Chapter 2
How Eye Gaze Feedback Changes Parent-Child Joint Attention in Shared Storybook Reading?

An Eye-Tracking Intervention Study

Jia Guo and Gary Feng

Abstract  Joint attention is critical for effective communication and learning during shared reading. There is a potential disassociation of attention when the adult reads texts while the child looks at pictures. We hypothesize the lack of joint attention limits children’s opportunity to learn print words. Traditional research paradigm does not measure joint attention in real-time. In the current study, three experiments were conducted to monitor parent-child joint attention in shared storybook reading. We simultaneously tracked eye movements of a parent and his/her child with two eye-trackers. We also provided real-time eye gaze feedback to the parent about where the child was looking at, and vice versa. Changes of dyads’ reading behaviors before and after the intervention were measured from both eye movements and video records. Baseline data showed little joint attention in the naturalistic parent-child shared reading. The real-time eye gaze feedback significantly increased parent-child joint attention and improved children’s learning.

2.1 Introduction

Joint attention, the capacity to coordinate attention with a social partner on a particular action or object, is essential for communication, visual search, problem solving, and many other collaborative activities (Brennan et al. 2008; Carletta et al. 2010; Nüssli et al. 2009; Richardson et al. 2007). Parent-child shared storybook reading is one of such activities. There is converging evidence that a key to learn print words is to engage children in a joint attention on texts during reading, i.e., children and adults must attend simultaneously to the target of learning and among themselves
(Ezell and Justice 2000; Gong and Levy 2009; Justice et al. 2006, 2008). However, prior eye-tracking studies showed that pre-reading children focus almost exclusively on pictures while parents read from print texts (Evans and Saint-Aubin 2005; Feng and Guo 2012; Justice et al. 2005, 2008). Additionally, most existing join attention regulation strategies such as finger pointing are adult-centered and limited, such that a parent regulates a child’s attention without accurate information about the child’s real-time attention state. We argue that the lack of joint attention and the limitation of traditional join attention regulations impede children’s learning of print words. And we conjecture that the ideal solution is to provide reading partners consistent, individualized, and real-time attention feedback.

The state of eye-tracking technologies allows us to show this feedback in real-time. We can show parents or children, a cursor on the computer screen that corresponds to the gaze location of the other person. The eye gaze feedback provides critical information that is missing in the traditional shared reading task.

First, the location of the eye gaze indicates the focus of attention at any given moment (Rayner 1998; Rayner et al. 2006). We expect that discovering children’s real-time attention state may trigger adults’ regulations of joint attention during shared book reading. The real-time eye gaze information is more instructional to the pre-reading children, who will see where and how grown-ups look when they read.

Second, having access to the other partner’s eye movements may change the dynamics of shared reading as well as greatly reduce the time and energy that partners spend on reengaging joint attention. The success or failure of a pedagogical attempt is immediately seen on the screen. Adults can give children more prompt and precise feedback when they watch children’s real-time eye movements.

Utilizing the advanced eye-tracking technique to measure and facilitate joint attention provides a new model of understanding the joint attentional interactions in shared book reading. As such, two aims are addressed in the current study. First, we seek to objectively measure the joint attention in shared storybook reading, by simultaneously tracking the eye gazes of both the parent and the child. While there are a handful of published studies looking at children’s eye gaze during reading, none has investigated the correlation and contingency between eye movements of children and parents. Using two eye trackers, we tracked dyads’ eye movements and measured their joint attention in real-time during shared book reading. The data and methodology will be useful to a wide array of researchers interested in joint attention and collaborative behaviors.

Second, we investigate whether two eye gaze based interventions enhance parent-child joint attention during reading. The interventions target the fact that partners in shared reading do not know where the other person is attending to. One intervention involves showing a moving cursor on the child’s monitor that indicates where the parent is looking. The other intervention shows the parent where the child is looking. With this critical piece of information, it is hypothesized that the dyad can better regulate their joint attention, which will facilitate children’s learning of print words.
2.2 General Method

2.2.1 Design and Analysis

Three experiments were conducted in the current study. Experiment 1 is the baseline in which we measured how much joint attention on print exists during the naturalistic shared book reading. We hypothesize that children rarely look at texts and therefore there is little joint attention on print between children and parents. This in turn implies limited print word learning during the reading sessions, as measured by children’s gains in the sight word recognition. Experiment 1 serves as the control condition for Experiments 2 and 3 in both of which we investigated whether the real-time feedback of eye gaze enhances parent-child joint attention and children’s word learning. We presented children how their parents read texts in Experiment 2 and showed parents their children’s real-time eye movements in Experiment 3. We hypothesize that the new paradigm will help dyads regulate joint attention and help children learn print words.

