Interference-based and capacity-based approaches to working memory in second language sentence processing

Interference-Based and Capacity-Based Approaches to Working Memory in Second Language Sentence Processing

Ian Cunnings

(University of Reading)

Abstract

The role that working memory may play in explaining potential differences between native and non-native sentence processing has been increasingly debated. In this chapter, I discuss how the conceptualisation of working memory is crucial to our understanding of its role in second language processing. In particular, I compare capacity-based approaches that focus on working memory resources and interference-based approaches that focus on memory encoding and retrieval. After reviewing evidence for both approaches to working memory, I argue that interference-based accounts provide a promising approach for examining the role of working memory in second language processing. Although I focus on non-native sentence processing, I also touch on related issues in second language acquisition.

Keywords: second language processing, second language acquisition, working memory, similarity-based interference, memory capacity

1. Introduction

There has long been consideration of the link between working memory, broadly defined as the ability to store and manipulate a limited amount of task-relevant information at one time (Baddeley, 2007; Cowan, 2017), and research in linguistics. Indeed, how grammaticality interacts with memory limitations has been debated since the early days of generative syntax (Chomsky & Miller, 1963; Hofmeister et al., 2013; Hofmeister & Sag, 2010; Sprouse, Wagers, & Phillips, 2012). One often cited example of how memory limitations interact with sentence comprehension is the decrease in acceptability observed as the number of centre-embeddings within a sentence increases, as in (1a–c).

(1) a. The cat was scared.

b. The cat the dog chased was scared.

c. The cat the dog the boy saw chased was scared.

The role that working memory plays in second language acquisition (SLA) and processing has also been widely debated (Cunnings, 2017a, 2017b; Harrington & Sawyer, 1992; Juffs & Harrington, 2011; Linck, Osthus, Koeth, & Bunting, 2014; McDonald, 2006; Wen, Mota, & McNeill, 2015). In research on native (L1) and non-native (L2) sentence comprehension, the similarities and differences between L1 and L2 processing are contested (e.g., Clahsen & Felser, 2006, 2018; Cunnings, 2017a, 2017b; Hopp, 2018, McDonald, 2006). An important question in this field has been whether putative differences between L1 and L2 processing can be neutralised when individual differences in working memory are considered (e.g., McDonald, 2006).

Crucial in examining the link between working memory and L2 sentence processing is the conceptualisation of working memory itself. In this chapter, I discuss how different ways of

conceptualising working memory lead to very different predictions about how it may modulate L2 sentence processing. I will contrast, in particular, "capacity-based" accounts of working memory during sentence processing with other "interference-based" models. Capacity-based models (Caplan & Waters, 1999; Daneman & Carpenter, 1980; Just & Carpenter, 1992) typically describe memory resource limitations during sentence processing in terms of the *amount* of information that can be stored in memory at one time. Alternatively, interference-based models (e.g., Lewis, Vasishth, & Van Dyke, 2006; McElree, 2000; McElree, Foraker, & Dyer, 2003; Van Dyke & Johns, 2012; Vasishth et al., 2019) focus instead on the *content* of information in memory, rather than the quantity per se. I argue that interference-based accounts, which focus on how information is encoded and retrieved from memory during real-time language processing, provide a well-motivated framework for comparing L1 and L2 sentence processing. Although I focus on L1 and L2 sentence processing, I also touch upon how this characterisation of working memory can help bridge research in L2 acquisition and L2 processing.

This chapter is outlined as follows. In Section 2, I discuss different accounts of working memory, with a focus on how capacity-based models and interference-based models can be applied to L2 sentence processing. Section 3 provides discussion of studies that have tested the capacity-based view, focusing on work that has used complex span tasks to measure individual differences in memory capacity, along with a discussion of how the interference-based account can be applied to research in L2 sentence processing. Section 4 provides a summary and conclusion.

2. Working Memory in L2 Sentence Processing

Although broadly defined as the ability to store and manipulate a limited amount of information, different theories conceptualise working memory in different ways. Some accounts assume a separate working memory component distinct from long-term memory. The multi-component model of working memory (e.g., Baddeley, 2007, 2015) is a highly influential model in this regard. Other accounts do not assume a separate working memory component, and instead posit that working memory is instead the activated portion of memory (e.g., Cowan, 1988, 2017). Others describe working memory in terms of executive attention, and the ability to allocate attentional resources to task-relevant information (e.g., Engle, 2002, 2018).

In research on sentence processing, these different conceptualisations of working memory make different predictions about how working memory should influence successful comprehension. Below, I discuss how these different approaches attempt to account for differences between L1 and L2 processing.

2.1 Capacity-Based Approaches to L2 Sentence Processing. Perhaps the dominant model of working memory in L2 sentence processing research is the capacity-based view (e.g., Harrington & Sawyer, 1992; McDonald, 2006). According to capacity-based accounts, individuals have a limited pool of cognitive resources that can be used to maintain task-relevant information. This view is most consistent with models that assume a separate working memory component that is distinct from long-term memory. From this perspective, the capacity of this working memory component is hypothesised to vary between individuals.

Daneman and Carpenter (1980) were amongst the first to propose that individual differences in working memory capacity are related to individual differences in language comprehension in L1 readers. Applying this to L2 learners, Harrington and Sawyer (1992)

argued that individual differences in working memory capacity predict individual differences in L2 reading. This approach predicts that working memory capacity is taxed to a greater extent during L2 processing than L1 processing, because L2 processing is slower, more effortful and generally more cognitively demanding. For example, Dekydtspotter and Renaud (2014) argued that generally slower processing in the L2 may lead to syntactic representations fading more quickly in a capacity-limited working memory. Such accounts would predict that L1 and L2 processing should be similar for L2 learners with high enough memory capacity or processing speed (e.g., McDonald, 2006).

