Applied Psycholinguistics (2020), **41**, 963–992
doi:10.1017/S0142716420000314

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

The processing of object–subject ambiguities in early second-language acquirers

Robyn Berghoff

Stellenbosch University
Corresponding author. Email: berghoff@sun.ac.za

(Received 26 March 2019; revised 24 April 2020; accepted 27 April 2020; first published online 09 July 2020)

Abstract
Studies of the second language (L2) processing of object–subject ambiguities (OSA), where the parser must revise an initially incorrect analysis of the input, have yielded two findings that the present paper addresses. First, L2 speakers are guided more strongly by plausibility information than L1 speakers in OSA processing. Second, L2 speakers generally do not perform reanalysis in real time, and their comparably poor performance on subsequent comprehension questions suggests ultimate failure to derive a grammatical parse. Because previous studies have targeted later L2 acquirers living outside of an L2 context, the present study tested whether the above two findings hold for childhood L2 acquirers with extensive naturalistic L2 exposure, hypothesizing that these factors would engender more nativelike processing. A self-paced reading task involving two kinds of OSA construction was conducted. Sensitivity to plausibility information differed across the first language and L2 groups per construction type. Within the L2 group, there is evidence that the early childhood acquirers initiated online reanalysis. Comprehension question accuracy did not differ across participant groups, suggesting that all participants succeeded in reanalysis to the same extent. The role of age of onset in L2 processing is discussed, as is the findings' relevance to current models of L2 processing.

Keywords: age of acquisition; ambiguity resolution; garden-path parsing; second language processing; subcategorization information

One of the necessary components of language acquisition is developing the ability to process linguistic input in real time. As syntactic representations of incoming input are constructed under conditions of uncertainty, without access to the entire sentence being available, language processing also occasionally demands on-the-fly reanalysis of incorrect interpretations. This is challenging, and sometimes fails even during first language (L1) processing (Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Patson, Darowski, Moon, & Ferreira, 2009). Examinations of mis- and reanalysis in second language (L2) processing therefore have much to tell us about the nature of the L2 system.

To date, a number of such examinations have been conducted with late L2 acquirers, and on the whole have observed marked differences between L1 and L2 processing of temporarily ambiguous sentences (Jacob & Felser, 2016; Juffs, 1998a, 1998b; Juffs & Harrington, 1996; Roberts & Felser, 2011; for an overview, see Papadopoulou, 2005). The processing of such constructions in childhood L2 acquirers has however received little attention. Models of L2 processing in which age of L2 acquisition (AoA) plays a significant role, for example, Clahsen and Felser's (2006a, 2006b, 2006c, 2017) Shallow Structure Hypothesis, predict native-like processing to be more likely among such L2 speakers. The Shallow Structure Hypothesis also allows for different sensitive periods for different processing phenomena, such that age effects may be observed even among childhood learners (see, e.g., Bosch, Veríssimo, & Clahsen, 2019; Veríssimo, 2018; Veríssimo, Heyer, Jacob, & Clahsen, 2018). In light of these considerations, the present study set out to determine whether childhood L2 acquirers process temporarily ambiguous input similarly to L1 speakers, and furthermore whether a distinction can be observed between the processing patterns of early and late childhood L2 acquirers.

L1 and L2 processing of object–subject ambiguities

During language comprehension, the individual must not only access the meaning of the lexical items they encounter but also compute a representation of the syntactic structure of the input via syntactic parsing. It is well established that syntactic representations are constructed incrementally, that is, upon encountering each subsequent word, the parser attempts to integrate it into the current representation (Marslen-Wilson, 1973). A phenomenon that has been used to investigate the incremental nature of parsing and its interaction with other contributors to linguistic interpretation is the garden-path sentence. Consider (1): the comma following *drank* clearly indicates that the noun phrase (NP) *the beer* marks the beginning of a separate clause, and so parsing presents minimal difficulty.

(1) While the men drank, the beer pleased all the customers.
(2a) While the men drank the beer pleased all the customers.
(2b) While the men sang the beer pleased all the customers.

In (2a), readers generally prefer an analysis in which *the beer* is the direct object of *drank*. Different sentence comprehension models explain this preference in different ways. The garden-path model holds that incoming material is preferentially attached inside the current clause (Frazier & Fodor, 1978). Constraint-based accounts attribute this preference to the fact that *drink* is more frequently transitive than intransitive and new clauses are typically signaled by disambiguating punctuation (MacDonald, Pearlmutter, & Seidenberg, 1994; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993). In online studies, the preference for the direct object analysis is reflected by longer reading times (RTs) at *pleased*, which is the earliest point in the sentence that serves to reject a direct object analysis of the preceding NP (Frazier & Rayner, 1982; Warner & Glass, 1987). What happens in the parsing of (2a) is that

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

the parser is led down the metaphorical garden path, initially favoring an analysis of the input that later turns out to be incorrect.

That parsing interacts with other linguistic aspects is indicated by the fact that the strength of the garden-path effect in (2) is modulated by how likely the temporarily ambiguous NP is as a direct object of *drank*. As such, the parser commits less strongly to an analysis in which *the beer* is the direct object of *sang* (cf. [2b]). Evidence in support of this claim is provided in an eye-tracking study by Pickering and Traxler (1998). Here, participants were found to have greater difficulty processing NPs that were implausible as direct objects of the preceding verb, as indicated by a larger number of regressions in the NP region in the implausible condition. The reverse effect was evident at the disambiguating region, where more processing difficulty was observed in the plausible condition. The latter finding indicates that the parser commits more strongly to a plausible than an implausible analysis, leading to greater processing difficulty when an initially plausible analysis must be abandoned.

Regarding the mechanisms involved in recovery from misanalysis, different models of sentence comprehension again make different proposals. One account holds that reranking of a selection of possible parses occurs, such that the initial parse is demoted and a grammatical alternative is promoted (Spivey & Tanenhaus, 1998), and another that the initial parse is repaired (Frazier & Clifton, 1998). In (2a), repair would entail detaching the NP from the verb phrase headed by *drank*, reanalyzing it as the subject of the main clause, and making the necessary case and thematic role modifications. In instances where the parser was strongly committed to the initial analysis, reanalysis may be sufficiently taxing that the reader is unable to abandon the incorrect parse completely (Christianson et al., 2001; Patson et al., 2009).

Although investigations of garden-path processing provide useful information on how the parser deals with misleading constructions, their utility also extends beyond this. By examining the (re)analysis processes and ultimate outcomes of garden-path comprehension, insight can be gained into the mechanisms employed in "normal" sentence comprehension as well. For this reason, studies of garden-path processing have proven valuable in comparisons between L1 and L2 sentence processing. Because they shed light on how parsing proceeds in each group, such studies contribute to addressing the central question of whether and under what circumstances L1 and L2 processing can converge. In this regard, three central findings have emerged.

The first is that, like L1 processing, L2 processing is also incremental. When processing temporarily ambiguous sentences, L2 speakers have been found to show sensitivity to the error signal provided by the disambiguating region (Jackson, 2008; Jacob & Felser, 2016; Jegerski, 2012; Juffs, 1998a, 1998b; Juffs & Harrington, 1996).

The second finding is that L2 parsing appears to be more sensitive to semantic and probabilistic information than L1 parsing. This is suggested by Roberts and Felser (2011), who examined the processing of two types of garden-path construction in L1 and L2 English (L1 Greek) speakers. The first type of construction had the form exemplified in (2). The second type, in which the temporarily ambiguous NP marks the beginning of a complement clause, is illustrated in (3).

(3a) The boy wrote the report would start an important debate.
(3b) The boy thought the report would start an important debate.

Roberts and Felser's (2011) L2 group showed significantly shorter RTs in the NP region when the NP was a plausible rather than an implausible object of the directly preceding verb in both construction types, whereas this effect only surfaced weakly in the complement-clause constructions for the native speakers. The heightened sensitivity to plausibility information observed in the nonnative group aligns with accounts of L2 processing (e.g., the Shallow Structure Hypothesis; Clahsen & Felser, 2006a, 2006b, 2006c, 2017), which propose that nonnative speakers are generally more reliant on semanticopragmatic information and underuse morphosyntactic cues in comparison with native speakers.

The third finding is that L2 speakers are typically less successful than L1 speakers in arriving at a correct analysis of temporarily ambiguous input. In terms of initiating reanalysis at the disambiguating region of a temporarily ambiguous construction, self-paced reading and eye-tracking studies have found that L2 speakers either do not attempt reanalysis or abandon attempts at reanalysis earlier than L1 speakers (Jacob & Felser, 2016; Roberts & Felser, 2011). In Roberts and Felser (2011) specifically, the L2 group only showed evidence of initiating reanalysis in the complement-clause constructions. This finding aligns with those of previous studies that have observed minimal processing difficulty in garden-path sentences of this type (Holmes, Kennedy, & Murray, 1987; Sturt, Pickering, & Crocker, 1999). The reduced reanalysis cost in constructions like (3) has been attributed to the fact that the temporarily ambiguous NP remains within the thematic domain of the preceding verb in the grammatical parse, and so the structural revisions required to derive this parse are relatively minor (Pritchett, 1992). In terms of the ultimate outcome of reanalysis processes—as tested by comprehension questions specifically probing the interpretation of the ambiguous region—L2 speakers have been found to perform more poorly than L1 speakers. This applies both to their responses to questions following ambiguous versus unambiguous sentences and to those following sentences where the direct object analysis had been plausible rather than implausible. Overall, their comparably poor performance suggests that L2 speakers are more susceptible to parsing breakdown in garden-path processing.