Joint attention was defined in the study as when the partners look simultaneously at (or near) the same visual object on a page (see the examples in Fig. 2.1).

The distance between the screen coordinates of a parent’s eye gaze and those of a child’s eye gaze was calculated at every 20 milliseconds for all reading sessions. To measure the real-time joint attention, we compared the average distance of two partners’ eye gaze locations with a cut-off value of 201.18 pixels which was determined for three reasons. First, when reading partners were asked to look at the same object on the screen in a pilot test, 80 percent of the calculated distance values were within 201.18 pixels. This result suggested when parents and children had joint attention on the screen, most of their eye gaze distance values were less than 201.18 pixels. Second, the visual angle which corresponds to the 201.18 pixels is about 10 degrees (Eyelink systems typically have 20 pixels/degree). The human fovea, where...
we have clear vision, is about 2 degrees. So the visual angle of 10 degrees is not a too small window size for a definition of joint attention. Third, the 201.18 pixels are close to the size of two 5-letter-long print words in pixels (the average length of a 5-letter-long word is 100 pixels). Therefore, we believe this is a very reasonable window to define joint attention in reading.

We determined the joint attention exists if the distance is smaller than 201.18 pixels and does not exist if the distance is larger than or equal to 201.18 pixels. The percentage of time when the distance of two partners’ eye gaze locations is smaller than 201.18 pixels represents how much joint attention the dyad has when reading together.

Video recordings of the parent-child reading interactions were transcribed and coded with InqScribe software. Adapting from the coding systems in previous studies (Chi et al. 2001; Ortiz et al. 2001; Whitehurst et al. 1988; Sulzby 1985), we have developed a coding system to analyze parent-child joint attention interactions. The inter-rater reliability analysis of the behavior coding was performed using 20% of the sample.

We hypothesized that (a) there is limited parent-child joint attention to texts in the naturalistic shared storybook reading, and (b) the eye gaze feedback facilitates joint attentional regulation and improves the learning of print words.

2.2.2 Participants and Materials

Altogether we recruited ninety-two parent-child dyads for this study. Thirty-seven dyads participated in Experiment 1; they also serve as the comparison group for the subsequent intervention experiments. Experiment 2 involved twenty-seven parent-child dyads. Experiment 3 involved twenty-eight dyads. All children participants were 4–5 year old English speakers who had no history of hearing, vision, or cognitive impairments. Parent participants were required to be person who reads most frequently with children at home. Three age appropriate storybooks were presented for dyads to read in all three experiments. Children’s sight word learning was measured before and after reading by asking children to name content words sampled from the storybooks.

2.2.3 Apparatus

Two contact-free eye trackers, a Tobii X50 and an Eyelink 1000 system, were used in the study. Both of the eye trackers are infrared-based remote eye tracking systems that make no contact with the participant. For each dyad, the parent was eye tracked by Tobii X50 and the child was eye tracked by Eyelink 1000. Two video recorders were used to record reading interactions among the dyad.

For each dyad, the parent and the child sat across a child-sized table at a 90 degree angle. One LCD monitor (1280 × 1024 pixels resolution) and Eyelink 1000
Fig. 2.2 The apparatus and experiment set-up

were put approximately 60 cm away from the child; while another LCD monitor (1280 × 1024 pixels resolution) and Tobii X50 were put approximately 60 cm away from the parent (see Fig. 2.2 for details of the set-up). Stimuli were presented simultaneously on both monitors. Stimulus presentation and eye movement calibration and recording were done using the Double Tracker program developed in our lab. The data were then exported offline for statistical analyses.

2.3 Experiment 1: Joint Attention in the Naturalistic Shared Book Reading

Experiment 1 serves as the baseline condition for the subsequent intervention experiments. For each reading dyad, both the parent and the child were eye tracked and video-taped in a naturalistic shared reading task. Before and after the reading trials, we measured children’s sight word recognition to determine they have learned new words.

Specifically, each parent read three books to his/her child in four reading trials (order counter-balanced among participants). They read one storybook in the first and fourth trial, and the other two different storybooks in the second and third trial. In the fourth trial the parent was asked to teach three words that the child did not recognize in the pre-test. To tease apart the impact of the instructions from the moving cursor in Experiment 2, children were asked to follow the parent’s eye gaze while listening to stories, even though they could not actually see the eye gaze on the screen. The average percentage of time children had joint attention with adults on texts in reading trial 1 (no word teaching) was compared to that in trial 4 (with word teaching). Children’s sight word recognition was measured before and after the reading.