Different approaches have described individual differences in working memory capacity in different ways. For example, Daneman and Carpenter (1980) hypothesised that sentence comprehension and other cognitive abilities utilise a shared domain-general working memory capacity, while others have argued that sentence processing has a dedicated, domain-specific pool of memory resources (e.g., Caplan & Waters, 1999). In addition to the question of whether sentence processing utilises a domain-general or domain-specific set of memory resources, another issue relates to precisely what types of information tax memory resources. Capacity-based approaches could be described in terms of the raw amount of information (i.e., words or sentences) that can be actively maintained at once, in terms of the number of different syntactic parses that may be maintained at once, or in terms of the number of different types of information sources that a reader can utilise at once (e.g., Daneman & Carpenter, 1980; Just & Carpenter, 1992; Just, Carpenter, & Keller, 1996). For example, Just and Carpenter (1992) claimed that high-capacity readers take multiple sources of syntactic and non-syntactic information into consideration when resolving syntactic ambiguity and are able to consider multiple possible parses of an ambiguous input. The hypothesis that high-capacity readers can

take into account multiple sources of information is reminiscent of research in SLA, where the Interface Hypothesis (Sorace, 2011; Sorace & Filiaci, 2006) proposes that L2 learners have difficulty integrating information between narrow syntax and other cognitive domains. In the most recent instantiation of the Interface Hypothesis, this difficulty is proposed to result from L2 difficulty in integrating multiple information sources during processing (Sorace, 2011). While L2 difficulties with anaphora resolution are well documented as evidence for this hypothesis (see Section 3.2), note that some studies have suggested that L2 learners may be *more* likely to take into consideration discourse-level information when resolving syntactic ambiguities than L1 readers (Pan & Felser, 2011; Pan, Schimke, & Felser, 2015). This is unexpected under a capacity-based view.

The discussion above highlights how different conceptualisations of memory 'capacity' lead to different predictions for L2 processing. In observing any L1/L2 differences, it is of course trivially true that L2 speakers can be said to lack some type of processing 'capacity' (in the broadest sense of the term) in comparison to L1 speakers. Precise characterisation of what is thought to tax memory 'resources' is required here to adequately test the capacity-based view of L2 processing (for critical discussion of the notion of resources, see Navon, 1984).

2.2 Interference-Based Approaches to L2 Sentence Processing. Other models of working memory, which typically do not posit a dedicated working memory component, focus on how attention is allocated to the encoding and retrieval of task-relevant information (e.g., Engle, 2002, 2018). Applying this type of model to sentence processing, successful language comprehension will involve encoding and retrieving information from memory that is relevant to the task of language comprehension.

The cue-based parsing framework has been developed as an account of these memory encoding and retrieval operations required for language comprehension (Lewis & Vasishth, 2005; Lewis et al., 2006; McElree, Foraker, & Dyer, 2003; Van Dyke & Johns, 2012; Vasishth et al., 2019). As an example of how memory encoding and retrieval is required during sentence processing, consider a filler-gap dependency as in (2a) (adapted from Cunnings & Sturt, 2018). Here, the displaced filler ('the book') needs to be interpreted as the direct object of the verb 'read,' even though they are non-adjacent (compare to 'The boy with the coffee very happily *read the book* before lunch'). According to cue-based parsing, a representation of 'the book' is encoded in memory when it is first encountered and then stored while the other constituents are processed (and themselves encoded in memory). At 'read', a representation of 'the book' needs to be retrieved from memory to correctly interpret the sentence.

(2) a. It was the book that the boy with the coffee very happily read before lunch.

b. It was the book that the boy with the magazine very happily read before lunch.

Memory retrieval within this framework is achieved via a direct-access memory retrieval mechanism that compares a set of retrieval cues against all items in memory in parallel, with the best matching item being retrieved (e.g., McElree, Foraker, & Dyer, 2003). In sentences like (2a), this will likely consist of a series of syntactic and semantic cues. For example, a syntactic cue, or set of cues, will be required to guide retrieval to the correct constituent in the syntactic structure. For simplicity sake, I refer to this here as [+OBJECT]. A semantic cue, or cues, will also guide retrieval to a constituent that can plausibly act as the direct object of the verb. In this case, and again simplifying somewhat, this might be [+READABLE]. Upon encountering the

verb, items in the sentence that match these retrieval cues become activated, and the item that becomes most activated is retrieved. In (2a), these combined cues uniquely identify 'the book' as the target for memory retrieval, which will become most activated. However, in (2b), although 'the book' provides the best match to this set of retrieval cues, 'the magazine' provides a partial match, as it matches the semantic but not syntactic cue and as such will become partially activated. Cue-based parsing predicts that this *distractor* constituent may be retrieved a proportion of the time, as a result of its partial match. This is known as similarity-based interference (Lewis, Vasishth, & Van Dyke, 2006; Van Dyke & Johns, 2012; Vasishth et al., 2019), as successful comprehension is dependent on the number of items that match a set of retrieval cues.

Different types of sentences will pose different challenges from this interference-based perspective. While syntactic and semantic retrieval cues may combine to resolve filler-gap dependencies, other types of retrieval cues will be required elsewhere. For example, anaphora resolution may involve cues related to a pronoun's number and gender or related to discourse prominence (see Section 3.2). Although different types of retrieval cues will thus be required for successful comprehension in different circumstances, the underlying prediction of this interference-based approach is that successful comprehension is dependent on how a reader or listener is able to encode and retrieve task-relevant information.

As there are different instantiations of the capacity-based view of sentence processing, there is also debate regarding the precise characterisation of interference-based accounts (for discussion, see Parker, Shvartsman, and Van Dyke, 2017; Vasishth et al., 2019). I do not attempt to examine this issue here. For present purposes, the main point is that these accounts all predict similarity-based interference to be a function of the similarity between a set of retrieval cues and

the number of constituents that match, or partially match, these cues. Thus, unlike the capacity-based view, which focuses on the raw *quantity* of information that needs to be maintained, interference-based accounts focus on the *quality* and *content* of information in memory as the primary determinant of successful comprehension (Van Dyke & Johns, 2012).

