Research Article

Children’s and Adolescents’ Processing of Temporary Syntactic Ambiguity: An Eye Movement Study

Paul E. Engelhardt

School of Psychology, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK

Correspondence should be addressed to Paul E. Engelhardt; p.engelhardt@uea.ac.uk

Received 11 April 2014; Revised 17 July 2014; Accepted 30 July 2014; Published 19 August 2014

Academic Editor: Glenda Andrews

This study examined the eye movements of 24 children and adolescents as they read sentences containing temporary syntactic ambiguities. Prior research suggested that children primarily use grammatical information when making initial parsing decisions, and they tend to disregard semantic and contextual information. On each trial, participants read a garden path sentence (e.g., While the storm blew the boat in the shed), and, afterwards, they answered a comprehension question (e.g., Did the storm blow the boat?). The design was 2 × 2 (verb type × ambiguity) repeated measures. Verb type was optionally transitive or reflexive, and sentences were ambiguous or unambiguous. Results showed no differences in first pass reading times at the disambiguating verb (e.g., sat). However, regressions did show a significant interaction. The unambiguous-reflexive condition had approximately half the number of regressions, suggesting less processing difficulty in this condition. Developmentally, we found that adolescents had significantly better comprehension, which seemed to be linked to the increased tendency to regress from the disambiguating word. Findings are consistent with the assumption that the processing architecture is more restricted in children compared to adolescents. In addition, results indicated that variance in ambiguity resolution was associated with interference control but not working memory.

1. Introduction

In the current study, we investigated sentence processing in children and adolescents. Garden path sentences contain a temporary syntactic ambiguity, in which a reader or listener is led into an incorrect interpretation, and then later material forces a revision of that earlier incorrect parse. Investigations of sentence processing have shown that children tend to rely on syntactic information, rather than semantic or contextual information when making initial parsing decisions. Much of this work used the visual world paradigm and spoken utterances. For example, Trueswell et al. [1] investigated 4- and 5-year-old children’s use of referential context to resolve ambiguous prepositional phrases, such as put the frog on the towel in the box. The location (or goal) interpretation of on the towel is typically preferred because it is syntactically simpler, and the verb put always requires a location of the putting action [2]. However, when there are two frogs in the display and one is on a towel, adult participants do not show evidence of pursuing the goal interpretation, which has been taken as evidence for a reduction in processing difficulty based on visual (or referential) context [3, 4]. Children, however, did not show evidence of using visual context to resolve the temporary ambiguity, and they made a substantial number of errors when carrying out these types of instructions (see also [5–8]).

There have been comparatively fewer sentence processing studies of reading in children. Traxler [9] investigated plausibility and verb selectional restrictions on syntactic ambiguity resolution using a self-paced reading task. He found that children (8–12 years old) did not use plausibility to avoid the garden path in a sentence, such as When Sue tripped the table fell over and the vase was broken. The implausibility occurs because it is not possible to trip an inanimate object. In the same study, Traxler also tested selectional restrictions (e.g., When Sue fell the policeman stopped and helped her up). Results showed that, in both types of sentences, children showed disruption of processing at the disambiguating verb (e.g., fell and stopped, resp.), which suggests that they interpreted the ambiguous noun phrase as a direct object despite the presence of contradictory (semantic) information. This finding indicates the use of a simplicity heuristic
(e.g., minimal attachment). If readers process a noun-verb-noun sequence, then parse that sequence as subject-verb-object, which is syntactically the simplest option.

In another relevant study, Joseph et al. [10] investigated children's processing of plausible and implausible sentences using eye tracking. The sentential paradigm, shown in examples (1a)–(1c), was adopted from Rayner et al. [11]. The sentences described an individual performing an action. Sentence (1b) is implausible compared to sentence (1a), because axes are not used to chop food, although axes are typically used to chop other things. Sentence (1c), in contrast, is anomalous because carrots cannot be inflated. Joseph et al. found that disruption occurred earlier when sentences were anomalous compared to implausible. Both children and adults showed equivalent disruption with the anomalous sentences. Children were, however, delayed in the detection of the implausible relation in sentence (1b). Based on these results, Joseph et al. concluded that children were similar to adults in their processing of thematic anomalies but had slower access to semantic information, which explains the delayed disruption for sentences such as (1b).

John used a knife to chop the large carrots for dinner.

(1b) John used an axe to chop the large carrots for dinner (implausible).

(1c) John used a pump to inflate the large carrots for dinner (anomalous).

Summing over existing literature, it seems that children (1) tend to rely on grammatical (category) information and structural simplicity when processing temporary ambiguities and (2) they have slower access to semantic information [1, 10]. Specifically, children interpreted a noun phrase following a verb as the direct object even when that interpretation was not semantically plausible. The combination of these findings suggests that the processing architecture of children is more restricted with respect to the types of information that can affect online/initial parsing decisions [12–14].