While the incremental nature of L2 parsing holds for L2 speakers in general, the other two findings reviewed above appear to vary in accordance with individual differences within L2 speakers. For one, in a study by Hopp (2015), those L2 speakers who scored higher on a measure of syntactic integration ability used both morphosyntactic and semantic information in their incremental analysis of the input, thus showing the more syntactically driven processing typically assumed to be characteristic of native speakers. In addition, in Roberts and Felser (2011), the faster L2 readers did seem able to initiate reanalysis online.

These results notwithstanding, variation *across* different L2 populations is comparatively underexamined within the sentence processing literature (although see Gerth, Otto, Felser, & Nam, 2017, for a cross-linguistic study of garden-path processing). Studies to date have primarily tested late L2 acquirers who have received limited naturalistic exposure to the L2, and little attention has been paid to the processing of early L2 acquirers who have received prolonged, extensive L2

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

exposure. With respect to online processing, recent findings indicate that different morphosyntactic phenomena may be subject to different sensitive periods (Veríssimo, 2018; see also Long, 1990; Werker & Hensch, 2015, for a similar idea with respect to phonological development). Further, regarding contact with the L2, it has been proposed that even if exposure commences outside of a particular sensitive period, "increased exposure may well lead to the increased automatization or entrenchment of grammatical processing routines" (Clahsen & Felser, 2017, p. 6). Recent findings suggest a correlation between exposure and processing nativelikeness: specifically, naturalistic L2 exposure seems to play a decisive role in engendering more syntactically rather than semantically driven L2 processing (Pliatsikas, Johnstone, & Marinis, 2017; Pliatsikas & Marinis, 2013). In light of these findings, it is plausible that early L2 acquirers who have received considerable naturalistic L2 exposure would show more nativelike garden-path processing. Given the observation of differing sensitive periods for different processing phenomena, however, age effects among these early L2 acquirers may also be observed.

The present study

The present paper reports on a self-paced reading study on garden-path processing in English. The participants were L1 Afrikaans–L2 English speakers and a control group of L1 English speakers. The study's primary aim is to determine whether garden-path processing in the L1 and L2 groups is similar. A secondary aim is to investigate whether the L2 participants show more sensitivity to semantic information than the L1 participants do. The question here is whether prolonged naturalistic L2 exposure commencing at an early age serves to reduce sensitivity to semantic cues during L2 processing.

The experiment reported on is a modified version of Roberts and Felser (2011), which is in turn based on Pickering and Traxler (1998). It focuses on the influence of plausibility information on the (re)analysis of constructions involving object–subject ambiguities (OSAs), where an NP that is initially taken to be the object of a preceding verb turns out to be the subject of a separate clause. In this regard, as indicated above, Pickering and Traxler (1998) found that (a) readers have greater difficulty processing a syntactically ambiguous region when the analysis they adopt is implausible rather than plausible, and (b) readers commit less strongly to an implausible analysis, which entails reduced reanalysis difficulty at the disambiguating region of the sentence.

As in Pickering and Traxler (1998) and Roberts and Felser (2011), the present study considers two kinds of OSA constructions. The first has the form exemplified in (4a), where a preposed adjunct clause is followed by a temporarily ambiguous NP that is the subject of the main clause. In the second type, exemplified in (4b), the temporarily ambiguous NP marks the beginning of a complement clause that is the argument of the verb. The plausibility of the direct object analysis of the temporarily ambiguous NP was manipulated across both construction types, such that the sections within parentheses in (4a) and (4b) represent the items in the implausible condition.

(4a) While the child climbed the ladder against the wall (the housewife in the kitchen) fell to the ground (adjunct-clause item).

(4b) The student wrote the report about the budget (the issue with the budget) would start an important debate (complement-clause item).

To enable a comparison between our results and those of previous studies, we employed the procedure used in Roberts and Felser (2011), as well as modified versions of their stimuli. The modification entailed extending the ambiguous region of the experimental items by adding a three-word phrase between the temporarily ambiguous noun and the disambiguating verb. Longer ambiguous regions have been found to result in stronger garden-path effects and thus greater processing difficulty (Christianson et al., 2001; Ferreira & Henderson, 1991; Tabor & Hutchins, 2004), and so Roberts and Felser (2011) shortened their ambiguous region to avoid placing undue demands on their L2 group. However, we assumed our highly proficient L2 participants would not have difficulty comprehending the garden-path sentences, and so we retained the modifying phrase used in Pickering and Traxler's (1998) stimuli. As such, we also extended the window during which a plausibility effect could be observed in the native-speaker group, allowing for the possibility that greater automaticity in L1 processing may delay responses to experimental manipulations (Kaan, Ballatyne, & Wijnen, 2015).

Our predictions were as follows: we expected greater similarity between our L1 and L2 groups' processing patterns, due to our L2 group's early L2 acquisition, prolonged and extensive L2 exposure, and high L2 proficiency. Specifically, we predicted (a) that the L2 group would not show heightened sensitivity to plausibility information compared to the L1 group; and (b) that the L2 speakers would show a reverse plausibility effect at the disambiguating region of both the adjunct-clause and complement-clause items, such that processing would be easier in the implausible than the plausible condition. The latter prediction was expected to result in similarly high accuracy scores on the comprehension questions in the L1 and L2 groups across both construction types.

Method

Participants

Thirty-three L1 Afrikaans–L2 English[1] speakers (mean age = 20.6 years, $SD = 1.98$ years) and 34 L1 English speakers (mean age = 20 years, $SD = 0.98$ years) participated in the study. The participants were resident in the Western Cape province of South Africa, where the languages with the largest proportions of L1 speakers are Afrikaans (approximately 50%), isiXhosa (approximately 25%), and English (approximately 20%; Census, 2011). Despite being the L1 of only approximately 9% of the population, English plays a dominant role in education, government, and commerce in South Africa (see, e.g., Posel & Zeller, 2016), with more than 80% of learners writing their secondary school exit examinations in English (Department of Basic Education, 2010; Plüddemann, 2015). Given the importance of English in education and South African society at large, the L2 participants had received extensive exposure to the language from a young age (on average, before age 5). At the time of testing, they were also enrolled at a university where the

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

majority of academic activity is conducted in English, and they reported receiving approximately equal exposure to both of their languages. As such, our L2 participants differ from those in Roberts and Felser (2011), who were first exposed to English in a classroom setting at an average age of 8 years and had spent an average of 2 years in an immersion environment. It should also be noted that our L1 group were also functionally bilingual, as they were raised in the bi-/multilingual South African context and at the time were enrolled at a bilingual university. Data on the L1 group's L2 exposure and proficiency is provided in Table A.1 in Appendix A.[2]

All participants had normal or corrected-to-normal vision. They received course credit for their participation in the experiment. The Language Experience and Proficiency Questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007) was administered to assess participants' language backgrounds, and a C-test consisting of three texts, each of which contained 20 gaps, was used to obtain a measure of English proficiency. Two of the texts in the C-test were developed by Keijzer (2007) in order to assess L1 attrition in L1 English speakers, and so are relatively difficult. They were therefore deemed appropriate for assessing the L2 speakers in the population under study. The third text was a modified version of a Wikipedia article about a South African university. The two groups scored comparably (L1: 75%; L2: 71%, $p < .001$, $d = 0.29$), but not at ceiling level. The background data of the L2 participants is summarized in Table 1.

As in Roberts and Felser (2011), the adjunct-clause items and complement-clause items were included in the same experiment, with the one type of construction serving as a filler for the other type. However, because structural differences prevent direct comparison of the processing of these two construction types, we present the results separately, beginning with those for the adjunct-clause items. Study materials are available on the IRIS database (Mackey & Marsden, 2015).

Results

Adjunct-clause items

Materials

The adjunct-clause garden-path items consisted of 20 sentence pairs containing preposed adjunct clauses. They were constructed from 14 optionally transitive verbs, each used no more than three times, which had comparable selectional restrictions in English and Afrikaans to avoid potential effects of lexical transfer. The critical nouns within each sentence pair were matched for syllable number and frequency, based on data from the Celex database (Centre for Lexical Information, 1993; see Table A.2 in Appendix A). Each noun appeared once in a plausible direct object condition and once in an implausible direct object condition. As in Roberts and Felser (2011), to verify that the plausibility manipulation worked as intended, eight English–Afrikaans bilinguals drawn from the same population as the participants in the main experiment, but who did not participate in the main experiment, conducted an offline rating task in which they rated each verb–noun pairing in terms of its plausibility. The participants were presented with sentences such as *The men drank the beer* and *The men sang the beer* and asked to rate them in terms of

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

Table 1. L2 participants' background data (standard deviation)

Age of L2 acquisition	4.5 (2.4)
English exposure (%)[a]	48 (15.1)
Afrikaans exposure (%)[a]	46.8 (14.7)
L2 speaking self-rating[b]	8.2 (1.4)
L2 reading self-rating	8.6 (1.2)
L2 spoken comprehension self-rating	8.8 (1.4)
C-test score (%)	72 (13.7)

[a]This is a self-reported estimate for current exposure at the time of testing.
[b]Self-ratings are on a scale from 1 to 10.

plausibility on a scale from 1 (*very plausible*) to 7 (*very implausible*). The plausible sentences received a mean rating of 1.37 and the implausible sentences a mean rating of 6.3. According to a *t* test, this is a significant difference ($p < .001$), indicating that the plausibility manipulation functioned as intended.