L1/L2 differences from this perspective would be described in terms of how L1 speakers and L2 learners encode and retrieve information from memory during processing (Cunnings, 2017a, 2017b). Cunnings proposed that L1/L2 differences can at least in part be explained in terms of how native and non-native speakers utilise different cues to guide memory retrieval. For example, if an L2 learner were to weigh the semantic cue over the syntactic cue more heavily than L1 readers in (2b), L2 readers may be more prone to interference from the distractor (‘the magazine’).

In teasing apart capacity-based and interference-based accounts of working memory in sentence processing, it is important to consider that differences in capacity and interference are often confounded (Van Dyke & Johns, 2012). For example, returning to (1), in (1a) the predicate ‘was scared’ is adjacent to the sentence subject (‘the cat’), while in (1c) they are separated by several other constituents. The capacity-based view would explain the comparative difficulty of (1c) in terms of a reader’s limited memory capacity being exceeded. However, as there are more potentially interfering constituents in (1c) than (1a), such that ‘the dog’ and ‘the boy’ could be subjects of ‘was scared’, cue-based retrieval can also explain the difficulty in (1c) in terms of interference. Thus, the two accounts both predict difficulty in (1c), but for very different reasons, and teasing apart the capacity-based and interference-based view can be difficult, as increasing memory load, like in longer dependencies, also increases interference. However, it is less clear

how a capacity-based view could explain differences between sentences like (2a) and (2b), as the number of words in each sentence is the same; it is the *content* of the sentences that differs.

Similar issues arise in disentangling capacity-based and interference-based approaches to working memory in L2 processing. Although some studies have been taken as evidence of capacity-based L2 limitations, whether these studies provide veridical evidence for capacity-based accounts of L2 processing over interference-based models is less clear. I argue that the interference-based account provides a well-specified framework that explicitly ties working memory operations with real-time sentence processing, and a fruitful way of conceptualising individual differences in L1 and L2 sentence processing.

3. Characterising and Measuring L2 Individual Differences

Important to the capacity-based view is the measurement of individual differences in memory capacity. Measuring individual differences in working memory is often achieved using one or more complex span task, which involve a participant remembering a list of items while completing a secondary task (see e.g., Conway et al., 2005; Mathy, Chekaf, & Cowan, 2018). A commonly used complex span task in L2 processing research is reading span, first developed by Daneman and Carpenter (1980), and adapted for L2 learners by Harrington and Sawyer (1992). In this task, participants must remember a series of words for subsequent recall while reading a series of sentences (which are often judged for grammaticality or plausibility). From a capacity-based view, this task is taken to measure the amount of information an individual is able to maintain at one time, with higher span readers presumably being able to hold more information in memory at once than lower span readers. From the perspective of L1/L2 differences during sentence processing, a higher span would suggest increased memory capacity and more cognitive resources that presumably can be utilised during L2 processing. As such, higher span L2 learners

may behave more nativelike. Below, I discuss how complex span tasks have been used to test this capacity-based view.

3.1 Memory Capacity in L2 Processing. Studies have examined the correlation between performance on complex span tasks and L2 sentence processing. Here, however, the results have been mixed. Some studies have tested the relationship between span scores and the processing of filler-gap dependencies (Dallas, DeDe, & Nicol, 2013; Dussias & Piñar, 2010; Felser & Roberts, 2007; Juffs, 2005; Miller, 2014). Dallas, DeDe, and Nicol (2013) examined sentences like (3), manipulating a displaced filler to be either a plausible ('which player') or implausible ('which football') direct object of the critical verb ('threatened'), using an event-related potential (ERP) paradigm. They expected implausible sentences to yield an increased N400 component (Kutas & Hillyard, 1980) compared to plausible sentences.

(3) The umpire asked which player/football the coach threatened before the game.

While L1 readers displayed the expected N400 effect, a group of Chinese L2 English learners did not. However, an individual differences analysis suggested that L2 learners with high English proficiency did show an N400. A composite score of memory span did not however correlate with the L2 learners' ERP responses. Other studies testing filler-gap dependencies using different paradigms also failed to observe an effect of memory span (Felser & Roberts, 2007; Juffs, 2005; Miller, 2014). One counterexample is Dussias and Piñar (2010), who examined filler-gap dependencies in a reading study. Here, high span L2 learners behaved nativelike, while low span L2 learners did not. The effects that Dussias and Piñar observed, however, were related to how filler-gap dependencies may be revised during processing, rather than how they are initially resolved. For example, in (4), readers may initially interpret 'who' to be the direct object of 'know' when it is in fact the subject of 'killed' and reading span was

found to correlate with this aspect of processing in the L2 group. This might suggest a particular role of memory capacity in reanalysis during L2 processing.

(4) Who did the policeman know killed the pedestrian?

This potential interpretation of these findings is complicated by the fact that studies examining other types of syntactic ambiguity have not found significant correlations between memory capacity and reanalysis during L2 processing (Juffs, 2004). Other studies have reported effects of memory span in some measures of ambiguity resolution but not others. For example, Hopp (2014) found significant correlations between reading span and L2 performance in an offline task on ambiguous relative clause attachment, but did not find significant correlations between reading span and relative clause attachment in an eye-tracking during reading task that examined online sentence processing.

Correlations between span scores and sentence processing have also been investigated in work examining morphosyntactic agreement. There is a long line of research in L2 processing of agreement phenomena, with some studies suggesting nativelike performance and others not (for review, see Cunnings, 2017a). Some researchers have examined the extent to which working memory may explain this discrepancy. Sagarra and Herschensohn (2010) tested Spanish sentences like (5), manipulating the extent to which an adjective (‘*famoso*’) agreed in either number or gender with a preceding noun (‘*prototipo*’).

