University of Reading

Parsing and working memory in bilingual sentence processing

Article

Accepted Version

Cunnings, I. ORCID: https://orcid.org/0000-0002-5318-0186 (2017) Parsing and working memory in bilingual sentence processing. Bilingualism: Language and Cognition, 20 (4). pp. 659-678. ISSN 1469-1841 doi: 10.1017/S1366728916000675 Available at https://centaur.reading.ac.uk/65915/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=10368492&fulltextType=RA&fileId=S1366728916000675
To link to this article DOI: http://dx.doi.org/10.1017/S1366728916000675

Publisher: Cambridge University Press

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

University of Reading

Reading's research outputs online

Running Head: Bilingual Sentence Processing

Parsing and Working Memory in Bilingual Sentence Processing

Ian Cunnings
University of Reading, UK

Address for correspondence:

Dr Ian Cunnings
School of Psychology and Clinical Language Sciences
University of Reading
Reading, UK
RG6 7BE

Email: i.cunnings@reading.ac.uk
Phone: +44 (0)118 378 5538

Abstract

A growing body of research has investigated bilingual sentence processing. How to account for differences in native (L1) and non-native (L2) processing is controversial. Some explain L1/L2 differences in terms of different parsing mechanisms, and the hypothesis that L2 learners adopt 'shallow' parsing has received considerable attention. Others assume L1/L2 processing is similar, and explain L1/L2 differences in terms of capacity-based limitations being exceeded during L2 processing. More generally, the role that working memory plays in language acquisition and processing has garnered increasing interest. Based on research investigating L2 sentence processing, I claim that a primary source of L1/L2 differences lies in the ability to retrieve information constructed during sentence processing from memory. In contrast to describing L1/L2 differences in terms of shallow parsing or capacity limitations, I argue that L2 speakers are more susceptible to retrieval interference when successful comprehension requires access to information from memory.

1. Introduction

Research investigating second language (L2) acquisition and bilingualism has increasingly examined the role of language processing in language acquisition and learning. This research in part has been motivated by the application of psycholinguistic theories to research in bilingual sentence comprehension. With regards to L2 processing, Clahsen and Felser's (2006a, 2006b) *shallow structure hypothesis* (SSH), which claims a reduced sensitivity to grammatical information during L2 processing, has been particularly influential. In contrast to the claims of qualitatively different parsing routines in the L1 and L2 made by the SSH, others argue that L1 and L2 sentence processing is fundamentally similar, and maintain that purported differences result from inefficient lexical access routines or from an increased burden on capacity-limited cognitive resources (McDonald, 2006; Hopp, 2006, 2010). More generally, the role that individual differences in working memory play in explaining individual differences in L2 attainment has also been widely debated (Juffs, 2004; Juffs & Harrington, 2011; Harrington & Sawyer, 1992; Linck, Osthus, Koeth, & Bunting, 2014; Sagarra, 2013; Wen, Mota, & McNeill 2015). Many researchers investigating the role of working memory in L2 acquisition and processing have been influenced by Baddeley and colleagues' model of working memory (e.g. Baddeley, 1986, 2000, 2007; Baddeley & Hitch, 1974), and assume a characterisation of individual differences in working memory in terms of capacity limitations. Recent research in psycholinguistics has however focused on the computational principles that underlie the working memory operations that subserve language comprehension (e.g. Lewis, Vasishth, & Van Dyke, 2006; Lewis & Vasishth, 2005; McElree, 2000, McElree, Foraker, & Dyer, 2003). Individual differences in such models are attributed to differences in the success of memory retrieval operations, rather than individual differences in processing capacity (Van Dyke & Johns, 2012; Van Dyke, Johns, & Kukona, 2014).

In this paper, I provide a critical overview of research investigating L2 sentence comprehension. I will argue that careful consideration of the working memory operations that

underlie sentence processing help provide a better characterisation of the similarities and differences between monolingual and bilingual sentence processing. Rather than implicating 'shallow' parsing or capacity-based limitations, I argue that L1/L2 differences that persist at high levels of proficiency can be characterised in terms of an increased susceptibility to interference during memory retrieval operations in L2 sentence processing. Although I will focus mainly on L2 learners, defined as speakers who have learnt a second language in adolescence or adulthood after an L1 has been fully acquired, in the General Discussion I also discuss how this approach to sentence processing can inform research in other bilingual populations. Below I first provide an outline of different accounts of L2 processing and the working memory operations that sub-serve language comprehension, before discussing how existing research comparing L1/L2 sentence processing is consistent with the hypothesis that accessing information constructed during parsing from memory is the primary source of difficulty during L2 sentence processing.

2. Theoretical approaches to L2 processing

The SSH (Clahsen & Felser, 2006a, 2006b) predicts that L2 speakers have difficulty computing fully specified syntactic structures during processing, and instead compute 'shallow' syntactic representations compared to L1 speakers. L2 speakers are argued to rely more heavily on semantic and pragmatic information to guide comprehension than is observed during L1 processing. Key evidence for this theory originally came from studies demonstrating that L2 speakers were sensitive to semantic and pragmatic information during parsing but construct apparently underspecified syntactic representations in the resolution of syntactic ambiguities (Felser, Roberts, Marinis, & Gross, 2003; Papadopoulou & Clahsen, 2002) and linguistic dependencies (Felser & Roberts, 2007; Marinis, Roberts, Felser, & Clahsen, 2005).

Another influential theory in the L2 literature is the *interface hypothesis* (Sorace, 2011), which claims that L2 speakers have difficulty in reaching native-like attainment in phenomena

that lie at the interface between syntax and other cognitive domains. Although originally based on evidence from offline experiments, the most recent instantiation of the IH (Sorace, 2011) claims that the difficulty associated with linguistic interfaces lies in L2 speakers' reduced ability to integrate syntactic and other information sources online during processing. Key evidence for this theory comes from studies investigating the production and comprehension of null and overt pronouns (e.g. Belletti, Bennati, & Sorace, 2007; Sorace & Filiaci, 2006).

Although the SSH and IH make different predictions regarding when L1/L2 differences may arise, both theories describe L1/L2 differences in terms of the *types* of information that may or may not be utilised during processing. Ullman (2001, 2005, 2015) describes L1/L2 differences in terms of differing reliance on declarative memory, which subserves lexical storage, and procedural memory, which subserves the computation of grammatical rules. While L1 speakers are hypothesised to utilise both declarative and procedural memory, Ullman claims L2 speakers rely more heavily on declarative memory, associatively memorising rule-governed properties of language that are subserved by procedural memory in L1 speakers. Proceduralisation of grammatical rules in the L2 is however predicted with sufficient exposure and proficiency. Evidence for this hypothesis has come from studies using neuroscientific techniques to test the hypothesised reliance on declarative knowledge and proceduralised syntactic parsing at different levels of proficiency (e.g. Bowden, Steinhaeur, Sanz, & Ullman, 2013).

While models such as the SSH predict qualitative differences between L1 and L2 processing, even at high levels of L2 proficiency, others claim that sentence processing in the L2 is qualitatively similar to L1 processing. McDonald (2006) and Hopp (2006, 2010), for example, assume that although L2 processing is cognitively more demanding than L1 processing, it is not qualitatively different. Such accounts predict that apparent differences between L1 and L2 processing can be neutralised if individual differences in cognitive resources or memory capacity are taken into account.

Below, I focus on the question of what *types* of information L2 speakers, at high levels of proficiency, utilise during sentence processing, and thus will not discuss hypothesised differences in the use of declarative and procedural memory. I will argue that existing evidence does not necessarily implicate 'shallow' L2 parsing (Clahsen & Felser, 2006a, 2006b), or a difficulty in integrating information from syntax and other cognitive domains during processing (Sorace, 2011). Differences in cognitive capacity also do not however explain L1/L2 differences. Rather, I argue that L1/L2 differences are better explained in terms of the memory encoding, storage and retrieval operations that subserve successful language comprehension.

3. Parsing and working memory

Baddeley's (1986, 2000, 2007) model of working memory has been influential in research investigating working memory in L2 acquisition and processing. Such theories can be considered tripartite models in that they assume a narrow *focus of attention*, a *working memory* component and a separate *long term memory* store. Individual differences are ascribed to capacity limitations in the amount of information that an individual is able to hold active in working memory at one time. For language comprehension, individual differences in working memory capacity are taken to influence how many words or sentences can be actively maintained, or the different types of syntactic and non-syntactic information sources that can be recruited at one time (e.g. Daneman & Carpenter, 1980; Just & Carpenter, 1992, Just, Carpenter, & Keller, 1996). Thus, capacity-based approaches describe L1/L2 differences in terms of the amount of information that L1 and L2 speakers are able to keep active at a given point in time (McDonald, 2006; Hopp, 2006, 2010).

In contrast, bipartite models assume no separate working memory component (for discussion, see McElree, 2006). Such models assume that the amount of information that can be activated in the focus of attention at one time is severely limited and much smaller than is measured in span tasks (Cowan, 2001; McElree, 2006). All other information needs to be retrieved

from memory. Rather than describing individual differences in terms of working memory capacity, such models emphasise the processes involved in memory encoding, storage and retrieval, and in switching attention between items brought in and out of focal attention.

In assuming a bipartite model, recent accounts of language processing have described sentence comprehension as involving skilled memory retrieval (e.g. Lewis et al., 2006; Lewis & Vasishth, 2005). This model makes a number of predictions about different aspects of sentence comprehension. Consider, for example, the resolution of linguistic dependencies, as in (1).

(1) It was the book that the boy read.

(1) involves a filler-gap dependency, in which successful interpretation requires the dislocated 'filler' ('*the book*') to be interpreted as the direct object of the verb '*read*'. In retrieval-based models of comprehension, a memory representation of the filler will be encoded when it is first encountered and then stored in memory as the other words in the sentences are processed (and themselves encoded as memory chunks). Upon reaching the verb, a memory retrieval operation will trigger retrieval of the filler as the direct object of the verb.

One factor that may influence retrieval in such cases is activation-based decay (Lewis et al., 2006; Lewis & Vasishth, 2005). After an item is first encoded in memory, its activation levels will decay as other items are encountered. Retrieval operations will subsequently boost activation however. Activation-based decay thus predicts distance effects in dependency resolution (e.g. Gibson, 1998), with more recent items being more easily retrieved. Another key factor that influences memory retrieval, that decay is very often confounded with, is similarity-based interference. A growing body of research has implicated a cue-based, direct access memory retrieval mechanism during language comprehension (e.g. Martin & McElree, 2008; 2009, 2011; McElree, 2000; McElree et al., 2003). In cue-based models, memory retrieval involves comparing

a set of retrieval cues against *all* items in memory in parallel. The item that provides the best match to the set of retrieval cues becomes highly activated and thus retrieved. However, as the retrieval cues are matched against all items in memory in parallel, an item that *partially* matches the retrieval cues may sometimes be retrieved. Similarity-based interference is thus a consequence of cue-based retrieval, and predicts that successful memory retrieval depends on the number of items in memory that match the cues at retrieval. Cue-based models predict that individual differences in sentence comprehension can be explained in terms of individual differences in susceptibility to interference (Van Dyke & Johns, 2012; Van Dyke, Johns, & Kukona, 2014).