The 20 adjunct-clause items were divided across two lists, such that participants saw only one version of each experimental item. All experimental items were followed by a comprehension question, where the correct answer was "yes" for one half of the questions and "no" for the other half. In so far as possible, the comprehension questions targeted a correct interpretation of the ambiguous NP, such that if the sentence was *While the child climbed the ladder against the wall fell to the ground*, the comprehension question would be *Did the child fall to the ground?* A correct response of "no" to such a question indicates that the participant has recovered an analysis where *the child* is not understood as the subject of the main clause. The experimental sentences were pseudorandomized and combined with the 20 complement-clause items and 44 filler sentences of different structural types. Two differently ordered versions of each list were compiled. Twelve of the fillers also contained preposed adjunct clauses where the direct object interpretation turned out to be correct. These were included to prevent participants from becoming aware of the experimental manipulation. One third of the fillers were followed by comprehension questions, again with the "yes" and "no" answers counterbalanced. Finally, 8 practice items were included at the beginning of the task.

Procedure

The self-paced reading experiment made use of the noncumulative moving window procedure (Just, Carpenter, & Woolley, 1982) and was administered using PsychoPy (version 1.85.2; Peirce, 2007) on a 15-inch laptop (resolution: 1366 × 768). All sentences were presented in a word-by-word fashion in black type (font: Consolas) on a pale blue background. The sentences all fit onto one line, and the end of each sentence was marked by a full stop.

Participants used the space bar to initiate the display of each successive word. The comprehension questions were answered using the "z" and "m" keys on the laptop. The experiment administrator provided verbal instructions to each participant, which were accompanied by written instructions on the screen. Participants

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

were given the opportunity to ask questions before and after completing the practice session. Three self-timed breaks were included in the experiment, so that participants could pause after each block of 22 sentences. The self-paced reading task took approximately 30 min, after which participants filled in the Language Experience and Proficiency Questionnaire and completed the C-test. An entire testing session lasted approximately 50 min.

Analysis

All analyses were conducted in the R environment for statistical computing (version 3.5.1; R Core Team, 2018). The code and data for the analyses are available on the Dataverse repository (Berghoff, 2019).

Question response accuracy was examined using a generalized linear mixed-effects model, fit using the lme4 package (version 1.1.17; Bates, Mächler, Bolker, & Walker, 2015). *P* values for the regression coefficients were obtained using the lmerTest package (version 2.0.36; Kuznetsova, Brockhoff, & Christensen, 2017).

For the RT analyses, in accordance with standard practice in sentence processing studies, we only analyzed data from trials where the comprehension question was answered correctly. We opted for minimal data trimming, in line with the recommendations in Baayen and Milin (2010). Thus, at the outset of the RT analysis, we only removed extreme values of more than 4000 ms. This affected 7.5% of the L1 English data and 8.6% of the L1 Afrikaans data.

Effects on RT were analyzed with linear mixed-effects models using the same package configurations reported for the question accuracy analyses above. In order to meet the normality requirement for linear regression, we ran analyses on the log-transformed RTs. We analyzed the data for each segment from the first word of the temporarily ambiguous NP (segment 6) onward (eight single-word segments in total for each item). At segments where there was an L1 Group $\times$ Plausibility interaction, we computed separate linear mixed-effects models for each group. To investigate possible differences across conditions and groups in the processing of the comprehension questions, we also examined the question response times.

For all mixed-effects analyses, a maximal approach was taken to determine the random effects structure (Barr, Levy, Scheepers, & Tily, 2013). Thus, for each model, we first tested a random-effects structure with random intercepts for subjects and items and by-participants and by-items random slopes for plausibility. Where the model would not converge, we simplified the random effects structure in line with the recommendations in Barr et al. (2013) until convergence was achieved. Model specifications are provided in the text.

Accuracy

The accuracy of both groups' responses to the experimental items was high, with the L1 English group answering 88.5% of the questions correctly, and the L1 Afrikaans group 88.8%. The generalized linear mixed-effects model computed to analyze the accuracy data included plausibility (implausible or plausible, contrast coded as −1 and 1) and group (L1 Afrikaans or L1 English, contrast coded as −1 and 1) as fixed effects. The maximal random effects structure justified by the design that would converge included random intercepts for participants and items, as well as by-participants and by-items random slopes for plausibility, without correlation between the plausibility random slope and the random intercept for items (see Barr et al., 2013, pp. 261–262).

The model results indicated that accuracy was significantly lower for the plausible versus the implausible items ($\beta = -0.56$, $SE = 0.24$, $z = -2.32$, $p = .020$). There was no main effect of group ($\beta = 0.00$, $SE = 0.12$, $z = -0.01$, $p = 1.00$), nor a Group × Plausibility interaction ($\beta = 0.10$, $SE = 0.10$, $z = 0.91$, $p = .36$).

RTs

For all RT models, plausibility (implausible or plausible; contrast coded as −1 and 1) and L1 group (L1 Afrikaans or L1 English, contrast coded as −1 and 1) were fixed effects. For all models except that for the segment 12 RTs, the maximal random effects structure justified by the design that would converge included random intercepts for participants and items and by-participants random slopes for plausibility. The segment 12 model included only random intercepts. Model outputs are provided in Table 2.

An example of an adjunct clause item divided into segments is provided below to serve as a reference for the presentation of the RT analyses and the overall RT pattern is illustrated in Figure 1.

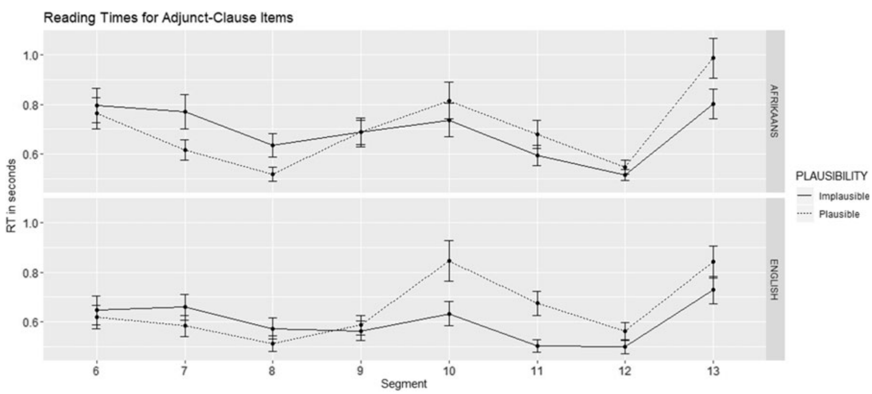

Figure 1. Reading times for adjunct-clause items. Error bars represent standard error.

While / the / men / drank / the / beer / imported / from / Europe / pleased / all / the / customers.
1 2 3 4 5 6 7 8 9 10 11 12 13

The reading speeds of the L1 English and L1 Afrikaans groups did not differ for the most part: the L1 English group were slightly faster than the L1 Afrikaans group only at segment 6, the first word of the temporarily ambiguous NP. At segment 7, which contained the first word of the phrase modifying the noun, there was a significant main effect of plausibility, where participants read faster when the NP was a plausible direct object of the verb in segment 4. The effect of plausibility carried over to segment 8. At this segment, there was also an interaction between L1 group and plausibility. Separate analyses indicated that whereas the L1 Afrikaans group read significantly faster in the plausible condition at this segment ($\beta = -0.08$, $SE = 0.02$, $t = -3.99$, $p < .001$), the plausibility effect was only a weak trend for the L1 English group ($\beta = -0.04$, $SE = 0.02$, $t = -1.73$, $p = .09$).[3]

Segment 10 included the main clause verb, which serves to indicate that the direct object analysis of the preceding NP is incorrect. At this segment, a marginal effect of plausibility appeared, where participants slowed down more upon encountering the disambiguating verb when the direct object analysis of the preceding NP had been plausible. This effect was also present at the following segment. At segment 11, in addition to the main effect of plausibility, there was also a significant L1 Group × Plausibility interaction. Separate analyses[4] revealed that the RTs of the L1 English group were significantly higher in the plausible condition ($\beta = 0.10$, $SE = 0.03$, $t = 3.56$, $p = .001$), whereas there was no effect of plausibility on the RTs of the L1 Afrikaans group ($\beta = 0.04$, $SE = 0.03$, $t = 1.20$, $p = .24$). No significant effects were observed at segment 12. At segment 13, the final segment, participants were slower in the plausible than the implausible condition. This likely reflects the heightened cost of reanalysis in the plausible condition, where participants took longer to settle on a final analysis of these items.

Response times across all comprehension questions were also examined.[5] Both groups were significantly slower in the plausible compared to the implausible condition ($\beta = 0.10$, $SE = 0.03$, $t = 3.50$, $p = .001$). There was no effect of L1 group on response times ($\beta = -0.02$, $SE = 0.04$, $t = -0.50$, $p = .62$), nor was there an L1 Group × Plausibility interaction ($\beta = 0.00$, $SE = 0.01$, $t = 0.01$, $p = .99$).

