Published in final edited form as:
Q J Exp Psychol (Hove). 2008 May ; 61(5): 708–723. doi:10.1080/17470210701400657.

Children's and adults' processing of anomaly and implausibility during reading: Evidence from eye movements

Holly S. S. L. Joseph,
University of Durham, Durham, UK

Simon P. Liversedge,
University of Southampton, Southampton, UK

Hazel I. Blythe,
University of Southampton, Southampton, UK

Sarah J. White,
University of Leicester, Leicester, UK

Susan E. Gathercole, and
University of York, York, UK

Keith Rayner
University of Massachusetts, Amherst, MA, USA

Abstract

The eye movements of 24 children and 24 adults were monitored to compare how they read sentences containing plausible, implausible, and anomalous thematic relations. In the implausible condition the incongruity occurred due to the incompatibility of two objects involved in the event denoted by the main verb. In the anomalous condition the direct object of the verb was not a possible verb argument. Adults exhibited immediate disruption with the anomalous sentences as compared to the implausible sentences as indexed by longer gaze durations on the target word. Children exhibited the same pattern of effects as adults as far as the anomalous sentences were concerned, but exhibited delayed effects of implausibility. These data indicate that while children and adults are alike in their basic thematic assignment processes during reading, children may be delayed in the efficiency with which they are able to integrate pragmatic and real-world knowledge into their discourse representation.

The present study investigated children's eye movement behaviour in relation to that of adults when reading sentences containing semantic implausibilities and anomalies. While there has been a good deal of research investigating eye movements and reading in adults (Rayner, 1998), there have been few studies into the nature of children's eye movements during reading. By the same token, while there has been a fair amount of research into how contextual or plausibility information influences children's syntactic processing, there has been little research investigating children's processing of plausibility information in itself, particularly using eye tracking as a methodology.

Studies by Rayner (1986) and McConkie et al. (1991) provided preliminary data regarding children's basic oculomotor behaviour during reading. Rayner (1986) compared the eye movements of children who were beginning to read and adults who were already proficient

Correspondence should be addressed to Holly Joseph, Department of Psychology, University of Durham, Durham, DH1 3LE, UK. E-mail: h.s.s.l.joseph@durham.ac.uk.

readers. Compared to proficient readers, beginning readers made longer and more frequent fixations (and consequently more saccades), made more frequent regressions, and had smaller perceptual spans. McConkie et al. (1991) examined the eye movements of children of different ages in an attempt to preliminarily map out the developmental trajectory of general aspects of oculomotor behaviour during reading. The frequency of fixations and regressions remained fairly constant between the ages of 7 and 10 years, although younger readers made longer fixations, shorter saccades, and more refixations than older readers.

It is noteworthy that while the Rayner (1986) and McConkie et al. (1991) studies provide important data regarding eye movement behaviour during reading in children, they did not investigate whether children of different ages, as well as adults, exhibit qualitatively different patterns of eye movements when they process sentences containing carefully controlled linguistic manipulations. Indeed, very few studies to date have set out to examine such questions (though see Blythe et al., 2005; Hyönä & Olson, 1995). While only a few studies have used eye movements to examine children's written language comprehension, there have been a number of studies that have used alternative methodologies to investigate the integration of context, pragmatic information, and real-world knowledge into ongoing meaning representations.

Trueswell, Sekerina, Hill, and Logrip (1999) investigated the influence of visual context on children's processing of syntactically ambiguous sentences, using the so-called visual world paradigm (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Four and five-year-old children listened to temporarily ambiguous sentences, such as "Put the frog on the napkin in the box", in which a prepositional phrase ("on the") could represent either a destination or a modifier to the preceding noun ("frog") while their eye movements were monitored. Visual context was also manipulated to favour one of the two possible interpretations of the prepositional phrase. Single referent contexts showed only one frog that was on a napkin, and therefore supported a destination interpretation of "on the napkin". In contrast, two referent contexts supported a modifier interpretation as two frogs were shown, one on a napkin, and one alone. Trueswell et al. found that children preferred the destination interpretation (the syntactically less complex alternative) regardless of visual or syntactic context. That is, in contrast to adults, children did not make use of contextual cues to disambiguate syntactically ambiguous structures. These findings seem to reflect a general insensitivity to pragmatic information during parsing in children. It may be that children are less adept than adults at integrating nonstructural information into their on-line language processing, although it is worth noting that Hurewitz, Brown-Schmidt, Thorpe, Gleitman, and Trueswell (2000) showed that children were able to use contextual cues in a language production task, indicating that the effects found in the Trueswell et al. study may be task-specific (see also Meroni & Crain, 2003, and Snedeker, Thorpe, & Trueswell, 2001, for comparable findings).

Nation, Marshall, and Altmann (2003) conducted an experiment in which they monitored children's eye movements to objects in a visual scene as they listened to spoken sentences such as 1a and 1b:

1a. Jane watched her mother eat the cake.

1b. Jane watched her mother choose the cake.

On hearing the target verb (*eat*) of Sentence 1a, children made fast anticipatory eye movements towards the target object on the screen (*cake*) when that object was the only edible entity. In contrast, on hearing Sentence 1b, children did not move their eyes until they heard the word (*cake*) when all entities in the visual display were "choosable". Nation et al. argued that these findings clearly indicate that children, like adults, are sensitive to verb selection restrictions and are able to integrate this with information extracted from the visual

context with the same accuracy and speed as do adults. That is, thematic role information is quickly activated and assigned on the basis of plausibility or thematic fit in children in the visual-world paradigm. This is a somewhat surprising finding because in the absence of syntactic ambiguity, it seems that there is no difference between adults and children in their ability to use contextual information to constrain ongoing processing.

Felser, Marinis, and Clahsen (2003) used a self-paced listening task to investigate children's reading of ambiguous relative clause sentences such as:

2. The doctor recognized the nurse of/with the pupils who was/were feeling very tired.

In sentences such as (2) above, while adults' preferences were influenced by the semantic properties of the preposition (*of*/*with*) adjoining the two potential antecedent noun phrases (*nurse* or *pupils*), children exhibited a preference for either the first or the second noun phrase dependent on their listening spans, but irrespective of the type of preposition involved. Felser et al. argued that children primarily rely on structural rather than semantic information during processing of modifier attachment ambiguities (see also Clahsen & Felser, 2006, and Sekerina, Stromswold, & Hestvik, 2004, for further evidence to support this claim).

