Parsing preferences and individual differences in non-native sentence processing: evidence from eye-movements

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Parsing preferences and individual differences in non-native sentence processing:

Evidence from eye-movements

Yesi Cheng

University of Reading, UK

Jason Rothman

UiT, The Arctic University of Norway, Norway

Universidad Nebrija, Spain

Ian Cunnings

University of Reading, UK

Address for correspondence:

Yesi Cheng

School of Psychology and Clinical Language Sciences

The University of Reading

Whiteknights Campus

Reading

RG6 7BE

Abstract

Using both offline and online measures, the present study investigates attachment resolution in relative clauses (RC) in English natives (L1) and non-natives (L2). We test how RC resolution interacts with linguistic factors and participant-level individual differences. Previous L1 English studies have demonstrated a low attachment preference and also an "ambiguity advantage", suggesting that L1ers may not have as strong a low attachment preference as is sometimes claimed. We employ a similar design to examine this effect in L1 and L2 comprehension. Offline results indicate that both groups exhibit a low attachment preference, positively correlated with reading span scores and with proficiency in the L2 group. Online results also suggest a low attachment preference in both groups. However, our data show that individual differences influence online attachment resolution for both native and non-natives; higher lexical processing efficiency correlates with quicker resolution of linguistic conflicts. We argue that the current findings suggest that attachment resolution during L1 and L2 processing share the same processing mechanisms and are modulated by similar individual differences.

Keywords: Non-native sentence processing; attachment resolution; eye-tracking during reading; individual differences

Introduction

The similarities and differences between native (L1) and non-native (L2) sentence processing are widely debated (e.g., Clahsen & Felser, 2006, 2018; Cunnings, 2017; Grüter & Rohde, 2013; Hopp, 2014; Kaan, 2014; McDonald, 2006). Some theories propose that L2 processing employs different parsing mechanisms from L1 processing (e.g., Clahsen & Felser, 2006; Jiang, 2004). Alternatively, others have argued that L1 and L2 processing employ similar mechanisms and that L2ers construct syntactic analyses in the same way as L1ers (e.g., Cunnings, 2017; Hopp, 2014; Kaan, 2014). Under such accounts, observable differences between groups are argued to be non-qualitative. Additionally, the extent to which L1 and L2 parsing are modulated by individual differences is also debated (e.g., Hopp, 2014; McDonald, 2006; Tanner et al., 2014). For example, Hopp (2014) argued that lexical automaticity plays a role in native-like performance by L2ers, with more efficient lexical processing leading L2ers to behave more like L1ers during sentence processing.

Relevant to this debate are studies that have examined how L1ers and L2ers resolve ambiguous relative clauses (RCs), using offline and online tasks with sentences like (1) (e.g., Carreiras & Clifton, 1999; Cuetos & Mitchell, 1988; Dussias, 2003; Felser et al., 2003b; Maia et al., 2007; Pan et al., 2015; Rothman, 2010; Scheepers et al., 2011).

(1a) The brother of the man who bought himself some books lived here.
(1b) We knew the brother of the man who bought himself some books.

In (1) the RC is embedded to modify a complex noun phrase (NP) ("the brother of the man"), which either serves as the syntactic subject (1a) or object (1b) of the sentence. In both sentences, the RC ("who bought himself…") is ambiguous, as it can be interpreted as referring back to either the local NP "the man" (low attachment) or the non-local NP "the brother" (high

3

attachment). English readers tend to prefer low attachment (e.g., Cuetos & Mitchell, 1988; Gibson et al., 1996). However, low attachment is not a universal tendency and attachment preferences vary across languages (e.g., Cuetos & Mitchell, 1988; Hemforth et al., 2015; Zagar et al., 1997). Additionally, whether L1 and L2 readers show similar attachment preferences has been contested, especially in the case where the L1 and L2 display distinct attachment preferences. While some studies suggest L2ers do not show any clear attachment preferences (e.g., Felser et al., 2003b; Omaki, 2005), others have argued that L2ers can behave similarly to L1ers, even when their L1 displays an opposite tendency and especially if individual differences in L2 processing are considered (e.g., Dekydtspotter et al., 2008; Hopp, 2014).

Linguistic factors also influence attachment preferences (e.g., Desmet et al., 2002; Desmet et al., 2006; Fodor, 2002). Most importantly for present purposes, some L1 studies suggest that the syntactic position of the constituent that the RC modifies influences attachment preferences of sentences like (1) (e.g., Hemforth et al., 2000; Hemforth et al., 2015). Whether L2ers are sensitive to such subtle differences, is, however yet to be systematically explored.

Against the above background, we conducted a study on relative cause (RC) attachment in L1and L2 English speakers, testing the syntactic position of the RC. To tease apart different accounts of L1 and L2 processing, we aimed to test the extent to which L2ers can process and interpret RCs as in (1) in a nativelike way. To examine whether individual differences influence how nativelike L2 processing can become in this domain (Hopp, 2014), we also investigated how individual differences in working memory, lexical processing and for L2ers proficiency, influence L1 and L2 RC resolution offline and during processing.

Relative clauses in L1 processing
A large literature has contested how parsing preferences influence the processing of RCs (e.g., Cuetos &Mitchell, 1988, and much subsequent literature), and two competing principles are

believed to influence low vs. high attachment. *Late Closure* (Frazier, 1979) or *Recency* (Gibson et al., 1996) predicts that new material is attached to the most recently processed constituent. On the other hand, *Predicate Proximity* (Gibson et al., 1996) holds that incoming material is preferably attached as close as possible to the head of a predicate. Therefore, readers who follow *Late Closure* or *Recency* favour low attachment of the RC to the local NP. Alternatively, readers who are guided by Predicate Proximity prefer high attachment to the non-local NP.

As mentioned above, English L1ers generally have a low attachment preference for ambiguous RCs as in (1) in offline tasks (e.g., Cuetos & Mitchell, 1988). In tasks that measure online processing, researchers have manipulated agreement features to force either high or low attachment, as in (2a) and (2b) respectively. These studies have typically shown shorter reading times at the disambiguating word ("was/were" in (2a/b)) for RCs that attach low than those that attach high in English (e.g., Felser et al., 2003b; Hopp, 2014; Omaki, 2014). However, the low attachment preference of English is not universal, and a high attachment preference has been attested in offline and online tasks in L1 studies of various other languages (e.g., Bidaoui et al., 2016; Brysbaert & Mitchell, 1996; Carreiras et al., 2004; Chernova & Chernigovskaya, 2015; De Vincenzi & Job, 1993; Maia et al., 2004; Papadopoulou & Clahsen, 2003).

(2a) The man blamed the brothers of the boy who were smiling all the time.
(2b) The man blamed the brothers of the boy who was smiling all the time.

Determining the robustness of the low attachment preference in L1 English is complicated by studies testing three different types of RCs together. For example, Van Gompel et al. (2005) compared reading times for sentences that force high attachment (3a), low attachment (3b) and a globally ambiguous condition (3c).

(3a) The governor of the province that will be retiring after the troubles is very rich.

(3b) The province of the governor that will be retiring after the troubles is very rich.

(3c) The bodyguard of the governor that will be retiring after the troubles is very rich.

Van Gompel et al.'s results suggested that, instead of showing a low attachment reading time advantage, the English L1ers they tested demonstrated an "ambiguity advantage", with shorter reading times for ambiguous sentences like (3c) than sentences where the RC was forced to attach either high (3a) or low (3b). Furthermore, the forced low attachment sentences were not significantly different from the high attachment ones in terms of reading times. These findings were interpreted as supporting an "unrestricted race model" of ambiguity resolution (Traxler et al., 1998; Van Gompel et al., 2000). The faster reading times for (3c) were taken to indicate that readers variably attached the RC either high or low across trials. (3c) is thus easiest because whichever attachment was initially computed at "that" will turn out to be plausible once the verb "retiring" is encountered. For both (3a) and (3b), whichever attachment is initially computed at "that" will be incorrect 50% of the time at the verb, leading to (3a/b) having equally longer reading times compared to (3c). Therefore, these results suggest that L1ers of English may not have as strong a low attachment preference as claimed, variably attaching instead to either available site. The fact that the ambiguous condition had faster reading times compared to the disambiguated conditions was also taken as evidence against the idea that the two possible interpretations in the ambiguous condition competed for activation in parallel as otherwise, Van Gompel et al. (2005) reasoned, the ambiguous condition should have had longer reading times due to competition.