Summary: Adjunct-clause items

Both groups showed a response to the plausibility manipulation, where the temporarily ambiguous NP was easier to process if it was plausible as a direct object of the preceding verb. However, the effect of plausibility lasted longer in the L1 Afrikaans group, persisting into segment 8. At the disambiguating region of the sentence, there is some evidence that both groups initiated reanalysis at segment 10, but the effect of plausibility on RTs only carries over to the following segment in the L1 English group. The groups did not differ in their overall performance on the comprehension questions, either in terms of the time taken to answer or accuracy.

Complement-clause items

Materials

The complement-clause items consisted of 20 sentence pairs, each containing an optionally transitive verb followed by a finite complement clause. Eleven verbs were selected, none of which occurred more than three times in the stimulus list. The critical nouns within each sentence pair were again matched for syllable length and frequency based on data from the Celex database (Center for Lexical Information, 1993; see Table A.3 in Appendix A), and each noun occurred once in a plausible direct object condition and once in an implausible direct object condition. In the same offline rating task discussed in relation to the adjunct-clause garden-path items, eight English–Afrikaans bilinguals who did not participate in the main experiment rated the plausibility of each verb–noun pair, with a rating of 1 indicating that the noun was a very plausible object of the verb, and a rating

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

Table 2. Model outputs, adjunct-clause items

	Segment 6 *beer*			Segment 7 *imported*			Segment 8 *from*			Segment 9 *Europe*			Segment 10 *pleased*			Segment 11 *all*			Segment 12 *the*			Segment 13 *customers*		
Parameter	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	t	*p*	Est. (*SE*)	*t*	*p*
Plausibility	−0.03 (0.05)	−0.60	.548	−0.15 (0.05)	−3.06	.002	−0.13 (0.04)	−3.36	.001	−0.01 (0.05)	−0.25	.804	0.13 (0.07)	1.81	0.070	0.14 (0.06)	2.61	.009	0.06 (0.04)	1.45	.146	0.14 (0.06)	2.55	.011
Group	−0.19 (0.11)	−1.78	.075	−0.12 (0.09)	−1.40	.163	−0.08 (0.07)	−1.10	.271	−0.15 (0.09)	−1.66	.097	−0.07 (0.10)	−0.70	0.482	−0.05 (0.06)	−0.83	.407	−0.02 (0.06)	−0.37	.708	−0.11 (0.08)	−1.34	.181
Plausibility × Group	0.05 (0.06)	0.88	.377	0.05 (0.05)	0.95	.344	0.09 (0.05)	1.99	.046	0.10 (0.06)	1.75	.081	0.11 (0.08)	1.36	0.175	0.14 (0.04)	3.04	.002	0.06 (0.04)	1.67	.094	0.03 (0.06)	0.53	.594

of 7 indicating that the pairing was implausible. The plausible items received a mean rating of 1.9 and the implausible items a mean rating of 4.9. A *t* test revealed that these are significantly different, $p < .001$, again indicating that the plausibility manipulation functioned as intended.

The 20 experimental items were divided across two lists, such that participants only saw one version of each. All experimental items were followed by a comprehension question with "yes" and "no" answers counterbalanced. These items were pseudorandomized and combined with the 20 adjunct-clause garden-path items and the 44 fillers to produce two differently ordered versions of each list. The fillers included 12 items where the direct object analysis was actually correct in order to disguise the experimental manipulation.

Analysis

As with the adjunct-clause items, we analyzed question response accuracy using a generalized linear mixed-effects model, with the same package configurations reported above.

RTs were analyzed for each segment from the first word of the temporarily ambiguous NP (segment 5) onward (nine single-word segments in total for each item). In addition, we analyzed the groups' RTs to the comprehension questions following the experimental items. Only RTs from trials where the comprehension question was answered accurately were analyzed, and RTs over 4000 ms were removed. This resulted in the removal of 7.5% of the data for the L1 Afrikaans group and 5.9% of the data for the L1 English group.

Effects on RT were analyzed using linear mixed-effects models with the same package configurations reported above. At segments where there was an L1 Group × Plausibility interaction, we computed separate linear mixed effects models for each group. Throughout, random effects structures were determined using a maximal approach (Barr et al., 2013).

Accuracy

The accuracy of both groups' responses to the comprehension questions was again high, with the L1 English group answering 88% of the questions correctly, and the L1 Afrikaans group answering 91.8% of the questions correctly. The generalized linear mixed-effects model fit to analyze question response accuracy included plausibility (implausible or plausible, contrast coded as −1 and 1) and group (L1 Afrikaans or L1 English, contrast coded as −1 and 1) as fixed effects, random intercepts for participants and items, and by-participants and by-items random slopes for plausibility (without correlation between the slopes and the intercepts; see Barr et al., 2013, pp. 261–262). The model results indicated no significant difference in performance across the groups ($\beta = -0.22$, $SE = 0.14$, $z = -1.56$, $p = .12$) or the conditions ($\beta = -0.08$, $SE = 0.25$, $z = -0.33$, $p = .74$), nor any L1 Group × Plausibility interaction ($\beta = 0.00$, $SE = 0.12$, $z = -0.02$, $p = .98$).

RTs

As in the adjunct-clause analysis, separate linear mixed-effects models were computed for each segment, with log-transformed RT as the dependent variable and L1 group (Afrikaans or English, contrast coded as −1 and 1) and plausibility (implausible or plausible, contrast coded as −1 and 1) as fixed effects. For all but two models, the maximal random effects structure that would converge included random intercepts for participants and items, as well as by-participants random slopes for plausibility. The segment 5 and 7 models converged with both sets of random intercepts, as well as by-items random slopes for plausibility. Model outputs are provided in Table 3.

An example of a complement-clause item divided into segments is provided below to serve as a reference for the presentation of the RT analyses and the overall RT pattern is illustrated in Figure 2.

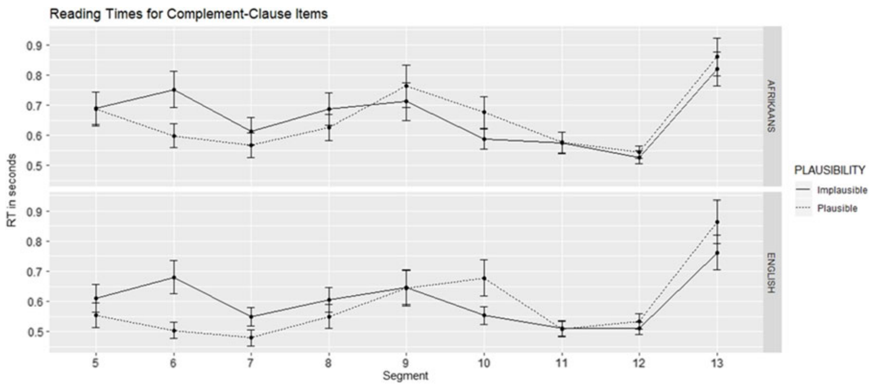

Figure 2. Reading times for complement-clause items. Error bars represent standard error.

The	student	wrote	the	report	about	the	budget	would	start	an	important	debate.
1	2	3	4	5	6	7	8	9	10	11	12	13

Overall RT did not differ between the L1 groups, as indicated by the absence of a main effect of L1 group at all segments. At segment 6, the first word of the phrase modifying the temporarily ambiguous noun, a significant main effect of plausibility was observed. Here, participants read faster when the noun was a plausible direct object of the verb in segment 3. This effect persisted at segment 7, where there was also a weak L1 Group × Plausibility trend. Separate analyses of the groups[6] indicated that the L1 English group were significantly faster in the plausible condition at this segment ($\beta = -0.12$, $SE = 0.04$, $t = -3.0$, $p = .004$), whereas the RTs of the L1 Afrikaans group did not differ across the plausible and implausible conditions ($\beta = -0.03$, $SE = 0.02$, $t = -1.62$, $p = .12$). At segment 8, participants still read faster in the plausible than the implausible condition, and the weak L1 Group × Plausibility interaction remained. The separate group analyses[7] again showed shorter RTs in the L1 English group in the plausible condition ($\beta = -0.12$, $SE = 0.05$, $t = -2.56$, $p = .01$), but no effect of plausibility in the L1 Afrikaans group ($\beta = -0.02$, $SE = 0.02$, $t = -1.1$, $p = .28$). There were no significant effects at segment 9, but at segment 10, which contained the disambiguating verb, a marginal main effect of plausibility was present. Participants were slower in the plausible than the implausible conditions here, indicating the increased difficulty of reanalysis

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

Table 3. Model outputs, complement-clause items

	Segment 5 *report*			Segment 6 *about*			Segment 7 *the*			Segment 8 *budget*			Segment 9 *would*			Segment 10 *start*			Segment 11 *an*			Segment 12 *important*			Segment 13 *debate*		
Parameter	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (SE)	*t*	*p*
Plausibility	−0.03 (0.04)	−0.96	.339	−0.20 (0.04)	−5.48	<.001	−0.09 (0.04)	−2.43	.015	−0.08 (0.04)	−2.11	.035	0.00 (0.05)	−0.03	.980	−0.10 (0.05)	−1.90	.058	0.01 (0.04)	0.24	.813	−0.03 (0.03)	−0.80	.421	−0.05 (0.05)	−1.00	.303
Group	−0.15 (0.10)	−1.56	.120	−0.13 (0.09)	−1.40	.168	−0.12 (0.07)	−1.55	.120	−0.13 (0.09)	−1.45	.148	−0.13 (0.09)	−1.42	.156	−0.06 (0.08)	−0.72	.471	−0.10 (0.07)	−1.66	.118	−0.04 (0.06)	−0.61	.540	−0.05 (0.08)	−0.62	.537
Plausibility × Group	−0.07 (0.04)	−1.64	.100	−0.06 (0.05)	−1.33	.184	−0.06 (0.04)	−1.70	.089	−0.07 (0.04)	−1.69	.092	−0.05 (0.06)	−0.92	.355	0.02 (0.06)	0.28	.781	−0.03 (0.04)	−0.70	.488	−0.01 (0.03)	−0.22	.826	0.05 (0.05)	0.82	.412

when the direct object analysis of the temporarily ambiguous NP had been plausible. There were no further significant effects at segments 10–13.