Although all three of the studies outlined above are interesting and have made important contributions to our understanding of children's online language processing, it would be unwise to generalize the findings to children's written language comprehension. One other study, although not using eye movement measures, has specifically examined the effect of plausibility information on syntactic processing during reading.

Traxler (2002) investigated the effect of subcategorization and plausibility information on syntactic ambiguity resolution in children (aged 8–12 years) using a self-paced reading task, in which children read sentences such as 3a–3c below.

3a. When Sue tripped the girl fell over and the vase was broken.

3b. When Sue tripped the table fell over and the vase was broken.

3c. When Sue fell the policeman stopped and helped her up.

He showed that children, like adults, were garden-pathed by sentences like (3a) but also by sentences like (3b) where they did not use plausibility information to help them when faced with a syntactic ambiguity (i.e., it is implausible to trip an inanimate object). He did, in his third experiment (see Sentence 3c), find a small negative correlation between intransitivity preference (how frequently a verb was intransitive as compared to transitive) and disruption to processing (as measured by total reading times). Nevertheless, in both Experiments 1 and 2, Traxler reliably found that children, in contrast to adults, did not make use of plausibility or subcategorization information during reading.

Taking the studies outlined above together, there do not appear to be any qualitative differences in the language-processing mechanisms of adults and children. Rather, children rely predominantly on structural information and disregard additional contextual information such as semantic or pragmatic fit, while adults make good use of such information, resulting in the observed differences between adults and children in the studies discussed. Interestingly, it is only in the Nation et al. (2003) study, which did not involve a manipulation of syntactic ambiguity, that children showed the same effects as adults (albeit in a different study).

The most relevant adult study to the present experiment is an investigation by Rayner, Warren, Juhasz, and Liversedge (2004) on the effects of plausibility on adults' reading

behaviour (though see also Braze, Shankweiler, Ni, & Palumbo, 2002; Marslen-Wilson, Brown, & Tyler, 1988; Ni, Fodor, Crain, & Shankweiler, 1998; Warren & McConnell, in press). Rayner et al. used sentences that described events in which an individual performed an action with an instrument. In each case, the verb had three thematic roles (see Sentences 4a–4c): an agent (*John*), an instrument1 (*knife*, *axe*, or *pump*) and a patient/theme (*carrots*).

4a. John used a knife to chop the large carrots for dinner.

4b. John used an axe to chop the large carrots for dinner.

4c. John used a pump to inflate the large carrots for dinner.

In all three sentences the instrument (*knife*, *axe*, or *pump*) could be plausibly used in conjunction with the main verb (*chop* or *inflate*). However, in the implausible condition (4b) the patient (*carrots*) was incongruous as the object of the verb (*chop*) given the particular instrument used (*axe*). That is, although axes are often used to chop things, and carrots are often chopped, an axe is not often used to chop carrots. To this extent, the sentence is implausible (though it does not describe a situation that is virtually impossible). By contrast, in the anomalous condition (4c), the patient (*carrots*) could not be used in conjunction with the verb (in this case *inflate*) since under ordinary circumstances carrots do not have the attribute of inflatability.2 A crucial difference between the implausible and anomalous conditions was that in the implausible condition the incongruity occurred due to the incompatibility of two objects involved in the event denoted by the main verb. By contrast, in the anomalous condition, the incongruity arose due to the sentential object being an inappropriate argument of the verb. Rayner et al. characterized the former manipulation as describing an implausible situation, whereas the latter manipulation describes an impossible, or extremely unlikely, event and was categorized as anomalous.

Disruption to processing occurred earlier when the sentences were anomalous rather than implausible. Specifically, for anomalous sentences disruption was apparent in first-pass reading time as soon as the critical word *carrots* was fixated, whereas in the implausible condition disruption was less immediate and only apparent for fixations made after a regressive saccade from the critical word. Rayner et al. (2004) suggested that the differential effects may be either due to the severity of the violation (i.e., how implausible it is perceived to be) or because, in most cases, anomalous violations can be detected on the basis of lexical information alone (a verb argument violation). Implausible violations, on the other hand, can only be detected at a later stage of processing after the semantic evaluation of the combination of a verb and the objects involved in the event it denotes.

To summarize, in Rayner et al.'s (2004) study, readers exhibited earlier disruption to processing of sentences containing anomalous than implausible thematic relations. On this basis, we also anticipated that adult readers in the present study would spend longer fixating the critical region of the anomalous than the implausible or control sentences during first-pass reading. Also in line with Rayner et al.'s findings, we predicted that adults would exhibit delayed disruption to processing for the implausible sentences than for the anomalous sentences, such that disruption may first occur within second-rather than first-pass fixations. Additionally, we took the opportunity to make direct comparisons between the adult data from Rayner et al.'s study and the adult data from our study in order to establish that the Rayner et al. results replicate. For this reason, we analysed the adult data separately as well as in conjunction with the child data.

[1]Note that the instrument (e.g., *knife*) in all experimental sentences is actually an argument of the verb *used* but will be conceptually understood by the reader to be the instrument used to carry out the action denoted by the infinitival verb (e.g., *chop*).
[2]The attribute of inflatability in carrots may be considered possible in certain circumstances such as cartoons.

In terms of the children, our predictions were a little more tentative. The vast majority of studies that have investigated whether children show a sensitivity to semantic information have manipulated semantic (visual or linguistic) context and have examined initial syntactic processing preferences. Semantic influences on processing are inferred on the basis of initial syntactic commitments. The only study that has examined children's on-line processing of plausibility information per se is that of Nation et al. (2003), and they found no differences in processing between adults and children. However, given that Nation et al.'s experiment was a nonreading experiment, with a highly constrained visual context that made verb arguments relatively predictable, it may be unwise to assume that children and adults read sentences containing implausibilities and anomalies in the same way. This is particularly the case given that we know that children are slower to process written language generally than are adults (Rayner, 1986). Thus, on the assumption that linguistic influences will occur with less immediacy in the eye movement record of children than of adults (Traxler, 2002), we anticipated that children might exhibit delayed disruption to processing when reading sentences containing anomalous and implausible thematic relations. Thus, while adults were expected to show first-pass effects of anomaly, we predicted that similar effects may be delayed for children. Specifically, we anticipated that effects for children would be spatially localized to words downstream from the critical word, or fixations made during second-pass reading (after a regressive saccade and later in the eye movement record than for adults). Similarly, we also anticipated delayed implausibility effects for children relative to those for adults.