In the current study, we also tested several other cognitive/executive abilities that have been assumed to be associated with sentence processing. We did so for two reasons. The first is that executive control is not fully developed in children compared to adults, and therefore, there may be processing differences based on age. The second is that we wanted to examine how sentence processing is affected by more general-purpose cognitive/executive control mechanisms. (Executive functions are assumed to be control mechanisms that regulate thoughts and action. They have been widely linked to PFC function, and the most commonly postulated executive functions are updating/monitoring working memory, switching flexibly between tasks/mental sets, and inhibiting dominant/prepotent responses [15–18].) For example, individual differences in working memory have sometimes been shown to be associated with sentence comprehension [19–25]. Working memory in sentence comprehension is typically measured with some version of a reading span task [26]. Caplan and Waters [27] concluded that the language comprehension system is isolated from domain-general cognitive abilities, such as working memory. Consistent with this view, DeDe et al. [28] found that online sentence processing is not affected by working memory capacity, but offline comprehension measures are affected by working memory capacity.

In contrast, Booth et al. [29] concluded that children's speed of comprehension was associated with short-term memory span. However, children with higher spans had slower reading times at the point of difficulty in object-relative sentences (e.g., The lawyer that the banker irritated filed a lawsuit), which suggests that individual differences in the ability to store words in short-term memory are related to the speed of sentence comprehension. The explanation offered was that high-span readers encoded information in short-term memory in order to facilitate processing at the difficult region of the sentence. In addition, Felser et al. [6] found that children, who had lower (listening) spans, had a preference for local attachments when a relative clause attachment site was ambiguous (e.g., The sister of the actress who shot herself on the balcony was under investigation). Swets et al. [30] found the opposite result in adults. Adults with low working memory showed a preference for high attachments, which those authors interpreted as evidence for differential chunking based on implicit prosody. Therefore, there is some evidence to suggest that individual differences in memory abilities, at least in children, are associated with both online and offline sentence comprehension measures, but overall, the evidence is mixed. In addition, the patterns between adults and children are also mixed.

Returning to sentence processing, the syntactic structures tested in the current work have been used in many previous studies with adults (see examples (2a) and (2b)). In these types of sentences, an ambiguous noun phrase (shown in italics) is initially interpreted as the direct object of the first verb. However, later information reveals that it is actually a subject and that the verb in the subordinate clause is either intransitive or reflexive. Christianson et al. [31] found that readers would often misinterpret these sentences as if the ambiguous noun phrase was both the object of the first verb and the subject of the second verb, an interpretation that is not grammatically permissible. Christianson et al. [31] concluded that the thematic roles that were initially and incorrectly assigned along the “garden path” often persisted into the final interpretation of the sentence. These effects have been replicated in many subsequent studies (e.g., [32–37]).

(2a) While Anna bathed, the baby played happily in the crib. (reflective)

(2b) While the man hunted, the deer ran through the woods. (optionally transitive)

In the current study, we were primarily interested in determining how children and adolescents process these types of ambiguities when reading. Temporary ambiguities are pervasive in language, and so efficient and accurate reading is dependent on the ability to revise ambiguities when they occur. Revision processes can be affected by several factors. One factor is the type of verb in the subordinate clause. When the verb is reflexive, participants have an easier time revising the misinterpretation compared to when
the verb is optionally transitive [38]. Reflexive verbs require the agent and patient to be the same person because the action is done to oneself. In addition, these verbs do not require a structural change from transitive to intransitive, as do the optionally transitive verbs. We expected children to show disruptions at the disambiguating word [9, 10], and, so, we examined first pass reading time, total reading time, and the number of regressions. We also investigated whether children recover from the misinterpretation or whether the misinterpretation “lingers” in the final interpretation by asking a comprehension question following each sentence [31].

Because this study tested children and adolescents, we used easier versions of the subject/object ambiguity than have been tested previously (see examples (3a) and (3b)). These sentences are easier than those used in earlier studies because the main clause explicitly negates the misinterpretation. In sentence (3a), the storm cannot blow the boat, if the boat is inside of a shed. In contrast, in sentence (2b), it is plausible that the deer ran through the woods while the man hunted the deer.

(3a) While the storm blew (,) the boat sat in the shed. (optionally transitive)

Question: Did the storm blow the boat?

(3b) As Sarah washed (,) the dog ran in the garden. (reflexive)

Question: Did Sarah wash the dog?

In summary, the eye movements of children and adolescents were monitored to investigate how they resolve temporary syntactic ambiguity. In garden path sentences, the critical word is called the disambiguation point (e.g., sat in example (3a)). At the point of disambiguation, the processing system may require several different processing operations in order to integrate that word with what has come previously. When sentences are disambiguated with a comma, reading times on the disambiguating verb are expected to be reduced. In addition, sentences with reflexive verbs should also be easier to process because the subject and object must refer to the same person. In contrast, with optionally transitive verbs, the subordinate clause verb must be revised from transitive to intransitive (i.e., the direct object is left unspecified). The current study also had two additional goals. Previous studies have only investigated sentence processing in children based on semantic implausibility and referential context [1, 9, 10]. In the current study, we examined adolescents, in addition to children. We were specifically interested in how processing tendencies change over the developmental period from childhood to adolescence. The second subgoal focused on the relationship between sentence processing measures and cognitive/executive function. Therefore, we collected several short-term/working memory measures and one measure of interference control. Recall, that there are mixed findings with respect to how differences in working memory are related to language comprehension [21–30].

2. Method

2.1. Participants. Participants were 24 children between the ages of 9 and 16 years (M = 13.58, SD = 2.47; 41.7% were male). All participants were native speakers of English and had normal or corrected-to-normal vision. Children were accompanied by a parent who completed a short demographic questionnaire and gave written informed consent. All participants were community recruited from local schools in the Newcastle/Gateshead area. Participants completed a battery of standardized tests to assess verbal short-term memory, verbal working memory, reading span, and interference control. Participants were reimbursed £10 for travel expenses.