Different types of interference effects have been reported during language comprehension. Agreement attraction (Pearlmutter, Garnsey, & Bock, 1999; Shen, Staub, & Sanders, 2013; Tanner, Nicol, & Brehm, 2014; Wagers, Lau, & Phillips, 2009) is one example. Consider (2), from Wagers et al. (2009).

(2a) The key to the cabinet(s) unsurprisingly was rusty from many years of disuse.

(2b) The key to the cabinet(s) unsurprisingly were rusty from many years of disuse.

In (2), the verb (‘*was*’/’*were*’) must agree in number with the head noun (*‘key*’) rather than the linearly closer ‘distractor’ (‘*cabinet(s)*’). Wagers et al. reported longer reading times after the verb in ungrammatical sentences like (2b) compared to grammatical (2a). However, the size of the ungrammaticality effect in (2b) was significantly reduced when the distractor matched the number features of the verb. No reliable differences were observed in grammatical sentences like (2a) however as a result of the number properties of the distractor.

This type of interference is sometimes called *facilitatory* interference as reading times in ungrammatical sentences are *faster* when a distractor partially matches the cues at retrieval. In (2b), the retrieval cues at the verb will include [+HEAD] and [+PLURAL]. No items fully match

this set of cues, as the phrasal head ('*key*') does not match the [+PLURAL] cue while the distractor is not [+HEAD]. When the distractor is plural ('*cabinets*'), while no item in memory fully matches the retrieval cues, both the phrasal head noun and the distractor provide a partial match. In this case, the plural distractor noun may sometimes be retrieved, leading to attenuation of the ungrammaticality effect.

In *inhibitory* interference, which has been observed for some dependencies, reading times are *slower* as a result of having multiple items in a sentence that match a set of retrieval cues (Badecker & Staub, 2002; Gordon, Hendrick, & Johnson, 2001; Van Dyke, 2007, Van Dyke & McElree, 2011). For example, in sentences like (3), from Van Dyke (2007), '*the resident'* needs to be retrieved as the subject of *'was complaining*'. Van Dyke reported longer reading times when the intervening distractor was animate (*'neighbor'*) compared to inanimate (*'warehouse*'). This was interpreted as indexing competition between the retrieval target and distractor, when both were animate subjects.[1] These results indicate that at a minimum [+SUBJECT] and [+ANIMATE] are utilised as retrieval cues in such dependencies.

(3) The worker was surprised that the resident who said that the neighbor/warehouse was dangerous was complaining about the investigation.

[1] Other accounts of inhibitory interference have been proposed that do not assume competition between items in memory (see e.g. Dillon, Mishler, Sloggett, & Phillips, 2013; Jäger, Benz, Roeser, Dillon, & Vasishth, 2015). I do not attempt to add to this debate here. For current purposes, it is sufficient to say that similarity-based interference can lead to inhibitory effects and longer reading times in grammatical sentences, in cases when multiple items match a set of retrieval cues.

Inhibitory interference is found in grammatical sentences when a retrieval target fully matches the cues to retrieval, such as 'the resident' in (3), but a distractor, such as 'the neighbor', also sufficiently matches the retrieval cues to cause competition. Facilitatory interference occurs in ungrammatical sentences where no item fully matches the cues at retrieval, but a distractor provides a partial match. Different types of interference effects thus provide evidence with regards to what types of information are utilised as memory retrieval cues during language processing. Within this framework, potential differences in dependency resolution in L1/L2 speakers can be characterised in terms of how different populations may utilise different memory retrieval cues.

Note that interference effects do not necessarily implicate 'shallow' parsing. For example, if agreement attraction effects were the result of a shallow parse, we would not expect to see an asymmetry in attraction effects between ungrammatical and grammatical sentences (Lewis & Phillips, 2015; Phillips, 2013). That is, if the parser constructed shallow parses, sentences like (2a) should be considered ungrammatical when the distractor mismatches the number features of the verb. However, such effects are not routinely observed.[2] Rather than shallow parsing, interference effects reflect the nature of retrieval operations in a 'noisy' memory architecture (Phillips, 2013).

[2] Some have reported such an effect in grammatical sentences where reading times are longer when the distractor *mismatches* the features of the verb (*'the key to the cabinets was…'*) compared to when they match (*'the key to the cabinet was…'*). This 'broken agreement' effect (Chen et al., 2007; Jiang, 2004; Pearlmutter et al., 1999) is not predicted under either facilitatory or inhibitory interference. However, Wagers et al. (2009) argued that such effects are usually only observed when the distractor is directly adjacent to the verb but not when additional lexical material intervenes, suggesting such effects may be an artefact indexing spillover processing following a plural noun.

Memory retrieval is also implicated in reanalysis processes during ambiguity resolution. A number of studies have reported that in garden-path sentences, an initially assigned sentence interpretation is not always successfully reanalysed (e.g. Christianson, Hollingworth, Halliwell & Ferreira, 2001). Christianson et al. (2001) reported that adult English speakers sometimes incorrectly interpret sentences like (4) as indicating that '*Anna dressed the baby*' despite the fact that that this direct object analysis, although initially preferred during incremental processing, is ultimately ungrammatical.

(4) While Anna dressed the baby that was small and cute spit up in the bed.

Such results might be taken as indicating that readers do not always construct fully specified syntactic analyses. While earlier formulations of 'good-enough' parsing described misinterpretation in such cases in these terms (e.g. Ferreira, Ferraro, & Bailey, 2002; Ferreira & Patson, 2007), more recent research has shown that these types of misinterpretations need not implicate an underspecified parse, but instead can be attributed to interference from the initially assigned sentence interpretation in memory, which is not fully erased (Slattery, Sturt, Christianson, Yoshida, & Ferreira, 2013).

4. Parsing and working memory in L2 sentence processing

Below I discuss how consideration of the memory retrieval operations described above can help characterise L1/L2 differences during sentence processing. Based on a review of studies investigating different linguistic phenomena, I argue that the existing evidence indicates that the primary source for L1/L2 differences relates to the ability to successfully retrieve information that has been constructed during sentence processing from memory. Rather than implicating shallow parsing, or describing L1/L2 differences in terms of processing capacity, I will argue that existing

research indicates that L2 learners do construct fully-specified syntactic parses, but that they are more susceptible to interference during memory retrieval than L1 speakers. Furthermore, this increased susceptibility to interference I will argue in part results from L2 learners implementing memory retrieval cues during dependency resolution differently to L1 speakers.

I begin below by discussing syntactic ambiguity resolution and reanalysis. I will argue that existing evidence does not implicate shallow L2 parsing during ambiguity resolution, nor do L2ers have a reduced capacity to take into account a wide variety of information sources during ambiguity resolution. Instead, L1/L2 differences persist during ambiguity resolution specifically when successful interpretation requires accessing and revising information previously constructed during processing. I will then discuss research on different linguistic dependencies. Here, I again will argue that apparently contrasting results in L2 learners' sensitivity to the role of syntactic information during dependency resolution can be explained in terms of increased susceptibility to interference during memory retrieval, but is difficult to explain in terms of shallow syntactic parsing. I will also argue that L2 learners do not have a difficulty in integrating information from syntax and interface domains (Sorace, 2011), but instead that existing findings indicate that L2 learners rely more heavily on discourse-based cues to memory retrieval than L1 speakers.

4.1 Syntactic ambiguity resolution and reanalysis

A large body of research has examined whether L2 learners utilise the same parsing principles during syntactic ambiguity resolution as L1 speakers (Felser et al., 2003; Frenck-Mestre & Pynte, 1997; Papadopoulou & Clahsen, 2003; Rah & Adone, 2010; Witzel, Witzel, & Nicol, 2012). I begin by discussing studies examining relative clause attachment ambiguities, as these were influential in motivating the SSH.

Felser et al. (2003) investigated sentences like (5), where the relative clause ('*who was reading a letter*') can modify either the first NP ('*the secretary*') or the second ('*the professor*').

Memory retrieval is implicated in the resolution of such ambiguities, as encountering the relative pronoun *who* triggers retrieval of a particular NP to which further incoming material is attached.

(5) The dean liked the secretary of/with the professor who was reading a letter.

When containing a thematic preposition such as *'with*', sentences such as (5) bias NP2 attachment while attachment preferences for the preposition *'of*' show cross-linguistic variation (Cuetos & Mitchell, 1998; Frazier & Clifton, 1996; Gilboy, Sopena, Clifton & Frazier, 1995). In a self-paced reading study that manipulated number agreement between the two NPs and the critical verb ('*was'*) as a diagnostic of attachment, Felser et al (2003) reported an NP2 attachment preference in both L1 and L2 English speakers for the preposition '*with*'. While the L1 English speakers also showed an NP2 preference for sentences containing the preposition *'of*', L2 speakers exhibited no clear preference. Papadopoulou and Clahsen (2003), and Dussias (2003), reported similar results of no clear attachment preference in L2 Greek and L2 Spanish respectively.

Felser et al. (2003) interpreted these results as indicating similar attachment preferences in L1 and L2 processing when lexico-semantic information is available, in this case a thematic preposition, but differences when attachment relies on structural information. While attachment in the L1 was interpreted as being guided by syntactic factors such as predicate proximity (Gibson, Pearlmutter, Canseco-Gonzalez & Hickok, 1996), which favours the most recent NP, L2 speakers showed no attachment preference, taken to be evidence of underspecified, shallow parsing (Clahsen & Felser, 2006a, 2006b). However, that L2ers exhibited no clear attachment preference suggests variable attachment to either NP1 or NP2 rather than necessarily implicating a shallow parse. Variable attachment to NP1 and NP2 is also observed in L1 processing (Traxler, Pickering & Clifton, 1998), and attachment preferences in such cases are known to be influenced by a variety of factors, including performance on memory span tests (Swets, Desmet, Hambrick & Ferreira,

2007) and task demands (Swets, Desmet, Clifton & Ferreira, 2008). Furthermore, recent studies have also shown clear attachment preferences, albeit not always in the same direction as L1 speakers, in L2 learners. Witzel, Witzel and Nicol (2012) reported an NP1 attachment preference for Chinese L2 learners of English, while the L1 English participants in their study demonstrated an NP2 preference. Hopp (2014) reported NP2 attachment preferences for both L1 English speakers and also L2 English speakers who had high levels of lexical automaticity. Pan, Schimke and Felser (2015) also reported that L2 English speakers' attachment preferences were reliably modulated by discourse context, with an NP1 preference in discourse contexts that biased NP1 attachment and an NP2 preference in NP2-biasing contexts, whereas the L1 speakers in their study showed no clear attachment preferences online. Pan et al.'s (2015) results suggest L2 speakers may be more sensitive to the role of discourse-context in the resolution of attachment ambiguities than L1 speakers, a finding that might not be expected under capacity-based approaches to L1/L2 differences (McDonald, 2006; Hopp, 2006, 2010) or the hypothesis that L2 speakers have difficulty integrating information from syntax and otsher domains (Sorace, 2011). Similar results were also reported by Pan and Felser (2011), who reported that L2 learners' online resolution of VP/NP attachment ambiguities were guided more by discourse context than L1 speakers.