The results in Experiment 1 supported our hypothesis that parent-child dyads did not have much joint attention on print and children did not learn many keywords from the pre-test to the post-test. Specifically, in trial 1 when word teaching was not
required, the average percentage of time children had joint attention with parents on print was 2.91%. In the reading trial 4, when adults were asked to teach children three words in the books, children significantly increased their joint attention on print to 6.41%, $t(36) = 2.48$, $p = .018$. Children’s average pre-test raw score of the sight word recognition were 1.81 (out of 10 words), which was lower than the post-test raw score of 2.19 (out of 10 words), $t(36) = 3.19$, $p = .003$. Children’s average number of learned new words was .38 words, measured by the difference of the pre- and post-test raw scores of the sight word recognition.

Taken together, children’s joint attention with parents on print in the naturalistic shared book reading trial were small. Children’s average word learning gain from the pre-test to the post-test was limited. These results were consistent with previous research findings (Evans and Saint-Aubin 2005; Evans et al. 2009; Feng and Guo 2012; Justice et al. 2005, 2008). When children read books with adults, they usually prefer pictures and avoid looking at texts. Since adults read from texts most of the time, children’s ignorance of texts would lead to little adult-child joint attention on print during the naturalistic shared book reading. Therefore, children can hardly improve their sight word learning from the shared book reading activity.

The naturalistic nature of the data set and the analyses we have done in Experiment 1 point to the need for the experimental evaluation of the relationship between children’s joint attention on texts and their word learning. A natural follow-up experiment would be to keep everything else (e.g., the reading materials, study set-up and procedures) the same, and use our newly developed eye tracking technology to let children see parents’ real-time reading eye movements during shared book reading.

2.4 Experiment 2: Externalizing Adults’ Visual Attention for Children in Shared Book Reading

Our pilot studies showed that most preschool children are unaware of the facts that we take for granted such as adults pay attention to print during shared book reading and they read texts from left to right. In Experiment 2, we investigate whether the externalization of adults’ reading processes helps children understand how adults read books as well as how this understanding helps children switch their own attention focus to print. Furthermore, we examine whether the increased joint attention on print promotes children’s word learning.

Experiment 2 involved showing a moving cursor on the children’s screen from reading trial 2 to 4; the moving cursor indicated the location of the parent’s eye gaze in real-time. Parents looked at a normal, static display of the page for all four reading trials. We ensured children and parents understood the gaze indicator using an iSpy-like game. Even the youngest children had no problem understanding the correspondence. Children’s sight word recognition was measured before and after the reading. In the fourth trial the parent was asked to teach three words that the child did not recognize in the pre-test. The average percentage of parent-child joint
attention on texts and children’s sight word learning outcomes in Experiment 2 were compared with those in Experiment 1.

The eye-tracking results showed that the intervention of eye gaze feedback for children significantly increased parent-child joint attention on texts. In trial 1 when no intervention was involved, the percentage of time children had joint attention with adults on texts was 5.35%; in trial 4 when children read the same book while watching their parents’ real-time eye gaze, children significantly increased the percentage of time they had joint attention on texts to 22.7%, $t(26) = -8.01, p < .001$.

To further compare the eye movement changes from the first to the fourth reading trial between Experiment 1 and 2, we did a repeated measures ANOVA using children’s average percentage of joint attention on texts as the dependent variable, the reading trial as the within-subjects independent variable (the first vs. the fourth reading trial), and whether children received the eye gaze feedback as the between-subjects factor (Experiment 1 vs. Experiment 2). The results showed that the main effect of the within-subjects variable was significant, $F(1, 62) = 70.8, p < .001$, suggesting that reading dyads on average significantly increased their joint attention on texts from the first to the fourth reading trial. The main effect of whether children received the eye gaze feedback was also significant, $F(1, 62) = 23.51, p < .001$. So was the interaction effect, $F(1, 62) = 31.24, p < .001$, indicating that the increase of percentage of joint attention on texts from the first to the fourth reading trial in Experiment 2 was significantly higher than that in Experiment 1 (see Fig. 2.3).

More print-directed joint attention resulted in more word learning: children learned on average 1.0 word, significantly higher than the 0.38 words children in Experiment 1 learned, $t(62) = 2.37, p = .02$. This result indicates that children who received the eye gaze feedback learned more words from the pre-test to the post-test than children who did not receive this eye gaze information.