2.2. Standardized Measures. Verbal short-term memory and verbal working memory. Participants completed three tests from the Automated Working Memory Assessment [39]. Tests were word recall, nonword recall, and listening recall. Word recall and nonword recall assess short-term memory span and require participants to recall lists of word and nonwords, respectively. The test started with one word and increased by one after the completion of four correct trials. The highest number tested was seven, and participants were required to recall the words in the correct order. The nonword recall task went up to a maximum of six, and both tests stopped if the participant responded incorrectly on three consecutive trials. For listening recall, participants listened to a sentence then responded as to whether the sentence made sense or not. Participants were also required to remember the final word in each sentence. At the end of a sequence of sentences, participants had to recall the final words. The practice trials consisted of one and two sentence sequences. The regular session trials started at one and went up to a maximum of six. The test finished when the participant got three consecutive trials incorrect. Higher scores on all three tests indicated better performance, and scores were age standardized.

2.2.1. Reading Span. For the reading span task, participants read a sentence out loud and remembered the final word. There were two practice trials that consisted of two sentence sequences. Participants then completed the experimental trials, which consisted of three trials at each span size 2–5. Each trial began with a drift correction dot that was presented on the left hand side of the computer screen, centered top-to-bottom. The experimenter initiated the trial, which consisted of the first sentence being presented. After the participant finished reading the first sentence out loud, the experimenter advanced the experiment to reveal the next sentence. At the end of a sequence of sentences, the word “recall” appeared on the screen, which was the participants cue to recall the final word of each sentence. Participants were free to recall the words in any order. Sentences were taken from technical report-CDRGI [40].

Participants responded orally into a microphone, and software automatically recorded responses in.wav format using an ASIO sound card. We calculated a reading span score.
in the following manner. For each span 2–5, we calculated the proportion of words correctly recalled, which results in a proportion for each of the four spans. We then fit a linear function to the data points using a simple regression with span as the predictor and proportion of words correctly recalled as the dependent variable. The standardized regression coefficient was used as proxy for reading span.

2.2.2. Interference Control. The ability to manage interference was assessed using the same worlds opposite worlds task from the Test of Everyday Attention for Children [41, 42]. Stimuli consisted of four cards that displayed a path of randomly sequenced digits “1’s” and “2’s.” The participant is asked to name the digits as quickly as possible, and the time to work through the path is recorded with a stop watch. In same worlds, the participant names the “1’s” and “2’s” as one and two, respectively. In opposite worlds, the participant does the reverse, calling “1’s” and “2’s” two and one, respectively. Each participant did two opposite world trials and two same world trials. We averaged the time for the two opposite world trials and the two same world trials, and then we subtracted the same average from the opposite average. Higher scores indicated less ability to manage interference.

2.3. Materials

2.3.1. Sentence Comprehension. Each participant read 59 sentences: three were for practice, 16 were garden path sentences like those shown in examples (3a) and (3b), and 40 were fillers. Half of the critical items were disambiguated with a comma. The fillers consisted of 16 main-subordinate structures, eight subordinate-main structures, and 16 coordination structures. All main-subordinate and subordinate-main structures had subordinate clauses that were transitive. The coordination structures consisted of two conjoined sentences. The remaining four fillers were subordinate-main structures in which the subordinate verb was intransitive. These sentences did not contain a comma. Over the entire experiment, the correct response for half of the trials was “yes” and for half of the trials was “no.” The critical sentences were rotated across two lists in a Latin square design with half the participants completing each list. For the critical items (see examples (3a) and (3b)), the ambiguous noun phrase was a plausible object of the verb.

2.4. Apparatus. Eye movements were recorded with an SR research Eyelink 1000 eye tracker sampling at 1000 Hz. Viewing was binocular, but only the position of the right eye was tracked. Stimulus presentation was programmed using SR research Experiment Builder software. The eye tracker and a 19” CRT display monitor (refresh rate of 140 Hz) were interfaced with a 3 GHz Pentium 4 PC, which controlled the experiment and logged the position of the eye throughout the experiment. For both the reading span task and the sentence comprehension task, participant’s heads were stabilized with the use of a chinrest.

2.5. Design and Procedure

2.5.1. Sentence Comprehension. At the start of each trial, participants were required to fixate a drift-correct dot. If the experimenter was satisfied with the calibration, he initiated the trial. The sentence appeared, and the initial letter of each sentence was in the same position, in terms of x and y coordinates, as the drift correct dot (i.e., on the left edge of the monitor and centered vertically). The entire sentence was presented on a single line on the screen. The participants read the sentence silently for as long as they liked and then pressed the spacebar on the keyboard to view the comprehension question. Participants pressed a “green” button to answer yes and a “red” button to answer no.

The design was 2 × 2 (verb type × ambiguity) repeated measures. Verb type was optionally transitive or reflexive, and ambiguity refers to whether the sentence was temporarily ambiguous or unambiguous. The disambiguated version had a comma following the intransitive verb. The dependent variables were reading times, regressions, and comprehension accuracy. For the eye movements, the main regions of analysis were the ambiguous noun phrase, the disambiguating verb, and the final noun phrase in the main clause. The length and frequency of the critical words in these three regions of interest are presented in Table 1. Statistical analyses revealed that there were no significant differences (all P’s > .24) for any of the three regions on any measure.