(5) *El ingeniero presenta el prototipo famoso/famosa/famosos en la conferencia*.
‘The engineer presented the prototype$_{\text{-MASC-SING}}$ famous$_{\text{-MASC-SING}}$ / famous$_{\text{-FEM-SING}}$ / famous$_{\text{-MASC-PLURAL}}$ at the conference.’

Spanish L1 speakers and intermediate but not beginner L2 learners showed longer reading times for ungrammatical sentences in which the adjective mismatched in either gender or

number with the preceding noun. Additionally, the size of the grammaticality effect was positively correlated with reading span scores in the intermediate L2 speakers for gender but not number agreement.

Coughlin and Tremblay (2013) investigated shorter and longer agreement dependencies with French clitics, to test the hypothesis that longer dependencies may pose a larger burden on working memory capacity. They tested sentences as in (6). In (6a/b), the clitic ('*le*'/'*les*') must agree in number with the sentence subject ('*Ce fruit*'/'*Ces fruits*'). In (6a) the distance between the clitic and subject is short, while in (6b) it is longer, due to an intervening constituent ('*avant l'entretien*').

(6) a. *Ce fruit/Ces fruits Marie le(s) mangera pour sa collation avant l'entretien*.
'The fruit/The fruits Marie it (them) will eat for her snack before the interview.'

b. *Ce fruit/Ces fruits avant l'entretien Marie le(s) mangera pour sa collation*.
'The fruit/The fruits before the interview Marie it (them) will eat for her snack.'

In a self-paced reading experiment, high proficiency but not intermediate proficiency L2 learners exhibited longer reading times following ungrammatical clitics in both short and long conditions. Numerically, the effect was larger in the shorter conditions. There was also a numerical tendency for higher reading span scores to correlate with larger grammaticality effects for plural clitics in the short condition, and for singular clitics in the long condition, though the relevant effects here were not statistically significant ($p = .083$ and .068 respectively). While this might be suggestive of an effect of reading span on L2 processing, these effects are not in the

predicted direction to be consistent with the hypothesis that longer dependencies impose a greater burden on memory capacity.

Foote (2011) investigated subject-verb number agreement and noun-adjective gender agreement in L2 learners of Spanish. The length of the dependency was also manipulated by including conditions with adjacent constituents, and a long condition in which another constituent intervened, as exemplified in (7a) and (7b) respectively. Both Spanish L1 speakers and L2 speakers exhibited longer reading times for sentences with ungrammatical compared to grammatical agreement, and the size of the grammaticality effect was smaller in the long conditions for both types of agreement. Reading span scores were however not found to correlate with grammaticality effects in this study.

(7) a. *Dicen que el libro blanco/blanca está en esa mesa*.

‘They say that the book$_{\text{-MASC}}$ white$_{\text{-MASC}}$ / white$_{\text{-FEM}}$ is on that table.’

(7) b. *El pollo del taco está rico/rica pero picante.*

‘The chicken$_{\text{-MASC}}$ of the taco is tasty$_{\text{-MASC}}$ / tasty$_{\text{-FEM}}$ but spicy.’

As such, while some studies have reported effects of memory span on the processing of gender agreement, these have not been consistently observed. Other studies have also reported similar effects of distance, however. For example, Keating (2010) examined noun-adjective gender agreement in Spanish, and found that while L1 readers showed grammaticality effects irrespective of distance, L2 learners showed grammaticality effects in short conditions only. Such results might be taken as supporting a capacity-based view of L2 processing, with L2 learners showing less sensitivity to agreement violations as more information needs to be maintained in memory. Here, however, dependency length and interference are confounded. In the studies reported by Coughlin and Tremblay (2013), Foote (2011) and Keating (2010), the

intervening constituents in the long conditions could all cause interference in computing the agreement dependency. For example, in (7b) the gender of the distractor constituent (‘*taco*’) may interfere when it matches the gender of the manipulated adjective (‘*rico*’). To fully tease apart the capacity-based view from interference, the morphological properties of the intervening constituent would need to be systematically manipulated (for further discussion, see Cunnings, 2017a: 667–668). Interference in agreement in L1 processing is well attested (e.g., Lago et al., 2015; Wagers, Lau, & Phillips, 2009), and further research is required here in L2 processing (see Tanner, Herschensohn, & Osterhout, 2012, for some preliminary evidence).

In sum, the evidence that individual differences in L2 sentence processing correlate with performance on complex span tasks is mixed. Although some studies have reported correlations between memory span and L2 processing, others have not. Additionally, some of the existing evidence is not able to adjudicate between capacity-based and interference-based accounts of working memory. Further research is needed to not only test the extent to which reading span scores correlate with L2 sentence processing, but to also tease apart effects of L2 memory capacity from interference.

3.2 Memory Interference in L2 Processing. As discussed above, from the perspective of cue-based parsing and interference-based accounts of working memory, individual differences in sentence processing relate to how information is encoded and retrieved from memory. This perspective, thus, focuses on the *quality* and *content* of representations in memory, rather than the *quantity* per se.

That the quality of memory representations may explain individual variation even in L1 sentence processing is exemplified by Van Dyke, Johns, and Kukona (2014). Van Dyke et al. examined filler-gap dependencies like (8), where ‘*the boat*’ needs to be retrieved from memory

and interpreted as the direct object of the main verb (‘fixed’/‘sailed’). Participants were tested in a ‘no memory load’ condition in which they simply read the sentences, and a ‘memory load’ condition in which they also had to remember a list of words (‘table’, ‘sink’, ‘truck’). The crucial point here being that the list of words in the memory load condition may interfere in resolving the filler-gap dependency in (8b) but not (8a), as in (8b) only the word list contains words that are a possible theme of the main verb (‘fixed’).

(8) a. It was the boat that the guy who lived by the sea sailed in two sunny days.

b. It was the boat that the guy who lived by the sea fixed in two sunny days.