A number of studies have examined how various linguistic factors influence attachment preferences, including whether attachment choices in L1ers are modulated by whether the complex NP the RC modifies is in subject position as in (1a) or object position as in (1b) (e.g.,

6

Hemforth et al., 2015; Kim & Christianson, 2017). Hemforth et al. (2015) found a stronger high attachment preference for NPs in object position in German and Spanish. Alternatively, the syntactic position did not influence attachment preferences in English and French, with English showing a general low attachment preference. However, this study tested offline preferences only, and did not test online processing.

Researchers have also examined how participant-level individual differences influence attachment resolution. This includes work examining the role of working memory, as measured by reading span tasks, although results have been mixed (e.g., Kim & Christianson, 2013; Kim & Christianson, 2017; Payne et al., 2014; Swets et al., 2007). Some studies reported that the low attachment preference in English increases as a function of higher reading span scores in either offline (Kim & Christianson, 2013; Omaki, 2005; Swets et al., 2007) or online measures (Kim & Christianson, 2013; Payne et al., 2014), suggesting that high-span individuals prefer attaching low more often than low-span individuals. However, results have been interpreted differently across these studies. Some researchers (e.g., Omaki, 2005) interpret these results from a resource limitation perspective whereas others (e.g., Kim & Christianson, 2013; Swets et al., 2007) see this as evidence for a chunking based account in which high-span and low-span readers chunk the relative clause differently, leading to different NPs being more salient.

Other studies observed the reverse pattern online (Felser et al., 2003a; Traxler, 2007) or no such effects offline (Felser et al., 2003a). For example, Traxler (2007) found that high-span individuals preferred high attachment while low-span readers preferred low attachment in an eye-tracking study. Traxler claimed that high-span individuals are more sensitive to discourse-salience and are drawn to the head NP rather than the subordinate NP, while low-span individuals may simply rely on linear distance.

Finally, Kim & Christianson (2017) and Payne et al. (2014) found that L1 readers with higher reading spans have greater difficulty processing globally ambiguous RCs than readers

with lower spans in an online self-paced reading task. Kim & Christianson (2017) interpreted this as indicating that high-span readers can hold the two potential interpretations in mind at the same time, which leads to competition between high and low attachment that is not found in lower span readers.

Relative clauses in L2 processing

The question of whether L2 learners show the same parsing preferences as L1 speakers has been widely examined. Some L2 studies found non-nativelike attachment preferences in L2ers (e.g., Dinçtopal-Deniz, 2010; Felser et al., 2003b; Fernandez, 2003; Omaki, 2005; Papadopoulou & Clahsen, 2003). For example, Felser et al. (2003b) tested sentences like (2) in an offline task and online self-paced reading task. They found that in both measures, the L2ers did not exhibit any structural preferences. These findings led to the conclusion that L2 processing is different from L1 processing when processing is guided by structural information.

Conversely, other studies have lent support to the claim that L2 parsing is guided by structure-based information in either offline or online tasks, or both (e.g., Bidaoui et al., 2016; Hopp, 2014; Witzel et al., 2012). For example, Hopp (2014) investigated online attachment preferences of advanced German speakers of L2 English using sentences like (2) during eye-tracking and found that the L2ers preferred to attach low, which was in line with the L1 results.

Less is known about how RC position modulates L2ers' attachment choices, and we are aware of only one study that has examined this issue. Kim and Christianson (2017) conducted a study with Korean learners of English using sentences like (1a/b) and found the L2ers' offline attachment preferences were not influenced by RC position. Syntactic position also was not found to influence online processing in self-paced reading. However, Kim and Christianson tested globally ambiguous sentences only, and did not test sentences disambiguated to high or low attachment.

Individual differences have also been examined in L2 attachment resolution. For example, Dekydtspotter et al. (2008) reported more nativelike attachment preferences as proficiency increased in an online study of L2 French. Hopp (2014) investigated the role of working memory, as measured by a reading span task, and lexical efficiency, tested using a lexical decision task, in both offline and online attachment resolution. Offline, increased reading span scores correlated with an increased low attachment preference, replicating some previous L1 English findings (e.g., Swets et al., 2007). The results from an eye-tracking experiment suggested that even though the L2 group as a whole did not show any clear attachment preferences for sentences like (4), where the ambiguity is not resolved immediately, a low attachment preference emerged when individual differences in lexical automaticity were considered, with the L2 individuals with high lexical automaticity preferring low attachment. Reading span scores, however, were not significantly correlated with online processing.

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

Finally, Kim & Christianson (2017) tested individual differences in English RCs during self-paced reading in Korean learners of English. They found that high span L2ers had longer reading times for sentences containing globally ambiguous RCs relative to those with low span scores. As similar patterns were observed when these L2ers processed their L1, Korean, Kim and Christianson concluded that high-span readers are more likely to consider both interpretations of ambiguous RCs in parallel than low-span readers, in both their L1 and L2.

The present study

The mixed findings in previous L2 studies suggest that L2 attachment resolution is influenced by a wide variety of factors. However, the role that the syntactic position of the RC may play

in influencing L2 attachment resolution has not been systematically examined. Furthermore, while existing studies have examined whether or not L1ers consider different potential attachment choices in parallel (e.g., Traxler et al., 1998), the extent to which L2 comprehension may involve parallel competition between multiple possible analyses has not been examined extensively. Although Kim and Christianson (2017) examined this issue in globally ambiguous RCs, to our knowledge, no existing published study has tested this issue by directly comparing the processing of ambiguous RCs to those disambiguated low and high.

As such, we employed a similar design to Van Gompel et al.'s (2005) L1 study. We tested RCs in subject and object position, as in (5) and (6) respectively. We tested globally ambiguous sentences, as in (5a/6a), and compared them to sentences that forced low attachment, as in (5b/6b), and sentences that forced high attachment, as in (5c/6c). Given that some previous L2 studies have reported no differences between high and low attachment conditions during processing (e.g., Felser et al., 2003b; Papadopoulou & Clahsen, 2003), our approach of including a globally ambiguous condition provides an important control for comparing low and high attachment sentences.

(5a) Subject-Modifying RC, Ambiguous

The brother of the man who accidentally hurt <u>himself</u> <u>yesterday afternoon</u> lived in town.

(5b) Subject-Modifying RC, Low Attachment

The sister of the man who accidentally hurt <u>himself</u> <u>yesterday afternoon</u> lived in town.

(5c) Subject-Modifying RC, High Attachment

The brother of the woman who accidentally hurt <u>himself</u> <u>yesterday afternoon</u> lived in town.

(6a) Object-Modifying RC, Ambiguous

We saw the brother of the man who accidentally hurt <u>himself</u> <u>yesterday afternoon</u>.

(6b) Object-Modifying RC, Low Attachment

We saw the sister of the man who accidentally hurt <u>himself</u> <u>yesterday afternoon</u>.

(6c) Object-Modifying RC, High Attachment

We saw the brother of the woman who accidentally hurt <u>himself</u> <u>yesterday afternoon</u>.

… Luckily it wasn't serious in the end.

If L1ers demonstrate a low attachment preference, as predicted by the garden path model (Frazier,1987), reading times should be longer in (5c/6c), where low attachment isn't possible, compared to conditions (5a/5b/6a/6b), where it is. Alternatively, if L1ers randomly attach either low or high, as predicted by the unrestricted race model (e.g., Van Gompel et al., 2000), reading times should be longer in disambiguating conditions (5b/5c/6b/6c) compared to ambiguous (5a/6a). Such findings would support the "ambiguity advantage". If L1ers consider both low and high attachment in parallel, longer reading times should be expected in (5a/6a) than (5b/5c/6b/6c), due to the competition between the two possible interpretations (e.g., MacDonald, 1994; Tabor & Tanenhaus, 1999).