Eye movement analyses focused on first pass reading time, total reading time, and the proportion of trials with a regression. First pass reading time is the sum of fixations from the moment the eyes first enter a region to the moment the eyes leave that region either in a forward or backward direction. The total time was calculated by summing the duration of all fixations that occurred in a particular interest area. We also examined the regressions out of an interest area, which are when a backwards eye movement (i.e., a right-to-left eye movement in English) is made from the current interest area to an earlier one. We calculated a proportion based on the number of trials per condition that contained a regression. We calculated the proportion of regressions, which included all regressions, whether they occurred on first, second, or third pass reading. Lower scores for all three measures indicate less processing difficulty. Analyses were also conducted on the proportion of correct responses to the comprehension questions. Higher scores indicate better performance. The entire experimental session (including all tasks) lasted approximately 60 minutes.

3. Results

The means for all standardized measures are presented in Table 2. For all of the following analyses, we first checked the data for outliers. Any value greater than three standard deviations from the mean in each condition was replaced with the mean in that condition. This affected less than 2% of the data. The results consist of two main sections. In the first, we examined results from a 2 × 2 (verb type × ambiguity) repeated measures ANOVA for both eye movements and comprehension accuracy. These analyses
were conducted with both subjects (F1) and items (F2) as random effects. The second set of analyses consisted of a series of multiple regressions to determine whether sentence processing and sentence comprehension were related to age, short-term/working memory, or interference control.

### 3.1. Eye Movements

#### 3.1.1. Ambiguous Noun Phrase. First pass reading times showed a significant main effect of verb type F(1,23) = 4.44, P < .05; F(2,1,7) = 3.64, P < .10, in which the optionally transitive verbs had higher reading times (see Figure 1(a)). Total reading time showed that both main effects were significant: ambiguity F(1,23) = 13.62, P < .01; F(2,1,7) = 15.10, P < .01 and verb type F(1,23) = 38.86, P < .01; F(2,1,7) = 18.73, P < .01. The ambiguous conditions and the optionally transitive verbs both resulted in higher reading times. The proportion of trials with a regression also showed two significant main effects: ambiguity F(1,23) = 5.95, P < .05; F(2,1,7) = 9.50, P < .05 and verb type F(1,23) = 10.98, P < .01; F(2,1,7) = 3.86, P < .10. Again, the ambiguous sentences and optionally transitive verbs had a higher proportion of trials with a regression, which suggests that both variables produce additive effects. In summary, there was a robust effect of verb type for all three measures, indicating that the optionally transitive verbs were more difficult to process compared to the reflexive verbs. Total reading time and regressions also showed an effect of ambiguity; when the sentence is ambiguous participants slowed down and made more regressions.

#### 3.1.2. Disambiguating Verb. The disambiguating verb is the key word in the sentence because in most cases it forces the comprehension system to revise an incorrect interpretation. For first pass reading time, there were no significant differences, which is consistent with the idea that children are somewhat slower and less effective readers (i.e., they do not immediately use the comma as a disambiguating signal). For total reading time, there was a marginal effect of ambiguity F(1,23) = 4.13, P = .054; F(2,1,7) = 1.69, P > .10, in which the ambiguous sentences had longer reading times (see Figure 2(b)). The interaction was not significant. However, there were larger differences between ambiguous and unambiguous sentences with reflexive verbs (i.e., ~130 ms), but almost no difference between ambiguous and unambiguous sentences was observed with optionally transitive verbs (i.e., ~6 ms).

Regressions showed a main effect of verb type F(1,23) = 4.80, P < .05; F(2,1,7) = 4.44, P < .07, in which there were more regressions with optionally transitive verbs. There was also a significant interaction between verb type and ambiguity F(1,23) = 4.47, P = .05; F(2,1,7) = 12.41, P < .01. The interaction was driven by a significant difference between the ambiguous and unambiguous conditions with reflexive verbs t(23) = 2.68, P < .05. Regressions were nearly twice as likely with the reflexive verbs in the ambiguous condition compared to the unambiguous condition, and performance was similar in both of the optionally transitive conditions. There was also a significant difference between the two unambiguous conditions t(23) = 3.48, P < .01, but the other two paired comparisons were not significant.

These results suggest two important facts about the resolution of temporary ambiguity in young readers. The first is that participants seemed to experience difficulty in the "unambiguous" condition with optionally transitive verbs, which may indicate that participants attached the ambiguous noun phrase, despite the comma. The second is that there was no difference between the two verb types when the sentence was ambiguous. This suggests that both types of verbs are processed equally at the disambiguating word. Perhaps, most surprisingly, this effect even occurred in total reading time, which is generally regarded as a late measure of processing.