Together, the existing evidence indicates variable attachment preferences in relative clause ambiguities for both L1 and L2 speakers that are influenced by a number of factors. Crucially, variable syntactic attachment indicates choice in ambiguity resolution rather than shallow parsing. While L2 learners may show increased sensitivity to the role of context in ambiguity resolution (Pan & Felser, 2011; Pan et al., 2015), these findings indicate L2 learners may be more sensitive to adopting pragmatically appropriate parses, rather than implicating shallow ones. These results also provide preliminary support for the hypothesis that L1/L2 differences can be described in terms of differences in how memory retrieval cues during language processing are implemented. For example, to the extent that relative pronoun interpretation relies on memory retrieval, the

finding that L2 learners show increased sensitivity to the role of discourse context in processing relative clauses (Pan et al., 2015) suggests L2 learners may rely more heavily on discourse-based cues to memory retrieval than L1 speakers.

A number of studies investigating the processing of temporarily ambiguous garden-path sentences have shown that L1 and L2 speakers initially adopt syntactically simple analyses of ambiguous sentences (Dussias & Cramer Scaltz, 2008; Frenck-Mestre & Pynte 1997; Hopp, 2015; Jackson, 2008; Jacob & Felser, 2016, Jegerski, 2012; Juffs, 1998; Juffs & Harrington, 1996; Rah & Adone, 2010; Roberts & Felser, 2011). Although the results of these studies indicate that both L1 and L2 speakers attempt reanalysis following garden-paths, fewer studies have investigated how successful these reanalysis processes are during L2 processing. Jacob and Felser (2016) investigated reanalysis in sentences like (6).

(6) While the gentleman was eating(,) the burgers were still being reheated in the microwave.

In (6), *'the burgers'* is temporarily ambiguous and may initially be interpreted as the direct object of *'eating*'. This initial analysis requires revision when it subsequently becomes clear that it is the subject of an embedded clause at *'were still'*. A comma was used to compare reading times in ambiguous and unambiguous sentences. Both L1 English speakers and German L2 English learners had longer reading at and after *'were still'* in ambiguous than unambiguous sentences, indicating that reanalysis was attempted in both groups. Both groups however sometimes answered a subsequent comprehension question incorrectly (*'Was the gentleman eating burgers'*), suggesting reanalysis was not always complete. The L2 learners were also more likely to answer the comprehension questions incorrectly than L1 speakers (see Pozzan and Trueswell, 2016 for

similar results using the visual world paradigm). These results suggest that initially assigned interpretations may be more likely to persist following reanalysis in L2 than L1 processing.

One interpretation of this finding might be that L2 learners do not construct a fully-specified parse following disambiguation, in line with the SSH. However, as discussed above, such results may not necessarily result from a shallow structure and instead may indicate interference from the initially assigned parse that is not fully erased. One way to distinguish these two accounts is to examine both the persistence of the initially-assigned direct object interpretation and the ability to assign the globally correct syntactic structure (see Slattery et al., 2013). Interestingly, while the results of Jacob and Felser (2016) indicate lower L2 comprehension accuracy following garden-paths, other studies have reported high comprehension accuracy rates for garden-path sentences in both L1 and L2 speakers (Hopp, 2015; Roberts & Felser, 2011). The comprehension questions in these latter studies however tapped whether readers assigned the globally correct syntactic analysis in which the temporarily ambiguous NP is interpreted as the subject of the embedded clause (e.g. *'When the girl was playing the boy made some funny noises' – 'Who made funny noises?'*; from Hopp, 2015). This asymmetry in L2ers' ability to accurately answer questions that tap the correct structural assignment of the ambiguous phrase as the subject of an embedded sentence, compared to lower accuracy rates for questions that tap persistence of the initially-assigned direct object analysis, indicates L2 learners can successfully assign the globally correct syntactic structure but exhibit increased persistence of the initially assigned interpretation. Although further research is required here, these results are compatible with the interference account of reanalysis difficulty, but is unexpected under shallow parsing.

In sum, I have argued that existing evidence indicates L2 learners do not lack capacity to consider multiple syntactic and non-syntactic information sources during syntactic ambiguity resolution, nor do existing results suggest shallow structures. Instead, the L1/L2 differences observed in reanalysis suggest increased interference from initially assigned interpretations that

are not fully erased. L2ers' sensitivity to the role of discourse in the resolution of ambiguous relative pronouns (e.g. Pan et al. 2015) suggests increased reliance on discourse-based cues to memory retrieval, which I discuss in more detail below.

4.2 Linguistic dependencies

The question of whether L2 learners process linguistic dependencies in the same way as L1 speakers has received considerable empirical attention. Before overviewing these studies below, I first describe the memory demands that different linguistic dependencies impose on language comprehension. A careful consideration of these issues indicates that memory retrieval is the key source of L1/L2 differences during dependency resolution.

One distinction to be made is between *backwards-looking* and *forwards-looking* dependencies (see also Felser, 2015). Reflexive anaphors, as in (7a), are backwards-looking dependencies in that encountering the reflexive '*himself*' triggers retrieval of a previously mentioned antecedent. Filler-gap dependencies, as in (7b), also involve retrieval, as the displaced filler ('*the book*') needs to be retrieved at the verb '*read*' and interpreted as its direct object. However, filler-gap resolution also includes a 'forwards-looking' anticipatory component because, unlike anaphora resolution where encountering a referent provides no clue as to whether a pronoun or reflexive will subsequently cue its retrieval at a later point in time, initially encountering the displaced filler in a filler-gap dependency triggers a forward-looking search for a corresponding 'gap' later in the sentence (in this case, at '*read*').

(7a) The boy from next door hurt himself.

(7b) It was the book that the boy read.

Such dependencies are constrained in different ways. Reflexives are constrained by Binding Principle A (Chomsky, 1981). In (8a), the reflexive '*himself*' must refer to the syntactically local antecedent '*the boy*' rather than non-local '*John*'. Filler-gap dependencies are constrained by island constraints (Ross, 1967). In (8b), the filler '*the book*' must be interpreted as the direct object of the verb '*bought*' rather than '*read*'.

(8a) John said that the boy from next door hurt himself.

(8b) It was the book that the boy who read at school bought.

These constraints differ in one sense in that Principle A restricts a backwards-looking dependency while island constraints restrict a forwards-looking one. However, the two constraints also differ along another important dimension. Binding constraints are an example of a constraint that restricts *what* information is to be retrieved once a memory retrieval operation is initiated. On the other hand, island constraints are an example of a constraint on *when* a memory retrieval operation can be initiated. In the case of (8b), retrieval of the direct object filler from memory is licensed by the grammar at the verb '*bought*' but not '*read*'. While '*what'* constraints thus restrict what information should be retrieved during dependency resolution, '*when'* constraints restrict when a memory retrieval operation can be initiated.

Sensitivity to constraints on *when* a memory retrieval operation can be initiated provides insight into the structural complexity of the parse currently being computed, as such constraints require computation of a certain degree of hierarchical structure as a pre-requisite to their application. For example, computation of the relative clause structure is a pre-requisite for the relative clause island constraint to ban a retrieval attempt of the direct object filler at the verb '*read*' in (8b). Observing a failure to apply constraints on *when* a dependency may be formed in a nativelike way would thus be compatible with shallow parsing, and would inform our

understanding of the syntactic complexity of the parse being computed during L2 processing. On the other hand, application of constraints on *what* information is retrieved provides insight with regards to which cues are utilised during memory retrieval, but need not necessarily implicate shallow parsing, depending on the observed pattern of results.

Felser (2015) recently argued that L2 learners behave similarly to L1 speakers in the resolution of forwards-looking but not backwards-looking dependencies. I argue however that the existing data are best described in terms of similarities in L1/L2 processing with regards to constraints on *when* dependencies can be formed, but differences with regards to *what* information is retrieved during dependency resolution. This contrast follows naturally from the hypothesis that L2 learners construct fully-specified syntactic parses but implement memory retrieval cues differently to L1 speakers, but is unexpected under a characterisation of L1/L2 differences in terms of shallow parsing or capacity-based limitations. I discuss in turn below how this account explains existing data from filler-gap dependencies, agreement processing and anaphora resolution.

Filler-gap dependencies

L1 gap-filling is an 'active' process, in that comprehenders do not wait until information unambiguously supports that a gap has been reached, but instead attempt to keep dependencies as short as possible (Frazier & Clifton, 1989; Stowe, 1986; Traxler & Pickering, 1996). L2 learners also adopt 'active' gap-filling (Dallas & Kaan, 2008; Williams, Mobius, & Kim, 2001; Williams, 2006). Consider (9), taken from Felser, Cunnings, Batterham and Clahsen (2012).

(9a) Everyone liked the magazine that the hairdresser read extensively and with such enormous enthusiasm about before going to the salon.

(9b) Everyone liked the shampoo that the hairdresser read extensively and with such enormous enthusiasm about before going to the salon.

Using a design previously used with L1 speakers (Traxler and Pickering, 1996), Felser et al. (2012) used a plausibility diagnostic to test active gap-filling in L1 and L2 processing. In (9) the ultimate gap-site is at *'about*', but the verb '*read*' also constitutes an earlier *potential* gap-site. Both fillers, '*the magazine*' and '*the shampoo*', are plausible at the ultimate gap-site but (9b) is implausible at the earlier gap-site. Felser et al. reported plausibility effects shortly after '*read*', with longer reading times in implausible (9b) than plausible (9a) for both L1 English speakers and German L2 learners of English, suggesting active gap-filling by both groups. Other studies have also reported plausibility effects in sentences like (9) as evidence of active gap-filling in L2 speakers from a variety of L1 backgrounds (Cunnings, Felser, Batterham & Clahsen, 2010; Dallas, DeDe, & Nicol, 2013; Kim, Baek, & Tremblay, 2015; Omaki & Schulz, 2011).