If L1 and L2 processing are qualitatively similar, the L2ers should behave like the L1ers. On the contrary, if they are different, L2ers should show different reading time patterns to the L1ers. For example, if L2ers do not show any clear preferences between low and high attachment (e.g., Felser et al., 2003b), then (5b/5c) and (6b/6c) should not differ. If this effect results from L2ers exhibiting variable attachment, then (5b/5c/6b/6c) should have longer reading times than ambiguous (5a/6a). Alternatively, if L2ers do not clearly resolve the RC at all during online processing, they may show no significant differences between any conditions.

Additionally, our study addresses another important gap in the literature by examining individual differences in this domain. To our knowledge, all other existing studies examining individual differences have tested either L1 speakers or L2 speakers but not both within the

same study. If L1 and L2 processing are similar, we would expect that individual differences should impact L1 and L2 processing similarly. As such, we examined how individual differences in reading span and lexical automaticity affect both offline and online ambiguity resolution in L1ers and L2ers. Based on prior findings, we predicted both L1/L2 individuals with higher reading spans should prefer low attachment more than those with low span scores for globally ambiguous sentences in offline comprehension (e.g., Hopp, 2014; Swets et al., 2007). For online processing, we expected L2ers with higher levels of lexical automaticity to behave more nativelike, and should show a stronger low attachment preference (Hopp, 2014). If L1 and L2 processing are qualitatively similar, lexical automaticity may influence L1 and L2 processing in a similar way. Finally, we also expected L2ers with higher proficiency to exhibit a more native-like pattern during attachment resolution (Dekydtspotter et al., 2008).

Method

Participants

The experiment was conducted with 66 English L1 speakers and 66 English L2 speakers with a first language that is reported to have a high attachment preference. This included Spanish, Italian, German, Dutch, French, Russian, Portuguese, Greek and Arabic. All participants were recruited from the University of Reading and surrounding areas. L1 participants were university students who participated in the study for course credit or a small payment. The L2 speakers were either students or were working in the local community at the time of testing. Their English proficiency was measured by a short version of the Oxford Placement Test. Their proficiency scores ranged from 30-60 out of 60 (mean = 48, SD = 0.93), representing intermediate to advanced level. All participants had normal or corrected to normal vision.

Materials

We combined an offline judgement and an online reading task to test participant's attachment preferences. The offline task consisted of 20 experimental items as in (7) and 40 fillers, which were pseudo-randomised in a latin-square design. We tested participants' offline attachment preferences for ambiguous RCs like (7a/b) in cases where the RC is in subject or object position.

(7a) Subject-Modifying RC

The brother of the man who often bought himself some books got married yesterday.

(7b) Object-Modifying RC

We met the brother of the man who often bought himself some books.

Question: Who bought himself some books?

Answer options: The brother/The man

Participants only saw one condition of each pair and, therefore, read 10 sentences like (7a) and 10 sentences like (7b). To indicate their attachment preferences, participants needed to answer comprehension questions such as "Who bought some books?" by choosing one of the two individuals shown in those sentences (i.e., either "the brother" or "the man"). The order of the options was counterbalanced. The choice of high attachment was coded as value 0 and low attachment was coded as 1. As such, a value towards 1 indicates a low attachment preference.

In the online reading task, we monitored participant's eye-movements as they read a series of texts. The materials for this task consisted of 36 experimental items like (5/6) and 80 fillers, randomised in a latin-square design. The experimental items contained a critical first sentence and a wrap-up sentence. The experimental items manipulated the position of the RC

such that the RC was either in subject position (5a-c) or object position (6a-c). Half of the items contained a masculine reflexive ("himself") and half a feminine reflexive ("herself"). The temporarily ambiguous RCs were disambiguated at the reflexive via gender match between the reflexive and the local or non-local NP. (5a/6a) are globally ambiguous as the reflexive matches the gender of both NPs. (5b/6b) are forced to attach low as the reflexive only matches the gender of the local NP whereas (5c/6c) are forced to attach high due to the gender match between the reflexive and the non-local NP. A full list of experimental materials for the offline and online tasks can be found in the supplementary materials.

We ran two additional tasks with both the L1 and L2 participants to investigate potential individual differences. To tap into working memory, a reading span task adapted from Daneman & Carpenter (1980) was administered to both groups. The participants read aloud sets of sentences presented one by one on a computer screen, as in (8) and (9). Half the sentences were grammatical and half ungrammatical. After each sentence participants judged whether it was grammatical or ungrammatical by pressing "1" or "2" on the keyboard. In the meantime, they had to memorise the final word of each sentence, underlined in (8) and (9). After the last sentence of a set, "RECALL" appeared onscreen and participants had to recall the to-be-remembered words, which were recorded by the experimenter. The set size increased from 2 sentences to 5 sentences as the experiment progressed. There were three sets at each set size, and participants completed all sets. The reading span score was calculated by dividing the total number of correctly recalled words by the total number of words that needed to be recalled.

(8) The young boy listened to music in his bedroom for <u>hours</u>. (Grammatical)

(9) *The old man picked up the phone his to spoke and <u>daughter</u>. (Ungrammatical)

A second individual differences task measured levels of lexical automaticity, using a lexical decision task adapted from Hopp (2014). There were 80 words in total, half of which were real English words and the other half were non-words following English phonotactics rules. The order in which the words were presented was randomised. Participants needed to decide as quickly as possible whether the word they saw was a real English word or not. They were instructed to rest two fingers from their preferred hand on the "1" and "2" keys, pressing "1" for real English words and "2" for non-words. Reaction times and accuracy were recorded. Following Hopp (2014), lexical automaticity was calculated by dividing the standard deviation of the reaction times to the real English words judged correctly by their average reaction times. For the L2 learners, a vocabulary screening task was administered at the end of the second session to test if they were familiar with the meaning (and gender) of the critical vocabulary (i.e. the noun phrases in the RCs).

Procedure

The study was conducted in two sessions at least 3 days apart. In the first session, participants completed the background questionnaire which provided information on language experience. This was followed by the main reading experiment where eye-movements were monitored by an SR Research Eyelink 1000 eye-tracker. Although viewing was binocular, the eye-movement record was recorded for the right eye only. The experiment began with a calibration procedure on a 9-point grid, and calibration was adjusted as needed between trials. The stimuli were presented onscreen in black letters. Before each sentence, a fixation marker appeared onscreen above the first word to be displayed. Upon fixating the marker, the sentence appeared. Participants were told to read as naturally as possible and to make sure they understood the sentences. All sentences were followed by a yes/no comprehension question that was answered by a push-button response. Comprehension questions did not probe attachment of the RC in

the critical sentences. Participants familiarised themselves with the procedure by first completing some practice trials before the main experiment. In the main experiment, experimental and filler items were pseudo-randomised in a latin-square design across six presentation lists that were completed by the same number of participants.

In the second session, participants first completed the reading span task and then the lexical decision task. Following that, they completed the offline attachment preference task. In addition, L2 speakers completed the proficiency test.

Data analysis

For the eye-tracking data, reading times were calculated at two regions of text. The reflexive region consisted of the critical reflexive, while the spillover region contained the rest of the clause (as exemplified in (5) & (6) using underlining). We calculated three reading time measures at each region. First-pass times summed the duration of fixations in a region entered from the left up until it was exited for the first time. Regression path times summed the duration of fixations starting when a region was first entered and up until but not including the first fixation in a region to the right. Total reading times refer to the duration of all fixations in a region regardless of when they occurred.