In summary, eye movement results for regressions (and to a lesser extent total reading times) are more or less equal in

---

**Table 1: Means and standard deviations for length and frequency of key words.**

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>KFWF</th>
<th>BVF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflexive verbs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguous noun (N − 1)</td>
<td>3.71(0.76)</td>
<td>76.33(72.29)</td>
<td>13.80(70.9)</td>
</tr>
<tr>
<td>Disambiguating verb</td>
<td>4.13(1.36)</td>
<td>312.67(207.03)</td>
<td>210.25(357.27)</td>
</tr>
<tr>
<td>Final noun</td>
<td>4.00(0.00)</td>
<td>95.17(60.96)</td>
<td>10.33(11.31)</td>
</tr>
<tr>
<td>Optionally transitive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguous noun (N − 1)</td>
<td>3.87(0.99)</td>
<td>87.63(89.82)</td>
<td>14.43(16.72)</td>
</tr>
<tr>
<td>Disambiguating verb</td>
<td>3.50(0.76)</td>
<td>215.63(167.59)</td>
<td>211.50(356.39)</td>
</tr>
<tr>
<td>Final noun</td>
<td>4.57(1.13)</td>
<td>77.00(40.56)</td>
<td>8.43(6.75)</td>
</tr>
</tbody>
</table>

*Note. Length is number of letters, KFWF is Kucera-Francis written frequency, and BVF is Brown verbal frequency.*

---

**Table 2: Mean, standard deviation, and range for all standardized measures.**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean (SD)</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word recall</td>
<td>23</td>
<td>100.97(14.14)</td>
<td>77</td>
<td>130</td>
</tr>
<tr>
<td>Nonword recall</td>
<td>23</td>
<td>93.43(11.39)</td>
<td>71</td>
<td>120</td>
</tr>
<tr>
<td>Listening recall</td>
<td>23</td>
<td>100.97(13.36)</td>
<td>77</td>
<td>144</td>
</tr>
<tr>
<td>Reading span</td>
<td>20</td>
<td>−18(0.16)</td>
<td>−0.77</td>
<td>−0.035</td>
</tr>
<tr>
<td>Interference control</td>
<td>23</td>
<td>2.70(1.45)</td>
<td>0</td>
<td>5.5</td>
</tr>
</tbody>
</table>

*Note. Technical problems resulted in the loss of reading span data from three participants.*
three of the four conditions. The condition that is significantly easier to process is the unambiguous condition with reflexive verbs. Therefore, it seems that participants needed both a disambiguating comma and an “easier” verb in order to avoid disruption of processing at the disambiguating word in the sentence.

3.1.3. Final Noun Phrase. First pass reading times showed a main effect of verb type \( F(1, 23) = 10.96, P < .01; F(2, 7) = 2.69, P > .10 \), in which there were longer reading times with the reflexive verbs compared to the optionally transitive verbs (see Figure 3(a)). For total reading time, there were no significant differences. The proportion of trials with a regression also showed a main effect of verb type \( F(1, 23) = 5.01, P < .05; F(2, 1, 7) = 2.83, P > .10 \). However, in this case, there were more regressions in the optionally transitive conditions, which is opposite to first pass reading times. These data suggest that people initially spend less time viewing the final noun and, then, are more likely to make a regression when the verb in the subordinate clause is optionally transitive. Importantly, there was no effect of ambiguity on any of the three measures, which suggests that there was no spillover or downstream consequences of the ambiguity past the disambiguating verb.

3.2. Comprehension Accuracy. The proportion of correct responses showed only a main effect of verb type \( F(1, 23) = 21.20, P < .05; F(2, 7) = 24.81, P < .01 \) (see Figure 4(a)). There were more incorrect responses with optionally transitive verbs compared to reflexive verbs. As mentioned previously, most earlier studies showed an interaction in comprehension accuracy, in which there were larger differences between the ambiguous and unambiguous sentences with reflexive verbs compared to optionally transitive verbs (e.g., [38, 41, 42]). We think that the absence of an interaction in the current study is due to the fact that the main clause cancels the garden path misinterpretation, which results in a substantially higher proportion of correct responses overall. As can be seen in Figure 4, there was an ambiguity trend in the data. Participants made slightly more incorrect responses when the sentences were ambiguous, and the difference between ambiguous and unambiguous sentences was larger for the optionally transitive verbs compared to reflexive verbs. We return to comprehension accuracy below in regard to age differences.

3.3. Multiple Regressions. We conducted a series of six multiple regressions to follow up the main findings reported above. More specifically, we wanted to follow up the effects
we observed in total reading time and the proportion of regressions at the disambiguating verb. We were also interested in comprehension accuracy, even though there was not a significant interaction in comprehension accuracy. To do so, we conducted multiple (backwards) regressions with six predictor variables (i.e., age, word recall, nonword recall, listening span, reading span, and interference control). In addition, we also focused these follow up analyses on the two ambiguous conditions. The bivariate correlations between all variables are presented in Table 3. (The predictor variables are 1–6 and the dependent variables are 7–12.).

Results showed that, across the six regressions, there were three variables that were retained (i.e., age, interference control, and nonword recall). However, they were not consistent in every analysis. Age was positively correlated with comprehension accuracy and negatively correlated with total reading time. This pattern is consistent with older participants being more efficient and accurate readers (i.e., the opposite of a speed-accuracy tradeoff). The regression analyses showed that age was a significant predictor with optionally transitive verbs for both total reading time and comprehension accuracy but was only significant with reflexive verbs on comprehension (see Table 4). These findings indicate that as children get older, they spend less time reading and are more likely to answer comprehension questions correctly.