Participants also completed a battery of individual differences measures, including IQ, reading span and measures of reading/listening comprehension. After partialling out shared variance with IQ, Van Dyke et al. found that reading times at the critical verb in (8b), in the memory load condition, correlated with measures of receptive vocabulary but not reading span. They took this as indicating that the quality of memory representations influences sentence processing, rather than memory capacity. Specifically, they argued that individuals with lower scores of receptive vocabulary may have lower quality memory representations that are more prone to activating irrelevant information. These lower quality representations then lead to higher amounts of interference during memory retrieval during sentence processing.

Although not couched within this framework of memory encoding and retrieval, some existing research in the L2 processing literature is also potentially consistent with the broad idea that the quality of memory representations is important for successful comprehension. Hopp (2018) proposed the Lexical Bottleneck Hypothesis. This predicts that L1/L2 differences are related to lexical knowledge, and that L2 processing can become nativelike if L2 learners have

robust lexical entries. Consider Hopp (2014), who conducted an eye-tracking during reading study on relative clause attachment in German learners of L2 English. Hopp tested sentences like (9), where the reflexive disambiguates interpretation of the relative clause to either a *local* or *non-local* noun phrase ('the boy' and 'the mother' respectively).

(9) The doctor examined the mother of the boy who had badly injured herself/himself with the knife.

L1 readers had longer reading times when the reflexive disambiguated towards the non-local noun phrase, replicating the long-standing observation of a local attachment preference in English (e.g., Cuetos & Mitchell, 1988). As a group, the L2 learners did not show any differences between conditions, replicating previous results showing no clear L2 attachment preferences during processing (Felser et al., 2003; but see Witzel, Witzel, & Nicol, 2012). However, a measure of lexical automaticity interacted with L2 processing, such that L2 learners with high levels of lexical automaticity showed a nativelike local attachment preference (see also Hopp, 2013, and Miller, 2014, for evidence of the role of lexical processing during L2 sentence processing).

Although suggestive of the importance of lexical representations in L2 processing, more research is required to test the extent to which L1/L2 differences can be reduced to differences in the quality of lexical representations. A key prediction here would be to test whether L1 processing is influenced by the same lexical factors as L2 processing. Such a finding would provide strong evidence of similarities between the relevant aspects of L1 and L2 processing once differences in lexical processing are accounted for. Direct L1/L2 comparisons in this regard are currently scant, however. Irrespective of this need for further research, the hypothesis that the quality of lexical representations is important for L2 sentence processing may be broadly in-line

with models of working memory that focus on the quality, rather than quantity, of information that needs to be encoded and retrieved from memory during comprehension.[1]

Given the importance of memory retrieval in cue-based parsing, another key predicted source of individual variation from the perspective of interference-based accounts is the set of retrieval cues that are utilised during sentence processing. Individual variation in the relative weightings of retrieval cues could give rise to individual differences in interference (Vasishth et al., 2019). For L2 processing, Cunnings (2017a) proposed that L2 learners may weight retrieval cues differently to L1 speakers, and argued in particular that L2 learners may weight discourse-based cues comparatively more highly than L1 speakers (see also Felser, 2016). It is beyond the scope of this chapter to provide an exhaustive review of how to apply this approach to different phenomena in the L2 sentence processing literature. Instead, I discuss below some recent studies on anaphora resolution to exemplify how this interference-based approach can be utilised to compare L1 and L2 sentence processing.

Felser and Cunnings (2012) investigated sentences like (10) in an eye-tracking during reading experiment with L1 English speakers and German learners of L2 English. They manipulated the extent to which a reflexive matched or mismatched in gender with an 'accessible' antecedent ('the soldier') and an 'inaccessible' antecedent ('James/Helen'). According to binding theory (Chomsky, 1981), only the accessible antecedent can be coindexed with the reflexive. From the perspective of cue-based parsing, anaphora resolution involves retrieving an antecedent from memory when the reflexive is encountered. Retrieval is likely to be

[1] The role of L2 lexical representations could potentially be described from a capacity-based view if greater lexical automaticity frees up processing resources for other aspects of sentence processing. Here, however, we require precise characterisations about how different resources are hypothesised to be allocated during processing (e.g., how are resources allocated to lexical access and syntactic parsing?). Precisely how efficient lexical processing 'frees up' resources for other aspects of sentence processing is currently not well specified, but precise characterisation of this issue is required to adequately test the capacity-based view.

guided by structural cues, that favour a syntactically local antecedent (e.g., [+LOCAL]), along with features of the reflexive (e.g., [+SINGULAR], [+MASCULINE] etc.), and potentially discourse-related factors (e.g., [+TOPIC]).

(10) a. James/Helen has worked at the army hospital for years. She/He noticed that the soldier had wounded himself while on duty in the Far East.

b. James/Helen has worked at the army hospital for years. She/He noticed that the soldier had wounded herself while on duty in the Far East.

Similar to some L1 studies (Dillon et al., 2013; Sturt, 2003), Felser and Cunnings (2012) found that L1 speaker's reading times were longer in (10b) than (10a), when the accessible antecedent mismatched in gender with the reflexive, suggesting rapid application of binding constraints. During first-pass processing of the reflexive, the L2 learners showed a different pattern and had longer reading times when the inaccessible antecedent mismatched the gender of the reflexive. L2 learners did show effects of the gender of the accessible antecedent in later measures of processing, such as rereading time, and evidenced knowledge of binding constraints in an offline task. The results however suggest that during initial processing, L2 learners weighted discourse-based cues over the syntactic requirement for a local antecedent. This finding might be in particular unexpected from a capacity-based view of L2 processing. If L2 processing was capacity limited, we might expect an even greater focus on the local antecedent in L2 processing than L1 processing, but this was not found (see Felser, 2019, for further discussion).

It should be noted that the extent to which anaphora resolution is susceptible to interference in L1 processing is widely debated. While some studies suggest L1 readers weight structural cues more highly than non-structural cues (Dillon et al., 2013), it is too strong to claim that L1 anaphora resolution is impervious to interference (Jäger et al., 2020). Further L1/L2

comparisons are required here to tease apart the relative weightings of different retrieval cues by L1 and L2 readers.