Trials in which a region was not fixated were treated as missing data. Fixations less than 80ms were combined with any neighbouring fixations if they were within one character of each other. All other fixations less than 80ms, as well as all those over 800ms, were removed. Due to a typographic error, the responses from one item in condition (6c) (object modifying RC, high attachment) were removed before analysis. Trials with excessive track loss were also removed, accounting for less than 1% of the data. Based on the vocabulary screening task, trials with critical vocabulary that the L2ers did not know were removed before analysis, accounting for less than 1% of the data.

Analysis was conducted using mixed-effects models (Baayen et al., 2008). For the offline task, a generalised mixed model was used containing sum coded (-1/1) fixed effects of Group (L1/L2), Position (subject RC/object RC) and their interaction. For the eye-movement data, reading times were log-transformed to minimise skew (see Vasishth & Nicenboim, 2016). Mixed models included sum coded fixed effects of Group (L1/L2) and Position (subject RC/object RC). The three-condition ambiguity manipulation involved two treatment-coded contrasts. One contrast (low attachment (LA)) compared the low attachment condition to the globally ambiguous condition, while the second contrast (high attachment (HA)) compared the high attachment condition to the globally ambiguous condition. In the case of any interactions between Position and the LA and/or the HA contrast, follow-up comparisons were conducted at the two levels of Position. For interactions with Group, additional analyses were conducted for each group separately.

All models were fit using the maximal random effects structure that converged (Barr, 2013; Barr et al., 2013). When the maximal model failed to converge, the random correlations were removed first. If the model still failed to converge, the random effect with the least variance was iteratively removed until the model converged.

We first conducted a main analysis as above to test for between-groups effects. We then conducted a series of additional analyses with three individual differences measures (working memory, lexical automaticity, and L2 proficiency) by adding each predictor separately into the maximal model, using the same method to achieve convergence. We analysed each individual differences measure separately to avoid issues related to multicollinearity. The three individual differences measures were included as centred, continuous predictors in each model, along with relevant interactions. The data and analysis code for our experiments is available at the Open Science Framework (OSF) website (https://osf.io/tvakf/).

Results

Individual differences measures and offline results

For reading span scores, the L1 (mean = 0.71, SD = 0.107) and L2 (mean = 0.72, SD = 0.117) groups did not differ significantly (estimate = 0.001, SE = 0.002, t = 0.801, p = .423), despite the fact that the task was presented in English. L1ers did however have significantly faster lexical automaticity than L2ers (L1 mean = 0.25, SD = 0.116; L2 mean = 0.34, SD = 0.155; estimate = 0.04, SE = 0.002, t = 16.41, $p < .001$).

The results from the offline task are shown in Table 1. The proportions here are descriptively all above 0.70 indicating a low attachment preference. Analysis revealed a significant effect of Position (estimate = 0.329, SE = 0.065, z = 4.99, $p < .001$), with the low attachment preference being stronger in object-modifying RCs than subject-modifying RCs in both groups. Neither the effect of group or interaction were significant (both z < 0.18, both p > 0.23).

Eye-tracking results

Accuracy to the comprehension questions was 95% for both groups (all participants scored above 82%), indicating all participants paid attention. A summary of the reading time data and statistical analysis is shown in Tables 2 and 3 respectively. For brevity, main effects of Group were found in each measure at each region, indicating slower reading times for the L2ers. We also do not discuss main effects of Position, or Position by Group interactions, below, as these are difficult to interpret on their own, but further interactions between Position, Group and LA or HA are informative about attachment preferences.

At the reflexive region, in first pass reading times we observed a significant interaction between Group and HA. Separate analyses on each group revealed that the L1 group showed significantly longer reading times for the high attachment than globally ambiguous condition

(estimate = 0.061, SE = 0.021, t = 2.92, *p* = .005), whereas the L2 group showed no significant differences (estimate = -0.007, SE = 0.02, t = -0.35, *p* = .728).

Moving onto regression path time, the HA effect was significant with longer reading times for high attachment RCs relative to ambiguous RCs. There were also numerical trends for an LA effect that differed across groups, with L1ers, but not L2ers, tending to show longer reading times in LA than ambiguous conditions. However, neither the LA effect nor the Group by LA interaction was significant.

In terms of total reading times, the HA effect was statistically significant with longer reading times for high attachment RCs than ambiguous ones. There was also a significant Group by HA interaction. Follow-up analyses revealed that the HA effect was present in both groups but with a larger effect in the L1 group (estimate = 0.221, SE = 0.034, t = 6.48, *p* < .001) compared to the L2 group (estimate = 0.12, SE = 0.03, t = 4.27, *p* < .001). This is illustrated in Figure 1.

At the spillover region, no significant effects were found for first pass reading time. For regression path time, there were three significant effects; the HA effect, the Group by HA interaction and the Position by HA interaction. To test the HA effect that was modulated by an interaction with Group, we conducted separate analyses for each group. These results found that both L1ers and L2ers had significantly longer reading times for high attachment RCs than the ambiguous baseline, but the effect was larger in the L1 group (estimate = 0.21, SE = 0.05, t = 4.63, *p* < .001) than the L2 group (estimate = 0.09, SE = 0.04, t = 2.18, *p* = .036). To examine the Position by HA interaction we tested HA effects in each position, collapsed across groups. These indicated that the HA effect was significant in both positions but with a smaller effect for subject position (estimate = 0.086, SE = 0.03, t = 3.33, *p* < .001) than object position (estimate = 0.231, SE = 0.045, t = 5.15, *p* < .001).

With respect to total reading times, a significant LA main effect was observed across the groups. Here, however, shorter reading times were found for low attachment RCs compared to globally ambiguous ones. The main effect of HA was also significant, with longer reading times for the high attachment condition than the baseline condition. This was, however, modulated by a significant three-way interaction between Group, Position and HA, which is illustrated in Figure 1. The follow-up analyses indicated that in subject position, The HA effect was significant in the L2 group (estimate = 0.105, SE = 0.038, t = 2.79, , p = .009), while the same numerical trend was not significant in the L1 group (estimate = 0.065, SE = 0.037, t = 1.73, p = .084). In object position, the HA effect was significant in the L1 group (estimate = 0.138, SE = 0.039, t = 3.53, p < .001), but not in the L2 group (estimate = 0.04, SE = 0.033, t = 1.19, p = .234).

Individual differences analysis: Offline task

Reading span scores, lexical automaticity and proficiency were included in separate models to examine their correlation with attachment preferences. In the reading span model, the main effect of Position was still significant (estimate = 0.337, SE = 0.071, z = 4.73, p < .001), as was the effect of reading span (estimate = 6.600, SE =1.553, z = 4.25, p < .001), in the absence of any further significant effects or interactions (all z < 0.12, all p > .24). The results here indicated that reading span scores were positively correlated with the low attachment preference in both groups (see Figure 2).

In the lexical automaticity model, apart from the significant effect of Position (estimate = 0.355, SE = 0.061, z = 5.76, p < .001), there was no significant effects or interactions (all z < 0.51, all p > 0.14). In the proficiency model, the effect of Position (estimate = 0.273, SE = 0.082, z = 3.31, p < .001) and the effect of proficiency (estimate = 0.085, SE = 0.028, z = 2.96, p < .01) were significant, but the Position by proficiency interaction was not significant

(estimate = 0.016, SE = 0.011, z = 1.54, *p* = .123). The proficiency model suggested language proficiency was positively correlated with the low attachment in the L2 group.

Individual differences analysis: Reading times

Lexical automaticity

At the reflexive region, there were numerical trends in first pass reading time for readers with high lexical automaticity to have shorter reading times, but the main effect of lexical automaticity was not significant (estimate = 0.237, SE = 0.121, t = 1.94, *p* = .052). We also observed numerical trends that the HA effect in both positions found in the main analysis in the L1 group was largely driven by those with high lexical efficiency, while the L2 group did not differ. However, the interaction between Group, Lexical Automaticity, Position and HA was not significant (estimate = -0.200, SE =0.103, t = -1.92, *p* = .054).