The second measure that was retained in the regression analyses was interference control (i.e., performance on the same worlds opposite worlds task). Recall that, in the opposite world trials, participants must monitor response conflict and suppress a competing response. In the opposite worlds trials, this entails inhibiting the prepotent response to name the digits according to their “correct” names. Performance on this task produced a significant effect with reflexive verbs on regressions and with optionally transitive verbs on comprehension accuracy. The correlations (and beta’s) were negative, suggesting that participants who were poorer at controlling interference were less likely to make a regression and less accurate on comprehension questions. The effect was strongest in the reflexive-ambiguous condition (see Tables 3 and 4). We think that this pattern makes sense, because, in this condition, there is a competition between two potential direct objects of the verb in the subordinate clause. The first is the ambiguous noun phrase, which ultimately turns out to be incorrect, and the second is the subject of the subordinate clause, which ultimately turns out to be correct. The final measure retained was nonword recall, which assesses short-term memory span. Here results showed negative correlations with total reading times, and the same pattern held for both types of verbs. This indicates that participants with higher memory spans spent less time reading the disambiguating
Figure 3: Results from the final noun phrase in the sentence. (a) shows the first pass reading times, (b) shows the total reading time, and (c) shows the proportion of trials with a regression. Error bars show the standard error of the mean.

Table 3: Bivariate correlations between age, executive function, and sentence comprehension measures.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Age</td>
<td>—</td>
<td>.04</td>
<td>−.26</td>
<td>.03</td>
<td>.33</td>
<td>−.14</td>
<td>−.19</td>
<td>−.32</td>
<td>.05</td>
<td>.41</td>
<td>.46</td>
<td>.58</td>
</tr>
<tr>
<td>(2) Word recall</td>
<td>—</td>
<td>.44*</td>
<td>.44*</td>
<td>−.24</td>
<td>.33</td>
<td>−.09</td>
<td>−.17</td>
<td>.07</td>
<td>−.01</td>
<td>.08</td>
<td>.26</td>
<td></td>
</tr>
<tr>
<td>(3) Nonword recall</td>
<td>—</td>
<td>.31</td>
<td>−.16*</td>
<td>.19</td>
<td>−.07</td>
<td>−.09</td>
<td>−.03</td>
<td>−.26</td>
<td>.13</td>
<td>.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Listening recall</td>
<td>—</td>
<td>−.12</td>
<td>—</td>
<td>.32</td>
<td>.23</td>
<td>.04</td>
<td>.43</td>
<td>.13</td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Reading span</td>
<td>—</td>
<td>−.39*</td>
<td>—</td>
<td>.26</td>
<td>−.09</td>
<td>−.52*</td>
<td>−.20</td>
<td>−.29</td>
<td>−.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) Interference control</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.68**</td>
<td>.60*</td>
<td>—</td>
<td>.06</td>
<td>.19</td>
<td>−.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) RA—total time</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.30</td>
<td>.14</td>
<td>.07</td>
<td>—</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8) OA—total time</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.03</td>
<td>.54**</td>
<td>—</td>
<td>.20</td>
<td>.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9) RA—regressions</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.39*</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10) OA—regressions</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11) RA—comprehension</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12) OA—comprehension</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


This finding is opposite to the one reported by Booth et al. [29]. Those authors found that higher spans had slower reading times at the point of difficulty in a sentence. Instead, our nonword recall results seem to dovetail more closely with the age effects. Higher performers spent less time reading the disambiguating word. There was one other significant effect. Nonword recall was retained in the comprehension analysis for optionally transitive verbs and the relationship was positive.

4. Discussion

In the current study, we found that after reading a garden path sentence participants were more likely to retain
the misinterpretation when the verb in the subordinate clause was optionally transitive. Overall, participants answered incorrectly 35% of trials with optionally transitive verbs and 11% of trials with reflexive verbs. There were no significant differences in comprehension accuracy based on ambiguity. However, the lack of an ambiguity effect is not altogether surprising and likely due to the fact that the main clause ruled out (or negated) the temporary misinterpretation in every sentence. Thus, this study produced higher overall accuracy rates than did previous studies (e.g., [31]). The eye movement measures did reveal significant interactions at the disambiguating verb. Specifically, there were fewer regressions and less time spent viewing the disambiguating verb in the unambiguous-reflexive condition. This indicates substantially less processing effort in this condition. Interestingly, the unambiguous condition with optionally transitive verbs patterned similarly to the two ambiguous conditions suggesting that participants experienced processing difficulty despite the presence of a comma, which should reduce garden path tendencies. These data indicate that participants, 9–16 years old, needed both a disambiguating comma and an easier verb in order to circumvent processing difficulty at the disambiguating word.

Recall that Traxler [9] and Joseph et al. [10] found that children had slower access (or slower integration) of different information sources during reading. We believe that our data is consistent with the conclusions from these two studies. In particular, we observed few differences between the four conditions in first pass reading times (see Figure 1(a)).

**FIGURE 4:** Results for comprehension accuracy. (a) shows the full sample, (b) shows the children, and (c) shows the adolescents. Error bars show the standard error of the mean.

**TABLE 4:** Standardized regression coefficients and \( t \) values for the significant predictor variables in the follow-up multiple regressions.