In a second experiment Felser et al. (2012) used a 'filled-gap' diagnostic (Stowe, 1986), as in (10).

(10a) Everyone liked the magazine that the hairdresser read quickly and yet extremely thoughtfully about before going to the beauty salon.

(10b) Everyone liked the magazine that the hairdresser read articles with such strong conclusions about before going to the beauty salon.

In (10), if *'the magazine'* is interpreted as the direct object of *'read'*, this should cause processing disruption at the following word in (10b), where it becomes clear that this verb has an overt direct object ('*articles'*), compared to (10a) when it does not. While the L1 speakers in Felser et al.'s (2012) study showed immediate sensitivity to this filled-gap diagnostic, the filled-gap effect appeared delayed compared to the plausibility manipulation for the L2 speakers. Note however, the filled-gap manipulation used by Felser et al. involved comparing conditions with

different lexical material directly following the critical verb, which may have influenced these results. More recently, Cunnings and Felser (2014) compared plausibility and filled-gap effects using a different filled-gap manipulation where lexical material was identical across conditions at the critical noun following the verb (e.g. '*Fred liked the book about which the journalist had written documents very passionately…*' vs. '*Fred liked the book which the journalist had written documents very passionately about…*'). They reported a similar time-course for both plausibility and filled-gap effects for L1 English speakers and L1 German L2 English speakers, indicating similar sensitivity to both diagnostics during L2 processing.

In addition to sentences like (9), Felser et al. (2012) also examined sentences like (11), where a dependency between the verb *'read'* and filler ('*the magazine'/'the shampoo'*) is ruled out as a result of the relative clause island constraint (Ross, 1967).

(11) Everyone liked the magazine/the shampoo that the hairdresser who read extensively and with such enormous enthusiasm bought before going to the salon.

In their original L1 study, Traxler and Pickering (1996) reported plausibility effects at the verb '*read'* in sentences like (9) but not in corresponding relative clause constructions as in (11), indicating filler-gap dependencies were not constructed inside the island clause. Felser et al. (2012) reported a similar contrast in plausibility effects in island and non-island environments for L1 and L2 processing. This indicates that both L1 and L2 speakers only attempted retrieval of the filler when this memory retrieval operation was licensed by the grammar.

Omaki and Schulz (2011) reported similar results for Spanish learners of English in a self-paced reading experiment (see also Aldwayan, Fiorentino and Gabriel, 2010, for similar results using a different diagnostic). One counter-example to this generalisation is an experiment reported by Kim et al. (2015). Kim et al. found that Spanish L2 learners of English showed native-like

sensitivity to the relative clause island constraint, while Korean learners temporarily violated the constraint, as suggested by a plausibility effect in both island and non-island conditions. They interpreted this as evidence of cross-linguistic transfer, resulting from the fact that Korean, but not Spanish, is a wh-in-situ language. Note however that this contrasts with the results of Cunnings et al. (2010), who reported that L2 English speakers with L1 Chinese, another wh-in-situ language, did apply island constraints in a native-like way in an eye-tracking experiment that used a similar plausibility manipulation. In sum, although the role of transfer in L2 island sensitivity requires further investigation, these results indicate nativelike application of the relative clause island constraint is possible during L2 processing.

The finding that L2 speakers respect island constraints during processing indicates some level of hierarchical structure building during L2 parsing. As noted by Felser et al. (2012), there is debate in the theoretical linguistics literature regarding whether island constraints constitute syntactic restrictions or whether they can be reduced to a processing-based account, which claims that dependencies are not formed inside island domains as a result of processing overload (compare Kluender & Kutas, 1993; Hofmeister & Sag, 2010; Phillips, 2013; Sprouse, Wagers, & Phillips, 2012; Yoshida, Kazanina, Pablos & Sturt, 2014). However, even under a processing account of island constraints, island sensitivity presupposes construction of some level of syntactic structure (Omaki & Schulz, 2011). Thus, island sensitivity in the L1 and L2 indicates similar levels of structural complexity during parsing, at least for these constructions.

While in (9)-(11) retrieval of the filler is triggered by an overt lexical item, in other cases dependencies are not triggered by lexical licensers. For example, in '*The manager who the consultant claimed that the new proposal had pleased will hire five workers tomorrow'*, although the ultimate gap-site is at '*pleased',* there is also an intermediate gap at the clause boundary before '*that'*. This intermediate gap is structurally rather than lexically mediated, as it is licensed by the spec-CP position in the syntactic structure of the embedded clause (e.g. Chomsky, 1995), rather

than on overt lexical item. Gibson and Warren (2004) reported that such sentences showed facilitated reading times at *'pleased'* compared to sentences without the intermediate gap (e.g. '*The manager who the consultant's claim about the new proposal had pleased will hire five workers tomorrow*'), suggesting intermediate retrieval of the filler at spec-CP increases its residual level of activation, facilitating subsequent retrieval at the verb. Marinis, Roberts, Felser and Clahsen (2005) replicated this facilitation effect in L1 speakers, but found similar reading times at the critical verb for L2 speakers for both types of sentences, suggesting L2 learners did not attempt intermediate retrieval at spec-CP. Marinis et al. interpreted this result as indicating shallow parsing such that L2 learners do not construct structurally-mediated gaps. More recently however, Pliatsikas and Marinis (2013) reported nativelike facilitation in such sentences in L2 speakers after extended naturalistic exposure to English, suggesting nativelike processing of structurally-defined gaps is possible. Similarly conflicting results have also been reported in other dependencies, such as indirect-object gaps (compare Felser & Roberts, 2007; Miller, 2014, 2015).

In sum, research on filler-gap dependencies has shown L1/L2 similarities with regards to *when* memory retrieval operations are initiated. Studies investigating island constraints indicate L2 learners only attempt memory retrieval operations when they are licensed by the grammar. While there has been conflicting results with regards to structurally-mediated gaps, they do appear to trigger memory retrieval operations in L2 learners with sufficient language exposure (Pliatsikas and Marinis, 2013), suggesting such learners do construct specified syntactic structures, and initiate memory retrieval operations accordingly, in a similar way to L1 speakers.

Studies on gap-filling in L2 processing have thus been important in establishing that L2 learners attempt retrieval operations at similar points in a sentence as L1 speakers. I argued that this would not be expected under shallow parsing, given that syntactic structure plays an important role in licensing when retrieval operations can be initiated, but is consistent with L2 learners constructing syntactically specified structures. The role that interference may play in memory

retrieval during filler-gap dependency resolution in L2 processing has not however been examined. Thus, while these results indicate sensitivity to the role that syntactic structure plays in guiding *when* a retrieval operation is initiated, it is necessary to examine other linguistic dependencies to examine the question of what role syntactic information plays with regards to cueing *what* information is retrieved by L1 and L2 speakers during dependency resolution.

Agreement

Many studies have investigated L2 processing of different morphosyntactic agreement features (Alemán-Bañón, Fiorentino, & Gabriele, 2014; Chen, Shu, Liu, Zheo, & Li, 2007; Coughlin & Tremblay, 2013; Foote, 2011; Foucart & Frenck-Mestre, 2012; Jiang, 2004, 2007; Keating, 2009, 2010; Sagarra & Herschensohn, 2010, 2011, 2013). However, few L2 studies have systematically examined the role that interference may play during agreement processing.

Tanner, Nicol, Herschensohn and Osterhout (2012) reported an ERP experiment investigating agreement attraction in sentences like (12) in L1 English speakers and Spanish L2 English learners. Both groups exhibited a P600 effect at the critical verb '*has/have*' in ungrammatical (12b) compared to grammatical (12a). This ungrammaticality effect was attenuated in (12b) when the linearly closer distractor (*'trophy/trophies*') was plural, indicating both L1 and L2 speakers are susceptible to retrieval interference during agreement processing. No differences however were observed in the grammatical conditions.

(12a) The winner of the big trophy/trophies has proud parents.

(12b) The winner of the big trophy/trophies have proud parents.

These results might be taken as evidence of shallow processing. However, as noted by Tanner et al., if attenuation of the ungrammaticality effect in the presence of a plural distractor in

(12b) was the result of a shallow parse in which the linear string '*... the big trophies have...*' is taken to be grammatical, then we would also expected to observe effects in grammatical sentences, such that the linear string *'... the big trophies has...*' should be considered ungrammatical. This pattern of results however was not observed. This asymmetry thus suggests retrieval interference as an explanation for these results, rather than shallow parsing.

Although other studies have not directly investigated retrieval interference during L2 agreement processing, the effect of dependency length has been examined, with longer dependencies in the L2 sometimes, but not always, exhibiting less sensitivity to agreement violations (e.g. Coughlin & Tremblay, 2013; Foote, 2011; Keating, 2009, 2010). Slower reading speed in the L2 may lead to enhanced effects of activation-based decay, which could lead to delayed or less robust retrieval of target items, and thus less robust violation effects, as dependency length increases. In support of this hypothesis, Kaan, Ballantyne and Wijnen (2015) reported nativelike sensitivity to number agreement violations for L2 speakers, but only when the L1 and L2 speakers were matched on reading speed.

Dependency length and the role of retrieval interference are very often confounded. As dependency length increases, so too the number of possible distractor items in memory will also likely increase, leading to an increased chance of retrieval interference the greater the length of a dependency. Consider the results of Keating (2010). Keating reported an eye-tracking experiment which examined adjective-noun gender agreement in L1 and L2 Spanish. Dependencies were either short (13a) or long (13b,c).

(13a) La tienda está abierta/*abierto los sábados y domingos por la tarde.

'The store-FEM is open-FEM/*open-MASC Saturdays and Sundays in the afternoon.'

(13b) El vestido de la muchacha es rosado/*rosada y tiene lunares blanco.

'The dress-MASC of the girl-FEM is pink-MASC/*pink-FEM and has white polka dots.'

(13c) La mochila de la estudiente está llena/*lleno de libros de texto.

'The backpack-FEM of the girl-FEM is full-FEM/*full-MASC of textbooks.'