A significant three-way interaction between Lexical Automaticity, Position and HA was observed for total reading times at the reflexive region (estimate = 0.314, SE = 0.153, t = 2.04, *p* = .041). This is illustrated in Figure 3. We can see that in subject position, individuals with high and with low levels of lexical automaticity behaved quite similarly, whereas in object position, the HA effect seemed larger for individuals with lower levels of lexical automaticity compared to the highly automatized participants.

At the spillover region, we only found a numerical trend for first pass reading times to be longer in less automatized individuals. However, the main effect of Lexical Automaticity was not significant (estimate = 0.308, SE = 0.175, t = 1.76, *p* = .079).

Reading span and proficiency

Regarding individual differences in reading span, we did not observe any statistically significant main effects or interactions of theoretical interest (all t < 1.83, all p > .07). In terms of proficiency, no effects of theoretical interest were significant (all t <1.65, all p > .101).

General Discussion

This study aimed to investigate parsing strategies in L1 and L2 RC attachment, more specifically whether attachment preferences were influenced by RC position and/or individual differences. The results showed that both L1 and L2 groups demonstrated a clear low attachment preference, modulated by syntactic position of the RC in both offline and online tasks. We also observed some interactions between individual differences and offline/online attachment preferences. The implications of our results are discussed in more detail below.

Attachment preferences and sentence processing in native and non-native speakers

In the offline task, both L1 and L2 groups preferred low attachment over high attachment for both types of RCs. The L1 result is generally in line with the L1 literature. The L2 results, however, differ from some previous findings where L2ers from an L1 that prefers high attachment demonstrated null preferences (e.g., Felser et al., 2003b; Kim & Christianson, 2017; Omaki, 2005; Papadopoulou & Clahsen, 2003) in offline measures. Even though some previous studies (e.g., Felser et al., 2003b; Omaki, 2005) tested L2ers in an English-speaking country, their participants might differ from those included in this study regarding the amount of exposure, the degree of and opportunity for usage of English, or English proficiency. Regardless, our results indicated that L2ers did not significantly differ from the L1ers in their offline attachment preferences.

For online processing, both groups exhibited a low attachment preference even though they differed during certain stages of processing. An absent LA effect and a clear HA effect

were observed in both groups across several measures, especially total reading times, suggesting a low attachment preference in L1ers and L2ers (see Figure 1). However, there is some evidence showing that the HA effect was delayed and smaller in the L2ers, even though the effect was in the same direction for both groups. During first pass reading time, only L1ers showed the HA effect, suggesting that they preferred low attachment whereas L2ers did not show any preferences in this measure. Both groups showed the HA effect in total reading time at the reflexive region, regression path time and total reading time at the spillover region, although the effect was larger for the L1ers in some measures. We argue that this suggests slower but not qualitatively different processing for L2ers, who showed the same HA effect as L1ers, albeit delayed and numerically smaller.

The online results are generally in line with accounts which predict a low attachment preference in L1 English (e.g., Felser et al., 2003b; Hopp, 2014). Our data do not fully replicate the "ambiguity advantage" observed in previous studies (e.g., Traxler et al., 1998; Van Gompel et al., 2005). In regression path times at the reflexive, there were numerical trends for the ambiguous condition to have shorter reading times than both disambiguated conditions, though the LA comparisons were not fully reliable here. Across the majority of measures, although longer reading times were observed for high attachment conditions, the globally ambiguous conditions did not have significantly longer reading times than the low attachment conditions. This might be partially supportive of the "ambiguity advantage", in that this pattern of results does not suggest competition in the ambiguous conditions. However, we did find some evidence of competition in one measure, namely total viewing times at the spillover region, where the ambiguous condition had longer reading times than the low attachment condition. This might indicate an initial preference for low attachment with delayed competition between the two attachment sites at the spillover region. We do not draw any strong conclusions about the ambiguity advantage here, but note that we did not find any L1/L2 differences in this regard.

Thus, our clearest finding across measures was the low attachment preference for both L1 and L2 readers.

The L2 online results corroborate findings that L2ers can show online structural preferences (e.g., Bidaoui et al., 2016; Hopp, 2014). However, in Hopp (2014), the L2 group (as a whole) only showed the low attachment preference for sentences that are immediately disambiguated by a copula but not for sentences that are disambiguated later, as used in our study. Our study thus also indicates L2ers can exhibit online attachment preferences in sentences with later disambiguation. Our results stand in stark contrast with other studies showing that L2ers failed to show structural preferences (e.g., Felser et al., 2003b; Omaki, 2005; Papadopoulou & Clahsen, 2003), which could be due to methodological differences. Most L2 studies that failed to attest attachment preferences employed self-paced reading, which does not allow information from rereading and later processing stages to be captured. Also, with self-paced reading, individuals with poorer memory might have impoverished representations of the previous text when they reach the critical region, while eye-tracking allows for more naturalistic reading. Future research is required here to examine how these methodological issues may influence attachment resolution in L1 and L2 processing.

Recall that syntactic position interacted with offline attachment resolution as both groups preferred attaching low more often when the RC modified the object of the main verb compared to when it modified the subject. This suggests that the low attachment preference was stronger for object-modifying RCs in L1 and L2ers alike. It could be the case that since the high attachment site is the subject and always first-mentioned in the subject-modifying RC, it receives more discourse salience than the low attachment site, leading to a slightly weaker low attachment bias in the subject-modifying RC conditions. The offline findings are inconsistent with Hemforth et al. (2015) and Kim & Christianson (2017), who did not find

effects of position in English RCs. However, both these studies had smaller samples than the current study (48 and 34 participants respectively, compared to 132 in our experiment).

The position effect was also attested during online processing. In regression path times at the spillover region, both groups showed larger HA effects in object-modifying RCs compared to subject-modifying RCs, suggesting a stronger low attachment preference for object-modifying RCs in both groups. We propose that high attachment in object-modifying RCs was more demanding due to the attenuated prominence of the high attachment site in these sentences. This interplay of syntactic position and attachment choices across the groups suggests that discourse-level information was processed similarly in L1 and L2 readers.

Individual differences in attachment preferences

Working memory as measured by reading span interacted with offline attachment preferences in both the L1 and L2 groups. The offline results suggested that participants with high span preferred low attachment more than those with low span, which is consistent with most previous offline studies (e.g., Kim & Christianson, 2012; Omaki, 2005; Swets et al., 2007; but see Felser et al., 2003a). However, our study did not aim to tease apart different accounts of this correlation, and we cannot distinguish between whether this suggests different prosodic processing strategies (e.g., Swets et al., 2007) or resource limitations (e.g., Omaki, 2005). Importantly for present purposes, we did not find significant differences between L1ers and L2ers in this regard, suggesting that individual differences in reading span influence L1 and L2 processing in a similar way. We also found L2 proficiency interacted with offline attachment preferences, suggesting L2ers with higher proficiency attached low more often than those with lower proficiency. The finding on proficiency is compatible with the online finding from Dekydtspotter et al. (2008) that participants with higher levels of proficiency showed a native-like attachment preference.

For online processing at the reflexive region, there was a trend during first-pass times that L1ers with fast lexical access exhibited the HA effect, but not L1ers with slower lexical efficiency or L2ers. This might suggest that lexical processing efficiency influences whether readers show attachment effects in early stages of processing, and it is compatible with claims that some aspects of lexical processing need to be complete before syntactic integration begins (e.g., Hopp, 2016; Tily et al., 2010).

During later stages of processing, lexical automaticity modulated L1 and L2 processing to a similar extent. Both highly and less automatized L1/L2 individuals preferred low attachment. However, differences appeared with regards to the high attachment condition particularly in object-modifying RCs, where the difficulty associated with high attachment was reduced in those individuals with higher lexical automaticity (see Figure 3). As discussed above, forcing high attachment RCs in object position seemed particularly difficult to parse, indexed by the stronger low attachment preference in object position from our offline and online findings. Hence, it could the case that individuals with faster lexical access overcame the linguistic conflict and completed reanalysis more quickly compared to individuals with slower access, when processing difficulty increased. Taken together, our results suggest that faster lexical access can facilitate attachment resolution in both L1 and L2 processing.