Method

Participants

A total of 24 adults and 24 children took part in the experiment. The adults were students at University of Durham. All were native English speakers with normal uncorrected vision and no known reading disabilities. All adult participants were naïve concerning the purpose of the study and received a payment of £5 per hour for taking part. The child participants were recruited through primary schools in the Durham area. The mean age of participants was 9 years and 6 months, with ages ranging from 7 years 0 months to 12 years 0 months. We chose this age group as we considered it unlikely that younger children would be able to read the sentences easily or to fulfil all the experimental requirements. In addition, the grammatical system is generally assumed to be complete by age 6 (Clahsen & Felser, 2006). We also considered that children older than 12 years old would approximate adults in their reading behaviour (Buswell, 1922; Rayner, 1998). All child participants were native English speakers with normal uncorrected vision and no known reading disabilities. Each child received a small gift in return for taking part in the study.

Apparatus

A Dual Purkinje Image eye-tracker was used to monitor participants' eye movements as they read sentences from a computer monitor at a distance of 80 cm. Participants' eye movements were monitored every millisecond to produce a sequence of fixations with start and finish times. Although only the right eye was monitored, participants read binocularly.

Materials

A total of 36 experimental items were constructed, and for each of these there were three versions (see Sentences 5a–5c below for control, implausible, and anomalous versions, respectively). We constructed three experimental lists according to a Latin square such that each list contained a different version of each item, and each list contained an equal number of control, implausible, and anomalous sentences.

5a. Robert used a trap to catch the horrible mouse that was very scared.

5b. Robert used a hook to catch the horrible mouse that was very scared.

5c. Robert used a radio to play the horrible mouse that was very scared.

As in Rayner et al.'s (2004) study, all items were designed so that the plausibility violation always occurred at the noun of the adjectival noun phrase (the critical target word) following the infinitival verb. The critical word (*mouse* in Sentences 5a–5c) was identical across conditions (as were the two immediately preceding words), which ensured that any effects observed at this point were unlikely to be due to factors other than the plausibility/anomaly manipulation. On average, there were no significant differences in the length ($t < 1$) or the frequency, $t(1, 35) = 1.7$, $p > .09$, of the infinitival verb across conditions (*catch* or *play*). Similarly, there was no significant difference in the frequency of the noun (*trap*, *hook*, or *radio*) denoting the instrument across conditions ($F < 1$). There was a significant difference in length, $F(1, 35) = 5.2$, $p < .01$, of this same noun. In the anomalous and control conditions, words were, respectively, 5.7 and 5.9 characters long on average, whereas in the implausible condition they were 6.5 characters. However, this difference is in a part of the sentence prior to the regions that we are interested in.

Questionnaires containing all 108 sentences were given to 19 undergraduate students (who did not take part in the eye tracking experiment). A total of 82 children also completed a 36-item norming questionnaire (the same as that given to adult participants but a third of the length to make the task easier for the children). The children were recruited from Year 3 (age 7–8) and Years 5 and 6 (ages 9–11) of three local primary schools and did not take part in the main experiment.

Analyses of adult and child ratings were carried out separately. Participants were asked to rate the sentences on a 5-point scale where 1 was "very strange", 3 was "a bit strange", and 5 was "not at all strange/normal". The mean ratings for each condition, together with the standard deviations, are shown in Table 1.

For the adults, there was a significant difference between ratings for the three groups of sentences, $F(2, 70) = 458.3$, $p < .001$; adults rated the anomalous sentences as being significantly stranger than the implausible sentences, $t(1, 35) = 6.843$, $p < .001$, the anomalous sentences as being significantly stranger than the control sentences, $t(1, 35) = 38.443$, $p < .001$, and the implausible sentences as being significantly stranger than the control sentences, $t(1, 35) = 20.721$, $p < .001$.

As with the adults, for the children, there was a significant difference between ratings for the three conditions, $F(2, 70) = 212.5$, $p < .001$; children reliably rated the anomalous sentences as stranger than the implausible sentences, $t(1, 35) = 2.7$, $p < .05$, the anomalous sentences as stranger than the control sentences, $t(1, 35) = 19.5$, $p < .001$, and the implausible sentences as stranger than the control sentences, $t(1, 35) = 16.9$, $p < .001$. There was no significant difference between the ratings of the younger and those of the older children, $F < 1$. These results indicate that in an off-line task children were able to perceive a difference between all three conditions and serve to validate the plausibility/anomaly manipulation employed in the present experiment across both participant populations.

Reading ability—All child participants completed the Wechsler Objective Reading Dimension (WORD; Rust, Golombok, & Trickey, 1992) as a measure of their reading ability. The WORD consists of three sections: Basic (word) Reading, Spelling, and Reading Comprehension. Children scored an overall average of 113 (mean = 100, *SD* = 15). All participants performed within or above the normal range (2 *SD*s below or above the mean) for their ages. The average reading age was 12.1 years.

Procedure

Experimental eye tracking sessions lasted between 20 and 30 minutes. Participants' head movements were minimized using forehead restraints and a dental bite bar. Children sat in a customized chair in order to position them correctly and comfortably. All participants were warned that some of the sentences were somewhat strange but that they should read the sentences as normally as possible and press a button when they had read the sentence in its entirety. Participants were also required to respond to a simple comprehension question by pressing a "yes" or "no" button after a third of all questions.

Results

All sentences were divided into five regions as shown in Example 5c (repeated here; see Appendix for a full list of the experimental sentences), three of which were of particular interest.

5c. /Robert used a radio to play/the horrible/mouse/that was/very shy./

The three regions of particular interest, and those we focus on exclusively, were the pretarget region (Region 2), which comprised the determiner and adjective (e.g., *the horrible*); the target word region (Region 3) comprising the noun (e.g., *mouse*); and the posttarget region (Region 4) comprising the two short words (or one long word), following the target noun (e.g., *that was*).

