic cues at the later retrieval stage to discriminate the target from the distractor. This possibility is very likely, especially for the sentences in the High Semantic, High Syntactic Interference condition, where semantic interference resolution may depend on using discriminative syntactic cues. Given that linguistic cues may be weighted differently depending on the characteristics of the material being read or different retrieval stages as discussed above, cue weighting might need to be incorporated in the cue-based retrieval theory to better understand memory mechanisms underlying sentence processing.

L1 and L2 question-response accuracy data also revealed interference effects. The patterns were similar to those of the online reading data, with high interference conditions eliciting lower comprehension accuracy compared to the low interference conditions. These observed interference effects on retrieval probability again provide supportive evidence for the cue-based retrieval account, and demonstrate that like more-skilled L1 readers, less-skilled L2 readers also use a cue-based retrieval mechanism in offline comprehension tasks.

What was interesting was that while no interaction between semantic and syntactic interference was found in the L2 online reading data, the L2 offline accuracy data revealed a reliable two-way interaction between the two fixed effects. This discrepancy between online and offline data could be due to task effects. During online reading task, participants were not required to resolve interference at retrieval, whereas, in the offline task, interference had to be resolved because participants were asked to verbally provide a correct target after reading a sentence. Unlike skilled L1 readers, less fluent L2 readers may have limited resources to compute syntactic relations during reading, because their resources may also be taxed by inhibiting their L1 and simultaneously activating less automatized L2. Because of their limited resources, it is likely that they may leave syntactic interference unresolved at retrieval until they are required to resolve it. In other words, less-skilled L2 readers may perform ‘good-enough’ processing, leaving the attachment underspecified. This explanation is supported by the fact that only the L2 online data failed to reveal the two-way interaction between the two fixed effects.

The evidence of interference effects observed in Experiments 1 and 2 suggests that sentence comprehension relies on a series of direct-access, cue-driven retrievals. However, given that the hierarchical structure of a sentence is encoded by the order of linguistic constituents, it is important to consider a possibility that a serial-search retrieval, which is a relatively slow search necessary for recovering order information (Gronlund et al. 1997; McElree 2001, 2006; McElree & Dosher 1993), may also play a role in sentence comprehension. That said, the cue-based retrieval model needs to incorporate how location-based functional relationships between constituents are computed during sentence processing, which should be the next goal for future research.

Factors influencing retrieval efficiency during sentence comprehension

Given compelling evidence for the cue-based retrieval model, the current research also investigated what factors influence target retrieval latency and probability during sentence comprehension. While Experiments 1 and 2 focused on features that are cued by the retrieval trigger (e.g., semantic and syntactic features of the subject in a subject-verb dependency), Experiments 3 and 4 explored a factor that is *not* cued by the retrieval trigger. Specifically, these two experiments examined semantic complexity effects of the target and the distractor on retrieval efficiency in L1 and L2 reading. The underlying assumption was that enhancing the quality of lexical representations of the target and/or the distractor through semantic elaboration increases the uniqueness of the target representation, facilitating target retrieval (Hofmeister, 2011; Hofmeister & Vasishth, 2014). Semantic complexity was manipulated by varying the number of modifiers for the target and the distractor, creating complex and simple semantic elaboration conditions for the target and the distractor respectively.

Based on the findings from Hofmeister and Vasishth (2014), complex semantic elaboration conditions were predicted to elicit reduced RTs at the retrieval sites and higher comprehension accuracy compared to the simple semantic elaboration conditions. However, the present results revealed a contrasting pattern, such that both L1 and L2 data demonstrated detrimental effects of semantic elaboration on retrieval latency and probability. Specifically, providing additional semantic information for the target and/or the distractor led to inflated RTs at the retrieval sites and lower comprehension accuracy. These observed semantic elaboration effects demonstrate that factors that are *not* cued by the retrieval trigger do indeed influence retrieval processes, but in ways that are predicted by the cue-based retrieval architecture that is currently proposed. Therefore, the current cue-based retrieval model needs to be further developed to incorporate factors that do *not* match the retrieval cues, but nevertheless influence retrieval efficiency in the model.

Although both L1 and L2 data revealed semantic complexity effects on retrieval efficiency, the direction of the effects was opposite to what was found in Hofmeister and Vasishth (2014). As mentioned earlier, Hofmeister and Vasishth manipulated semantic complexity for the target and the distractor that appeared *before* the target. However, the issue with their stimuli was that the critical sentences also included another distractor that shared the same features with the retrieval cues, and this appeared *after* the target, creating both proactive and retroactive interference. According to the cue-based retrieval model, linguistic chunks in memory decay as a function of time and prior retrievals, and thus having two distractors, which shared the same features with the retrieval cues before and after the target, could have been a confound in their study. The sentences in the current research included only one distractor after the target, allowing us to examine retroactive interference on retrieval efficiency.

The following question is then why providing additional semantic information for the target and/or the distractor increased retrieval interference, instead of facilitating retrieval processes. As cue-based retrieval claims, if the probability of retrieving a correct target greatly depends on the distinctiveness of the representation itself, we should have seen facilitative effects of semantic elaboration on retrieval latency and probability. One possibility, consistent with the activation-based framework, is that processing additional modifiers may increase spurious activation of irrelevant linguistic information at retrieval. Although the modifying words used in the current experiments were likely to describe attributes of the target and the distractor in real-world, it is possible that activations of these various modifiers spread through their semantically associated networks (i.e. spreading activation, Anderson, 1995) and generated spurious activations at retrieval, increasing retrieval interference.