Our findings in terms of lexical automaticity are not entirely consistent with Hopp (2014). For sentences that were not immediately disambiguated, as tested in our study, Hopp (2014) found that attachment preferences only emerged for L2ers with high levels of lexical automaticity, while L2ers with low lexical automaticity didn't show any significant differences between conditions. However, our results present a different pattern where the individuals with lower levels of lexical automaticity exhibited larger effects. These apparently opposing findings might be due to differences at the group level. Specifically, the L2 group as a whole in Hopp (2014) did not show any attachment preferences for non-locally disambiguated

sentences, while the L2 group in our study did. It could thus be that the L2ers in Hopp (2014) did not have as efficient lexical processing as those tested in our study at the group level. This would not be so odd when one considers the context of testing. Whereas Hopp's participants were tested in a non-native English environment (Germany) our participants were tested in native immersion (the United Kingdom).

Our online findings did not replicate our offline working memory and proficiency effects or any previously reported effects in relation to working memory or proficiency. The null effects of reading span during online processing fail to replicate some previous findings (e.g., Payne et al., 2014; Traxler, 2007), but they are consistent with Hopp (2014), who also did not observe effects of reading span during online reading. One possibility is that the spread of reading span scores in our study was less than in Payne et al., (2014) and Traxler (2007), potentially making reading span effects more difficult to observe. Note however that our combined L1/L2 sample is larger than previous studies. Therefore, sample size may also play a role in these mixed findings. Further replications with large samples are required to tease these issues apart.

Finally, although we tested only L2ers whose L1 has previously been reported to have a high attachment preference, our L2 group did comprise speakers from different L1 backgrounds. The two largest subgroups included L1 Italian speakers (n = 16) and L1 Spanish speakers (n = 11), while smaller subgroups of below 10 participants included L1 Dutch, German, French, Portuguese, Russian, Greek and Arabic speakers. Since these groups are not large enough for inferential analysis, especially for the eye-tracking data, we only descriptively examined how each subgroup behaved in the offline task. Here, all subgroups descriptively indicated a low attachment preference by choosing low attachment more than 65% of the time in the offline task. Thus, we do not believe a single subgroup was responsible for our findings. However, given differences in the number of participants across the subgroups and in the

strength of attachment biases across languages, this may have affected our results. Hence, future research will need to examine how specific L1 backgrounds may influence L2 attachment resolution.

Conclusion

We argue that our results suggest quantitative rather than qualitative differences between L1ers and L2ers in attachment resolution. Both groups showed similar parsing preferences online, though the effects were smaller and delayed in L2ers. Also, our results suggest that the syntactic position of the RC influences ambiguity resolution, which we argued results from differences in discourse salience between RCs in subject or object position. Importantly, this effect of discourse salience appeared to modulate L1 and L2 processing in a similar way. L1 and L2 attachment resolution also interacted with individual differences in a largely similar way across groups, with better lexical processing efficiency being related to quicker linguistic conflict resolution in both L1 and L2 processing. In sum, our results suggest similar processing strategies in ambiguity resolution are possible in L1 and L2 comprehension, and highlight the importance of considering both linguistic factors and individual differences when examining the similarities and differences between L1 and L2 sentence processing.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*, 328.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Bidaoui, A., Foote, R., & Abunasser, M. (2016). Relative Clause Attachment in Native and L2 Arabic. *International Journal of Arabic Linguistics*, *2*(2), 75-95.

Brysbaert, M., & Mitchell, D. C. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *The Quarterly Journal of Experimental Psychology Section A*, *49*(3), 664-695.

Carreiras, M., & Clifton, C. (1999). Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory & Cognition*, *27*(5), 826-833.

Carreiras, M., Salillas, E., & Barber, H. (2004). Event-related potentials elicited during parsing of ambiguous relative clauses in Spanish. *Cognitive Brain Research*, *20*(1), 98-105.

Chernova, D., & Chernigovskaya, T. V. (2015). Syntactic Ambiguity Resolution in Sentence Processing: New Evidence from a Morphologically Rich Language. In *EAPCogSci*.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(1), 3-42.

Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, *40*(3), 693-706.

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

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

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

Dekydtspotter, L., Donaldson, B., Edmonds, A. C., Fultz, A. L., & Petrush, R. A. (2008). Syntactic and prosodic computations in the resolution of relative clause attachment ambiguity by English-French learners. *Studies in Second Language Acquisition*, *30*(4), 453-480.

Desmet, T., Brysbaert, M., De Baecke, C. (2002). The correspondence between sentence production and corpus frequencies in modifier attachment. *The Quarterly Journal of Experimental Psychology: Section A*, *55*(3), 879-896.

Desmet, T., De Baecke, C., Drieghe, D., Brysbaert, M., & Vonk, W. (2006). Relative clause attachment in Dutch: On-line comprehension corresponds to corpus frequencies when lexical variables are taken into account. *Language and Cognitive Processes*, *21*(4), 453-485.

De Vincenzi, M., & Job, R. (1993). Some observations on the universality of the late-closure strategy. *Journal of Psycholinguistic Research*, *22*(2), 189-206.

Dinçtopal-Deniz, N. (2010). Relative clause attachment preferences of Turkish L2 speakers of English. *Research in Second Language Processing and Parsing*, *53*, 27.

Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learners: Some effects of binguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition*, *25*(4), 529-557.

Felser, C., Marinis, T., & Clahsen, H. (2003a). Children's processing of ambiguous sentences: A study of relative clause attachment. *Language Acquisition*, *11*(3), 127-163.

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003b). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, *24*(3), 453-489.

Fernández, E. M. (2003). *Bilingual Sentence Processing: Relative clause attachment in English and Spanish*. Amsterdam: Benjamins.

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. *North East Linguistic Society*, *32*, 113–132.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies.* Unpublished doctoral dissertation, University of Connecticut.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559–586). Hillsdale, NJ: Erlbaum.

Gibson, E., Pearlmutter, N., Canseco-González, E., & Hickok, G. (1996). Cross-linguistic attachment preferences: Evidence from English and Spanish. *Cognition*, *59*, 23–59.

Grüter, T., & Rohde, H. (2013). L2 processing is affected by RAGE: Evidence from reference resolution. In *the 12th conference on Generative Approaches to Second Language Acquisition (GASLA)*.

Hemforth, B., Fernandez, S., Clifton Jr, C., Frazier, L., Konieczny, L., & Walter, M. (2015). Relative clause attachment in German, English, Spanish and French: Effects of position and length. *Lingua*, *166*, 43-64.

Hemforth, B., Konieczny, L., Seelig, H., & Walter, M. (2000). Case matching and relative clause attachment. *Journal of Psycholinguistic Research*, *29*(1), 81-88.

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

Hopp, H. (2016). The timing of lexical and syntactic processes in second language sentence comprehension. *Applied Psycholinguistics*, *37*(5), 1253-1280.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*(4), 603-634.

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different?. *Linguistic Approaches to Bilingualism*, *4*(2), 257-282.

Kim, J. H., & Christianson, K. (2013). Sentence complexity and working memory effects in ambiguity resolution. *Journal of Psycholinguistic Research*, *42*(5), 393-411.

Kim, J. H., & Christianson, K. (2017). Working memory effects on L1 and L2 processing of ambiguous relative clauses by Korean L2 learners of English. *Second Language Research*, *33*(3), 365-388.

Maia, M., Costa, A., Fernández, E., & Lourenço-Gomes, M. D. C. (2004). A compreensão de orações relativas ambíguas em Português Brasileiro e Europeu: Um estudo comparativo. *Revista ABRALIN*, *3*(1), 2.

Maia, M., Fernández, E. M., Costa, A., & Lourenço-Gomes, M. D. C. (2007). Early and late preferences in relative clause attachment in Portuguese and Spanish. *Journal of Portuguese Linguistics*, *5*(6), 227-250.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*(2), 157-201.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*(3), 381-401.