In their responses to the comprehension questions,[8] there were no significant differences in RTs across groups ($\beta = 0.00$, $SE = 0.04$, $t = -0.02$, $p = .98$) or across conditions ($\beta = 0.04$, $SE = 0.04$, $t = 1.16$, $p = .25$), nor was there an L1 Group × Plausibility interaction ($\beta = -0.02$, $SE = 0.01$, $t = -1.20$, $p = .23$).

Summary: Complement-clause items

As with the adjunct-clause items, both groups showed an effect of plausibility, where the temporarily ambiguous NP was read more quickly when it was plausible as a direct object of the preceding verb. In contrast to the adjunct-clause results, however, the plausibility effect was more persistent in the L1 English speakers for these items. At the point in the sentence signalling the need for reanalysis, both groups slowed down slightly more in the plausible compared to the implausible condition. This effect was marginal and did not persist to the following segments, suggesting that reanalysis of these items did not pose much difficulty for the participants. The groups' comprehension question accuracy and response times again did not differ: they were equally successful in recovering the correct interpretations of the experimental items, and the duration of this recovery process did not differ according to L1 status.

Analysis of age effects

As noted earlier in the paper, L2 speakers with earlier ages of L2 acquisition (AoA) are typically assumed to be more likely to resemble L1 speakers in their processing behavior. Specifically in terms of L2 grammatical processing, nativelike performance has been found in L2 speakers with an AoA of 5–6 years or lower (Granena & Long, 2013; Veríssimo et al., 2018; Weber-Fox & Neville, 1996), with L1–L2 differences subsequently increasing as AoA increases. As a follow-up to the main analyses presented above, we investigated whether there were AoA influences on processing within our L2 group, whose AoAs ranged from 1 to 10 years. We computed linear mixed-effects models for each segment of interest in both construction types. In all models, the dependent variable was log-transformed RT. Fixed effects were plausibility (implausible or plausible, contrast coded as −1 and 1), AoA (centred around the mean), and the interaction of these two variables. Random intercepts were included for participants and items. For the adjunct-clause models, the maximal random effects structure justified by the design that would converge included by-participants random slopes for plausibility for all but three segments. At segment 10, by-participants and by-items random slopes for plausibility were included. At segments 7 and 11, no random slopes were included. For the complement-clause items, the segment 5–7 models included only random intercepts. By-participants random slopes for plausibility were included in the remaining models.

For the complement-clause items, there were no significant main effects of AoA, nor any interactions between AoA and plausibility, at any of the segments. However, for the adjunct-clause items, there was a significant main effect of AoA at segment 10, the disambiguating verb, and a Plausibility × AoA trend at segment 12. To further examine this AoA effect, a median split was performed on the adjunct-clause data for the L2 speakers, yielding one group with an AoA below 5

($n = 16$; AoA range = 1–4 years), and one group with an AoA equal to or greater than 5 ($n = 17$; AoA range = 5–10 years). Of importance, a t test indicated that the two groups did not differ in proficiency, as measured by their C-test scores ($t = -1.28$, $p = .21$), and thus AoA and L2 proficiency are not confounded here.

We ran separate linear mixed effects models per group for each segment of the adjunct-clause items, with log-transformed RT as the dependent variable, Plausibility as a fixed effect (implausible or plausible, contrast coded as −1 and 1), and random intercepts for participants and items. The models for the earlier acquirers included by-participants random slopes for plausibility at all but three segments (7, 9, and 12), where segments 7 and 12 only had the random intercepts included in the other models, and segment 9 only had by-participants random intercepts. The models for the later acquirers had by-participants random slopes for plausibility at segments 10, 12, and 13, by-items random slopes for plausibility at segments 6 and 7, both by-items and by-participants random slopes for plausibility at segment 9, and only random intercepts at segments 8 and 11.

The group of late childhood acquirers showed the main effects of plausibility observed in the L2 group as a whole, where they read the temporarily ambiguous NP (segments 7 and 8) more quickly in the plausible compared to the implausible condition. However, no effect of plausibility was observed at any of the remaining segments, thus suggesting that reanalysis was not attempted at or following the disambiguating region of the sentence.

In the group of early childhood acquirers, main effects of plausibility were also seen at segments 7 and 8. At segment 10, the onset of the disambiguating region, there was a significant reverse effect of plausibility, which persisted to the following two segments. These effects indicate that participants experienced greater processing difficulty at the disambiguating region when the initial analysis had been plausible, which is the same pattern that was observed in the L1 English group.

Overall, then, it appears that the group of early childhood L2 acquirers was able to initiate syntactic analysis online. As in the comparison between the L1 and L2 groups, however, performance on the comprehension questions did not differ significantly across the early and late childhood L2 acquirers ($\beta = 0.09$, $SE = 0.16$, $z = 0.58$, $p = .56$);[9] nor did the time taken to answer the questions ($\beta = -0.20$, $SE = 0.28$, $t = -0.73$, $p = .47$).[10] This indicates that the late childhood L2 acquirers were equally able to recover a correct parse of the temporarily ambiguous constructions without incurring an additional time cost.

Discussion

Previous investigations of L2 OSA processing have found L2 speakers (a) to be more sensitive to plausibility information than L1 speakers and (b) to have greater difficulty than L1 speakers in recovering a correct parse of the input. In the present study, the results regarding sensitivity to plausibility information were mixed, with L2 speakers showing greater sensitivity than L1 speakers in the adjunct-clause condition, and the reverse being true in the complement-clause condition. Regarding garden-path recovery, we observed a clear slowdown at the disambiguating region of the plausible adjunct-clause items in the L1 speakers. Among the L2 speakers, we found an effect of AoA, where the group of early childhood acquirers (AoA < 5) showed a slowdown effect comparable to that found in the L1 speakers. In the late childhood acquirers (AoA ≥ 5), no such effect was present. Furthermore, in contrast to previous findings (e.g., Jacob & Felser, 2016), the L1 and L2 groups did not differ in the accuracy of their responses to the comprehension questions, and so the non-native speakers were evidently not more susceptible to parsing breakdown.

The interaction between L1 group, plausibility, and construction type observed in our data is not present in Roberts and Felser (2011), where only the L2 speakers showed a strong response to the plausibility manipulation, and where this response was present across both types of construction. In our data, there was an initial strong response to the plausibility of the temporarily ambiguous NP in the L1 and the L2 group. However, although both groups were sensitive to the plausibility manipulation regardless of the construction type, they differed in each showing a stronger reaction to plausibility information in a particular construction type.

Findings from the monolingual processing literature offer some insight into why the L1 speakers' sensitivity to the plausibility manipulation may have differed across the two construction types. A close examination of the adjunct- and complement-clause items reveals that they differed in terms of the subcategorization information (Chomsky, 1965) provided by the initial verb. Subcategorization information specifies whether a verb can take only an NP complement, only a sentential complement, either type of complement, or no complement at all. Among verbs of the penultimate sort, information on verb bias (the tendency for a verb to take a particular kind of complement) can also serve to generate expectations about upcoming input. For example, while "hit" can only take an NP complement, and "proved" can only take a sentential complement, "wrote" can take either type: both "John wrote the manuscript" and "John wrote the manuscript had been destroyed in the fire" are possible (Garnsey, Pearlmutter, Myers, & Lotocky, 1997, p. 59). Information on the frequency with which "wrote" occurs in direct object versus sentential complement constructions in a particular speech community can be used to classify it in terms of subcategorization bias.

Verb bias has been found to interact with plausibility information (Garnsey et al., 1997), such that plausibility only has a clear effect under certain verb bias conditions. For example, in Garnsey et al. (1997), readers only showed sensitivity to the plausibility of a temporarily ambiguous NP as a direct object of the preceding verb in the absence of clear verb bias information. A possible explanation for this finding is that when verb bias information is available, readers adopt the corresponding structural analysis of the temporarily ambiguous section, and so rely less on the plausibility of the NP to guide parsing. However, in the absence of verb bias information, neither a direct object nor a sentential complement analysis is favored by the available cues, and the reader may rely more heavily on plausibility information in deciding which parse to adopt.

In the present study, clear differences in subcategorization biases are evident across the adjunct-clause and complement-clause conditions. All but two of the verbs in the former condition ("called" and "painted") can only take an NP-complement (see Table A.4 in Appendix A); arguably, the two verbs that constitute the exceptions also have a heavy direct object bias. These verbs therefore provide strong cues in favor of a direct object analysis of the following NP. In contrast, in the complement-clause items, none of the verbs are restricted to taking a particular kind of complement. While some of the verbs may have been biased in this regard,[11] what is important is that the cue provided by the subcategorization information is considerably weaker in this condition.