Alternatively, the additional modifiers could have increased processing/memory load during sentence processing and caused difficulty at retrieval because sentences with additional modifiers in fact included more words up until the verb compared to those without modifiers. This hypothesis is in line with the capacity-based approach, however, which finds little additional support in the present results, in that this emphasizes the storage component of memory. That said, future research should further investigate whether or not different memory operations are also employed for different types of syntactic structures.

Another interesting finding was that while no interaction between target and distractor complexity was found in the L1 online reading time data, the L2 online data revealed a reliable two-way interaction in regression-path duration in the post-verb region. The interaction pattern was that the target complexity effects were greater in the simple distractor conditions than in the complex distractor conditions, indicating that target semantic complexity effects were modulated by the quality of lexical representation of the distractor. The observed L1-L2 difference suggests that L2 readers may be more susceptible to interference from the distractor with semantic elaboration than L1 readers. Given that less-skilled L2 readers may have lower quality lexical representations compared to skilled L1 readers, their target and distractor representations are more likely to be noisier. Thus, adding additional modifiers for the distractor could have increased L2 readers' spurious activations of irrelevant linguistic information at retrieval, resulting in retrieval operations that are less efficient and more subject to interference.

The current research explored the influence of the semantic component of lexical representations (i.e. semantic complexity) on retrieval efficiency, and the findings in Experiments 3 and 4 demonstrated detrimental effects of semantic elaboration on retrieval processes. Given that the quality of lexical representations is determined not only by semantic features, but by other features, such as orthographic and phonological information, future research should examine cues associated with other components of lexical representations to better understand the role of lexical representation quality on retrieval interference.

In fact, recent research has investigated whether phonological overlap between distractors and a retrieval target causes retrieval interference during sentence comprehension, but the findings are mixed. For example, Acheson and MacDonanld (2011) demonstrated that phonological similarity within a RC (e.g., *The* ***baker*** *that the* ***banker*** *sought bought the house*) resulted in slower reading times and lower comprehension question-response accuracy compared to when there was no phonological overlap within the RC (e.g., *The* ***runner*** *that the* ***banker*** *feared bought the house*). Kush, Johns, and Van Dyke (2015), on the other hand, reported no effects of phonological similarity on retrieval processes. By using a memory-load paradigm in a self-paced reading task, they manipulated the words in the memory load, such that the words either rhymed or did not rhyme with the target word (*boat*) in an object cleft clause (e.g., Rhyme Memory Load: *coat, vote, note*; No Rhyme Memory Load: *table, sink, truck*; Sentence: *It was the **boat** that the guy who drank some hot coffee sailed on two sunny days*). Their results showed a detrimental effect of phonological overlap at the encoding region (i.e., longer reading time at *that the guy* in the rhyme condition compared to the no-rhyme condition), but no effect was found at the critical verb region.

Findings of the role of semantic complexity and phonological cues in retrieval processes are unclear as shown in discrepancies between the current research and Hofmeister and Vasishth (2014), as well as between studies examining the role of phonological overlap on retrieval interference. Given that the cue-based retrieval model does not make predictions about the cues that are *not* cued by the retrieval trigger, data from future work that examines these factors will not only be critical to further developing the cue-based model, but also will enhance our understanding of memory operations underlying sentence comprehension.

The relationships between retrieval ability and individual difference factors

Successful reading comprehension depends not only on the characteristics of the material being read, but also on people's individual cognitive differences and language-related skills. Thus, the current research examined how individual difference factors interact with the experimental manipulations and influence retrieval processes during sentence comprehension. The following section discusses the influence of readers' WM capacity and reading speed on retrieval interference and semantic elaboration effects during L1 and L2 sentence comprehension.

Working memory capacity. The current research examined the interaction between L1 and L2 readers' WM capacity and the effects of retrieval interference (in Experiments 1 and 2), and semantic elaboration (in Experiments 3 and 4).

The most interesting finding was that L1 readers' WM capacity showed no interaction with retrieval interference and semantic elaboration effects in their online reading time data. On the other hand, the L2 reading time data showed moderating effects of WM capacity on these fixed effects, such that L2 readers with lower WM capacity were more susceptible/sensitive to retrieval interference and semantic elaboration effects compared to higher span L2 readers.

No moderating effects of L1 readers' WM capacity on retrieval interference are consistent with the findings from Van Dyke et al. (2014), which demonstrated that after partialling out variability associated with intelligence, English L1 speakers' WM capacity was no longer a strong predictor of comprehension. This finding is critical in that it provides supportive evidence for a foundational assumption of the cue-based retrieval model, namely that sentence processing relies on a cue-driven retrieval operation, rather than fixed WM capacity, and suggests that individual differences in WM capacity may not be a strong predictor of reading ability.

The next important question is what WM processes are associated with retrieval ability and explain the observed differences between L1 and L2 readers. The most likely candidate that would be consistent with the cue-based retrieval framework is attentional control, which is responsible for sustaining focus on relevant information and suppressing irrelevant information (Farmer, Misyak, & Christiansen, 2012). Compared to L1 readers, L2 readers may have a higher proportion of low-quality lexical representations, whose linguistic features are not fully specified, and thus as Malko et al. (2016) claimed, their L2 input is noisier than that of L1 readers. This noisier input makes it difficult for L2 readers to retrieve the correct bundle of features when retrieval is required, and attentional control may be recruited during retrieval to direct focus to relevant feature bundles. Furthermore, given that L2 readers also need to inhibit their automatized L1 while simultaneously processing L2, it is very likely that attentional control is involved in L2 retrieval processes. In contrast, skilled L1 readers have a higher portion of high-quality lexical representations, and thus their lower-level linguistic processes (e.g., lexical access) are automatized and very efficient. Consequently, their retrieval processes would be less susceptible to interference and thus rely less on cognitive resources like attentional control.