Omaki, A. (2005). *Working memory and relative clause attachment in first and second language processing* (Doctoral dissertation).

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

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*(4), 501-528.

Payne, B. R., Grison, S., Gao, X., Christianson, K., Morrow, D. G., & Stine-Morrow, E. A. (2014). Aging and individual differences in binding during sentence understanding: Evidence from temporary and global syntactic attachment ambiguities. *Cognition*, *130*(2), 157-173.

Rothman, J. (2010). On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics in Language Teaching*, *48*(2-3), 245-273.

Scheepers, C., Sturt, P., Martin, C. J., Myachykov, A., Teevan, K., & Viskupova, I. (2011). Structural priming across cognitive domains: From simple arithmetic to relative-clause attachment. *Psychological Science*, *22*(10), 1319-1326.

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, *136*(1), 64.

Tabor, W., & Tanenhaus, M. K. (1999). Dynamical models of sentence processing. *Cognitive Science*, *23*(4), 491-515.

Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition*, *17*(2), 277-293.

Tily, H., Fedorenko, E., & Gibson, E. (2010). The time-course of lexical and structural processes in sentence comprehension. *The Quarterly Journal of Experimental Psychology*, *63*(5), 910-927.

Traxler, M. J. (2007). Working memory contributions to relative clause attachment processing: A hierarchical linear modeling analysis. *Memory & Cognition*, *35*(5), 1107-1121.

Traxler, M. J., Pickering, M. J., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, *39*(4), 558-592.

Van Gompel, R. P., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *52*(2), 284-307.

Van Gompel, R. P., Pickering, M. J., & Traxler, M. J. (2000). Unrestricted race: A new model of syntactic ambiguity resolution. In A. Kennedy (Ed.), *Reading as a Perceptual Process* (pp. 621-648). Amsterdam, The Netherlands: Elsevier.

Vasishth, S. & Nicenboim, B. (2016). Statistical methods for linguistic research: Foundational Ideas – Part I. *Language and Linguistics Compass*, *10*, 349-369.

Witzel, J., Witzel, N., & Nicol, J. (2012). Deeper than shallow: Evidence for structure-based parsing biases in second-language sentence processing. *Applied Psycholinguistics*, *33*(2), 419-456.

Zagar, D., Pynte, J., & Rativeau, S. (1997). Evidence for early closure attachment on first pass reading times in French. *The Quarterly Journal of Experimental Psychology Section A*, *50*(2), 421-438.

Table 1. *Low attachment preferences in the offline task (standard errors in parentheses)*

	Native Speakers	Non-Native Speakers
Subject Modifying RC	0.73 (0.01)	0.77 (0.01)
Object Modifying RC	0.82 (0.01)	0.82 (0.01)

Table 2. *Summary of reading times in milliseconds (standard errors in parentheses)*

	First Pass Time		Regression Path Time		Total Viewing Time	
	Native Speakers	Non-Native Speakers	Native Speakers	Non-Native Speakers	Native Speakers	Non-Native Speakers
Reflexive Region						
Subject Modifying RC, Globally Ambiguous	226 (5)	282 (7)	266 (10)	366 (18)	389 (13)	481 (15)
Subject Modifying RC, Low Attachment	241 (5)	275 (6)	311 (16)	351 (19)	397 (15)	470 (15)
Subject Modifying RC, High Attachment	251 (7)	275 (7)	326 (17)	387 (17)	501 (19)	578 (20)
Object Modifying RC, Globally Ambiguous	235 (6)	271 (6)	327 (20)	339 (13)	380 (12)	491 (17)
Object Modifying RC, Low Attachment	239 (5)	286 (7)	382 (45)	352 (13)	406 (15)	491 (17)
Object Modifying RC, High Attachment	251 (7)	276 (6)	379 (29)	355 (13)	523 (19)	543 (20)
Spillover Region						
Subject Modifying RC, Globally Ambiguous	413 (11)	485 (12)	496 (25)	555 (16)	757 (30)	849 (28)
Subject Modifying RC, Low Attachment	397 (11)	488 (13)	522 (27)	587 (22)	735 (27)	818 (25)
Subject Modifying RC, High Attachment	398 (11)	488 (12)	598 (35)	614 (25)	803 (28)	924 (27)
Object Modifying RC, Globally Ambiguous	428 (14)	523 (15)	990 (55)	1292 (70)	654 (23)	809 (25)
Object Modifying RC, Low Attachment	412 (14)	507 (14)	997 (67)	1226 (63)	612 (23)	765 (25)
Object Modifying RC, High Attachment	418 (14)	504 (14)	1628 (103)	1532 (88)	753 (29)	815 (24)

Table 3. *Summary of the statistical analysis*

	Reflexive Region			Spillover Region		
	Estimate (SE)	*t*	*p*	*Estimate (SE)*	*t*	*p*
First Pass Time						
Group	**0.085 (0.017)**	**5.19**	**< .001**	**0.094 (0.023)**	**4.08**	**< .001**
Position	0.001 (0.009)	0.18	.856	0.009 (0.014)	0.68	.494
LA	0.023 (0.013)	1.73	.083	-0.026 (0.015)	-1.68	.092
HA	0.026 (0.013)	1.92	.055	-0.017 (0.016)	-1.08	.277
Group * Position	-0.011 (0.010)	-1.06	.290	0.014 (0.014)	0.99	.321
Group * LA	-0.017 (0.013)	-1.26	.206	0.024 (0.016)	1.48	.139
Group * HA	**-0.034 (0.014)**	**-2.45**	**.016**	0.017 (0.016)	1.04	.299
Position * LA	0.001 (0.013)	0.03	.970	-0.004 (0.015)	-0.25	.798
Position * HA	0.001 (0.013)	0.04	.967	0.002 (0.016)	0.14	.883
Group * Position * LA	0.017 (0.013)	1.25	.209	-0.003 (0.015)	0.20	.835
Group * Position * HA	0.013 (0.013)	0.98	.324	-0.012 (0.016)	0.75	.447
Regression Path Time						
Group	**0.092 (0.025)**	**4.21**	**< .001**	**0.108 (0.028)**	**3.79**	**< .001**
Position	0.016 (0.015)	1.10	.271	**0.277 (0.026)**	**10.57**	**< .001**
LA	0.033 (0.019)	1.66	.097	-0.001 (0.022)	-0.03	.975
HA	**0.071 (0.020)**	**3.42**	**<.001**	**0.159 (0.023)**	**6.90**	**< .001**
Group * Position	**-0.028 (0.013)**	**-2.09**	**.035**	0.021 (0.025)	0.84	.402
Group * LA	-0.039 (0.019)	-1.96	.0504	0.003 (0.022)	0.17	.868
Group * HA	-0.028 (0.020)	-1.34	.179	**-0.064 (0.023)**	**-2.82**	**.004**
Position * LA	0.002 (0.019)	0.11	.907	-0.023 (0.022)	-1.04	.299
Position * HA	-0.012 (0.019)	-0.65	.513	**0.074 (0.024)**	**3.04**	**.003**
Group * Position * LA	0.024 (0.020)	1.20	.229	0.004 (0.022)	0.18	.855
Group * Position * HA	0.005 (0.022)	0.23	.816	-0.036 (0.023)	-1.61	.108
Total Viewing Time						
Group	**0.119 (0.027)**	**4.30**	**< .001**	**0.102 (0.029)**	**3.50**	**< .001**
Position	-0.002 (0.013)	-1.15	.881	**-0.041 (0.014)**	**-2.84**	**.004**
LA	0.007 (0.019)	0.39	.694	**-0.039 (0.017)**	**-2.23**	**.025**
HA	**0.167 (0.022)**	**7.61**	**< .001**	**0.084 (0.020)**	**4.10**	**< .001**
Group * Position	-0.001 (0.013)	-0.08	.935	0.026 (0.014)	1.83	.066
Group * LA	-0.025 (0.019)	-1.29	.197	0.007 (0.017)	0.40	.683
Group * HA	**-0.052 (0.022)**	**-2.37**	**.019**	-0.015 (0.019)	-0.82	.408
Position * LA	0.019 (0.019)	0.99	.845	-0.020 (0.017)	-1.17	.240
Position * HA	0.004 (0.022)	0.20	.828	0.001 (0.017)	0.05	.958
Group * Position * LA	-0.004 (0.019)	-0.22	.166	0.004 (0.017)	0.22	.818
Group * Position * HA	-0.028 (0.021)	-1.40	.324	**-0.036 (0.017)**	**-2.05**	**.040**