All participants performed at a minimum of 75% accuracy on the comprehension measure. Trials were excluded if (a) the participant did not fixate three or more regions in total, (b) the participant did not fixate the first two regions, or (c) the participant did not fixate two of the three regions of interest (Regions 2, 3, and 4). Furthermore, a proportion of the child participants did not complete all the trials due to fatigue or tracker loss. On this basis, 37 trials were eliminated (4.4% of all trials) from the adult data, and 129 trials (14.9% of all trials) were eliminated from the children's data (exclusions were equally distributed across conditions). In addition, fixations less than 80 ms were combined with fixations on adjacent letters or else eliminated if they were not within three characters of another fixation. Fixations longer than 1,200 ms were also excluded.

The following eye movement measures were computed: first-fixation duration (the duration of the first fixation in a region); gaze duration[3] (the sum of all fixations in a region before the eyes leave the word either to the right or to the left); go-past reading time (the sum of all the temporally contiguous fixations in a region, including any regressive eye movements to the left of the region, until the point of fixation progressed to the region to the right); and total reading time (the sum of all fixations in a region). Repeated measures analyses of variance (ANOVAs), 2 (group: adults vs. children) × 3 (plausibility: anomalous vs. implausible vs. control), were carried out for analyses of the pretarget region, the target region, and the posttarget region.

Pretarget region

Table 2 shows the reading time measures for the pretarget region. There were no effects of the plausibility manipulation on the duration of the first fixation, Fs < 1, but there was a reliable effect of group, $F_1(1, 46) = 13.05$, $p < .005$; $F_2(1, 33) = 131$, $p < .001$, which, unsurprisingly, showed that children made longer first-fixation durations than did adults. There was no reliable interaction between group and plausibility (Fs < 1.1; ps > .3). In gaze

[3]For regions larger than a single word, gaze duration is often referred to as first-pass reading time (Rayner, 1998). We use both terms throughout this article.

duration, there was a suggestion of an effect of the plausibility manipulation (a difference of 33 ms), but the effect was not reliable, $F_1(2, 92) = 2.94$, $p = .058$; $F_2(2, 66) = 1.55$, $p = .2$. There was no reliable interaction (Fs < 1). Once again, there was an effect of group, $F_1(1, 46) = 14.95$, $p < .001$; $F_2(1, 33) = 129$, $p < .001$. Children exhibited longer gaze durations than did adults. These results replicate those of Rayner et al. (2004) who found no first-pass (or gaze duration) effects of plausibility, although they did find a numerical difference between the anomalous condition and the implausible and control conditions.

Finally, total reading time showed a highly significant effect of plausibility, $F_1(2, 92) = 21.87$, $p < .001$; $F_2(2, 70) = 16.4$, $p < .001$. Participants spent longer overall reading the anomalous than the control sentences, $t_1(47) = 5.59$, $p < .001$; $t_2(35) = 4.96$, $p < .001$, but there was no difference in reading time between the implausible and control sentences (ts < 1.4, $p > .1$). Once again, there was also a significant effect of group, $F_1(1, 46) = 12.96$, $p < .005$; $F_2(1, 35) = 87.08$, $p < .001$, with longer total reading times for children than for adults. Finally, there was an interaction between plausibility and group that was statistically significant by items and marginal by participants, $F_1(2, 92) = 2.79$, $p = .07$, $F_2(2, 70) = 4.88$, $p < .05$. To explore this interaction, the adult and child data were analysed separately. Adults showed effects of anomaly and implausibility in total reading time in the pretarget region, $F_1(2, 46) = 19.09$, $p < .001$; $F_2(2, 70) = 13.44$, $p < .001$; they spent longer reading the pretarget region in the anomalous than in the control condition, $t_1(23) = 6.82$, $p < .001$; $t_2(35) = 5.74$, $p < .001$, and in the implausible than in the control condition, $t_1(23) = 2.98$, $p < .01$; $t_2(35) = 2.84$, $p < .01$. Given that this measure includes second-pass fixations, it indicates that adults ultimately detected both the anomaly and the implausibility. The results were somewhat different for the children. Like adults, they showed a reliable effect of the plausibility manipulation in total reading time, $F_1(2, 46) = 11.22$, $p < .001$; $F_2(2, 70) = 11.41$, $p < .001$; but only the difference between the anomalous and control conditions was reliable, $t_1(23) = 3.75$, $p < .005$; $t_2(35) = 3.73$, $p < .005$. Total reading times were no different in the implausible than in the control condition (ts < 1). Thus, while adults showed effects of both the anomaly and implausibility manipulations, children were influenced only by the anomaly manipulation. Rayner et al. (2004) found an anomaly effect, but not an implausibility effect, in total reading time with their adult participants. We have no immediate explanation of why this effect occurred in our study but not in that of Rayner et al. However, since the effect occurred in the relatively late measure of total reading time, it should not be taken as an indication of initial detection of anomaly or implausibility.

In summary, although first-pass (gaze duration) effects in the pretarget region were not reliable, a 33-ms difference in gaze durations between the anomalous and control conditions was observed (16 ms longer than that in the Rayner et al., 2004, study). In total reading time reliable effects of anomaly and implausibility were obtained for adults; however, for children the only reliable effects obtained were for anomalies. As indexed by longer total reading times, children showed no sensitivity to the implausible sentences. Finally, we also consistently found effects of group in every measure, showing that children took longer to read sentences than did adults.

Target region

Table 3 shows the reading-time measures for the target region. As per Rayner et al. (2004), for first-fixation duration, despite a numerical trend in the predicted direction (anomaly effect = 13 ms; implausibility effect = 5 ms), there was not a reliable effect of plausibility (Fs < 1.8, ps > .1). There was a significant effect of group, $F_1(1, 46) = 11.61$, $p < .005$; $F_2(1, 35) = 51.47$, $p < .001$, with longer first fixations for children than for adults, but no interaction (Fs < 1). For gaze duration, we again found an effect of plausibility that was marginal by participants but not by items, $F_1(2, 92) = 2.62$, $p = .07$; $F_2(2, 70) = 1.58$, $p > .2$. There was an effect of group, $F_1(1, 46) = 15.6$, $p < .001$; $F_2(1, 35) = 117$, $p < .001$, but no

reliable interaction (*F*s < 1). When the adult data were analysed separately they did exhibit a reliable effect of plausibility in gaze duration, $F_1(2, 46) = 4.57$, $p < .05$; $F_2(2, 70) = 4.31$, $p < .05$. This effect replicates the results from the Rayner et al. study. It appears that that the increased variability introduced by the child data is the cause of the lack of reliability of the plausibility effect overall, as both adults and children did make longer fixations in the anomalous than in the control conditions.