While only L2 online reading time data showed a reliable interaction between L2 readers' WM capacity and the effects of interference and semantic elaboration on retrieval, both L1 and L2 offline accuracy data showed the moderating effects of WM capacity on these fixed effects. The overall interaction patterns varied in terms of experimental manipulations (retrieval interference vs. semantic complexity) and populations of interest (skilled L1 vs. less-skilled L2 readers). The findings in Experiments 1 and 2 showed that overall lower-span L1 and L2 readers seem to be more susceptible to retrieval interference. On the other hand, in terms of the relationship between WM capacity and the ability to use additional semantic information for efficient retrieval, L1 and L2 readers showed different interaction patterns. The two-way interaction between L1 readers' WM capacity and distractor elaboration effects was driven by higher span L1 readers' higher question-response accuracy for the Simple Distractor conditions. The three-way interaction in the L2 data was driven by the higher-span L2 readers' higher accuracy for the Simple Target conditions than the Complex Target conditions when there was a complex distractor. Again, this L1-L2 difference could be due to L2 readers' less efficient lower-level reading processes (e.g., lexical access), possibly increasing their reliance on attentional control functions of WM.

Reading speed. Motivated by Perfetti's (1985) Verbal Efficiency theory, which suggests that successful reading comprehension depends on the efficiency of word-level processing, and the supportive evidence for a strong correlation between word reading and comprehension (Hess & Radtke, 1981; Jackson & McClelland, 1979; Perfetti & Hogaboam, 1975), the current study examined the association between reading speed (an average RT per word) and retrieval ability in L1 and L2 sentence comprehension.

Both L1 and L2 online and offline data revealed the moderating effects of reading speed on retrieval interference and semantic elaboration, which suggests the importance of word decoding skills in successful sentence comprehension. What was very interesting was that L1 and L2 were alike in terms of interaction patterns between reading speed and retrieval interference, such that slower L1 and L2 readers were more susceptible to retrieval interference in both online and offline data. In particular, the fact that during online reading, the reading speed measure interacted with only semantic interference is interesting because it provides clear evidence for the previous claim that faster readers with extensive language experience would be more sensitive to semantic cues (Nicenboim et al., 2016; Traxler et al., 2012), and that they would be more efficient at discriminating lexical items in memory, leading to successful target retrieval during reading (Van Dyke & Johns, 2012; Van Dyke & Shankweiler, 2012).

In terms of interaction between reading speed and elaboration effects, faster L2 readers were sensitive to target elaboration effects, such that they showed shorter reading times for the simple targets. In the offline accuracy data, both L1 and L2 readers showed a two-way interaction between reading speed and distractor elaboration effects, such that slower L1 and L2 readers were more accurate in responding to comprehension questions in the Simple Distractor conditions. The finding that *slower* L1 and L2 readers were more sensitive to distractor semantic elaboration and accurate in responding to comprehension questions for the simple distractor is unexpected because faster readers who may be more sensitive to semantic information were predicted to show greater semantic elaboration effects compared to slower readers. This unexpected pattern could be because although slower readers may encode semantically elaborated information, faster readers may preferentially not encode some unnecessary semantic information (e.g., adjunct phrases) as a reading strategy to reserve resources for other required computations during sentence processing (cf. Frazier & Clifton, 1997). The fact that reading speed only interacted with semantic elaboration for the distractor, which is *not* the to-be-retrieved target, may support this explanation.

Skilled versus less-skilled reading

One of the important goals of this dissertation was to understand whether the ability to use a cue-driven retrieval operation differentiates skilled versus less-skilled reading. Thus, the current research compared reading patterns between L1 and L2 readers. The following demonstrates the observed L1-L2 differences in sensitivity to retrieval interference and suggests the most likely predictor that may explain comprehension variability. Also, the final section reviews theories of L2 sentence processing, particularly, the Shallow Structure Hypothesis (Clahsen & Felser, 2006) and the Good Enough processing theory (Christianson, 2016; Christianson et al., 2001; Christianson et al., 2006; Christianson et al., 2010, Ferreira, 2003; Ferreira et al., 2001; Ferreira & Patson, 2007; Lim & Christianson, 2013a, 2013b; Patson et al., 2009; Swets et al., 2008), in relation to the findings of the current research.

The observed retrieval interference effects in both L1 and L2 data suggest that not only skilled readers, but also less-skilled readers employ an efficient cue-driven operation during sentence comprehension, and thus the ability to employ the cue-driven retrieval may *not* determine skilled- versus less-skilled reading. The immediate question is then what characterizes relatively more-skilled or less-skilled readers? The findings that the L2 online data failed to show interaction between semantic and syntactic interference (whereas the L1 online revealed the two-way interaction), and that only L2 online data revealed moderating effects of WM capacity on retrieval interference and semantic elaboration point towards the hypothesis that L1-L2 differences may arise due to individual differences in the quality of lexical representations. Less-skilled L2 readers may have a higher portion of low-quality of lexical representations whose linguistic features are underspecified, and this makes their linguistic input noisy and increases spurious activation of irrelevant information during sentence processing. Less-skilled readers thus may need to rely more on cognitive resources like attentional control of WM to direct their focus on the correct feature bundles and inhibit irrelevant information. The fact that no moderating effects of WM capacity were found in the L1 online data, but found in the L2 online data may support this hypothesis. This view is consistent with Perfetti's Verbal Efficiency (2007, 2011) and the findings from Van Dyke et al. (2014), emphasizing the quality of readers' mental representations for successful sentence comprehension.