L1 speakers had longer reading times at the critical adjective when it mismatched in gender with the sentence subject, irrespective of length. The L2 learners only showed this effect for short but not long dependencies. A crucial difference between the short and long conditions is not just the length of the dependency however but also that the long condition includes an intervening distractor noun that bears gender marking that could cause interference. In (13b), the distractor ('*muchacha*') has a different gender to the head noun but crucially matches the gender of the adjective in the *ungrammatical* condition. In (13c), the distractor (*'estudiente'*) matches the gender of the head noun, and the gender of the adjective in the *grammatical* condition. Keating (2010) included sentences of both types in his study (although this factor was not systematically varied). A combination of *facilitatory* and *inhibitory* retrieval interference could explain the L2ers' reduced sensitivity to gender violations in sentences likes (13b,c). In (13b), *facilitatory* interference would be a factor in reducing reading times in *ungrammatical* sentences, as the distractor provides a partial match to the gender cues of the ungrammatical adjective. In (13c), although ungrammatical sentences should usually lead to longer reading times, reading times in the *grammatical* condition, where multiple items match the gender of the adjective, could be elevated as a result of *inhibitory* interference. For the L1 speakers, these two competing pressures did not cancel out the ungrammaticality effect for sentences like (13b,c), but if L2 learners are more susceptible to interference, this may explain their insensitivity to violations in the long dependency condition.

In sum, the results of Tanner et al. (2012) show L1 and L2 speakers are susceptible to retrieval interference. Moreover, the specific pattern of results obtained is compatible with the interference account of attraction effects, but would not be expected under shallow parsing.

Apparent L2 insensitivity to agreement violations in other studies, such as Keating (2010), at least in part may be explained in terms of L2 learners' increased sensitivity to retrieval interference. While further examination of the role of interference is clearly required, these results are indicative of memory retrieval being a cause of difficulty during L2 agreement processing. Further research to systematically evaluate the role of retrieval interference in agreement during L2 processing would be a promising avenue of future research.

Anaphora resolution

A number of studies have investigated the application of binding constraints on anaphora resolution during sentence processing. Cue-based models of memory retrieval (Lewis et al., 2006; Lewis & Vasishth, 2005) predict anaphora resolution should be susceptible to interference, such that in sentences like (14a/b), syntactically 'inaccessible' antecedents ('*the man'*) that match the gender cues of a reflexive may be retrieved some proportion of the time, even though the only grammatically 'accessible' antecedent is '*the boy'* (Chomsky, 1981).

(14a) The man said the boy introduced himself.

(14b) The boy who greeted the man introduced himself.

The extent to which the application of binding constraints is susceptible to interference has been widely debated in the L1 literature (Badecker & Straub, 2002; Clackson, Felser, & Clahsen, 2011; Cunnings & Felser, 2013; Cunnings & Sturt, 2014; Dillon et al., 2013; Jäger et al., 2015; Nicol & Swinney, 1989; Xiang, Dillon, & Phillips, 2009; Sturt, 2003). For example, in two eye-tracking experiments, Sturt (2003) factorially manipulated gender congruence between a reflexive and two antecedents in sentences similar to (14a,b) (e.g. '*The surgeon who treated Jonathan/Jennifer had pricked himself/herself with a used syringe needle*'). Sturt found reading

times at the reflexive were longer when the syntactically 'accessible' antecedent mismatched the gender of the reflexive, but reading times in early eye-movement measures were not affected by the gender of the syntactically 'inaccessible' antecedent. This suggests that the syntactically accessible antecedent was initially retrieved. Although the question of whether reflexive resolution is impervious to interference during L1 processing is controversial, results such as these suggest syntactic constraints constitute more highly weighted cues to antecedent retrieval than gender congruence (Dillon et al., 2013).

Felser et al. (2009) examined reflexive resolution in L1 English speakers and Japanese L2 English learners. They reported the results of an eye-tracking experiment that tested sentences like (15), manipulating the gender of the syntactically inaccessible antecedent.

(15) John/Jane noticed that Richard had cut himself with a very sharp knife.

Felser et al. reported longer first-pass reading times at the reflexive for the L2 but not L1 speakers when the inaccessible antecedent matched the gender of the reflexive. This effect was not however observed in similar sentences where the inaccessible antecedent was placed in a less prominent position such that it no longer c-commanded the reflexive (e.g. '*It was clear to John/Jane that Richard had cut himself…*'). The L2 group's antecedent choices in an offline task were also similar to the L1 speakers.

Felser et al.'s results can be interpreted as indexing cue-based competition between the accessible antecedent and a discourse-salient inaccessible antecedent when both matched the gender cues of the reflexive during L2 processing. Felser and Cunnings (2012) further examined reflexive resolution in German L2 English speakers. Following Sturt (2003), they factorially manipulated gender congruence between a reflexive and two antecedents in texts like (16). In each

condition, '*the soldier'* is the only syntactically accessible antecedent according to binding theory, even though the syntactically inaccessible antecedent (*'James'/'Helen'*) is discourse prominent.

(16a) James/Helen has worked at the army hospital for years. He/She noticed that the soldier had wounded himself/herself while on duty in the Far East.

(16b) James/Helen has worked at the army hospital for years. The soldier that he/she treated on the ward wounded himself/herself while on duty in the Far East.

In Experiment 1, exemplified by (16a), Felser and Cunnings found that the L2 learners had longer reading times during first-pass processing at the reflexive when the inaccessible antecedent mismatched it in gender, while the L1 speakers had longer reading times when the accessible antecedent mismatched the gender of the reflexive. The L2 learners also showed a similar pattern of results during first-pass processing at the reflexive in Experiment 2, exemplified by (16b), where again reading times were longer when the inaccessible antecedent mismatched in gender with the reflexive. In both experiments, the L2 speakers exhibited sensitivity to the gender of the accessible antecedent, with longer reading times following gender mismatches, in reading time measures that index later stages of processing, such as rereading times of the reflexive. The L2 learners also exhibited nativelike performance of binding Principle A in an offline judgement task.

As with Felser et al. (2009), the results of Felser and Cunnings (2012) indicate temporary violation of binding constraints during L2 processing. The results of Experiment 2, as in (16b), when the inaccessible antecedent was linearly closer to the reflexive than the accessible antecedent, might be taken as evidence of shallow parsing, as early stages of processing suggested consideration of the linearly closest antecedent only. This preference for the linearly closest antecedent might also be expected under capacity-based approaches to L2 processing. However, if L2 speakers only considered the linearly closest antecedent, as a result of either a shallow-parse

or capacity limitations, then this would predict in Experiment 1, as in (16a), that L2 learners' reading times should have been influenced by the gender of the accessible antecedent, as in this experiment it was linearly closer. This however was not the case. Instead, I argue that this specific pattern of results, not expected under the two accounts discussed above, is better explained in terms of retrieval interference.

As I argued that the specific pattern of results observed for agreement attraction suggests retrieval interference rather than shallow processing, I also argue that these results for L2 speakers do not implicate shallow L2 parsing, but instead indicate a type of cue-based retrieval interference, with L2 learners relying more heavily on discourse-based retrieval cues than is observed in L1 processing. More specifically, I assume the retrieval cues at the reflexive constitute (at minimum) a cue for a clause-mate antecedent, a cue for subject (and/or topic) antecedents, and cues derived from the gender/number properties of the reflexive. For L1 speakers, the syntactic locality requirement for a clause-mate antecedent constitutes a highly weighted retrieval cue (Dillon et al., 2013). For L2 speakers, the preference for a subject/topic seems to be a more heavily weighted cue than in L1 speakers, to the extent that syntactically inaccessible but discourse-prominent subject antecedents are sometimes retrieved. The fact that the inaccessible antecedent was mentioned twice in the discourse in (16), might also have raised its residual levels of activation, increasing its chance of retrieval. Thus, I argue that these results indicate different weightings in the application of retrieval cues in L1 and L2 processing, rather than shallow parsing.

The question of whether the application of binding Principle B is susceptible to interference has also been debated in the L1 literature (Clifton, Kennison, & Albrecht, 1997; Badecker & Straub, 2002; Chow, Lewis, & Phillips, 2014; Kennison, 2003; Nicol & Swinney, 1989). Patterson, Trompelt and Felser (2014) reported an eye-tracking study investigating the application of Principle B by German L2 learners of English in sentences like (17), where the matrix subject is the only syntactically accessible antecedent (Chomsky, 1981).

(17a) John remembered that Mark had taught him a new song on the guitar.

(17b) John remembered than Jane had taught him a new song on the guitar.

(17c) Jane remembered that John had taught him a new song on the guitar.

Patterson et al. (2014) reported longer reading times at and shortly after the pronoun in sentences like (17c), where the accessible antecedent mismatches the gender of the pronoun, compared to (17a,b), where the accessible antecedent matches the gender of the pronoun, for both L1 and L2 speakers. No significant differences in reading times, for either group, were observed between conditions (17a) and (17b), which manipulated the gender of the syntactically local, inaccessible antecedent. This suggests both groups retrieved the syntactically accessible non-local antecedent only upon encountering the pronoun. In a second experiment, Patterson et al. studied so-called 'short-distance' pronoun constructions (e.g. '*Barry saw Gavin place a gun near him on the ground with great care'*), where the pronoun can refer to both syntactically local and non-local antecedents (Reinhart & Reuland, 1993). In this experiment, which had three conditions similar to (17), they found that reading times for L2 speakers again were longer when the non-local antecedent mismatched in gender with the pronoun, but the gender of the local antecedent did not significantly influence reading times. The L1 speakers showed a different pattern, with longer reading times when the local antecedent mismatched the gender of the pronoun, suggesting consideration of this antecedent in contexts where it is syntactically licensed. Taken together, Patterson et al. interpreted their results as indicating that upon encountering the pronoun, L2 speakers preferentially retrieved the discourse-prominent matrix subject.

While the results of Patterson et al. suggests sensitivity to Principle B in L2 processing, as they also note, this pattern of results would also be expected if L2 speakers rely more heavily on discourse-based cues to antecedent retrieval. Together, the results for reflexives and pronouns are

compatible with the hypothesis that discourse-based cues (e.g. [+SUBJECT], [+TOPIC]) are more highly weighted retrieval cues during L2 processing than syntactic locality constraints (see also Felser, in press).

The hypothesis that discourse-based cues to anaphora resolution are more highly weighted during L2 processing also has implications for research in L2 acquisition of null and overt pronouns. In null subject languages like Italian, exemplified in (18) (from Sorace, 2011), the null subject pronoun in (18a) most likely refers to the current discourse topic ('*La vecchietta*'), while the overt pronoun in (18b) will likely index a topic-shift to '*la ragazza'*.

(18a) La vecchietta salute la ragazza quando *pro* attraversa la strada.

(18b) La vecchietta salute la ragazza quando *lei* attraversa la strada.

'The old woman greets the girl when *ø/she* crosses the road.'

A common finding in the L2 literature is that while the interpretation of null pronouns can become nativelike, L2 speakers show an increased tendency to interpret overt pronouns as maintaining the current discourse topic compared to L1 speakers (e.g. Belletti, Bennati, & Sorace, 2007; Sorace & Filiaci, 2007; for a similar contrast in languages with personal and demonstrative pronouns, see Ellert, 2013; Wilson, 2009). This contrast in L2 pronoun resolution was instrumental in the formulation of the interface hypothesis (Sorace, 2011; Sorace & Filiaci, 2006). As discussed above, in the most recent instantiation of the IH, Sorace (2011) attributes non-nativelike acquisition at linguistic interfaces to difficulty in integrating information from syntactic and non-syntactic sources during processing.