In the go-past measure, there was an effect of the plausibility manipulation, which was reliable by items but not by participants, $F_1(2, 92) = 1.46$, $p > .2$; $F_2(2, 70) = 3.61$, $p < .05$; go-past reading time was longer in the anomalous than in the control condition, $t_1(47) = 2.36$, $p < .05$; $t_2(35) = 2.53$, $p < .05$, but not longer in the implausible than in the control condition (*t*s < 1.1, *p*s > .2). There was an effect of group, $F_1(1, 46) = 14.4$, $p < .001$; $F_2(1, 35) = 64.5$, $p < .001$, with children having longer go-past reading times than adults. There was no reliable interaction ($F < 1$). Again, this pattern of data replicates that of Rayner et al. (2004) who also obtained no significant effect of implausibility in the go-past measure in the target region (although like us they did obtain a numerical trend).

As in Rayner et al. (2004), total reading time was influenced by the plausibility manipulation, $F_1(2, 92) = 13.28$, $p < .001$; $F_2(2, 70) = 15.05$, $p < .001$; readers spent longer reading the anomalous than the control sentences, $t_1(47) = 4.34$, $p < .001$; $t_2(35) = 3.73$, $p < .005$, but showed no difference in total reading time between the implausible and the control sentences (*t*s < 1). There was a main effect of group, $F_1(1, 46) = 8.37$, $p < .01$; $F_2(1, 35) = 41.57$, $p < .001$, with longer total reading times for children than for adults. There was no interaction between group and the plausibility manipulation (*F*s < 1.2, *p*s > .3).

In summary, in the target region, there were first-pass (gaze duration) effects of the plausibility manipulation with adults looking 24 ms longer at the target word in the anomalous than in the control condition, children looking 11 ms longer, and both adults (42 ms) and children (78 ms) taking longer to get past the target word in the anomalous than in the control condition. There were also reliable anomaly effects in total reading time. There were no effects of implausibility in the target region. These findings fully replicate those of Rayner et al. (2004).

Posttarget region

Table 4 shows the reading time measures for the posttarget region. In first-fixation duration, once again there was no effect of plausibility (*F*s < 1) and no interaction (*F*s < 1). There was an effect of group, reliable by items but not by participants, $F_1(1, 46) = 2.15$, $p = .15$; $F_2(1, 34) = 16.09$, $p < .001$. In gaze duration, there was no effect of plausibility (*F*s < 1.8, *p*s > .1). There was a reliable effect of group, $F_1(1, 46) = 7.61$, $p < .01$; $F_2(1, 34) = 61.25$, $p < .001$, but no interactive effect (*F*s < 1). These effects differ from those reported by Rayner et al. (2004) who found a significant effect of anomaly in both first-fixation duration and gaze duration in the posttarget region. We suspect that our failure to obtain reliable effects in this region for these measures is a consequence of the additional variability within our data set introduced by the children.

In go-past time, there was a significant effect of plausibility, $F_1(2, 92) = 11.23$, $p < .001$; $F_2(1.6, 68) = 9.50$, $p < .005$ (see Figure 1); go-past times were longer in the anomalous than in the control condition, $t_1(47) = 3.97$, $p < .001$; $t_2(34) = 3.42$, $p < .005$, but there were no reading-time differences between the implausible and control conditions (*t*s < 1.3, *p*s > .2). There was an effect of group, $F_1(1, 46) = 8.84$, $p < .01$; $F_2(1, 34) = 56.9$, $p < .001$, and an interaction between group and plausibility, marginal by items but not by participants, $F_1(2, 92) = 1.72$, $p = .19$; $F_2(2, 68) = 2.35$, $p = .10$. Note that to directly compare our adult data with those from the Rayner et al. (2004) study, we analysed the data from the adults

separately. Go-past times were longer for adults in the anomalous than in the control condition, $t_1(23) = 3.89$, $p < .005$; $t_2(34) = 3.44$, $p < .005$, and importantly longer in the implausible than in the control condition (although marginal by participants), $t_1(23) = 1.84$, $p = .079$; $t_2(34) = 2.24$, $p < .05$. This is the first point at which adults showed implausibility effects, and our data replicate Rayner et al.'s findings.

Finally, in total reading time there was an effect of the plausibility manipulation, $F_1(2, 92) = 10.83$, $p < .001$; $F_2(2, 70) = 7.26$, $p < .005$; participants spent longer reading the anomalous than the control sentences, $t_1(47) = 5.03$, $p < .001$; $t_2(35) = 3.30$, $p < .005$, and longer reading the implausible than the control sentences, $t_1(47) = 2.84$, $p < .01$; $t_2(35) = 2.74$, $p < .05$. There was also an effect of group, $F_1(1, 46) = 5.59$, $p < .05$; $F_2(1, 35) = 47.55$, $p < .001$, but no interaction (Fs < 1.2, ps > .3). This finding indicates that while earlier reading-time measures showed differential effects of plausibility, ultimately both adults and children exhibited disruption to processing in the implausible condition. This is the first indication of an implausibility effect in children and shows that children, as well as adults, are able to discriminate between both anomalous and implausible thematic violations during reading. Again, this result replicates that of Rayner et al. (2004) who also obtained reliable anomaly and implausibility effects in total time in the posttarget region.

To summarize the results in the posttarget region, go-past time revealed disruption to processing as a result of the plausibility manipulation. While adults spent longer reading in the anomalous and the implausible sentences than in the control sentences (as per Rayner et al., 2004), children showed a difference between the anomalous and control conditions only. This differential pattern of effects for adults and children is the same pattern as that observed in the total reading-time measure in the pretarget region. In contrast, total reading time in the posttarget region revealed both anomaly and implausibility effects for both adults and children, indicating that both participant groups had detected the implausible thematic violation in the latter stages of reading the sentences.