In Experiment 2, L2 readers showed sensitivity to not only semantic cues, but also syntactic cues in both online and offline data, and this provides evidence *against* the Shallow Structure Hypothesis (Clahsen & Felser, 2006), which claims that L2 readers are not able to use syntactic information and that they therefore construct shallow structural representations, relying more on semantic and pragmatic information during sentence processing. Rather, the observed

L2 readers' reading patterns can be better captured by Good Enough processing because L2 sentence processing did indeed proceed along both semantic and syntactic routes (Lim & Christianson, 2013a, b). Regarding Cunning's (2016) claim that less-skilled L2 readers may be more susceptible to retrieval interference than L1 readers during reading, the current research may not be able to accurately evaluate how susceptible L2 readers are to similarity-based retrieval interference in comparison to L1 readers, unless we can be sure that L2 readers generated the correct retrieval cues and were able to fully specify features of the target and distractor during reading. However, the finding that only the L2 online data revealed interactions between target and distractor semantic elaboration in Experiments 3 and 4 may suggest that L2 readers are more susceptible to interference from the distractor than L1 readers. Important to note, however, is that the primary source of L1-L2 differences in reading performance may be attributed to the differences in the quality of lexical representations between L1 and L2 readers, rather than the differences in sensitivity to similarity-based interference during reading because individual differences in vocabulary knowledge may influence retrieval efficiency during reading.

Conclusion

Given that the defining property of natural language is the ability to establish non-adjacent relationships, this dissertation investigated how readers retrieve linguistic representations outside focal attention during sentence comprehension, and what factors influence these retrieval processes. The findings of the current research provide evidence that sentence comprehension relies on a direct-access, cue-driven retrieval operation, and suggest that attentional control of WM may be involved in retrieval processes, as the cue-based retrieval model claims that what is limited is attentional resources, rather than capacity *per se.* Also, the data presented in the current research revealed the detrimental effects of semantic elaboration on retrieval efficiency. Given that the quality of lexical representation is determined not only by semantic features, but also by other features, such as orthographic and phonological information, future study should further investigate whether overlap of these features between distractors and a retrieval target influences retrieval latency and probability to better understand the role of quality of lexical representations on retrieval interference.

References

Abney, S. P., & Johnson, M. (1991). Memory requirements and local ambiguities of parsing strategies. *Journal of Psycholinguistic Research*, *20*, 233-250.

Acheson, D. J., & MacDonald, M. C. (2011). The rhymes that the reader confused the meaning: phonological effects during on-line sentence comprehension. *Journal of Memory and Language*, *65*, 193-207.

Adams, E. J., Nguyen, A. T., & Cowan, N. (2018). Theories of Working Memory: Differences in Definition, Degree of Modularity, Role of Attention, and Purpose. *Language, Speech, and Hearing Services in Schools*, *49,* 340–355.

Anderson, J. R. (1995). *Cognitive psychology and its implications* (4th ed.). New York: W. H. Freeman.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111,* 1036-1060.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Baddeley, A. D. (1966). The Capacity for generating information by randomization. *Quarterly Journal of Experimental* Psychology, 18, 119-129.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4,* 417-423.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1-29.

Baddeley, A., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.) *The psychology of learning and motivation: Advances in research and theory*, *8*, 47-89. New York:

Academic Press.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278

Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 317–33.

Bialystok, E., & Miller, B. (1999). The problem of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition*, *2*, 127-145.

Blackwell, A., & Bates, E. (1995) Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation*. Journal of Cognitive Neuroscience*. *7*, 228–257.

Carpenter, P. A., Miyake, A., & Just, M. A. (1994). *Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging*. San Diego, CA: Academic Press.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*, 77-94.

Caplan, D. & Waters, G.S. (2013). Memory mechanisms supporting syntactic comprehension. *Psychonomic Bulletin and Review*, *20*, 243-268.

Christianson, K. (2016). When language comprehension goes wrong for the right reasons: Good-Enough, underspecified or shallow language processing. *The Quarterly Journal of Experimental Psychology*, *69*, 817 –828.

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

Christianson, K., Luke, S. G., & Ferreira, F. (2010). Effects of plausibility on structural priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 538–544.

Christianson, K., Williams, C. C., Zacks, R. T., & Ferreira, F. (2006). Younger and older adults' "Good-Enough" interpretations of garden-path sentences. *Discourse Processes*, *42*, 205–238.

Clahsen, H. & Felser, C. (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences*, *10*, 564-570.

Clark, S. E., & Gronlund, S. D. (1996). Global Matching Models of Recognition Memory: How the Models Match the Data. *Psychonomic Bulletin & Review, 3*, 37-60.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, *3*, 75-84.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R.W. (2005). Working memory span tasks: a methodological review and user's guide. *Psychonomic Bulletin and Review*, *12*,769–786.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Brain and Behavioral Sciences*, *24*, 87–114.

Cowan, N. (2006). *Working memory capacity.* New York: Psychology Press.

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

Cunnings, I. (2016). Parsing and Working Memory in Bilingual Sentence Processing. *Bilingualism: Language and Cognition*, 1-20.

Daneman, M., and Carpenter, P. (1980). Individual differences in WM and reading. *Journal of*

Verbal Learning and Verbal Behavior, *19*, 450–466.

Dillon, B. A. short discourse on reflexives: A reply to Cunnings (2016). *Bilingualism: Language and Cognition*.

Dosher, B. A. (1979). Empirical approaches to information processing: Speed-accuracy tradeoff or reaction time. *Acta Psychologica, 43*, 347-359.

Dosher, B.A., & McElree, B. (2002). Memory search: Retrieval processes in short-term and long-term recognition. In J. H. Byrne (Ed.) *Learning & Memory*. New York: Gale group.

Dussias, Paola 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.