The claim that L2 learners weight discourse-based cues more heavily than L1 speakers can also potentially provide an alternative account of the long-debated controversy in SLA research on null and overt pronouns. In null subject languages, null pronouns typically refer to the current discourse topic, while overt pronouns index a topic-shift to a less prominent antecedent. Studies have shown that L2 speakers interpret null pronouns in a similar way to L1 speakers, but additionally allow overt pronouns to refer to topic antecedents, rather than indexing a topic shift (e.g., Belletti, Bennati, & Sorace, 2007; Sorace & Filiaci, 2006). These results motivated the Interface Hypothesis (Sorace, 2011; Sorace & Filiaci, 2006), which states that L2 learners have difficulty integrating information between syntax and other cognitive domains. However, these results are also consistent with the claim that L2 speakers weight discourse-based cues that favour prominent antecedents more heavily than L1 speakers.

Although overt pronouns may cause difficulty for L2 learners in a null subject L2, they do not always appear to cause difficulty for learners of a non-null subject L2, where overt pronouns refer to the current discourse topic. Cunnings, Fotiadou, and Tsimpli (2017) investigated the processing of overt pronouns as in (11) using the visual world paradigm. They found that upon hearing the pronoun, both L1 and L2 listeners' gaze across a visual display quickly focused on the discourse prominent subject antecedent ('Mr Smith') rather than the object ('Peter'). This was despite the fact that the L2 learners' L1 was a null subject language in which overt pronouns index a topic shift. This finding would be unexpected under a capacity-based view, which may favour the linearly closer object antecedent ('Peter'). However, it is

compatible with the hypothesis that L2 learners strongly weight discourse-based cues to anaphora resolution.

(11) After Mr Smith spoke to Peter by the till in the shop, he paid for the expensive ice cream that looked tasty.

Although Cunnings (2017a) emphasised the relative weighting of discourse-based cues in L2 processing, further research is required to test the extent to which this claim generalises to other linguistic phenomena. A focus on discourse-based cues may be pertinent for anaphora, where discourse prominence is a key factor in determining coreference, but whether this is the case for other aspects of sentence processing is a matter for future research. For example, it might be that discourse prominence plays less of a role in the processing of morphosyntactic agreement, but L1/L2 differences may still arise if L1 speakers and L2 learners weigh morphological and syntactic cues differently to guide retrieval. What is required here is careful consideration and hypotheses about what retrieval cues a particular group of L2 learners may utilise. The hypothesis that L2 learners may weight discourse-based cues more highly than L1 speakers is one starting point to examine this issue, but further research is required to elucidate how L1 speakers and L2 learners weigh different retrieval cues during processing.

3.3 Retrieval Cues in Processing and Acquisition. This account of working memory, that focuses on the quality of representations in memory and the cues utilised to guide memory retrieval, would predict that instead of describing individual variation in L2 processing in terms of memory capacity, the primary predictor of individual differences in both L1 and L2 processing should be related to linguistic experience. For example, robust lexical representations require adequate linguistic experience of the relevant vocabulary. Additionally, the relative weightings of different retrieval cues by L1 speakers and L2 learners would result from the

differing degrees of linguistic exposure to the relevant cues. From this perspective, L2 acquisition can in part be considered a process of learning to weight the relevant retrieval cues in an appropriate way to guide successful memory retrieval during sentence processing.

This focus on a learners' linguistic experience echoes long-debated questions in the L2 acquisition literature on the role of proficiency and L1 transfer. For example, one might hypothesise that for reflexives in sentences like (10), L2 learners whose L1 allows long-distance reflexives (e.g., Japanese or Chinese) might weight retrieval cues differently to L2 learners whose first language requires local binding as in English. Note that the first language of the L2 leaners tested by Felser and Cunnings (2012), German, has binding constraints similar to English. However, these learners still seemed to weight retrieval cues to initially favour the discourse prominent non-local antecedent. Felser, Sato, and Bertenshaw (2009) reported that Japanese L2 learners also temporarily violate binding constraints on reflexives during processing, though they tested slightly different stimuli to Felser and Cunnings. Direct comparisons between L2 learners whose L1 allows local binding compared to those that allow non-local binding are required here.

Transfer might also play a role in resolving pronouns as in (11). Although the results of Cunnings et al. (2017) are not suggestive of transfer, which would have favoured less prominent 'Peter' as the antecedent of the pronoun over 'Mr Smith', other studies have suggested a role of L1 transfer in processing and interpreting overt pronouns in a non-null subject L2 (Contemori, Asiri, & Perea Irigoyen, 2019; Roberts, Gullberg, & Indefrey, 2008). Further research is required to examine the extent to which an L2 learners' first language may influence how they weight different cues to memory retrieval during L2 processing.

3.4 Task Performance and Characterising Individual Differences. Although measures such as reading span are often assumed to be measures of working memory capacity, the precise characterisation of what such tasks measure is debated. For example, it has long been questioned whether reading span is a measure of memory capacity or linguistic experience (e.g., Kidd, Donnelly, & Christiansen, 2018; MacDonald & Christiansen, 2002). Indeed, Farmer et al. (2017) recently demonstrated that linguistic experience influences reading span performance in L1 speakers. The extent to which L2 performance on the reading span test is related to memory capacity or L2 experience has also been debated (e.g., Juffs & Harrington, 2011).

One way to potentially overcome this confound with L2 proficiency is to test L2 learners on the reading span test in both their L1 and L2 and calculate a composite reading span score from both tests (Juffs & Harrington, 2011). Another possibility is to use a complex span task, such as operation span (Turner & Engle, 1989), which is less dependent on language proficiency. More broadly, a number of researchers have advocated computing composite scores across several complex span tasks rather than testing a single measure (e.g., Waters & Caplan, 2003). Additionally, the way that span tasks are administered and scored can also influence the extent to which they correlate with other tasks (Leeser & Sunderman, 2016). To date, these issues have not been systematically examined in L2 sentence processing. Further research is required here to more stringently test the role that span tasks and memory capacity may play in L2 processing.