<table>
<thead>
<tr>
<th>Criterion variable</th>
<th>Predictor(s)</th>
<th>( \beta )</th>
<th>( t ) value (21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflexive—ambiguous</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total reading time</td>
<td>Nonword recall</td>
<td>–.41</td>
<td>–2.09*</td>
</tr>
<tr>
<td></td>
<td>Interference</td>
<td>–.52</td>
<td>–2.56*</td>
</tr>
<tr>
<td></td>
<td>Comprehension</td>
<td>.45</td>
<td>2.30*</td>
</tr>
<tr>
<td>Optionally transitive—ambiguous</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total reading time</td>
<td>Nonword recall</td>
<td>–.44</td>
<td>–2.23*</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>–.43</td>
<td>–2.17*</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>.63</td>
<td>3.91**</td>
</tr>
<tr>
<td></td>
<td>Comprehension</td>
<td>.39</td>
<td>2.36*</td>
</tr>
</tbody>
</table>

Note. **\( P < .01 \), *\( P < .05 \).
In Figures 1(b) and 1(c), there was a clear pattern for total reading time and regressions in which both main effects were significant, and so the general pattern is similar in (b) and (c). In contrast, the pattern in first pass reading times is different, and the difference is primarily due to the fact that the unambiguous-reflexive verb condition is quite high (i.e., the rightmost bar in (a)). If one now considers the pattern in Figure 2, some interesting similarities emerge. Panels (b) and (c) show an interaction driven by low values in the unambiguous-reflexive verb condition. Panel (a), in contrast, shows no significant differences between any of the four conditions. Again, the difference comparing (b) and (c) to (a) is that the unambiguous-reflexive verb condition is high, just as it is in Figure 1(a). What this shows is that first pass reading times do not show any effects of the experimental manipulations. In all cases, participants' first pass reading times do not show sensitivity to the presence of a comma, which should be a reliable signal to syntactic structure. Thus, in all four conditions, participants appear to be initially attaching the ambiguous noun phrase as the object of the subordinate verb, which is similar to what Traxler [9] reported for implausible relations.

This tendency to attach the ambiguous noun phrase to the subordinate verb in all conditions may be due to a combination of factors, such as less experience with punctuation and complex syntactic structures, a reliance on syntactic parsing heuristics, and/or a failure to apply the correct (implicit) prosody to the sentence. One of the most widely debated topics in sentence comprehension is whether initial parsing is based on processing heuristics (i.e., minimal attachment and late closure) or a constraint-satisfaction process [13, 43–46]. The Good Enough Theory makes similar predictions regarding the use of heuristics in comprehension [47–49] (i.e., when processing demands are high and cognitive resources low, the processing system will employ fast and frugal decision making [50–52].) Our results seem to be more consistent with the former as participants had a strong tendency to interpret the noun phrase following the subordinate verb as the direct object. This pattern of results is consistent with minimal attachment, which assumes that the processing system will initially build the simplest structure possible.

A second possibility is implicit prosody. The implicit prosody hypothesis suggests that readers project a default prosodic contour onto text during silent reading and that it influences readers' syntactic analysis [53]. One explanation for the fact that the comma did not seem to have an effect on parsing in this study is that participants do not project the correct prosody on to the sentence. The “correct” prosodic contour for a subordinate-main structure, such as While the storm blew the boat sat in the shed, is to place a prosodic break in between the subordinate verb (e.g., blew) and the ambiguous noun phrase (e.g., the boat) [33, 54–56]. It may be the case that younger readers do not automatically associate the presence of a comma with the placement of prosodic boundary, which leads to online garden path effects even in unambiguous sentences.

In summary, the findings from this study are consistent with previous work insofar as child and adolescent readers only showed differences between the experimental conditions in late eye movement measures, which suggests that they are slower and less efficient processors of written language. The other main finding from this study was that participants appeared to show garden path effects in the unambiguous condition with optionally transitive verbs. We attributed this effect to the application of parsing heuristics (i.e., minimal attachment) and/or a failure to apply the correct implicit prosody. Another possibility, which we think is unlikely, is that children lack explicit training on the use of commas in written language. The primary reason that we think it is unlikely is that the comma clearly has an effect on total reading times and regressions at the ambiguous noun phrase and regressions at the disambiguating word, which suggests that the comma is not systematically ignored. Moreover, many (if not most) university-level students fail to mark subordinate/main clause boundaries with a comma, which is suggestive evidence that people are not given instruction at this level of specificity. In the next sections, we discuss the main findings of the second set of analyses (i.e., the multiple regressions).

4.1. Age. Age was retained as a significant predictor in total reading time, comprehension accuracy, and the proportion of trials with a regression. Moreover, the relationships were substantially stronger for optionally transitive as compared to reflexive verbs. (In Figure 4(b), we present the means for children, 13 years and younger (N = 10), and in Figure 4(c), we present the means for adolescents, 14 years and older (N = 14).) The correlations in Table 3 revealed positive correlations between age and comprehension accuracy, suggesting that as children get older they become better at avoiding garden path misinterpretations. Age also produced a significant effect on the number of regressions and total reading times, but only with optionally transitive verbs. Again, the correlations revealed that as age increased participants were more likely to make a regression. The reverse occurred for total reading time. Rayner [57] found that children made longer fixations and more regressions compared to adult readers. However, we do not believe that our data is inconsistent with what Rayner found, because our analyses focused on the disambiguating word in a garden path sentence.