It is also worth noting here that there were no significant correlations between reading age, or chronological age, and any of the significant effects described (all ps > .5). It can be concluded therefore that although there was a large age range in the child group, the effects described above do not appear to be modulated by age.4

Overall, the earliest point at which reliable disruption to processing was observed in response to the anomalous thematic violation was in the target region in gaze duration. Although failing to reach statistical significance in the overall analyses, adults exhibited a reliable effect of anomaly, and children showed a numerical trend in the predicted direction. Reliable implausibility effects were observed only in the much later measure of total reading time in the posttarget region (although adults showed implausibility effects in the go-past measure in the posttarget region as per Rayner et al., 2004). We therefore observed, as predicted, increased immediacy in the detection of anomalous as compared to implausible thematic violations. Interestingly, we did not observe reliable differences in the time course of anomaly detection between adults and children. Although adults exhibited stronger effects than children, both adults and children exhibited disruption to processing in the anomalous condition in the target region. There were, however, differences in time course of the implausibility detection: In go-past time in the posttarget region, adults' processing was disrupted by both the implausibility and anomaly while children showed only anomaly effects. This same pattern was observed in total time in the pretarget region, with an

[4]It is interesting to note that analyses comparing younger children (7–9 years old) to older children (10–11 years old) revealed that although younger children made longer fixations overall and took longer to read the sentences, no differential anomaly or implausibility effects were observed between groups.

implausibility effect for children found only for total time in the posttarget region. It appears that although children are delayed in their processing of implausibility as compared to adults, both adults and children showed increased immediacy in their anomaly as compared to their implausibility detection and exhibited disruption to processing in response to both kinds of manipulation.

Discussion

There were both commonalities and differences in the time course of processing of the different kinds of thematic relations in adults and children. It was predicted that adults would exhibit earlier disruption to processing of sentences containing anomalous than implausible thematic relations. This prediction was met as adults exhibited disruption to processing during first pass in the anomalous condition in the target region, but showed only later effects of implausibility. In line with Rayner et al. (2004) this finding shows that anomalous thematic relations are detected more immediately than implausible thematic relations (see also Warren & McConnnell, in press). It appears that the increased severity of an anomalous (as compared to an implausible) thematic violation induces more immediate and substantive disruption in the eye movement record.

We also predicted that disruption to processing would be delayed in children when reading sentences containing anomalous and implausible thematic relations, as compared to adults. This prediction was only partially met. Although the adults showed reliable first-pass effects of anomaly, and those observed for the children failed to reach statistical significance, it seems clear that both groups did show consistent and substantial reading-time differences between the anomalous and control sentences during first pass. We argue, then, that the time course of processing associated with thematic anomaly detection is similar in adults and children. This finding seems to fit well with the Nation et al. study (2003), which found that the time course for identifying a plausible (as compared to anomalous distractors) verb argument was immediate and the same for child and adult participants.

Importantly, however, while children did not differ from adults in their anomaly detection, they did appear to be delayed relative to adults in their implausibility detection. While disruption to processing in adults was observed in the implausible condition compared with the control condition for go-past reading times in the posttarget region (and in total time in the pretarget region), children showed no similar difference in reading times between these two conditions in these measures. Only in the total reading times in the posttarget region did children first show a reliable effect of implausibility. These results indicate that although children do detect implausible thematic violations during reading, they are delayed in doing so relative to adults. Presumably, an anomalous thematic violation may be detected through the recognition of an illegal combination of linguistic constituents. By contrast, the detection of a thematic implausibility relies both on thematic processing and on the integration of real-world knowledge and pragmatic information with respect to the discourse representation that is being developed on an ongoing basis by the reader. If this is the case, then it appears that children are similar to adults in terms of basic thematic processing during comprehension, but are less efficient than adults in the integration of real-world knowledge into the discourse representation. It may be for this reason that children are delayed in their detection of thematic implausibilities relative to adults. Such an interpretation sits well with previous research into children's on-line language processing using different methodologies that have found that children are less able to use contextual information to guide parsing (e.g., Trueswell et al., 1999).

In summary, the key findings from the present study are as follows. First, to a significant extent, we replicated the findings of Rayner et al. (2004) for adult participants (see also

Warren & McConnell, in press). Clear anomaly and implausibility effects were obtained with a similar time course. We also demonstrated sensitivity to thematic implausiblities and anomalies in both adults and children both in an off-line rating task and on-line in eye movements during reading. Additionally, the eye movement records from the children showed thematic anomaly effects that were as immediate as those for adults, whereas thematic implausibility effects were delayed relative to adults. These data suggest that while children and adults are similar in terms of basic thematic assignment processes that occur during reading, they differ in the efficiency with which they are able to integrate pragmatic and real-world knowledge into the discourse representation.

Acknowledgments

This research was supported by an Economic and Social Research Council studentship to the first author, Leverhulme Trust Grant No. F/00128/AB and Biotechnology and Biological Sciences Research Council Grant No. 12/S19168 to the second author, and HD17426 from the US National Institute of Health to the sixth author. We thank Martin Fischer and two anonymous reviewers for their helpful comments on an earlier draft.

References

Blythe HI, Liversedge SP, Joseph HSSL, White SJ, Findlay JM, Rayner K. The binocular co-ordination of eye movements during reading in children and adults. Vision Research. 2006; 46:3898–3908. [PubMed: 16879851]

Braze D, Shankweiler D, Ni W, Palumbo LC. Readers' eye movements distinguish anomalies of form and content. Journal of Psycholinguistic Research. 2002; 31:25–44. [PubMed: 11924838]

Buswell, GT. Fundamental reading habits: A study of their development. Chicago: University of Chicago Press; 1922.

Clahsen H, Felser C. Grammatical processing in language learners. Applied Psycholinguistics. 2006; 27:3–42.

Cooper RM. The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory and language processing. Cognitive Psychology. 1974; 6:84–107.

Felser C, Marinis T, Clahsen H. Children's processing of ambiguous sentences: A study of relative clause attachment. Language Acquisition. 2003; 11:127–163.

Hurewitz F, Brown-Schmidt S, Thorpe K, Gleitman LR, Trueswell JC. One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. Journal of Psycholinguistic Research. 2000; 29:597–626. [PubMed: 11196065]

Hyönä J, Olson RK. Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1995; 21:1430–1440.