To examine how the eye gaze intervention changes parent-child reading interactions, we included thirty-six dyads in Experiment 1 (data from one dyad were excluded due to poor video quality) and twenty-seven dyads in Experiment 2 in the video coding and analysis. The inter-rater reliability was 0.79 ($p < .01$), which

![Fig. 2.3 Percentages of parent-child joint attention on print from reading trial 1 (no intervention) to trial 4 (with intervention) between Experiment 1 and 2](image-url)
The frequency of the behavior of children’s reading texts along with parents from reading trial 1 (no intervention) to trial 4 (with intervention) between Experiment 1 and 2 could be claimed as good levels of agreement according to previous research (Lan-dis and Koch 1977; Ortiz et al. 2001). We did repeated measures ANOVAs using the average frequency of each coded behavior (time per minute) as the dependent variable, the reading trial as the within-subjects independent variable (no intervention trial 1 vs. with intervention trial 4), and whether children received the eye gaze feedback as the between-subjects factor (Experiment 1 vs. Experiment 2). The results revealed children did respond more to parents’ reading strategies when seeing parents’ real-time eye scanning patterns. For example, children in both experiments increased the occurrences of the behavior of reading texts along with parents from the first to fourth reading trial ($F_{within}(1, 61) = 23.09, p < .001$), but children who received the eye gaze feedback showed a significantly larger increase ($F_{between}(1, 61) = 4.34, p = .041; F_{interaction}(1, 61) = 4.31, p = .042$, see Fig. 2.4).

Overall, the comparisons between Experiment 1 and 2 indicated that with the eye gaze direction more tightly tied to the focus of joint attention, children saw an external representation of reading processes unfolding in real-time. When children heard adults read texts while simultaneously looking at the corresponding words, they had the best opportunity to learn the correspondence between the sound, spelling, and meaning of the words. Furthermore, the real-time eye gaze feedback helped adults efficiently draw children’s attention to those target words and children also responded more to parents’ word teaching attempts. Children’s increased responses to parents’ pedagogical efforts, as well as the increased joint attention on print and the improved understanding of reading processes further promoted children’s print learning.

One limitation for Experiment 2 is that only children had the opportunities of knowing where adults look at on books but adults were still blind to children’s attention states. As a trade-off to this limitation, in Experiment 3 we show parents where children pay attention to during shared reading.
2.5 Experiment 3: Informing Parents of Children’s Real-Time Visual Attention in Shared Book Reading

Parents control the reading activity in the traditional shared reading paradigm, but they have little knowledge about where their children are paying attention to and whether their reading strategies are effective. We argue that informing parents of children’s real-time visual attention can help parents regulate joint attention in shared book reading.

In Experiment 3, we presented each parent a moving cursor on the screen from reading trial 2 to 4. The moving cursor indicated children’s eye gaze location as children read. For all four reading trials children looked at a normal, static display of the page. Parents were encouraged to utilize the eye gaze information to regulate children’s attention. To tease apart the impact of the instructions from the moving cursor in Experiment 2, children were asked to follow the parent’s eye gaze while listening to stories, even though they could not actually see the eye gaze in Experiment 3. Dyads’ eye movements and reading interactions during reading sessions were recorded and analyzed. Children’s sight word recognition was measured before and after the reading. In the fourth trial the parent was asked to teach three words that the child did not recognize in the pre-test. The average percentage of parent-child joint attention on texts and children’s sight word learning outcomes in Experiment 3 were compared with those in Experiment 1. We predict that seeing children’s real-time visual attention makes parents adjust their reading strategies accordingly. The parents’ changed reading behaviors would enhance the efficiency of joint attention regulation and therefore significantly increase the time children spend scanning texts in reading. Children can learn more words due to the increased print exposure.

Positive intervention effects were also found in Experiment 3. The percentage of joint attention on texts increased from 3.48 % in reading trial 1 (no intervention trial) to 12.87 % in reading trial 4 (with intervention trial), \( t(27) = -5.05, p < .001 \). To further compare the eye movement changes from the first to the fourth reading trial between Experiment 1 and 3, we did a repeated measures ANOVA using children’s average percentage of joint attention on texts as the dependent variable, the reading trial as the within-subjects independent variable (the first vs. the fourth reading trial), and whether parents received the eye gaze feedback as the between-subjects factor (Experiment 1 vs. Experiment 3). The results showed that the main effect of the within-subjects variable was significant, \( F(1, 63) = 31.72, p < .001 \), suggesting that reading dyads on average significantly increased their joint attention on texts from the first to the fourth reading trial. The main effect of the eye gaze feedback was also significant, \( F(1, 63) = 5.35, p = .024 \). So was the interaction effect, \( F(1, 63) = 6.62, p = .012 \). The results indicated that the increase of percentage of joint attention on texts from the first to the fourth reading trial in Experiment 3 was significantly higher than that in Experiment 1 (see Fig. 2.5).