However, a general over-reliance on discourse-based cues to antecedent retrieval during L2 pronoun resolution can also explain this pattern of results. If L2 learners rely heavily on discourse-based retrieval cues, then this would predict similar behaviour between L1 and L2

speakers for pronominal forms (e.g. null subjects) that trigger retrieval of the current discourse topic, but not for pronouns that cue a topic shift (e.g. overt subjects in null subject languages), where L2 learners would be predicted to show a weaker preference for such forms to index a topic shift. As the IH was formulated to account for findings at linguistic interfaces, it is not clear how it would explain the L1/L2 differences in core syntactic dependencies, like reflexives, reported by Felser et al. (2009) and Felser and Cunnings (2012). The combined results discussed above however follow naturally from the hypothesis that L2 anaphora resolution involves an overreliance on discourse-based retrieval cues.

Another study to consider in this domain is an eye-tracking experiment reported by Roberts et al. (2008). Roberts et al. investigated Dutch sentences like (19), in L1 speakers and L2 learners with L1 German or Turkish.

(19a) De werknemers zitten in het kantoor. Terwijl Peter aan het werk is, eet hij een boterham.

'The workers are in the office. While Peter is working, he is eating a sandwich.'

(19b) Peter en Hans zitten in het kantoor. Terwijl Peter aan het werk is, eet hij een boterham.

'Peter and Hans are in the office. While Peter is working, he is eating a sandwich.'

Roberts et al. reported two main findings. Firstly, upon encountering the pronoun during reading, both groups of L2 learners, but not the L1 Dutch participants, had longer reading times in (19b), when multiple antecedents matched the gender of pronoun, compared to (19a), when there was only one gender matching antecedent in the discourse. Sorace (2011) interpreted these results as being in line with the IH, suggesting difficulty in integrating syntactic and discourse-level information. However, as discussed above, studies examining other types of ambiguous

sentences have shown that L2 learners can successfully integrate syntactic and discourse-level information to guide parsing (e.g. Pan et al. 2015), indicating L2 learners do not have a general difficulty integrating information from multiple sources. An alternative explanation is that this effect can be interpreted as inhibitory interference, indexing competition between the two antecedents in (19b), when multiple items in memory match the gender cues of the pronoun (similar to Felser et al., 2009, for reflexives).

Roberts et al. also reported an effect of L1 transfer in participants' interpretive preferences. For sentences like (19b), both L1 Dutch speakers and German L2 learners almost always interpreted the pronoun as referring to the discourse-prominent antecedent *'Peter*', while the Turkish L2 learners interpreted it as referring to *'Hans'* approximately 50% of the time. Roberts et al. took this as evidence of L1 transfer, as in Turkish, a null subject language, overt pronouns index a topic shift. These results suggest that while L2 learners may rely more heavily on discourse-based cues to pronoun resolution than L1 speakers, particular L2 groups may transfer the properties of discourse-based retrieval cues from their L1 to the L2. The role of transfer may not be absolute however. Cunnings, Fotiadou and Tsimpli (under review) reported a visual world paradigm study investigating subject pronoun resolution in English in sentences like (20).

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

Both L1 English speakers, and Greek L2 learners of English, interpreted the subject pronoun as referring to the sentence subject and current discourse topic ('*Peter'*), as evidenced by both eye-movements during listening and their ultimate interpretive preferences. If transfer played a role during processing, this finding might be unexpected given that in Greek, a null-subject language, overt pronouns index a topic shift (Papadopoulou, Peristeri, Plemenou, Marinis, &

Tsimpli, 2015), which in this case would bias interpretation to the object antecedent ('*Mr Smith*'). Cunnings et al.'s (under review) results however suggest that L2 learners can overcome these L1 biases, at least in a language such as English that does not have both null and overt pronouns. Further research is needed in investigating how discourse-based cues to anaphora resolution may be transferred during L2 processing.

Finally, while the above discussion has focused on *backwards-looking* anaphoric dependencies, in cataphoric reference, pronouns trigger a *forwards-looking* search for an antecedent (Van Gompel & Liversedge, 2003). The antecedent search in such cases is however restricted by binding Principle C (Chomsky, 1981), which disallows coreference between a pronoun and an antecedent that it c-commands. Consider (21), adapted from Kazanina, Lau, Lieberman, Yoshida and Phillips (2007).

(21a) He/She chatted with some fans while the quarterback signed autographs for the kids, but Steve/Carol wished the charity event would end soon.

(21b) His/Her managers chatted with some fans while the quarterback signed autographs for the kids, but Steve/Carol wished the charity event would end soon.

In (21a) the subject pronoun c-commands '*the quarterback'* while in (21b) the possessive pronoun does not. Thus, Principle C should rule out a dependency between this noun and the subject pronoun in (21a), but not between it and the possessive pronoun in (21b). Note that this restriction provides another example of a constraint on *when* a memory retrieval operation may be initiated, rather than a constraint on *what* information is to be retrieved. Specifically, Principle C restricts the initiation of a memory retrieval operation (of the pronoun) to when a referent is encountered that is outside of the pronouns binding domain. Thus, sensitivity to Principle C requires construction of the relevant binding/scope domain as a pre-requisite to its application.

This contrasts with the constraints on anaphora discussed above, which constrain *what* information should be retrieved once a reflexive or pronoun triggers a retrieval.

In a series of self-paced reading experiments, Kazanina et al. (2007) reported longer reading times for L1 English speakers after '*the quarterback'* in sentences like (21b) when the possessive pronoun mismatched in gender with this noun compared to when there was a gender match. This gender mismatch effect was not observed in sentences like (21a). Kazanina et al. interpreted their results as indicating that Principle C restricts cataphoric pronoun resolution during processing.

Although cataphoric pronoun resolution has not been widely studied during L2 processing, two unpublished experiments have investigated Principle C in L2 speakers (Bertenshaw, 2009; Rodriguez, 2008). Bertenshaw (2009) conducted an eye-tracking experiment with L1 English speakers and L2 learners with L1 Japanese. Using materials similar to Kazanina et al., Bertenshaw reported that both L1 and L2 English speakers made more regressions out of the critical noun (equivalent to the '*the quarterback*') in sentences similar to (21b) when there was a gender mismatch compared to when there was a gender match, but this effect was not observed in sentences similar to (20a). This suggests Principle C restricted the initiation of memory retrieval operations during both L1 and L2 processing. Rodriguez (2008) also reported sensitivity to Principle C in a self-paced reading experiment with Spanish and Chinese L2 English speakers. Although further research is required in this area, these results suggest nativelike application of Principle C during L2 processing. As with the results on filler-gap dependencies and island constraints discussed above, this suggests similarly complex parses during L1 and L2 processing.

In sum, these results indicate L2 learners implement antecedent retrieval cues during anaphora resolution differently to L1 speakers. In particular, L2 learners weight discourse-based cues to retrieval more heavily than L1 speakers. I have argued that this pattern of results need not suggest shallow parsing, or a difficulty in integrating information from multiple sources. The

results of studies examining Principle C indicate similarities in L1 and L2 processing with respect to constraints on *when* memory retrieval operations can be initiated, similar to other linguistic dependencies. L1/L2 differences appear to persist in anaphora resolution specifically with regards to constraints which cue *what* information needs to be retrieved during dependency resolution.

5. General discussion

The above overview indicates both similarities and selective differences between L1/L2 processing. I have argued that L2 learners construct parses of similar structural complexity to L1 speakers, but that differences between L1 and L2 processing result from differences in how information constructed during parsing is retrieved from memory. Below I compare this account of L1/L2 differences to other approaches of L2 sentence processing. I will also discuss how this cue-based approach to sentence processing may inform our understanding of sentence processing in other bilingual populations.

Shallow L2 parsing and cognitive capacity

I have argued that L2 learners at high levels of proficiency do not routinely adopt shallow parses. I argued that studies examining syntactic ambiguity resolution highlight the variable nature of attachment preferences, in both L1 and L2 processing, rather than shallow processing. Although L2 learners are sensitive or more sensitive to discourse-level and semantic information during ambiguity resolution than L1 speakers (e.g. Felser & Roberts, 2011; Pan & Felser, 2011; Pan et al., 2015), such findings do not in themselves indicate shallow parsing and instead suggest L2 learners are more sensitive to constructing pragmatically-appropriate parses. L1/L2 differences are observed during reanalysis, which crucially relies on accessing and revising information in memory. Importantly, existing studies (e.g. Felser & Roberts, 2011; Jacob & Felser, 2016; Pozzan & Trueswell, 2016; Hopp, 2015) suggest the difficulty here is not related to an inability to assign

a globally correct syntactic structure, which might be expected under shallow parsing, but instead is limited to difficulty in fully erasing the initially assigned interpretation from memory.

L2 learners also show considerable, but selective, sensitivity to the role of syntactic information in guiding the resolution of linguistic dependencies. L2 learners behave similarly to L1 speakers with regards to *when* particular retrieval operations are initiated during sentence processing (e.g. Bertenshaw, 2009; Cunnings et al., 2012; Omaki & Schulz, 2011; Pliatsikas & Marinis, 2013; Rodriguez, 2009). Sensitivity to the role that syntactically-mediated gaps, island constraints and binding Principle C play in licensing when particular memory retrieval operations can be initiated indicates similar levels of syntactic sophistication in the parses constructed during L1/L2 processing. It remains an open question whether L1/L2 differences may emerge here if sufficiently complex syntactic structures are examined, but the existing evidence suggests L2 learners behave similarly to L1 speakers with regards to the role that syntax plays in licensing the initiation of memory retrieval operations during language comprehension.

Note also that the L1/L2 similarities that are observed in dependencies, such as gap-filling and cataphoric pronoun resolution, that involve a forwards-looking anticipatory search might be unexpected under recent accounts which claim L2 learners do not predict upcoming content like L1 speakers (see Kaan, 2014 for review). Irrespective of potential differences in prediction in L1 and L2 processing, the results of studies examining forwards-looking dependencies are difficult to explain in terms of differences in predictive behaviour in L1 and L2 speakers.