It is unlikely that a single construct related to working memory and/or language experience can explain all L1/L2 differences in sentence processing, and different aspects of language comprehension may correlate with different individual differences. For example, although individual differences in memory encoding and retrieval may be relevant when examining linguistic dependencies, a different set of memory operations are likely required for

other aspects of comprehension, such as the reanalysis processes involved in resolving temporary syntactic ambiguity. Consider (12), where the reader may initially interpret ‘the baby’ as the direct object of the subordinate clause verb (‘dressed’) when it is in fact the subject of the main verb. Even L1 readers are known to have difficulty with such sentences, sometimes maintaining the initially assigned interpretation (‘Anna dressed the baby’) after reanalysis (Christianson et al., 2001). Recent studies have indicated that reanalysis causes particular difficulty for L2 learners (Jacob & Felser, 2016; Pozzan & Trueswell, 2016; for in-depth discussion, see Fujita, 2019).

(12) After Anna dressed the baby played in the cot.

Recent research in L1 processing has eschewed memory capacity as a predictor of reanalysis success in sentences like (12) in favour of cognitive control, the ability to switch attention during cases of information conflict (Kan et al., 2013; Novick et al., 2014). Woodard, Pozzan, and Trueswell (2016) recently provided evidence that individual differences on cognitive control tasks correlate with reanalysis ability during L1 language development, but how cognitive control may influence reanalysis in L2 processing has not yet been systematically examined.

In sum, further research is required to tease apart the capacity-based and interference-based accounts of L2 processing. In testing purported measures of memory capacity, it is important to tease apart capacity from language experience, and to consider how different aspects of sentence processing rely on different memory operations.

4. Conclusion

While the interaction between sentence processing and working memory has long been debated, recent approaches to working memory in language comprehension that focus on memory encoding and retrieval allow for greater integration between work in second language acquisition and non-native sentence processing. While capacity-based views of working memory may have been dominant in research in L2 processing, I argue that a clearer focus on the memory encoding and retrieval operations required during real-time sentence processing would allow for new insight into key debate in L2 acquisition and processing, and a better understanding of the similarities and differences between L1 and L2 processing.

References

Baddeley, A. (2007). *Working Memory, Thought, and Action*. Oxford, UK: Oxford University Press.

Baddeley, A. (2015). Working memory in second language learning. In Zhisheng, W., Morta, M. & McNeil, A. (Eds). *Working memory and second language acquisition and processing: Theories, research and commentaries* (pp. 17–28). Bristol, UK: Multilingual Matters.

Belletti, A., Bennati, E., & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language & Linguistic Theory*, *25*(4), 657–689.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*(1), 77–94.

Chomsky, N. (1981). *Lectures on government and binding: The Pisa lectures* (7th ed). Berlin ; New York: Mouton de Gruyter.

Chomsky, N., & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In R. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology, Volume 2* (pp. 269–323). New York: Wiley.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, *42*(4), 368–407.

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

Clahsen, H., & Felser, C. (2018). Some notes on the Shallow Structure Hypothesis. *Studies in Second Language Acquisition, 40,* 693–706.

Contemori, C., Asiri, O., & Perea Irigoyen, E. D. (2019). Anaphora resolution in L2 English: An analysis of discourse complexity and cross-linguistic interference. *Studies in Second Language Acquisition*, *41*(5), 971–998.

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 & Review*, *12*(5), 769–786.

Coughlin, C. E., & Tremblay, A. (2013). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics*, *34*(3), 615–646.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104,* 163–191.

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170.

Cuetos, F., & Mitchell, D. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition*, *30*(1), 73–105.

Cunnings, I. (2017a). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, *20*(4), 659–678.

Cunnings, I. (2017b). Interference in native and non-native sentence processing. *Bilingualism: Language and Cognition*, *20*(4), 712–721.

Cunnings, I., Fotiadou, G., & Tsimpli, I. (2017). Anaphora resolution and reanalysis during L2 sentence processing: Evidence from the visual world paradigm. *Studies in Second Language Acquisition*, *39*(4), 621–652.

Cunnings, I., & Sturt, P. (2018). Retrieval interference and semantic interpretation. *Journal of Memory and Language*, *102*, 16–27.

Dallas, A., DeDe, G., & Nicol, J. (2013). An event-related potential (ERP) investigation of filler-gap processing in native and second language speakers. *Language Learning*, *63*(4), 766–799.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

Dekydtspotter, L., & Renaud, C. (2014). On second language processing and grammatical development: The parser in second language acquisition. *Linguistic Approaches to Bilingualism, 4*, 131–165.

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*(2), 85–103.

Dussias, P. E., & Piñar, P. (2010). Effects of reading span and plausibility in the reanalysis of wh-gaps by Chinese-English second language speakers. *Second Language Research*, *26*(4), 443–472.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23.

Engle, R. W. (2018). Working memory as executive attention: A revisit. *Perspectives on Psychological Science*, *12*, 190–193.

Farmer, T. A., Fine, A. B., Misyak, J. B., & Christiansen, M. H. (2017). Reading span task performance, linguistic experience, and the processing of unexpected syntactic events. *Quarterly Journal of Experimental Psychology*, *70*(3), 413–433.

Felser, C. (2016). Binding and coreference in non-native language processing. In A. Holler & K. Suckow (Eds.), *Empirical perspectives on anaphora resolution.* (pp. 229–253). Berlin: De Gruyter.

Felser, C. (2019). Structure-sensitive constraints in non-native sentence processing. *Journal of the European Second Language Association*, *3*(1), 12–22.

Felser, C., & Cunnings, I. (2012). Processing reflexives in a second language: The timing of structural and discourse-level constraints. *Applied Psycholinguistics*, *33*(3), 571–603.

Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research*, *23*(1), 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*(3), 453–489.