Ericsson, k. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211-245.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

Farmer, T. A., Misyak, J. B., & Christiansen, M. H. (2012). Individual differences in sentence processing. In M. J. Spivey, K. McRae, & M. F. Joanisse (Eds.), *Cambridge handbooks in psychology. The Cambridge handbook of psycholinguistics* (pp. 353-364). New York, NY, US: Cambridge University Press.

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language*, *54*. 541–553.

Fedorenko, E., Gibson, E., & Rohde, D. (2007). The nature of working memory in linguistic, arithmetic and spatial integration processes. *Journal of Memory and Language*, 56, 246–269.

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

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

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164 –120.

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research*, *30*, 3 –20.

Ferreira, F., & Patson, N. D. (2007). The "Good Enough" approach to language comprehension. *Language and Linguistics Compass*, *1*, 71 –83.

Fernandez, E. M. (2006). How do language learners build syntactic structure on-line? *Applied Psycholinguistics*, *27*, 59-64.

Foote, R. (2010). Age of acquisition and proficiency as factors in language production: Agreement in bilinguals. *Bilingualism: Language and Cognition*, *13*, 99−118.

Foraker S, McElree B. (2007) The role of prominence in pronoun resolution: Active versus passive representation. *Journal of Memory and Language*, *56*, 357–383.

Frazier, L., & *Clifton* Jr, C. (*1997*). Construal: Overview, motivation, and some new evidence. *Journal of Psycholinguistic Research*, *26*, 277-297.

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*, 673-679.

Gallo, D., Meadow, N., Johnson, E., & Foster, K. (2008). Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. *Journal of Language and Memory, 58*, 1095–1111.

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

Glaser, Y. G., Martin, R. C., Van Dyke, J. A., Hamilton, A. C., & Tan, Y. (2013). Neural basis of semantic and syntactic interference in sentence comprehension. *Brain Language*, *126*, 314–326.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*, 3-10.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory load interference in syntactic processing. *Psychological Science*, *13*, 425–430.

Gronlund, S. D., Edwards, m. B., & Ohrt, D. D. (1997). Comparison of the retrieval of item versus spatial position information. *Journal of Experimental Psychology: Learning, Memory, & cognition*, *23*, 1261-1274.

Hess, T. M., & Radtke, R. C. (1981). Processing and memory factors in children's reading comprehension skill. *Child Development*, *52*, 479-488.

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Process*. *26*, 376–405.

Hofmeister, P., & Vasishth, S. (2014). Distinctiveness and encoding effects in online sentence comprehension. *Frontiers in Psychology*. *5*:1237.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, *22*, 369-397.

Hopp, H. (2014). Working memory effects in the L2 processing of ambiguous relative clauses. *Language Acquisition*, *21*, 250-278.

Hoshino, N., Dussias, P.E., & Kroll, J.F. (2010). Processing subject-verb agreement in a second language depends on proficiency. *Bilingualism: Language and Cognition*, 13, 87−98.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434-446.

Jacob, G., Lago, S., & Patterson, C. (2016). L2 processing and memory retrieval: Some empirical and conceptual challenges. *Bilingualism: Language and Cognition*, 1-3. doi:10.1017/S1366728916000948

Jackson, M., & McClelland, J. (1979). Sensory and cognitive determinants of reading speed. *Journal of Experimental Psychology: General*, *108*, 151-181.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329-354.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *98*, 122–149.

Just, M. A., Carpenter, P. A., & Keller, T. A. (1996). The capacity theory of comprehension: New frontiers of evidence and arguments. *Psychological Review*, *103*, 773–780.

Just, M. A., & Varma, S. (2002). A hybrid architecture for working memory: Reply to MacDonald and Christiansen (2002). *Psychological Review*, *109*, 55-65.

Just, M. A., & Varma, S. (2007). The organization of thinking: What functional brain imaging reveals about the neuroarchitecture of complex cognition. *Psychological Review*, *109*, 55-65.

Kaan, E. Susceptibility to interference: Underlying mechanisms, and implications for prediction. *Bilingualism: Language and Cognition*. doi: http://dx.doi.org/10.1017/S1366728916000894

Kamide, Y., & Mitchell, D. C. (1997). Relative clause attachment: Non-determinism in Japanese parsing. *Journal of Psycholinguistic Research*, *26*, 247-254.

Kilborn, K. (1991). Selective impairment of grammatical morphology due to induced stress in normal listeners: Implications for aphasia. *Brain and Language*, *41*, 275-288.

Kim, J. H. & Christianson, K. (2013). Sentence complexity and working memory effects in ambiguity resolution. *Journal of Psycholinguistic research*, *42*, 393-411.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*, 15-47.

King, J. W., & Just, M. A. (1991). Individual differences in syntactic parsing: The role of working memory. *Journal of Memory and Language, 30*, 580–602.

Kliegl, R., Masson, M. E. J., Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, *18*, 655–681.

Kohonen, T. (1984). *Self-organization and associative memory.* Berlin: Springer-Verlag.

Kush, D., Johns, C. and Van Dyke, J. 2015. Identifying the role of phonology in sentence-level reading difficulty. *Journal of Memory and Language*, *79*, 18-29.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?!. *Intelligence*, *14*, 389–433.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive science*, *29*, 375–419.

Lewis, R. L., Vasishth, S., and Van Dyke, J. A. (2006). Computational principles of WM in sentence comprehension. *Trends in Cognitive Sciences*. *10*, 447–454.

Li, P., Zhang, F., Tsai, E., Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*, 673-680.

Lim, H. L., & Christianson, K. (2013a). Second language sentence processing in reading for comprehension and translation. *Bilingualism: Language and Cognition*, *16*, 518–537.