The L1/L2 differences that are observed during the resolution of linguistic dependencies I have argued are primarily restricted to the issue of *what* information needs to be retrieved during dependency resolution, and are best described in terms of cue-based retrieval interference rather than shallow parsing. This is in-line with recent proposals that interference is a primary determinant of successful language comprehension (Van Dyke & Johns, 2012; Van Dyke et al., 2014). The specific pattern of results obtained in L1 and L2 agreement attraction (Tanner et al.,

2012) are unexpected under shallow parsing but follow naturally from the retrieval interference account. The observed effects in anaphora resolution (e.g. Cunnings & Felser, 2012) I have also argued can be explained in terms of an increased reliance on discourse-based cues to antecedent retrieval, but would be unexpected under a characterisation of shallow parsing in terms of a preference for linearly recent antecedents.

I have also argued that capacity limitations cannot provide a full account of L1/L2 differences. The observation that L2 learners show increased sensitivity to discourse-level information during ambiguity resolution (Pan & Felser, 2011; Pan et al., 2015), for example, would not be expected if L2 processing is capacity-limited compared to L1 processing. Increased competition during anaphora resolution when local and distant antecedents match the gender properties of a pronoun have also been reported during L2 processing (Felser et al., 2009; Roberts et al., 2008). If L2 processing is capacity-limited compared to L1 processing, we might expect *attenuated* effects of multiple-matching antecedents during L2 anaphora resolution, rather than the opposite pattern that is observed. These competition effects, as evidence of inhibitory interference, naturally follow however from the hypothesis that L2 processing is more susceptible to effects of retrieval interference. It is also not clear how slowed processing or lexical access times could explain the contrasting interference effects observed for L1 and L2 speakers. Indeed, reading times were similar for both L1 and L2 speakers in the study reported by Felser and Cunnings (2012), yet L1/L2 differences were still observed. Of course, it is difficult to conclude that capacity limitations play no role in L1/L2 differences, but the overview above exemplifies that a precise characterisation of the memory encoding, storage and retrieval operations involved in sentence processing can provide a promising framework to describe L1/L2 differences.

Individual differences measures, such as reading span (Daneman & Carpenter, 1980; Harrington & Sawyer, 1992), play an important role in capacity-based approaches to L2 processing. Cue-based approaches to parsing emphasise an approach to individual differences that

emphasises the quality, rather than quantity, of representations in memory (Van Dyke et al. 2014). Different individual differences measure may be related to the different memory operations involved in language comprehension. For example, the memory access and revision processes required for garden-path recovery may be dissociable from the memory retrieval operations involved in linguistic dependencies. In the L1 processing literature, successful garden-path reanalysis has been argued to rely on cognitive control, the ability to switch attention during information conflict, and correlates with performance on tasks that tap conflict monitoring (Kan, Teubner-Rhodes, Drummey, Nutile, Krupa, & Novick, 2013; Novick, Hussy, Teubner-Rhodes, Harbison, & Bunting, 2014; Vuong & Martin, 2014). Little is currently known with regards to individual differences in memory retrieval during linguistic dependency resolution, and it is possible that different types of inhibitory and facilitatory interference effects have different sources. Van Dyke et al. (2014) recently found inhibitory interference effects correlated with individual differences measures that tap the quality of representations stored in memory. Whether or not other types of interference correlate with these or other measures is currently not known. Further research is required to systematically examine how individual differences may influence different memory operations during L2 language comprehension, and the question of whether L2 processing can become fully nativelike if individual differences are taken into account.

Sentence processing in other bilingual populations

While non-native sentence processing has received growing empirical attention, fewer studies have investigated the phenomena discussed above in bilingual comprehension more broadly. From the perspective of cue-based parsing, one question in bilingual language comprehension is whether the weightings of retrieval cues in one language influence how similar cues are weighted in the other languages that a bilingual speaker may know.

Some studies have investigated whether knowledge of an L2 influences relative clause attachment in the L1. In a self-paced reading study, Dussias (2003) reported that Spanish L1 speakers with L2 English demonstrated an online preference for NP2 attachment in Spanish, even though Spanish has an NP1 preference (Cuetoes & Mitchell, 1988). In a follow-up eye-tracking study, Dussias and Sagarra (2007) reported an NP1 attachment preference for monolingual Spanish speakers and L1 Spanish speakers with limited exposure to L2 English, while a group with extensive exposure to English demonstrated an NP2 preference. These results suggest that experience of a second language may influence how cues to relative clause attachment are weighted in the L1.

Research in anaphora resolution has also investigated bilingual populations other than L2 learners. One preliminary study investigating the application of Principle A in adult English by late L2 learners and early child bilinguals was reported by Kwon, Cunnings and Lesmana (2013). In a self-paced reading study which factorially manipulated gender agreement between a reflexive and two antecedents (e.g. *'John/Mary claimed that the pilot injured himself/herself quite badly during the journey'*), they found evidence of interference from the linearly-distant inaccessible antecedent in both the early and late bilinguals. The early bilinguals in this study knew Mandarin Chinese and English, and as such this interference effect could be a result of transfer of the long-distance binding cues of the Chinese reflexive *ziji*.

The interpretation of null and overt pronouns is one area that has been examined in L2 speakers and other bilingual populations. For example, as L2 learners over-extend interpretation of overt pronouns in a null-subject L2 to refer to the current discourse topic (Belletti et al., 2007; Sorace & Filiaci, 2007), Tsimpli, Sorace, Heycock and Filiaci (2004) reported that L1 Italian speakers under attrition from L2 English also tended to interpret overt pronouns as referring to the current discourse topic more often than monolingual Italian speakers in a picture verification task. This suggests that the retrieval cues that bias a topic antecedent in L2 English may have influenced

the topic-shift cues of overt pronouns in L1 Italian. Recent results from an eye-tracking experiment investigating null and overt pronoun resolution in Spanish monolinguals and bilinguals under attrition from L2 English reported by Chamorro, Sorace and Sturt (2015) suggest that such attrition effects may be attenuated following re-exposure to the L1, indicating potential flexibility with regards to the weighting of different cues to memory retrieval during language comprehension.

In sum, existing research on bilingual sentence processing outside of the non-native context suggests that the retrieval cues adopted during memory retrieval operations in one language may influence how retrieval cues are weighted in the other language(s) that a bilingual speaker may know. While it is not entirely clear how shallow parsing or capacity-based approaches to L2 processing might explain effects observed in a bilingual speakers' L1, and indeed these theories were not intended to account for such data, the cue-based approach provides a framework to examine potential differences between L1 and L2 speakers, and other bilingual populations as well. Further research is required to investigate how knowledge of multiple languages may influence how memory retrieval cues are weighted in the different languages that a bilingual speaker may know.

6. Conclusions

Successful sentence and discourse comprehension crucially relies on the ability to encode, store and retrieve information from memory. A precise characterisation of these processes can help provide a way of explaining the differences observed during sentence processing in different populations of speakers. This echoes similar claims regarding the roles of working memory and cognitive control/executive functions more generally in bilingualism and multilingual language acquisition, that a precise characterisation of memory and cognitive control processes, and a clear understanding of how performance on individual differences tasks relate to specific aspects of language comprehension, is crucial to our understanding of the similarities and differences

between monolingual and bilingual sentence processing and cognition more generally (Hartsuiker, 2015; Linck et al., 2014).

I have argued that a key determinant of L1/L2 sentence processing differences relates to variability in the ability to retrieve information constructed during online parsing from memory, and in particular that existing evidence suggests that L2 sentence processing is more susceptible to effects of similarity-based retrieval interference. Results from studies investigating anaphora resolution suggest this increased susceptibility to interference results from L1 and L2 speakers differently weighting syntactic and discourse-level cues to memory retrieval. Systematic examination of interference effects in other linguistic dependencies is required to gain a better understanding of the role that memory retrieval and interference plays in explaining L1/L2 differences, and further research is required to investigate whether these L1/L2 differences in sentence processing can be overcome. The role that individual differences in memory encoding, storage and retrieval may play in sentence processing in bilingual populations other than L2 learners is also currently poorly understood and warrants further investigation. By gaining a better understanding of these processes we can hopefully arrive at a better understanding of the similarities and differences in sentence processing and comprehension in different bilingual populations.

References

Aldwayan, S., Fiorentino, R. & Gabriele, A. (2010). Evidence of syntactic constraints in the processing of wh-movement: A study of Najdi Arabic learners of English. In VanPatten, B., & Jegerski, J. (eds.). *Research in Second Language Processing and Parsing*, pp. 65-86. Amsterdam: John Benjamins.

Alemán-Bañón, J., Fiorentino, R., & Gabriele, A. (2014). Morphosyntactic processing in advanced second language (L2) learners: An event-related potential investigation of the effects of L1-L2 similarity and structural distance. *Second Language Research*, 30, 275-306.

Baddeley, A. (1986). *Working memory*. Oxford: OUP.

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

Baddeley, A. (2007). *Working Memory, Thought, and Action.* Oxford: OUP.

Baddeley, A., & Hitch, G. (1974). Working memory. In Bower, G. (ed.). *The Psychology of Learning and Motivation; Advances in Research and Theory*, New York: Academic Press.

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 748–769.

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

Bertenshaw, N. (2009). *The application of binding constraints by Japanese L2 learners of English*. Unpublished PhD dissertation. University of Essex, UK.

Bowden, H., Steinhauer, K., Sanz, C., & Ullman, M. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, 51, 2492-2511.

Chen, L., Shu, H., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject-verb agreement in L2 learning. *Bilingualism: Language and Cognition*, 10, 161-174.

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

Clackson, K., Felser, C., & Clahsen, H. (2011). Children's processing of reflexives and pronouns in English: Evidence from eye-movements during listening. *Journal of Memory and Language*, 65, 128-144.

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

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

Chamorro, G., Sorace, A., & Sturt, P. (2015). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition. *Bilingualism: Language and Cognition*. Available online, DOI: http://dx.doi.org/10.1017/S1366728915000152

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

- (1995). *The Minimalist Program*. Cambridge: MA. MIT Press.

Chow, W-Y., Lewis S., & Phillips, C. (2014). Immediate sensitivity to structural constraints in pronoun resolution. *Frontiers in Psychology*, 5, 630. DOI: http://dx.doi.org/10.3389/fpsyg.2014.00630

Clifton, C., Kennison, S., & Albrecht, J. (1997). Reading the words her, his, and him: Implications for parsing principles based on frequency and structure. *Journal of Memory and Language*, 36, 276-292.

Coughlin, C. & Tremblay, A. (2013). Proficiency and working memory based explanations for non-native speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistis*¸34, 615-646.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-185.

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

Cunnings, I. & Felser, C. (2013). The role of working memory in the processing of reflexives. *Language and Cognitive Processes*, 28, 188-219.

Cunnings, I. & Felser, C. (2014). *Plausibility and filled gap effects in native and non-native sentence processing*. Poster presented at *AMLaP* 2014, University of Edinburgh, UK.