Felser, C., Sato, M., & Bertenshaw, N. (2009). The on-line application of binding Principle A in English as a second language. *Bilingualism: Language and Cognition*, *12*(4), 485–502.

Foote, R. (2011). Integrated knowledge of agreement in early and late English–Spanish bilinguals. *Applied Psycholinguistics*, *32*(1), 187–220.

Fujita, H. (2019). *Reanalysis processes in native and non-native language comprehension* (Unpublished Doctoral Thesis). University of Reading.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*(1), 25–38.

Hofmeister, P., Jaeger, T. F., Arnon, I., Sag, I. A., & Snider, N. (2013). The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes*, *28*(1–2), 48–87.

Hofmeister, P., & Sag, I. A. (2010). Cognitive constraints and island effects. *Language*, *86*(2), 366–415.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33–56.

Hopp, H. (2014). Working memory effects in the L2 processing of ambiguous relative clauses. *Language Acquisition*, *21*(3), 250–278.

Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Language*, *17*, 5–27.

Jacob, G., & Felser, C. (2016). Reanalysis and semantic persistence in native and non-native garden-path recovery. *Quarterly Journal of Experimental Psychology*, *69*(5), 907–925.

Jäger, L. A., Mertzen, D., Van Dyke, J. A., & Vasishth, S. (2020). Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study. *Journal of Memory and Language*, *111*, 104063.

Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society*, *102*(2), 199–225.

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, *21*(2), 121–151.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, *44*(2), 137–166.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 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*(4), 773–780.

Kan, I. P., Teubner-Rhodes, S., Drummey, A. B., Nutile, L., Krupa, L., & Novick, J. M. (2013). To adapt or not to adapt: The question of domain-general cognitive control. *Cognition*, *129*(3), 637–651.

Keating, G. D. (2010). The effects of linear distance and working memory on the processing of gender agreement in Spanish. In B. VanPatten & J. Jegerski (Eds.), *Language Acquisition and Language Disorders*, *53*, 113–134.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, *22*(2), 154–169.

Kutas, M., & Hillyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & Phillips, C. (2015). Agreement attraction in Spanish comprehension. *Journal of Memory and Language*, *82*, 133–149.

Leeser, M., & Sunderman, G. (2016). Methodological implications of working memory tasks for L2 processing research. In G. Granena, D. Jackson, & Y. Yilmaz (Eds.), *Cognitive Individual Differences in Second Language Processing and Acquisition* (pp. 89–104). Philadelphia: John Benjamins Publishing Company.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419.

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

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*(4), 861–883.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35–54.

Mathy, F., Chekaf, M., & Cowan, N. (2018). Simple and complex working memory tasks allow similar benefits of information compression. *Journal of Cognition*, *1*(1), 31.

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*(3), 381–401.

McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, *29*(2), 111–123.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*(1), 67–91.

Miller, K. (2014). Accessing and maintaining referents in L2 processing of wh-dependencies. *Linguistic Approaches to Bilingualism*, *4*, 167–191.

Navon, D. (1984). Resources - A theoretical soup stone? *Psychological Review*, *91*, 216–234.

Novick, J. M., Hussey, E., Teubner-Rhodes, S., Harbison, J. I., & Bunting, M. F. (2014). Clearing the garden-path: Improving sentence processing through cognitive control training. *Language, Cognition and Neuroscience*, *29*(2), 186–217.

Pan, H.-Y., & Felser, C. (2011). Referential context effects in L2 ambiguity resolution: Evidence from self-paced reading. *Lingua*, *121*(2), 221–236.

Pan, H.-Y., Schimke, S., & Felser, C. (2015). Referential context effects in non-native relative clause ambiguity resolution. *International Journal of Bilingualism*, *19*(3), 298–313.

Parker, D., Shvartsman, M., & Van Dyke, J. A. (2017). The cue-based based retrieval theory of sentence comprehension: New findings and new challenges. In L. Escobar, V. Torrens, & T. Parodi (Eds.), *Language Processing and Disorders* (pp. 121–144). Newcastle: Cambridge Scholars Publishing.

Pozzan, L., & Trueswell, J. C. (2016). Second language processing and revision of garden-path sentences: A visual word study. *Bilingualism: Language and Cognition*, *19*(3), 636–643.

Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 Influence and General Learner Effects. *Studies in Second Language Acquisition*, *30*(3), 333–357.

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, *120*(8), 2022–2039.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–33.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, *22*(3), 339–368.

Sprouse, J., Wagers, M., & Phillips, C. (2012). A test of the relation between working-memory capacity and syntactic island effects. *Language*, *88*(1), 82–123.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, *48*(3), 542–562.

Tanner, D., Herschensohn, J., & Osterhout, L. (2012). Electrophysiological markers of interference and structural facilitation in native and nonnative agreement processing. In A. Biller, A. Chung, & A. Kimball (Eds.), *Proceedings of the 36th Boston University Conference on Language Development* (pp. 294–606). Somerville: Cascadilla.

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

Van Dyke, J. A., & Johns, C. L. (2012). Memory interference as a determinant of language comprehension: Interference in comprehension. *Language and Linguistics Compass, 6*(4), 193–211.

Van Dyke, J. A., Johns, C. L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition, 131*(3), 373–403.

Vasishth, S., Nicenboim, B., Engelmann, F., & Burchert, F. (2019). Computational models of retrieval processes in sentence processing. *Trends in Cognitive Sciences, 23*(11), 968–982.

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language, 61*(2), 206–237.

Waters, G. & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers, 35*, 550–564.

Wen, Z., Mota, M. B., & McNeill, A. (Eds.). (2015). *Working memory in second language acquisition and processing*. Bristol ; Buffalo: Multilingual Matters.

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.

Woodard, K., Pozzan, L., & Trueswell, J. C. (2016). Taking your own path: Individual differences in executive function and language processing skills in child learners. *Journal of Experimental Child Psychology, 141*, 187–209.