The strong initial reaction to plausibility observed in the L1 group aligns with previous findings that readers attempt to integrate a temporarily ambiguous NP as a direct object of a preceding verb regardless of the verb's bias (Traxler, 2002; Traxler, Pickering, & Clifton, 1998; Van Gompel, Pickering, & Traxler, 2001). However, Traxler (2005) also provides evidence that verb bias information is accessed rapidly as parsing proceeds, which would account for the rapid change in the strength of the plausibility effect in our data. Following Garnsey et al. (1997), the presence of clear verb bias information in the adjunct-clause condition may have led the L1 group to rely less on plausibility information here, whereas the absence of verb bias information in the complement-clause condition elicited the opposite response.

Although some studies have found L2 speakers to be able to both learn verb bias information and employ it during online parsing (Dussias & Cramer Scaltz, 2008; Jegerski, 2012; although cf. Frenck-Mestre & Pynte, 1997, Lee, Lu, & Garnsey, 2013; Qian, Lee, Lu, & Garnsey, 2018, for alternative findings), previous experiments with L2 speakers have not explicitly manipulated plausibility in conjunction with verb bias. Thus, while it is evident that L2 speakers *can* use verb bias information during online processing, it is not clear whether they do so preferentially, given the simultaneous availability of plausibility information.[12] The pattern of results observed in the present study could be explained if the L2 speakers did employ verb subcategorization information to guide parsing, but weighted the plausibility information more heavily than the L1 speakers did, thus experiencing greater processing disruption if the temporarily ambiguous NP was implausible given their expectations regarding the upcoming input. For example, supposing that the L2 group also accessed the subcategorization information provided by the verb in the adjunct-clause condition, they may, upon encountering a subsequent NP, have favored an analysis in which this NP was the direct object of the verb. Subsequently, assuming heightened sensitivity to plausibility information in this group, in accordance with the findings of Roberts and Felser (2011), they may have reacted strongly when the NP was implausible in this role. In the complement-clause condition, the absence of unequivocal subcategorization information may have prevented strong expectations regarding the upcoming input from being generated. In this case, there would be a reduced clash between the reader's expectations and the plausibility of the input in this condition, leading to a mitigated response to implausible NPs.

The speculative nature of this explanation is acknowledged, but it suggests potentially fruitful avenues for further investigation of the relative weightings of verb bias versus plausibility information in L1 and L2 processing. In particular, the fact that both the early and late L2 acquirers showed the same differential sensitivity to plausibility information across the two construction types—as indicated by the absence of a Plausibility × AoA interaction at the temporarily ambiguous NP in Tables 4 and 5—may suggest that AoA is not the sole determinant of L2 speakers' prioritization of plausibility information. Further investigation into this issue should begin with a verb bias norming study similar to Kennison (1999).

The remainder of this section addresses the second main finding of this study, namely, the difference in behavior at the disambiguating region across the L1 and L2

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

groups. One of the present study's predictions was that given our L2 participants' early AoA, extensive naturalistic L2 exposure, and high L2 proficiency, they would exhibit processing patterns similar to those of the L1 group. This prediction was borne out at the disambiguating region of the complement-clause items, where both the L1 and L2 speakers showed only a marginal slowdown in the plausible condition. The mild nature of the processing disruption observed at this point is in line with the findings of previous studies with monolingual speakers, and, as discussed above, has been attributed to the fact that the structural revisions required to arrive at a grammatical parse of garden-path constructions of this sort occur within the thematic domain of the predicate. In spite of this, L2 speakers with a later AoA and limited naturalistic L2 exposure have been found to slow down significantly at the disambiguating region of complement-clause OSA constructions (Roberts & Felser, 2011), which reflects difficulty in revising their initial parse. Our results, however, indicate that nativelike processing of these constructions is possible in L2 speakers of the profile examined here.

The study's prediction of L1–L2 convergence in reanalysis behavior was not entirely borne out for the adjunct-clause items, where more substantial structural revisions are required to derive a grammatical parse. Specifically, the L1 group exhibited a significantly stronger slowdown at the disambiguating region in the plausible condition compared to the L2 group, reflected in the Plausibility × Group interaction at segment 11. Closer scrutiny of the L2 group revealed an effect of AoA at this region: whereas the early childhood acquirers (AoA < 5) exhibited longer RTs here in the plausible condition, the late childhood acquirers (AoA ≥ 5) did not. This suggests that only the early childhood acquirers initiated reanalysis of these constructions online. The finding that the later L2 acquirers did not initiate reanalysis aligns with the findings of Roberts and Felser (2011).

The present data do not allow for strong claims to be made regarding the precise AoA at which differences between online reanalysis behavior in L1 and L2 speakers begin to emerge. However, the difference between the processing patterns of the early and late childhood L2 acquirers in our study is in accordance with previous observations of a distinction between L2 acquirers with AoAs on either side of approximately 6 years of age, which has been cited as the offset of the period in which ultimate attainment of nativelike L2 morphosyntax is most likely (Flege, Yeni-Komshian, & Liu, 1999; Granena & Long, 2013; Hyltenstam, 1992; Johnson & Newport, 1989; Long, 1990). Studies of online processing specifically have observed greater L1–L2 differences in L2 speakers with an AoA of 4 to 5 years and up (Veríssimo et al., 2018; Weber-Fox & Neville, 1996).[13] Regardless of the exact AoA at which processing nativelikeness becomes less likely, it is nonetheless striking that age effects are observed in the present sample of L2 speakers, who on the whole have had extensive exposure to the L2 from a young age.

An additional novel finding of the present study is that the nonnative group as a whole performed within the native-speaker range on the offline measure of comprehension. This is in contrast to some previous findings (e.g., Jacob & Felser, 2016) where, in response to comprehension questions like those of the present study that probed the interpretation of the temporarily ambiguous NP, the L2 group scored significantly lower than the L1 group. In this regard, the L1 and L2 groups were equally accurate in their responses to the comprehension questions following

Table 4. Age effects on L2 processing of adjunct-clause items

	Segment 6 *beer*			Segment 7 *imported*			Segment 8 *from*			Segment 9 *Europe*			Segment 10 *pleased*			Segment 11 *all*			Segment 12 *the*			Segment 13 *customers*		
Parameter	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	t	*p*
Plausibility	−0.03 (0.03)	−1.10	.270	−0.08 (0.02)	−3.80	.000	−0.09 (0.02)	−3.93	.000	−0.03 (0.03)	−1.0	.318	0.03 (0.04)	0.87	0.387	0.03 (0.03)	1.13	.258	0.01 (0.02)	0.56	.575	0.06 (0.04)	1.76	.078
AoA	0.13 (0.08)	1.60	.109	0.06 (0.06)	0.90	.369	0.06 (0.05)	1.12	.263	0.08 (0.06)	1.32	.188	0.14 (0.07)	2.07	0.038	0.06 (0.05)	1.30	.194	0.04 (0.04)	0.92	.358	0.06 (0.07)	0.86	.391
Plausibility × AoA	−0.01 (0.02)	−0.40	.687	0.00 (0.02)	0.06	.954	0.01 (0.02)	0.51	.607	0.00 (0.02)	0.17	.863	0.00 (0.02)	0.18	0.860	−0.02 (0.02)	−1.33	.184	−0.02 (0.01)	−1.73	.084	0.00 (0.02)	0.02	.984

Table 5. Age effects on L2 processing of complement-clause items

	Segment 5 *report*			Segment 6 *about*			Segment 7 *the*			Segment 8 *budget*			Segment 9 *would*			Segment 10 *start*			Segment 11 *an*			Segment 12 *important*			Segment 13 *debate*		
Parameter	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	t	*p*	Est. (SE)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*
Plausibility	0.00 (0.02)	0.13	.894	−0.08 (0.02)	−4.22	.000	−0.03 (0.02)	−1.53	.126	−0.02 (0.02)	−1.05	.292	0.01 (0.03)	0.41	0.680	0.04 (0.03)	1.50	.134	0.00 (0.02)	0.16	.873	0.02 (0.01)	1.13	.258	0.01 (0.03)	0.37	.712
AoA	0.11 (0.07)	1.57	.117	0.06 (0.07)	0.81	.416	0.09 (0.05)	1.66	.096	0.07 (0.07)	1.11	.266	0.09 (0.07)	1.35	0.178	0.06 (0.06)	1.09	.276	0.06 (0.05)	1.12	.265	0.04 (0.04)	1.00	.315	0.07 (0.06)	1.27	.203
Plausibility × AoA	0.01 (0.02)	0.85	.396	−0.02 (0.01)	−1.13	.257	0.02 (0.01)	1.42	.155	0.02 (0.02)	1.12	.264	0.03 (0.02)	1.32	0.187	0.01 (0.02)	0.26	.793	0.01 (0.02)	0.75	.455	0.01 (0.01)	0.99	.324	0.02 (0.02)	1.23	.219

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

Table 6. Later L2 acquirers, adjunct-clause items

	Segment 6 *beer*			Segment 7 *imported*			Segment 8 *from*			Segment 9 *Europe*			Segment 10 *pleased*			Segment 11 *all*			Segment 12 *the*			Segment 13 *customers*		
Parameter	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*
Plausibility	−0.04 (0.03)	−1.51	.132	−0.09 (0.03)	−3.23	.001	−0.07 (0.02)	−3.26	.001	−0.03 (0.04)	−0.67	.503	0.00 (0.05)	0.07	.941	0.01 (0.04)	0.25	.806	−0.02 (0.03)	−0.63	.531	0.05 (0.04)	1.36	.174