Parents became more effective facilitating children’s word learning. Children learned 1.25 words, significantly higher than the word learning gain in Experiment 1 (0.38 words), \( t(63) = 3.83, p < .001 \). This result further confirmed that children
learned more words from the pre-test to the post-test when parents used the eye gaze feedback to effectively direct children’s attention to words.

The behavioral coding and analysis also supported our hypotheses. Thirty-six dyads in Experiment 1 (data from one dyad were excluded due to poor video quality) and twenty-eight dyads in Experiment 3 were included in the comparisons for the changes of the frequencies of the behaviors from reading trial 1 (no intervention) to reading trial 4 (with intervention) between Experiment 1 and Experiment 3. The inter-rater reliability was 0.78 ($p < .01$). We did repeated measures ANOVAs using the average frequency of each coded behavior (time per minute) as the dependent variable, the reading trial as the within-subjects independent variable (no intervention trial 1 vs. with intervention trial 4), and whether parents received the eye gaze feedback as the between-subjects factor (Experiment 1 vs. Experiment 3).

The results revealed when informed with children’s real-time eye gaze locations, parents increased the frequencies of regulating joint attention and teaching words. For example, parents in both experiments increased the occurrences of behavior of asking children to look at specific words (e.g., “Can you help me find the word ‘cat’ on the screen?”) from the first to fourth reading trial ($F_{within}(1, 62) = 62.95, p < .001$), but parents who received the eye gaze feedback showed a significantly larger increase ($F_{between}(1, 62) = 15.32, p < .001; F_{interaction}(1, 62) = 15.32, p < .001$, see Fig. 2.6).

Compared to parents in Experiment 1, parents in Experiment 3 significantly increased the occurrences of the behavior of providing specific feedback (e.g., adults saying “Yes, you are looking at the right place.”; “No, you are not looking at the place I want you to look.”) from the first to fourth reading trial ($F_{between}(1, 62) = 26.57, p < .001; F_{interaction}(1, 62) = 26.57, p < .001; F_{within}(1, 62) = 27.63, p < .001$, see Fig. 2.7).

Parents’ behavior changes induced children’s more frequent verbal responses. Although children in both experiments increased the occurrences of the behavior
of asking or answering print-related questions from the first to fourth reading trial ($F_{within}(1, 62) = 54.59, p < .001$), children in Experiment 3 showed a significantly larger increase ($F_{between}(1, 62) = 22.71, p < .001; F_{interaction}(1, 62) = 22.64, p < .001$, see Fig. 2.8).

The above comparisons between Experiment 1 and Experiment 3 suggested when parents received the real-time eye gaze feedback, parents had better opportunities to observe their child’s attention state and fine-tune their strategies to increase child interest and participation. Children who experienced positive direction, coaching, and correction more easily attended to and internalized the knowledge parents attempted to teach them, and developed the interest and motivation to sustain their learning. These changes in turn provided more teachable moments for parents.
2.6 Conclusion

The current study measured parent-child joint visual attention in real-time, which allows us to go beyond prior research that focuses exclusively on the child in shared reading, and study shared reading as a joint attentional interaction involving dynamic transactions between partners and real-time cognitive strategies within individuals. Building on prior research, we found that pre-reading children had limited joint attention with their parents in the naturalistic shared reading paradigm. Parents had little information about where children were attending to, and children had even less idea about how adults actually read. This resulted in a poorly regulated joint attentional interaction when it comes to learning print words.

More importantly, our eye gaze interventions successfully remedied the joint attentional structure by leveraging the eye-tracking technology in the shared reading. By providing real-time feedback of the partner’s visual attention, we demonstrated significant improvements in the amount of joint attention on print and changes in parental attentional regulation strategies during reading. More interestingly, children did not simply look at the moving cursor or print words, but actually read and processed the words. This was shown by increased word learning by children, along with children’s changes of concept of reading processes. These results suggest that by providing a critical piece of information—namely, where the partner is looking—we can facilitate the regulation of joint attention and improve children’s learning of print words. Our intervention targets limitations in joint attention regulation in the traditional shared reading practice, but it is not specific to reading. The data and methodology of this study would also be useful to a wide array of research topics on collaborative learning activities. To the extent learning involves joint attention (e.g., in math tutoring), the eye gaze feedback can be an effective aid for learning. Additionally, the scenario of this research is quite similar to online collaborative activities where partners focus on the common content on different screens (e.g., online gaming, collaborative search, etc.). Insights into such behaviors and mental processes may help design better multimedia software products and web applications.
References