Cunnings, I., Batterham, C., Felser, C. & Clahsen, H. (2010). 'Constraints on L2 learners' processing of wh- dependencies: Evidence from eye-movements' in VanPatten, B. & Jegerski, J. (Eds.), *Second Language Processing and Parsing I,* pp 87-110. Amsterdam, John Benjamins.

Cunnings, I., Fotiadou, G. & Tsimpli, I. (under review). Grammatical roles and gender agreement in L2 pronoun resolution: Evidence from the visual world paradigm. Ms., University of Reading.

Cunnings, I. & Sturt, P. (2014). Coargumenthood and the processing of reflexives. *Journal of Memory and Language*, 75, 117-139.

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, 766-799.

Dallas, A. & Kaan, E. (2008). Second language processing of filler-gap dependencies in late learners. *Language and Linguistics Compass*, 2, 372-388.

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

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

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

Dussias, P., & Cramer Scaltz, T. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologica*, 128, 501–513.

Dussias, P. & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 10, 101-116.

Ellert, M. (2013). Resolving ambiguous pronouns in a second language: A visual-world eye-tracking study with Dutch learners of German. *International Review of Applied Linguistics*, 51, 171-197.

Felser, C. (2015). Native vs. non-native processing of discontinuous dependencies. *Second Language*, 14, 5-19.

Felser, C. (in press). Binding and coreference in non-native language processing. To appear in Holler, A. & Suckow, K. (eds.), *Empirical perspectives on anaphora resolution: Information structural evidence in the race for salience*. Berlin: De Gruyter.

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

Felser, C., Cunnings, I., Batterham, C., & Clahsen, H. (2012). The timing of island effects in nonnative sentence processing. *Studies in Second Language Acquisition*, 34, 67-98.

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.

Felser, C. & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research*, 23, 9-36.

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, 485-502.

Ferreira, F., Bailey, K., & Ferraro, V. (2002). Good enough representations in language comprehension. *Current Directions in Psychological Science*, 11, 11-15.

Ferreira, F., & Patson, N. (2007). The good enough approach to language comprehension. *Language and Linguistics Compass*, 1, 71-83.

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

Foucart, A. & Frenck Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, 66, 226-248.

Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, 4, 93-126.

Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge: MIT Press.

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

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

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickock, G. (1996). Recency preferences in the human sentence processing mechanism. *Cognition*, 59, 23–59.

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

Gilboy, E., Sopena, J., Clifton, C., & Frazier, L. (1995). Argument structure and association preferences in Spanish and English compound NPs. *Cognition*, 54, 131–167.

Gordon, P., Hendrick, R., & Johnson, M. (2001). Memory interference during sentence processing. *Psychological Science*, 13, 425–430.

Harrington, M. and Sawyer, M. (1992). Working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14, 25-38.

Hartsuiker, R. (2015). Why it is pointless to ask under which specific circumstances the bilingual advantage occurs. *Cortex*, 73, 336-337.

Hofmeister, P., and Sag, I. (2010). Cognitive constraints and island effects. *Language*, 86, 366-415.

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

Hopp, H. (2010). Ultimate attainment in L2 inflectional morphology: Performance similarities between non-native and native speakers. *Lingua*, 120, 901-931.

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

Hopp, H. (2015). Individual differences in the second language processing of object-subject ambiguities. *Applied Psycholinguistics*, 36, 129-173.

Jackson, C. (2008). Proficiency level and the interaction of lexical and morphosyntatic information during L2 sentence processing. *Language Learning*, 58, 875-909.

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

Jäger, L., Benz, L., Roeser, J., Dillon, B & Vasishth, S. (2015). Teasing apart retrieval and encoding interference in the processing of anaphors. *Frontiers in Psychology*, 6, 506. DOI: http://dx.doi.org/10.3389/fpsyg.2015.00506

Jegerski, J. (2012). The processing of subject–object ambiguities in native and near-native Mexican Spanish. *Bilingualism: Language and Cognition*, 15, 721-735.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*, 603–634.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, 57, 1-33.

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

Juffs, A. (1998). Main verb vs reduced relative clause ambiguity resolution in second language sentence processing. *Language Learning*, 48, 107–147.

Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language processing research. *Language Learning*, 46, 283–323.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 Learning. *Language Teaching: Reviews and Studies*, 42, 137-166.

Just, M. & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *The Psychological Review*, 99, 122-149.

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

Kaan, E. (2014). Predictive sentence processing in L2 and L2. What is different? *Linguistic Approaches to Bilingualism*, 4, 257-82.

Kaan, E., Ballantyne, J., & Wijnen, F. (2015). Effects of reading speed on second-language sentence processing. *Applied Psycholinguistics*, 36, 799-830.

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

Kazanina, N., Lau, E., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backward anaphora. *Journal of Memory and Language*, 56, 384-409.

Keating, G. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning*, 59, 503-535.

Keating, G. (2010). The effects of linear distance and working memory on the processing of gender agreement in Spanish. In B. VanPatten & J. Jegerski (Eds.), *Research in second language processing and parsing*, pp. 113-134. Amsterdam: John Benjamins.

Kennison, S. (2003). Comprehending the pronouns her, him, and his: Implications for theories of referential processing. *Journal of Memory and Language*, 49, 335-352.

Kim, E., Baek, S., & Tremblay, A. (2015). The role of island constraints in second language sentence processing. *Language Acquisition*, 22, 384-416.

Kluender, R., & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, 8, 573-633.

Kwon, N., Cunnings, I., & Lesmana, M. (2013). *Time Course of Reference Resolution by Early and Late Bilinguals*. Poster presented at *AMLAP*, Aix-Marseille Université, France.

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

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

Lewis, S. & Phillips, C. (2015). Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research*, 44, 27-46.

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

Marinis, T., Roberts, L, Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, 27, 53-78.

Martin, A., & McElree, B. (2009). Memory operations that support language comprehension: Evidence from verb-phrase ellipsis. *Journal of Experimental Psychology: Learning, Memory, Cognition*, 35, 1231-1239.

Martin, A., & McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language*, 58, 879-906.

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

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

McElree, B. (2006). Accessing recent events. In Ross, B. (ed.), *The Psychology of Learning and Motivation*, Volume 46. San Diego: Academic Press.

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

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

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

Miller, A. (2015). Intermediate traces and intermediate learners: Evidence for the use of intermediate traces during sentence processing in second language French. *Studies in Second Language Acquisition*, 37, 487-516.

Nicol, J. & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistics Research*, 18, 5-20.

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

Omaki, A. & Schulz, B. (2011). Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition*, 33, 563-588.

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

Pan, H., Schimke, S. & Felser, C. (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.

Papadopoulou, D., Peristeri, E., Plemenou, E., Marinis, T. & Tsimpli, I. (2015). Pronoun ambiguity resolution in Greek: Evidence from monolinguals adults and children. *Lingua*, 155, 98-120.

Patterson, C., Trompelt, H., Felser, C. (2014). The online application of binding condition B in native and non-native pronoun resolution. *Frontiers in Psychology*, 5, 147. DOI: http://dx.doi.org/10.3389/fpsyg.2014.00147

Pearlmutter, N., Garnsey, S., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41, 427–456.

Phillips, C. (2013). Some arguments and non-arguments for reductionist accounts of syntactic phenomena. *Language and Cognitive Processes*, 28, 156-187.

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

Pliatsikas, C. & Marinis, T. (2013) Processing empty categories in a second language: When naturalistic exposure fills the (intermediate) gap. *Bilingualism: Language and Cognition*, 16, 167-182.

Rah, A. & Adone, D. (2010). Processing of the reduced relative clause versus main verb ambiguity in L2 learners at different proficiency levels. *Studies in Second Language Acquisition*, 32, 79-109.

Reinhart, T. & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry*, 24, 657-720.

Roberts, L. & C. Felser (2011). Plausibility and recovery from garden-paths in second-language sentence processing. *Applied Psycholinguistics*, 32, 299-331.

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, 333-357.

Rodriguez, G. (2008). *Second language sentence processing: Is it fundamentally different?*. Unpublished PhD dissertation. University of Pittsburgh, USA.

Ross, J. (1967). Constraints on variables in syntax. Unpublished doctoral dissertation, Massachusetts Institute of Technology, USA.

Sagarra, N. (2013). Working memory in second language acquisition. In Chapelle, C. (ed.) *The Encyclopaedia of Applied Linguistics*. Oxford: Wiley-Blackwell.

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, 2022-2039.

Sagarra, N. & Herschensohn, J. (2010). Proficiency and animacy effects on L2 gender agreement processes during comprehension. *Language Learning*, 120, 2022-2039.

Sagarra, N. & Herschensohn, J. (2013). Processing of gender and number agreement in late Spanish bilinguals. *International Journal of Bilingualism*, 17, 607-627.

Shen, E., Staub, A., & Sanders, L. (2013). Event-related brain potential evidence that local nouns affect subject-verb agreement processing. *Language and Cognitive Processes, 28*, 498-524.

Slattery, T., Sturt, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language*, 69, 104-120.

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

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

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

Stowe, L. (1986). Parsing wh- constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, 1, 227-245.

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

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

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

Tanner, D., Nicol, J., & Brehm, L. (2014). The time course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language,* 76, 195-215.

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., & Pickering, M. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, 35, 542-562.

Traxler, M., Pickering, M., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, 39, 558-592.

Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism*, 8, 257-277.

Ullman, M. (2001). The neural basis of lexicon and grammar: The declarative/procedural model. *Bilingualism: Language and Cognition*, 4, 105-122.

Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In Sanz, C. (ed.), *Mind and Context in Adult Second Language Acquisition*, pp. 141-178. Washington, Georgetown University Press.

Ullman, M. (2015). The declarative/procedural model: A neurobiologically motivated theory of first and second language. In VanPatten, B. & Williams, J. (eds.), *Theories in Second Language Acquisition: An Introduction (2nd edition)*. New York: Routledge.

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

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

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

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

Van Gompel, R. & Liversedge, S. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29, 128-139.

Vuong, L. & Martin, R. (2014). Domain-specific executive control and the revision of misinterpretations in sentence comprehension. *Language, Cognition and Neuroscience*, 29, 312-325.

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

Wen, Z., Mota, M., & McNeill, A. (2015). *Working Memory in Second Language Acquisition and Processing*. Multilingual Matters.

Williams, J. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition*, 9, 71-88.

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

Wilson, F. (2009). *Processing at the syntax-discourse interface in second language acquisition.* Unpublished PhD thesis, University of Edinburgh.

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

Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain & Language*, 108, 40-55.

Yoshida, M., Kazanina, N., Pablos, L., & Sturt, P. (2014) On the origin of islands. *Language, Cognition and Neuroscience*, 29, 761-770.