Lim, H. L., & Christianson, K. (2013b). Integrating meaning and structure in L1-L2 and L2-L1 translations. *Second Language Research*, *29*, 233-256.

Lim, H. L., & Christianson, K. (2014). Second language sensitivity to agreement errors: Evidence from eye movements during comprehension and translation. *Applied Psycholinguistics*, *36*, 1283–1315.

Malko, A., Ehrenhofer, L., & Phillips, C. Theories and frameworks in second language processing. *Bilingualism: Language and Cognition*. doi: https://doi.org/10.1017/S1366728916001000

Marinis, T., Roberts, L, Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Essex Research Reports in Linguistics*, *45*, 43-79.

Martin, A. E., & McElree, B. (2009). Memory Operations That Support Language Comprehension: Evidence From Verb-Phrase Ellipsis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1231-1239.

Martin, A. E., & McElree, B. (2011). Direct-access retrieval during sentence comprehension: Evidence from Sluicing. *Journal of Memory and Language*, *64*, 327-343.

McDonald, J.L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, 21, 395-423.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*, 381–401.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*, 35–54.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24,* 56–98.

McElree, B. (2000). Sentence comprehension is mediated by content addressable memory. *Journal of Psycholinguistic Research*, *29*, 111–123.

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 817–835.

McElree, B. (2006). Accessing recent events. *Psychology of Learning and Motivation*. *46*, 155–200.

McElree, B. & Dosher, B. A. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General*, *122*, 291-315.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*, 67–91.

Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorder: Making normal adults perform like aphasic patients. C*ognitive Neuropsychology*, *11*, 671–717.

Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In D. R. Luce, R. R. Bush, & E. Galanter (Eds.) *Handbook of mathematical psychology*, Vol. II. New York: John Wiley.

Miyake, A., Carpenter, P. A., & Just, M. A. (1995). Reduced resources and specific impairments in normal and aphasic sentence comprehension. *Cognitive Neuropsychology, 12*, 651-679.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related: A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*, 621–640.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78*, 679-684.

Nairne, J. S. (2002). Remembering over the short-term: the case against the standard model. *Annual Review of Psychology*, *53*, 53–81.

Nicenboim, B., Logacev, P., Gattei, C., & Vasishth, S. (2016). When high-capacity readers slow down and low-capacity readers speed up: Working memory and locality effects. *Frontiers in Psychology*. 7:280. doi: 10.3389/fpsyg.2016.00280

Nicenboim, B. & Vasishth, S. (2018). Models of retrieval in sentence comprehension: A computational evaluation using Bayesian hierarchical modeling. *Journal of Memory and Language*, *99*, 1-34.

Nicol, J., & Greth, D. (2003). Production of subject–verb agreement in Spanish as a second language. *Experimental Psychology, 50*, 196–203.

Oberauer, K. (2002). Access to information in WM: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory and Cognition*. *28*, 411–421.

Omaki, A. (2016). Linking learning and parsing in bilingual sentence processing. *Bilingualism: Language and Cognition*.

Öztekin, I., &McElree, B. (2006). Proactive interference slows recognition by eliminating fast assessments of familiarity. *Journal of Memory & Language*, *57*, 126–149.

Öztekin, I., & McElree, B. (2007). Proactive interference slows recognition by eliminating fast assessments of familiarity. *Journal of Memory and Language*, *57*, 126–149.

Pan, H.-Y., S. Schimke & C. Felser (2015). Referential context effects in non-native relative clause ambiguity resolution. *International Journal of Bilingualism*, *19*, 298–313.

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, *25*, 501-528.

Parker, D., Shvartsman, M., Van Dyke, J. A. (2017). *The cue-based retrieval theory of sentence comprehension: New findings and new challenges*. In Escobar, L., Torrens, V., Parodi, T. (eds.) Language Processing and Disorders. Newcastle: Cambridge Scholars Publishing.

Patson, N. D., Darowski, E. S., Moon, N., & Ferreira, F. (2009). Lingering misinterpretations in garden-path sentences: Evidence from a paraphrasing task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 280 –285.

Perception Research Systems. (2007). *Paradigm Stimulus Presentation*, Retrieved from http://www.paradigmexperiments.com

Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.

Perfetti, C. A. (2011). Phonology is critical in reading but a phonological deficit is not the only source of low reading skill. In S. A. Brady, D. Braze, & C. A> Fowler (Eds.), *Explaning individual differences in reading: Theory and evidence (*pp. 153-171). New York:

Psychology Press.

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and comprehension skills. *Journal of Educational Psychology*, *67*, 461-469.

Pickering, M.J. & van Gompel, R.P.G. (2006). Syntactic parsing. In M.J Traxler & M. Gernsbacher. (Eds.), *Handbook of Psycholinguistics*, 2nd Ed. (pp. 455-503), Amsterdam: Academic Press.

Reed, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science*, *181*, 574-576.

Reed, A. V. (1976). The time course of reconition in human memory. *Memory & Cognition*, *4*, 16-30.

Scherag, A., Demuth, L., Rosler, F., Neville, H. J., & Roder, B. (2004). The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects. *Cognition*, *93*, 97-108.

Swan, D., & Goswami, U. (1997a). Phonological Awareness Deficits in Developmental Dyslexia and the Phonological Representations Hypothesis. *Journal of Experimental Child Psychology*, *66*, 18-41.

Swan, D., & Goswami, U. (1997b). Picture naming deficits in developmental dyslexia: The phonological representations hypothesis. *Brain and Language*, *56*, 334-353.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. Jr. (2008). Underspeci fication of syntactic ambiguities: Evidence from self-paced reading. *Memory and Cognition*, 36, 201 –216

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, *136*, 64-81.