Table 7. Earlier L2 acquirers, adjunct-clause items

	Segment 6 *beer*			Segment 7 *imported*			Segment 8 *from*			Segment 9 *Europe*			Segment 10 *pleased*			Segment 11 *all*			Segment 12 *the*			Segment 13 *customers*		
Parameter	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*	Est. (*SE*)	*t*	*p*
Plausibility	−0.02 (0.04)	−0.55	.583	−0.08 (0.03)	−2.78	.005	−0.10 (0.04)	−2.85	.004	−0.02 (0.02)	−0.76	.447	0.07 (0.04)	2.04	.042	0.06 (0.03)	1.88	.060	0.04 (0.02)	1.97	.049	0.07 (0.05)	1.49	.135

the experimental items, and there was also no difference in accuracy across the early and late childhood L2 acquirers. The pattern observed across the board was lower accuracy in the plausible compared to the implausible condition. Evidently, then, all participants were equally able (or unable, as the case may have been) to recover a correct parse of the adjunct-clause items. In terms of RTs to the comprehension questions, both the L1 and L2 group were significantly slower in responding to the questions following the plausible items. However, this delay was not extended among the late childhood L2 acquirers. It therefore appears that, although the late childhood acquirers did not seem to initiate reanalysis of the plausible items online, they were still able to arrive quickly at a grammatical parse of these constructions.

This finding suggests that even for late childhood L2 acquirers, there is some benefit to extensive naturalistic L2 exposure and high L2 proficiency in facilitating the revision of structural analyses. While online reanalysis may still not occur in particularly difficult garden-path sentences, it seems that L2 speakers of this profile are nonetheless able to perform rapid offline reanalysis, ultimately leaving them on equal footing with native speakers and early childhood L2 acquirers in terms of the final outcome of comprehension.

Conclusion

The results of the present study have several implications. First, they suggest that L1–L2 differences in the online processing of structurally complex garden-path sentences may be found even among childhood L2 acquirers, with differences in L1–L2 convergence across early and late childhood acquirers being observed. The difference in online processing between relatively earlier and later L2 acquirers observed here, however, was not reflected in their offline performance, as the two groups performed equivalently on the comprehension questions following the experimental items, which probed the interpretation of the temporarily ambiguous region of the sentence. Second, contra certain models of L2 processing, L2 speakers are not always more sensitive than L1 speakers to semantic information. To explain the interaction between construction type, L1 versus L2 status and plausibility observed in our study, we have suggested an interplay between two information sources—verb bias and plausibility—that may be weighted differently across L1 and L2 speakers. Further research should aim to shed light on this interplay in non-native language processing.

Acknowledgments. The author would like to acknowledge the work of Leah Roberts and Claudia Felser in developing the sentences upon which the current study's stimuli were based and thank them for making these sentences publically available. The author would also like to thank Emanuel Bylund for valuable discussion and feedback on earlier versions of this manuscript.

Notes

1. A reviewer points out that English and Afrikaans differ in that Afrikaans is a verb-second (V2) language, but V-final in embedded clauses and simple clauses with modals. Conceivably, parsing routines from the V2 L1 may then affect L2 parsing. However, previous studies have found limited evidence for transfer of L1 parsing routines in OSA processing. For example, Havik, Roberts, van Hout, Schreuder, and Haverkort (2009) tested subject- and object-relative clause processing in L1 German learners of L2 Dutch. Both are V2 languages, and monolingual speakers of both languages have been shown to have a subject-relative preference. In L2 Dutch, however, the German speakers showed no such preference, suggesting that L1 parsing strategies were not employed.

2. As a reviewer points out, the fact that our L1 group knew another language sets them apart from the monolingual native speaker control groups typically employed in sentence processing studies. This practice of employing monolingual control groups, like the notion of the "native speaker," has been extensively criticized (see, e.g., Birdsong & Gertken, 2013; Cook, 1999; Ortega, 2011; Rampton, 1990). The bilingualism of our L1 group mitigates to some extent the concern about comparing bilingual L2 speakers to monolingual L1 speakers and also moves away from the idealized "native speaker" construct. The most suitable comparison group for this study would however have been bilingual speakers who acquired both languages from birth (see, e.g., Bylund, Hyltenstam, & Abrahamsson, 2013 for discussion of this point).

3. Both models included by-participants random slopes for plausibility, in addition to random intercepts for subjects and items.

4. No random slopes were included in these models, as neither model would converge with them.

5. The model included fixed effects of L1 group and plausibility and random intercepts for participants and items.

6. The L1 Afrikaans model included by-items random slopes for plausibility. The L1 English model did not converge with any random slopes.

7. The L1 English model included by-participants and by-items random slopes for plausibility. The L1 Afrikaans model would not converge with any random slopes.

8. The model included fixed effects of L1 group and plausibility, random intercepts for participants and items, and by-participants random slopes for plausibility.

9. The maximal random effects structure that would converge included random intercepts for participants and items.

10. The maximal random effects structure that would converge included random intercepts for participants and items.

11. Based on Garnsey et al.'s (1997) ratings, four of the complement-clause verbs would be classified as having a sentential-complement bias ("confessed," "believed," "proved," and "suggested"), and three as having a direct-object bias ("warned," "wrote," and "confirmed"). Again, though, the point is that, in contrast to the adjunct-clause verbs, none of the complement-clause verbs that were used was restricted to taking a particular kind of complement.

12. As noted by a reviewer, the extent to which verb bias is semantic or syntactic in nature is debatable. It is thus not clear whether current models of L2 processing such as the Shallow Structure Hypothesis that deal with semantic versus syntactic information would predict L2 speakers to prioritize plausibility information over verb bias information. We speculate here, based on past research, only that L2 speakers may generally be more sensitive to plausibility information than L1 speakers.

13. As a reviewer points out, the age effect observed here could also be due to increased L1 entrenchment among the later L2 acquirers (see, e.g., Hernandez, 2013; MacWhinney, 1987, 2005).

References

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, **3**, 12. doi: 10.21500/20112084.807

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**, 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67.* doi: 10.18637/jss.v067.i01

Berghoff, R. (2019). Replication data for: The processing of object–subject ambiguities in early L2 acquirers. Harvard Dataverse, V2. https://doi.org/10.7910/DVN/DZ6XJ

Birdsong, D., & Gertken, L. M. (2013). In faint praise of folly: A critical review of native/non-native speaker comparisons, with examples from native and bilingual processing of French complex syntax. *Language, Interaction and Acquisition*, **4**, 107–133.

Bosch, S., Veríssimo, J., & Clahsen, H. (2019). Inflectional morphology in bilingual language processing: An age-of-acquisition study. *Language Acquisition*, **26**, 339–360. doi: 10.1080/10489223.2019.1570204

Bylund, E., Hyltenstam, K., & Abrahamsson, N. (2013). Age of acquisition effects or effects of bilingualism in second language ultimate attainment? In G. Granena & M. Long (Eds.), *Language learning & language teaching. Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 69–102). Amsterdam: Benjamins.

Census. (2011). *Census in brief*. Pretoria: Statistics South Africa.

Centre for Lexical Information. (1993). *The Celex lexical database*. Nijmegen: Max Planck Institute for Psycholinguistics.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, **42**, 368–407. doi: 10.1006/cogp.2001.0752

Clahsen, H., & Felser, C. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, **27**, 107–126. doi: 10.1017/S0142716406060206

Clahsen, H., & Felser, C. (2006b). Grammatical processing in language learners. *Applied Psycholinguistics*, **27**, 3–42. doi: 10.1017/S0142716406060024

Clahsen, H., & Felser, C. (2006c). How native-like is non-native language processing? *Trends in Cognitive Sciences*, **10**, 564–570. doi: 10.1016/j.tics.2006.10.002

Clahsen, H., & Felser, C. (2017). Some notes on the Shallow Structure Hypothesis. *Studies in Second Language Acquisition*, **24**, 1–14. doi: 10.1017/S0272263117000250

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly*, **33**, 185–209.

Department of Basic Education. (2010). *The status of the language of learning and teaching (LoLT) in South African public schools: A quantitative overview*. Pretoria: Author.

Dussias, P. E., & Cramer Scaltz, T. R. (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. doi: 10.1016/j.actpsy.2007.09.004

Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language*, **30**, 725–745.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, **41**, 78–104.

Frazier, L., & Clifton, C. (1998). Sentence reanalysis, and visibility. In *Reanalysis in sentence processing* (pp. 143–176). New York: Springer.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, **6**, 291–325.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, **14**, 178–210.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology A*, **50**, 119–148. doi: 10.1080/027249897392251

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, **37**, 58–93. doi: 10.1006/jmla.1997.2512

Gerth, S., Otto, C., Felser, C., & Nam, Y. (2017). Strength of garden-path effects in native and non-native speakers' processing of object–subject ambiguities. *International Journal of Bilingualism*, **21**, 125–144. doi: 10.1177/1367006915604401

Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, **29**, 311–343. doi: 10.1177/0267658312461497

Havik, E., Roberts, L., van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning*, **59**, 73–112.

Hernandez, A. E. (2013). *The bilingual brain*. Oxford: Oxford University Press.