Note: LA = Low Attachment, HA = High Attachment

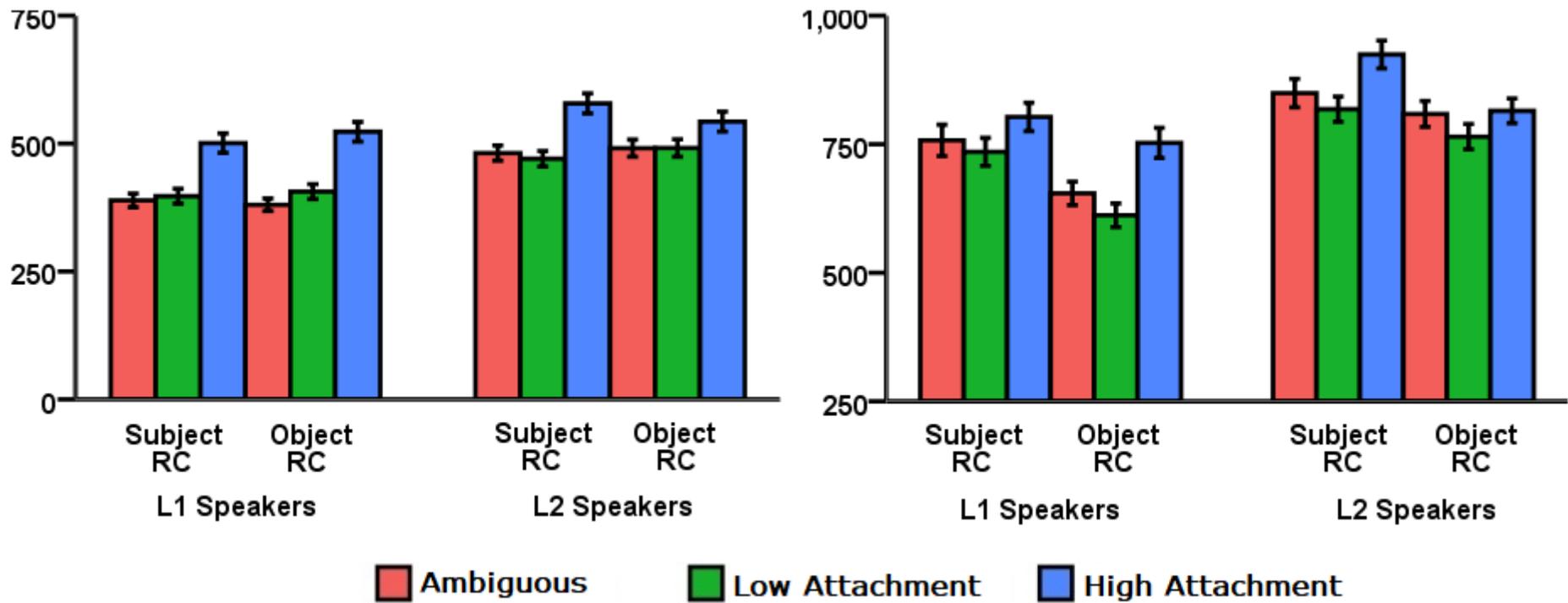Figure 1. *Total viewing times at the reflexive region (left) and spillover region (right)*

Figure 2. *Interaction between reading span and attachment preferences for each position and group.*

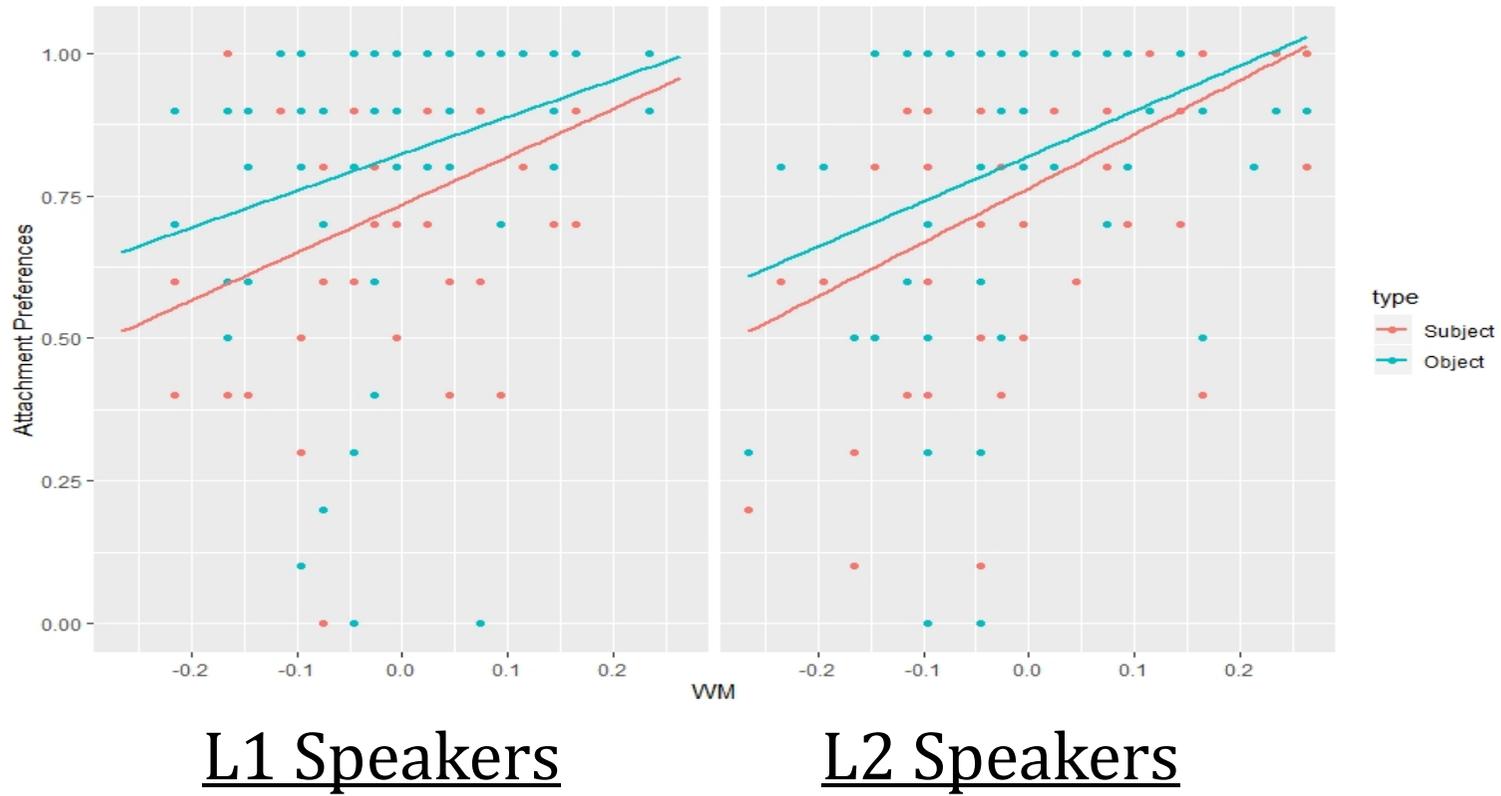## L1 Speakers   L2 Speakers

Figure 3. *Interaction between lexical automaticity and total reading times for each position across groups.*