Tabor, W., Galantucci, B., and Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language*, *50,* 355–370.

Tan, Y., Martin, R. C., & Van Dyke, J. A. (2017) Semantic and syntactic interference in sentence comprehension: A comparison of working memory models. *Frontiers in Psychology*, 8:198. doi: 10.3389/fpsyg.2017.00198

Tanner, D., Nicol, J., Herschensohn, J., & Osterhout, L. (2012). Electrophysiological markers of interference and structural facilitation in native and nonnative agreement processing. In Biller, A., Chung, A., & Kimball, A. (eds.), *Proceedings of the 36th Boston University Conference on Language Development*, pp. 594–606. Somerville: Cascadilla.

Traxler, M. J., Long, D. L., Johns, C. L., Zirnstein, M., Tooley, K. M., & Jonathan, E. (2012). Individual differences in eye-movements during reading: working memory and speed-of-processing effects. *Journal of Eye Movement Research*, *5*, 1–16.

Troyer, M., Hofmeister, P., & Kutas, M. (2016) Elaboration over a discourse facilitates retrieval in sentence processing. *Frontiers in Psychology*, *7*, 374.

Turner, M. T., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*,127–154.

Underwood, B. J., & Keppel, G. (1962). An evaluation of two problems of method in the study of retention. *American Journal of Psychology*, *75*, 1–17.

Underwood, B. J., and G. Keppel. 1962. An evaluation of two problems of method in the study of retention. *American Journal of Psychology*, *75*. 1–17.

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014).Working memory and fluid intelligence: capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26.

Van Dyke, J. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 407–430.

Van Dyke, J., and Johns, C. L. (2012). Memory interference as a determinant of language comprehension. *Language and Linguistics Compass*, *6*, 193–211.

Van Dyke, J., Johns, C. L., & Kukona, A. (2014). Low WM capacity is only spuriously related to poor reading comprehension. *Cognition*, *131*, 373–403.

Van Dyke, J., and Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: a cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*, 285–316.

Van Dyke, J., and McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, *55*, 157–166.

Van Dyke, J., and McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, *65*, 247–263.

Van Dyke, J.A. & Shankweiler, D. (2012). *From Verbal Efficiency Theory to Lexical Quality: The Role of Memory Processes in Reading Comprehension*. In M.A. Britt, S.R. Goldman & J-F Rouet (Eds), Reading: From Words to Multiple Texts (pp. 115-131). Routledge, Taylor & Francis Group.

Verhaeghen, P., Cerella, J., & Basak, C. (2004). A working memory workout: How to change to size of the focus of attention from one to four in ten hours or less. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1322-1337.

Watkins, O.C., & Watkins, M. J. (1975). Build-up of proactive inhibition as a cue overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, *104*, 442–52.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review*, *72*, 89–104.

Wickelgren, W. A. (1973). The long and the short of memory. *Psychological Bulletin*, *80*, 425-438.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychological*, *41*, 67-85.

Witzel, J., Witzel, N., & Nicol, J. (2012). Deeper than shallow: Evidence for structure-based parsing biases in L2 sentence processing. *Applied Psycholinguistics*, *33*, 419-456.

Wolf, M. & Bowers, P. G. (1999). The Double-Deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, *91*, 415-438.

Appendix A: Example of experiment sentences

[Experiments 1 and 2]

The *resident* (who was living near the dangerous warehouse/neighbor // who said that the warehouse/neighbor was dangerous) last month *had complained* about the investigation.

The client (who had arrived after the important meeting/visitor // who implied that the meeting/visitor was important) that day was waiting in the office.

The resident (who was living near the dangerous warehouse/neighbor // who said that the warehouse was dangerous) last month had complained about the investigation.

The teacher (who was designing the new curriculum/specialist // who realized that the curriculum/specialist was new) last night will come to the office.

The opponent (who was fighting the corrupt government/ governor // who had claimed that the government/ governor was corrupt) for nearly three years should be arrested immediately.

The manager (who liked the clever show/producer // who said that the show/producer was clever) at the opening ceremony could negotiate a good deal.

The player (who played in the tough competition/ competitor // who signaled that the competition/ competitor was tough) before this game will make the shot.

The actor (who was starring in/with the inspiring play/director // who said that the play/director was inspiring) after/at the press conference will speak for an hour.

The instructor (who had looked for the prepared resume/student // who assumed that the resume/student was prepared) last semester will find out the truth.

The passenger (who was sitting in the new seat/driver // who commented that the seat/driver was new) on the bus was talking on the phone.

The physicist (who had admired the amazing calculation/chemist // who shouted that the calculation/chemist was amazing) at the conference was making too much noise.

The attorney (who was questioning the unusual motion/witness // who commented that the motion/witness was unusual) in the courtroom was exaggerating quite a bit.

The candidate (who was attacked by the dishonest commercial/senator // who charged that the commercial/senator was dishonest) in the newspaper was losing the race.

The publicist (who had paid for the brilliant painting/painter // who assumed that the painting/painter was brilliant) at the first meeting will cancel the exhibit.

The judge (who had criticized the questionable evidence/witness // who decided that the evidence/witness was questionable) recently had misunderstood the facts.

The secretary (who was complaining about the unreasonable policy/director // who complains that the policy/director is unreasonable) on TV is quitting next month.

The director (who disliked the outrageous performance/performer // who exclaimed that the performance/performer was outrageous) in the movie had wanted to quit.

The receptionist (who had called during the important meeting/client// who knew that the meeting/client was important) last time had forgotten the company's policy.

The surgeon (who had operated with/on the difficult tool/patient // who denied that the tool/patient was difficult) last night complained about the procedure.