Marslen-Wilson WD, Brown C, Tyler LK. Lexical representations and language comprehension. Language and Cognitive Processes. 1988; 3:1–16.

McConkie, GW.; Zola, D.; Grimes, J.; Kerr, PW.; Bryant, RB.; Wolff, PM. Children's eye movements during reading. In: Stein, JF., editor. Vision and visual dyslexia. London: Macmillan; 1991.

Meroni, L.; Crain, S. On not being led down the kindergarten path. In: Beachley, B.; Brown, A.; Coulin, F., editors. Proceedings of the 27th Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press; 2003. p. 531-544.

Nation K, Marshall CM, Altmann GTM. Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. Journal of Experimental Child Psychology. 2003; 86:314–329. [PubMed: 14623215]

Ni W, Fodor JD, Crain S, Shankweiler D. Anomaly detection: Eye movement patterns. Journal of Psycholinguistic Research. 1998; 27:515–540. [PubMed: 9750312]

Rayner K. Beginning readers and the perceptual span in beginning and skilled readers. Journal of Experimental Child Psychology. 1986; 41:211–236. [PubMed: 3701249]

Rayner K. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin. 1998; 124:372–422. [PubMed: 9849112]

Rayner K, Warren T, Juhasz BJ, Liversedge SP. The effect of plausibility on eye movements in reading. Journal of Experimental Psychology: Learning, Memory and Cognition. 2004; 30:1290–1301.

Rust, J.; Golombok, S.; Trickey, G. WORD: Wechsler Objective Reading Dimensions manual. London: The Psychological Corporation; 1992.

Sekerina IA, Stromswold K, Hestvik A. How do adults and children process referentially ambiguous pronouns? Journal of Child Language. 2004; 31:123–152. [PubMed: 15053087]

Snedeker, J.; Thorpe, K.; Trueswell, JC. On choosing the parse with the scene: The role of visual context and verb bias in ambiguity resolution. In: Moore, J.; Stenning, K., editors. Proceedings of the 23rd Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum Associates; 2001. p. 964-969.

Tanenhaus MK, Spivey-Knowlton MJ, Eberhard K, Sedivy J. Integration of visual and linguistic information in spoken language comprehension. Science. 1995; 286:1632–1634. [PubMed: 7777863]

Traxler MJ. Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: Evidence from self-paced reading. The Quarterly Journal of Experimental Psychology. 2002; 55A:75–96. [PubMed: 11873857]

Trueswell J, Sekerina I, Hill N, Logrip M. The kindergarten-path effect: Studying online sentence processing in young children. Cognition. 1999; 73:89–134. [PubMed: 10580160]

Warren T, McConnell K. Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. Psychonomic Bulletin & Review. in press.

APPENDIX

A comprehensive list of the sentences used in the experiment

For each set of three, the first sentence is the anomalous condition, the second is the implausible condition, and the third is the plausible (control) condition.

Beatrice used a towel to dry the important programme on the computer.

Beatrice used a key to open the important programme on the computer.

Beatrice used a password to open the important programme on the computer.

The man used the formula to explain the beautiful boat after the trip.

The man used the shoelace to tie up the beautiful boat after the trip.

The man used the rope to tie up the beautiful boat after the trip.

Robert used a radio to play the horrible mouse that was very scared.

Robert used a hook to catch the horrible mouse that was very scared.

Robert used a trap to catch the horrible mouse that was very scared.

Justin used a needle to sew the spotted Dalmatian that he was walking.

Justin used a joystick to control the spotted Dalmatian that he was walking.

Justin used a muzzle to control the spotted Dalmatian that he was walking.

Jenny used a hose to water the small butterfly flying past.

Jenny used a mousetrap to catch the small butterfly flying past.

Jenny used a net to catch the small butterfly flying past.

Sarah used a fork to eat the fresh water extremely carefully.

Sarah used a purse to carry the fresh water extremely carefully.

Sarah used a bucket to carry the fresh water extremely carefully.

Matthew used a brush to sweep the bright star in the sky.

Matthew used a microscope to watch the bright star in the sky.

Matthew used a telescope to watch the bright star in the sky.

Dad used a fork to eat the purple flowers in the garden.

Dad used a sword to protect the purple flowers in the garden.

Dad used a fence to protect the purple flowers in the garden.

Todd used a hammer to nail the heavy shopping from Tesco.

Todd used a helicopter to carry the heavy shopping from Tesco.

Todd used a trolley to carry the heavy shopping from Tesco.

Ben used a car to climb the highest branch of the tree.

Ben used a map to reach the highest branch of the tree.

Ben used a ladder to reach the highest branch of the tree.

The witch used a cloth to polish the special liquid for the magic potion.

The witch used a basket to hold the special liquid for the magic potion.

The witch used a bowl to hold the special liquid for the magic potion.

John used a straw to drink the carrots for dinner last night.

John used an axe to chop the carrots for dinner last night.

John used a knife to chop the carrots for dinner last night.

The man used a feather to tickle the thin spaghetti yesterday evening.

The man used a kettle to boil the thin spaghetti yesterday evening.

The man used a pan to boil the thin spaghetti yesterday evening.

The woman used the expensive video to record the birthday present yesterday.

The woman used the fluffy towel to wrap the birthday present yesterday.

The woman used the pretty paper to wrap the birthday present yesterday.

Phillip used the match to light the smelly cheese from Italy.

Phillip used the scissors to cut the smelly cheese from Italy.

Phillip used the knife to cut the smelly cheese from Italy.

Mum used a CD player to hear the dirty dishes in the sink.

Mum used a hoover to clean the dirty dishes in the sink.

Mum used a sponge to clean the dirty dishes in the sink.

The man used a submarine to attack the front porch for the party.

The man used a toothbrush to clean the front porch for the party.

The man used a mop to clean the front porch for the party.

Mum used the music to calm the hot beans for dinner.

Mum used the toothpick to serve the hot beans for dinner.

Mum used the spoon to serve the hot beans for dinner.

Dad used the fork to eat the overgrown grass in the garden.

Dad used the scissors to cut the overgrown grass in the garden.

Dad used the lawnmower to cut the overgrown grass in the garden.

Mum used a spoon to feed the chocolate cake in the oven.