Holmes, V. M., Kennedy, A., & Murray, W. S. (1987). Syntactic structure and the garden path. *Quarterly Journal of Experimental Psychology*, **39**, 277–293.

Hopp, H. (2015). Individual differences in the second language processing of object–subject ambiguities. *Applied Psycholinguistics*, **36**, 129–173. doi: 10.1017/S0142716413000180

Hyltenstam, K. (1992). Non-native features of near-native speakers: On the ultimate attainment of childhood L2 learners. In R. J. Harris (Ed.), *Advances in psychology. Cognitive processing in bilinguals* (pp. 351–368). Amsterdam: Elsevier.

Jackson, C. (2008). Proficiency level and the interaction of lexical and morphosyntactic 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. *Quarterly Journal of Experimental Psychology*, **69**, 907–925. doi: 10.1080/17470218.2014.984231

Jegerski, J. (2012). The processing of subject–object ambiguities in native and near-native Mexican Spanish. *Bilingualism: Language and Cognition*, **15**, 721–735. doi: 10.1017/S1366728911000654

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, **21**, 60–99.

Juffs, A. (1998a). Main verb versus reduced relative clause ambiguity resolution in L2 sentence processing. *Language Learning*, **48**, 107–147.

Juffs, A. (1998b). Some effects of first language argument structure and morphosyntax on second language sentence processing. *Second Language Research*, **14**, 406–424.

Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language sentence processing. *Language Learning*, **46**, 283–323. doi: 10.1111/j.1467-1770.1996.tb01237.x

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, **111**, 228.

Kaan, E., Ballatyne, J. C., & Wijnen, F. (2015). Effects of reading speed on second-language sentence processing. *Applied Psycholinguistics*, **36**, 799–830. doi: 10.1017/S0142716413000519

Keijzer, M. (2007). *Last in first out? An investigation of the regression hypothesis in Dutch emigrants in Anglophone Canada.* PhD dissertation, Vrije Universiteit Amsterdam.

Kennison, S. M. (1999). American English usage frequencies for noun phrase and tensed sentence complement-taking verbs. *Journal of Psycholinguistic Research*, **28**, 165–177.

Kuznetsova, A., Brockhoff, P. B., & Christensen, Rune H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*. doi: 10.18637/jss.v082.i13

Lee, E.-K., Lu, D. H.-Y., & Garnsey, S. M. (2013). L1 word order and sensitivity to verb bias in L2 processing. *Bilingualism: Language and Cognition*, **16**, 761–775. doi: 10.1017/S1366728912000776

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, **12**, 251–285. doi: 10.1017/s0272263100009165

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, **101**, 676.

Mackey, A., & Marsden, E. (Eds.). (2015). *Advancing methodology and practice: The IRIS repository of instruments for research into second languages*. London: Routledge.

MacWhinney, B. (Ed.). (1987). *Mechanisms of language acquisition*. London: Routledge.

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford: Oxford University Press.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research*, **50**, 940. doi: 10.1044/1092-4388(2007/067)

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, **244**, 522.

Ortega, L. (2011). SLA after the Social Turn: Where cognitivism and its alternatives stand. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 167–180). London: Routledge.

Papadopoulou, D. (2005). Reading-time studies of second language ambiguity resolution. *Second Language Research*, **21**, 98–120. doi: 10.1191/0267658305sr254oa

Patson, N. D., Darowski, E. S., Moon, N., & Ferreira, F. (2009). Lingering misinterpretations in garden-path sentences: Evidence from a paraphrasing task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **35**, 280–285. doi: 10.1037/a0014276

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, **162**, 8–13. doi: 10.1016/j.jneumeth.2006.11.017

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24**, 940–961. doi: 10.1037//0278-7393.24.4.940

Pliatsikas, C., Johnstone, T., & Marinis, T. (2017). An fMRI study on the processing of long-distance wh-movement in a second language. *Glossa: A Journal of General Linguistics*, **2**, 1052. doi: 10.5334/gjgl.95

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. doi: 10.1017/S136672891200017X

Plüddemann, P. (2015). Unlocking the grid: Language-in-education policy realisation in post-apartheid South Africa. *Language and Education*, **29**, 186–199. doi: 10.1080/09500782.2014.994523

Posel, D., & Zeller, J. (2016). Language shift or increased bilingualism in South Africa: Evidence from census data. *Journal of Multilingual and Multicultural Development*, **37**, 357–370. doi: 10.1080/01434632.2015.1072206

Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. Chicago: University of Chicago Press.

Qian, Z., Lee, E.-K., Lu, D. H.-Y., & Garnsey, S. M. (2018). Native and non-native (L1-Mandarin) speakers of English differ in online use of verb-based cues about sentence structure. *Bilingualism: Language and Cognition*, **16**, 1–15. doi: 10.1017/S1366728918001001

R Core Team. (2018). R: A language and environment for statistical computing. Austria, Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rampton, M. (1990). Displacing the "native speaker": Expertise, affiliation, and inheritance. *ELT Journal*, **44**, 97–101.

Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics*, **32**, 299–331. doi: 10.1017/S0142716410000421

Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24**, 1521.

Spivey-Knowlton, M. J., Trueswell, J., & Tanenhaus, M. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology*, **47**, 276–309.

Sturt, P., Pickering, M. J., & Crocker, M. W. (1999). Structural change and reanalysis difficulty in language comprehension. *Journal of Memory and Language*, **40**, 136–150.

Tabor, W., & Hutchins, S. (2004). Evidence for self-organized sentence processing: Digging-in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **30**, 431.

Traxler, M. J. (2002). Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: Evidence from self-paced reading. *Quarterly Journal of Experimental Psychology: Section A*, **55**, 75–96.

Traxler, M. J. (2005). Plausibility and verb subcategorization in temporarily ambiguous sentences: Evidence from self-paced reading. *Journal of Psycholinguistic Research*, **34**, 1–30. doi: 10.1007/s10936-005-3629-2

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

Van Gompel, R. P., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, **45**, 225–258.

Veríssimo, J. (2018). Sensitive periods in both L1 and L2: Some conceptual and methodological suggestions. *Bilingualism: Language and Cognition*, **21**, 932–933. doi: 10.1017/S1366728918000275

Veríssimo, J., Heyer, V., Jacob, G., & Clahsen, H. (2018). Selective effects of age of acquisition on morphological priming: Evidence for a sensitive period. *Language Acquisition*, **25**, 315–326.

Warner, J., & Glass, A. L. (1987). Context and distance-to-disambiguation effects in ambiguity resolution: Evidence from grammaticality judgments of garden path sentences. *Journal of Memory and Language*, **26**, 714–738.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, **8**, 231–256. doi: 10.1162/jocn.1996.8.3.231

Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology*, **66**, 173–196. doi: 10.1146/annurev-psych-010814-015104

Appendix A

Table A.1. L1 participants' background data (standard deviation)

Age of L2 acquisition	5.7 (3)
English exposure (%)[a]	74 (14.6)
Afrikaans exposure (%)[a]	21.9 (15.5)
L2 speaking self-rating[b]	6.2 (1.5)
L2 reading self-rating	6.6 (1.6)
L2 comp. self-rating	7.4 (1.4)
C-test score (%)	75.3 (9.2)

[a]This is a self-reported estimate for current global exposure at the time of testing.
[b]Self-ratings are on a scale from 1 to 10.

Table A.2. Syllable length and frequency information for critical nouns, adjunct-clause items (Center for Lexical Information, 1993)

Pair number	Noun	Syllable length	Log frequency
1	boat	1	1.88
	horse	1	2.12
2	ladder	2	1.20
	housewife	2	1.15
3	tree	1	2.28
	fish	1	2.21
4	coffee	2	1.96
	flower	2	1.97
5	milk	1	2.0
	dog	1	2.06
6	beer	1	1.69
	song	1	1.70
7	cake	1	1.53
	truck	1	1.56
8	boy	1	2.54
	car	1	2.55
9	ice-cream	2	0.70
	puppy	2	0.85
10	plane	1	1.80
	song	1	1.70

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

Table A.3. Syllable length and frequency information for critical nouns, complement-clause items (Center for Lexical Information, 1993)

Pair number	Noun	Syllable length	Log frequency
1	issue	2	2.16
	report	2	2.16
2	mother	2	2.68
	problem	2	2.70
3	book	1	2.64
	girl	1	2.64
4	doctor	2	2.26
	story	2	2.36
5	boss	1	1.34
	crime	1	1.20
6	magazine	3	1.81
	professor	3	1.93
7	poem	2	1.40
	tutor	2	1.42
8	answer	2	2.12
	teacher	2	2.21
9	lady	2	2.05
	theory	2	2.15
10	king	1	2
	truth	1	2.13

https://doi.org/10.1017/S0142716420000314 Published online by Cambridge University Press

Table A.4. Verb biases for complement- vs. adjunct-clause constructions

Construction	Verb
Complement clause	confessed
	warned
	cautioned
	wrote
	resolved
	believed
	read
	advised
	proved
	confirmed
	suggested
Adjunct clause	sailed
	climbed
	painted
	drank
	played
	baked
	parked
	called
	walked
	polished
	rode
	ate
	flew
	sang

Cite this article: Berghoff, R. (2020). The processing of object–subject ambiguities in early second-language acquirers. *Applied Psycholinguistics* **41**, 963–992. https://doi.org/10.1017/S0142716420000314