The criminal (who had shot at the hidden car/man // who saw that the car/man was hidden) in the garage was lying in the bushes.

The teller (who was working in the boring office/boss // who felt that the office/boss was boring) in the back building will quit the job.

The intern (who had studied with the terrible manual/scientist // who denied that the manual/scientist was terrible) the whole month was leaving the university.

The employer (who had criticized the old document/employee // who remarked that the document/employee was old) at the company meeting will regret the nasty comment.

The manager (who had lied about the cheap car/associate // who admitted that the car/associate was cheap) in the end will resign before the summer.

The doctor (who had worked with/on the aging instrument/patient // who had realized that the instrument/patient was aging) for many years will transfer soon.

The businessman (who was complaining about the rude delay/passenger // who complained that the delay/passenger was rude) this afternoon should call the airline.

The merchant (who had sold the creative artwork/artist // who said that the artwork/artist was creative) in the exhibition was lying about the price.

The trainer (who had criticized the health drink/consultant // who asked whether the drink/consultant was healthy) in the gym will act more professionally.

The professor (who had cited the controversial opinion/expert // who disagreed that the opinion/expert was controversial) in this new book was moving to California.

The editor (who had interviewed for the prestigious position/citizen // who recognized that the position/citizen was prestigious) yesterday was receiving a reward.

The owner (who had worked on the unusual landscape/landscaper// who liked that the landscape/landscaper was unusual) last summer was handing over the business.

The critic (who had liked the interesting painting/painter // who had thought that the painting/painter was interesting) for years will buy something.

The cashier (who had screamed about the dangerous fire/robber // who screamed that the fire/robber was dangerous) in the lobby was looking for the exit.

[Experiments 3 and 4]

The (smiling German) client who implied that the (young celebrity) visitor was important that day was waiting in the office.

The (paranoid Chicago) resident who said that the (mysterious foreign) neighbor was dangerous last month had complained about the investigation.

The (careful biology) teacher who realized that the (experienced computer) specialist was new last night will come to the office.

The (clever political) opponent who had claimed that the (honest Republican) governor was corrupt for nearly three years should be arrested immediately.

The (friendly business) manager who said that the (charismatic music) producer was clever at the opening ceremony could negotiate a good deal.

The (talented basketball) player who acknowledged that the (powerful veteran) opponent was tough before this game will make the shot.

The (funny sitcom) actor who said that the (visionary film) director was inspiring at the press conference will speak for an hour.

The (lenient physics) instructor who assumed that the (responsible sophomore) student was prepared last semester will find out the truth.

The (rude teenaged) passenger who commented that the (slow French) driver was new on the bus was talking on the phone.

The (loud research) physicist who shouted that the (ingenious lab) chemist was amazing at the conference was making too much noise.

The (biased state) attorney who commented that the (genuine character) witness was unusual in the courtroom was exaggerating quite a bit.

The (aggressive presidential) candidate who charged that the (liberal US) senator was dishonest in the newspaper was losing the race.

The (unreliable magazine) publicist who assumed that the (bizarre abstract) painter was brilliant at the first meeting will cancel the exhibit.

The (cold federal) judge who decided that the (vague adolescent) witness was questionable recently had misunderstood the facts.

The (law firm) secretary who complains that the (mean museum) director is unreasonable on TV is quitting next month.

The (impatient old) director who exclaimed that the (clumsy young) performer was outrageous in the movie had wanted to quit.

The (disorganized hotel) receptionist who knew that the (notable Canadian) client was important last time had forgotten the company's policy.

The (skilled brain) surgeon who thought that the (emergency room) patient was difficult last night complained about the procedure.

The (merciless drug) criminal who saw that the (scared old) man was hidden in the garage was lying in the bushes.

The (talkative bank) teller who felt that the (introverted American) boss was boring in the back building will quit the job.

The (new medical) intern who denied that the (disciplined resident) scientist was terrible the whole month was leaving the university.

The (arrogant factory) employer who remarked that the (easygoing production) employee was old at the company meeting will regret nasty comment.

The (frustrated company) manager who admitted that the (lazy senior) associate was cheap in the end will resign before the summer.

The (kind family) doctor who had realized that the (nursing home) patient was aging for many years will transfer soon.

The (unhappy international) businessman who complained that the (noisy airplane) passenger was rude this afternoon should call the airline.

The (cunning sales) merchant who said that the (experimental Norwegian) artist was creative in the exhibition was lying about the price.

The (moody Olympic) trainer who asked whether the (established British) consultant was healthy in the gym will act more professionally.

The (intellectual college) professor who agreed that the (unconventional Southern) author was controversial in this new book was moving to California.

The (humble newspaper) editor who recognized that the (heroic rural) citizen was prestigious yesterday was receiving a reward.

The (picky company) owner who liked that the (skilled garden) landscaper was unusual last summer was selling the business.

The (wealthy art) critic who had thought that the (British modern) painter was interesting for years will buy something.

The (terrified restaurant) cashier who screamed that the (tall armed) robber was dangerous in the lobby was looking for the exit.

Appendix B: Language background survey results of L2 English participants

	Experiment 2	**Experiment 4**
Current age (in years)	23	25
Years of L2 use	8	9
Gender of participants	Female: 21 (out of 40 participants)	Female: 28 (out of 42 participants)
Native language	Chinese (50%), Spanish (12.5%), Hindi (12.5%), Indonesian (10%), Korean (5%), Others (10%)	Chinese (59.5%), Korean (30%), Spanish (4.8%), Others (5.7%)
Mean TOEFL score	102 (out of 120)	103 (out of 120)
Mean cloze task score	30 (out of 40)	29 (out of 40)