Mum used a crane to put the chocolate cake in the oven.

Mum used an oven glove to put the chocolate cake in the oven.

Richard used a ghost to scare the different heights of children in the class.

Richard used a stopwatch to measure the different heights of children in the class.

Richard used a ruler to measure the different heights of children in the class.

The woman used a mobile phone to ring her valuable books while she was away.

The woman used a jewellery box to store her valuable books while she was away.

The woman used a cardboard box to store her valuable books while she was away.

The man used a pencil to write the expensive bottle of wine in the restaurant.

The man used a tin opener to open the expensive bottle of wine in the restaurant.

The man used a corkscrew to open the expensive bottle of wine in the restaurant.

Liz used her coffee mug to drink her friend’s phone number yesterday.

Liz used her compass to find her friend’s phone number yesterday.

Liz used her address book to find her friend's phone number yesterday.

The prince used a microphone to sing the lovely princess from the dragon.

The prince used a floppy disk to save the lovely princess from the dragon.

The prince used a sword to save the lovely princess from the dragon.

The boy used a stepladder to climb the enormous fish for dinner.

The boy used a baseball glove to catch the enormous fish for dinner.

The boy used a fishing rod to catch the enormous fish for dinner.

The girl used a trumpet to play the tiny picture for her mother.

The girl used a roller to paint the tiny picture for her mother.

The girl used a brush to paint the tiny picture for her mother.

The shop assistant used a spade to dig the food around the supermarket.

The shop assistant used a wallet to carry the food around the supermarket.

The shop assistant used a trolley to carry the food around the supermarket.

The farmer used a saw to cut the dirty pigsty in his farm.

The farmer used a duster to clean the dirty pigsty in his farm.

The farmer used a hose to clean the dirty pigsty in his farm.

The waiter used a seed to grow the fresh milk in the teacup.

The waiter used a bucket to pour the fresh milk in the teacup.

The waiter used a jug to pour the fresh milk in the teacup.

Dad used a blue pen to write his shaggy beard in the morning.

Dad used a blunt knife to shave his shaggy beard in the morning.

Dad used a sharp razor to shave his shaggy beard in the morning.

The student used a piano to play the difficult word he didn't understand.

The student used a roadmap to find the difficult word he didn't understand.

The student used a dictionary to find the difficult word he didn't understand.

Santa Claus used a teaspoon to eat his Christmas sleigh around the world.

Santa Claus used a door handle to pull his Christmas sleigh around the world.

Santa Claus used a reindeer to pull his Christmas sleigh around the world.

Gemma used her frying pan to cook her handsome boyfriend on Valentine's Day.

Gemma used her alarm clock to ring her handsome boyfriend on Valentine's Day.

Gemma used her mobile phone to ring her handsome boyfriend on Valentine's Day.

Dad used a kitchen broom to brush his family on holiday to Greece.

Dad used a wheelbarrow to take his family on holiday to Greece.

Dad used a small plane to take his family on holiday to Greece.

Damien used a stepladder to climb the nasty rat in his house.

Damien used a machine gun to kill the nasty rat in his house.

Damien used a mousetrap to kill the nasty rat in his house.

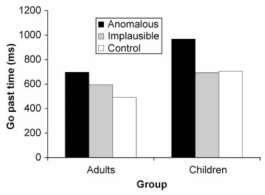

Figure 1.
Go-past reading times for adults and children in the posttarget region in the anomalous, implausible, and control conditions.

Table 1

Adults' and children's mean ratings on a 1–5 scale for sentences in the anomalous, implausible, and control conditions in an off-line questionnaire

	Mean rating		
		Children	
Condition	**Adults (n = 19)**	**7–8 yrs (n = 41)**	**9–11 yrs (n = 41)**
Anomalous	1.55 (0.51)	1.98 (0.74)	1.71 (0.51)
Implausible	2.44 (0.68)	2.04 (0.98)	2.02 (0.72)
Control	4.83 (0.16)	4.44 (0.43)	4.41 (0.39)

Note: Standard deviations in parentheses.

Table 2

Mean fixation times in the pretarget region for anomalous, implausible, and control conditions

Group	Condition	First-fixation duration	Gaze duration	Go-past time	Total reading time
Adults	Anomalous	245 (65)	371 (161)	433 (224)	572 (295)
	Implausible	241 (75)	366 (153)	424 (210)	512 (285)
	Control	238 (59)	338 (139)	405 (281)	451 (254)
Children	Anomalous	285 (133)	551 (360)	734 (817)	873 (580)
	Implausible	287 (109)	539 (320)	642 (466)	677 (429)
	Control	292 (107)	518 (318)	624 (461)	678 (466)

Note: All values in ms. Standard deviations in parentheses.

Table 3

Mean fixation times in the target region for anomalous, implausible, and control conditions

Group	Condition	First-fixation duration	Gaze duration	Go-past time	Total reading time
Adults	Anomalous	273 (92)	324 (141)	395 (226)	507 (311)
	Implausible	263 (75)	299 (111)	369 (211)	428 (250)
	Control	258 (81)	300 (110)	353 (185)	394 (244)
Children	Anomalous	317 (126)	428 (264)	604 (602)	645 (473)
	Implausible	311 (122)	406 (204)	541 (512)	522 (364)
	Control	306 (129)	417 (254)	526 (457)	552 (440)

Note: All values in ms. Standard deviations in parentheses.

Q J Exp Psychol (Hove). Author manuscript; available in PMC 2009 March 30.

Table 4

Mean fixation times in the posttarget region for anomalous, implausible, and control conditions

Group	Condition	First-fixation duration	Gaze duration	Go-past time	Total reading time
Adults	Anomalous	273 (100)	376 (232)	697 (736)	593 (377)
	Implausible	273 (115)	388 (251)	595 (565)	566 (429)
	Control	263 (110)	373 (237)	491 (399)	500 (300)
Children	Anomalous	292 (115)	481 (316)	968 (1040)	751 (500)
	Implausible	297 (119)	494 (332)	693 (550)	669 (424)
	Control	297 (147)	466 (314)	705 (735)	611 (461)

Note: All values in ms. Standard deviations in parentheses.

Q J Exp Psychol (Hove). Author manuscript; available in PMC 2009 March 30.