As we mentioned in the Introduction, garden path sentences almost always draw readers into a misanalysis, and, therefore, making a regressive eye movement at the error signal is a good thing to do in order to reprocess the sentence. Regressions at a disambiguating word, for the most part, are thought to arise from an inability to integrate the current word with the previous context, and as a result, participants must go back and reread the sentence [14, 57]. Returning to the issue of whether children make more regressions compared to adults, we interpret Rayners' results as showing that children make more regressions compared to adults in places where it is not necessarily beneficial to do so. In contrast, our results show that adolescents are more likely to make a regression, specifically, in those places where it is beneficial for comprehension.
4.2. Interference Control. The ability to control interference also produced a significant effect on the number of regressions at the disambiguating word when the subordinate verb was reflexive. The correlations, in Table 3, generally reveal negative relationships between interference control and the proportion of regressions and comprehension accuracy. With interference control, higher scores indicated poorer performance or poorer ability to manage a competing response. Therefore, a negative correlation suggests that individuals with less ability to manage interference were less likely to make a regression. A relationship between interference control and the ambiguous-reflexive verb condition is quite interesting. We expected participants to attach the ambiguous noun phrase as the direct object of the subordinate verb and then the disambiguating verb signals that the ambiguous noun phrase is a subject and not an object. With reflexive verbs, the revision processes are easier because instead of making the subordinate verb intransitive, reflexives are revised to make the direct object implicitly coreferential with the subject of the subordinate clause. Therefore, in the case of reflexive verbs, there are two entities that compete to fulfill the object role (i.e., the ambiguous noun phrase and the subject of the subordinate clause). We believe that the relationship between regressions and interference control is likely due to the competition between these two potential entities for the object role of the subordinate verb. Participants that fail to make regressions are less likely to resolve this competition. The pattern of correlations is similar for comprehension accuracy.

4.3. Short-Term/Working Memory. In the current study, we observed several significant relationships between nonword recall, a short-term memory task, and the dependent measures. Participants with higher spans spent less time reading the disambiguating word and had better comprehension for sentences containing optionally transitive verbs. As mentioned above, these results are counter to the findings from Booth et al. [28]. We did not observe any significant effects with the two working memory measures (i.e., listening recall and reading span). As with all null effects, it is impossible to draw definitive conclusions, and this is especially true in the current study given the size of the sample. Recall also that the reading span data were lost from three participants, which means that power is even lower for this particular measure (see Table 3). However, because significant relationships between garden path sentence comprehension and working memory have been found in past research, we feel it is important to discuss the issue, whilst bearing in mind the sample and loss of data on one of the working memory measures. Christianson et al. [32] found that older participants with higher reading span scores were more likely to answer comprehension questions correctly with reflexive verbs, and they proposed that it was due to revision processes in some way taxing working memory capacity (see also [58–60]). An alternative view was proposed by Caplan and Waters [26], and they argued that online sentence processing is independent of working memory resources [27, 61].

The current results seem to be consistent with the assumptions made by Caplan and Waters, because our results indicated that neither reading span nor listening span was predictive of reading times, regressions, or comprehension. However, this might be due to the fact that comprehension accuracy was substantially higher in our study compared to previous reports, which introduces the possibility of a ceiling effect, especially with adolescent participants. A second possibility is that the current study tested relatively short sentences. Engelhardt et al. [34] tested the long versions of these types of ambiguities (e.g., While the woman cleaned the dog that was big and brown ran in the garden). Moreover, the main clause did not exclude the garden path misinterpretation with the optionally transitive verbs. Results for 13–16-year-old children in that study (N = 63) showed that comprehension accuracy was 19% correct for optionally transitive verbs and 32% correct for reflexives. Moreover, comprehension accuracy for sentences with reflexive verbs was related to working memory (r = .29, P < .05). Thus, working memory resources seem to be more important for comprehension when sentences are long and the misinterpretation persists for a longer period of time.

5. Conclusions

The main goal of this study was to investigate how children and adolescents resolve temporary syntactic ambiguity when reading. Our results revealed two main findings. The first was that participants were garden pathed even when the sentences were unambiguous. We attributed this effect to a reliance on syntactic parsing heuristics or the failure to apply the correct prosodic contour to the sentence. However, the number of regressions and the late reading time measure showed that processing difficulty was reduced when the verb was easy (i.e., reflexive) and there was a comma separating the subordinate and main clauses. The second main finding concerns the developmental trends in the data. We found that older participants were more likely to make regressions at the disambiguating verb and they were more likely to answer comprehension questions correctly. To our knowledge, this is the first eye movement study that examined syntactic ambiguity resolution in both child and adolescent participants. Previous studies [1, 9, 10] only tested children in the 8–12-year-old range. Therefore, the primary contributions of this study are that we (1) tested a novel age group and (2) tested garden path reanalysis with a novel type of lexical information. Some verbs tested in this study were difficult to revise (i.e., optionally transitive), whereas others, because of their lexical properties (i.e., reflexives), were easier to revise. Another strength of the study is that these types of garden path structures have been tested in both younger (18–35 years) and older (70–80 years) adults. The current study makes a contribution insofar as we now have a more complete picture of processing tendencies covering the entire lifespan. Across development from childhood to adolescence, there is a tendency to produce more regressions at the disambiguating word and a greater tendency to obtain the correct interpretation of temporarily ambiguous sentences.
Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The author would like to thank Emma Stevenson and Ruth H. Davies for recruiting participants and for helping collect the data. The author would also like to thank Alison Brown and Rebecca Jennings for assistance with the data analysis.

References


