Chen, W. et al. (Eds.) (2016). Proceedings of the 24th International Conference on Computers in Education. India: Asia-Pacific Society for Computers in Education

Effects of Embodied Interaction on Reading Comprehension in a Multimodal Environment

Ya-Wen CHENG[a], Wei-Chieh FANG[b], Wei CHANG[a], Li-Chun LIN[a] & Nian-Shing CHEN[a*]
[a] *Department of Information Management, National Sun Yat-sen University, Kaohsiung, Taiwan*
[b] *Department of Education, Washington University in St Louis, MO, USA*
[*] nschen@mis.nsysu.edu.tw

Abstract: The study examined the effects of embodied interaction on reading comprehension in a multimodal environment. Specifically, this study compared how textual, visual or gestural modes affect young children's reading comprehension. 91 Chinese EFL learners were assigned to one of the three modes to complete a reading task. During the reading task, they made several predictions by selecting a text description, selecting a picture or doing gestures that best match their predictions. Results showed that the three modes led to comparable level of reading comprehension. However, when reading ability was taken into account, the high level readers' comprehension was negatively affected by the gesture mode while the low level readers' comprehension was not. The findings suggest that visuals and gestures can be integrated into a multimodal reading environment to make reading an active learning experience. However, in designing an interactive digital storybook involving gestures, it is recommended that individual differences in reading ability be addressed to maximize its effect. Suggestions for future studies are discussed.

Keywords: Gesture-based interactions, multimodal texts, E-book, embodiment, reading comprehension

1. Introduction

A multimodal learning environment is normally filled with visual, aural and textual input in a digital format. The rich affordances of such a digital learning environment are beneficial to young readers (Dalton et al., 2011). Reading is not an automatic process as readers need to "visualize, infer, predict, conceptualize and imagine" the ideologies conveyed in the text (Walsh, 2006). In multimodal texts, compared with print-based texts where meaning-making mainly relies on understanding the words (Walsh, 2006), readers are able to use various senses to comprehend the texts. Studies looking into the effects of technology-mediated reading comprehension found that digital tools and learning environments positively affect reading comprehension (Dalton et al., 2011). Multimodal texts have been found to engage readers, activate various schemata (Yanguas, 2009), encourage imagination (Maureen, 2006), facilitate understanding of information (Coiro, 2003). However, little research has looked into the how kinesthetic involvement such as gestures affects reading comprehension in a multimodal learning environment.

2. Literature Review

Typical multimodal studies in the field of reading have explored visual and audio modes. However, a few researchers have argued that reading is not a passive learning activity. Rather it can be an embodied activity, in which comprehension can arise from simulating the linguistic content by using body systems of perception, action and emotion (Glenberg, 2011). This notion has received support from theories of embodied cognition, which proposed that cognitive processes are rooted in the body's interactions with the world (Wilson, 2002; Barsalou, 2008). One such theory is simulation theory of language comprehension, which posits that language comprehension can benefit from mapping abstract symbols of language such as texts onto embodied experiences. These embodied experiences are encoded in memory as perceptual symbols during previous interactions with objects in authentic situation. Thus, not only language symbols but also affordances deriving from interaction with an object can reactivate the representation of embodied experiences in memory, which in turn can lead to better text comprehension.

With the advancement of gesture-based technology, a multimodal environment has the potential to involve integrated modes of senses such as visual, tactile, hearing and kinesthetic senses at the same time (Chen & Fang, 2014). In line with the embodied approach, a few studies in educational technology have attempted to explore whether computer-guided physical involvement can lead to enhanced learning during reading. Glenberg et al. (2009) found that physical manipulation of real objects and manipulation of digital image can facilitate young children's comprehension on a short narrative story describing a farm life (e.g., the farmer brings hay to the horse). First- and second-graders, who were first asked to read a text and then to manipulate toys to correspond to what they read, comprehended the story better than those who merely read the text twice. It was also found that manipulation of digital images on the computer could lead to the same effect on reading comprehension. These insights suggest that manipulation of digital content can help to (1) map words to meaningful representation and (2) map the syntactic relations to actions during reading.

Another study by Homer et al. (2014) examined whether reading a Kinect-based storybook would distract or enhance story comprehension and vocabulary learning during reading. In their study, kindergarten and first grade students received three interventions: reading a storybook, reading a digital storybook or interacting with a digital storybook in a Kinect-based literacy game. Those in the interactive digital condition used gestures or movement to interact with the story content including target vocabulary and plot points. It was found that the three types of storybook interventions led to comparable levels of story comprehension, suggesting that interactive digital storybook did not enhance nor distract readers. However, the interactive digital storybook did lead to better vocabulary learning. Based on the review here, it appears that the effects of physical engagements on reading comprehension are mixed. More studies are needed to look into how technology can engage learners physically during reading and how they affect reading comprehension.

This study attempts to include the physical mode in a multimodal reading environment and to compare the effects of physical mode with the effects of textual and visual modes on reading comprehension. As young readers usually fail to understand the whole text while they are able to decode text (Oakhill et al., 2003), this study focuses on how to help young readers build cohesive understanding of the text. To this end, a prediction making strategy, an instructional strategy commonly used to activate top-down processing to facilitate general reading comprehension (Ajideh, 2003), was implemented in a multimodal reading environment.

The following questions guided the current study:

1. What are the effects of making prediction in the textual, visual or gestural mode on reading comprehension?
2. Do the effects of making a prediction in different modes vary depending on the existing levels of reading ability?

3. Method

3.1 Participants

A total of 91 Chinese EFL learners, aged 9 to 10, were recruited from a science summer camp at the National Science and Technology Museum in Southern Taiwan. They were 3 to 6 graders who were assigned to the textual, visual and gestural conditions using stratified sampling to ensure all graders were balanced across all conditions. They had an average of 3.8 years of English learning experience.

3.2 Digital Storybook

A storybook titled I Entered a Wonderland was adapted from an English textbook for secondary school students in Taiwan. The story was a modified version with a plot different from the original story, which was considered rather unfamiliar to our participants. The word choice and grammatical complexity were tailored to the language level of the participants. The story was audio recorded by a native American speaker.

Three digital versions were created using Visual Studio in Windows. The interactive version of the Kinect-based story can be run on Microsoft operating system with a Kinect unit. The application allows participants to use gestures to navigate the application and do gestures related to the story lines in a natural user interface. The other two digital versions, textual and gestural ones, can be run on the internet browsers. Readers can interact with the digital content in the graphic user interface.

2.3 Experimental and Application Design

An experiment with a between-subjects design was conducted to compare the effects of making predictions in the interactive textual, visual and gestural modes. A multimodal application was developed to allow readers to select a text description, select a picture or act out a scenario that corresponds to their prediction. The application begins with a reading section followed by a prediction section for the three groups. An additional practice section was provided only for the gestural group. The following is the application design:

2.3.1 Practice Section

The purpose of the practice section is to familiarize participants with the nine gestures they might be using in the prediction section (see figure 1). This section is only available to the gestural group.

Figure 1. Practice section.

2.3.2 Reading Section

During the reading section, the screen shows story lines and pictures. When a story line is being read to the participants, it is displayed in red. At the end of the reading section, the application pauses and shows a blank to prompt readers to make prediction. The user interface is the same for the three conditions.

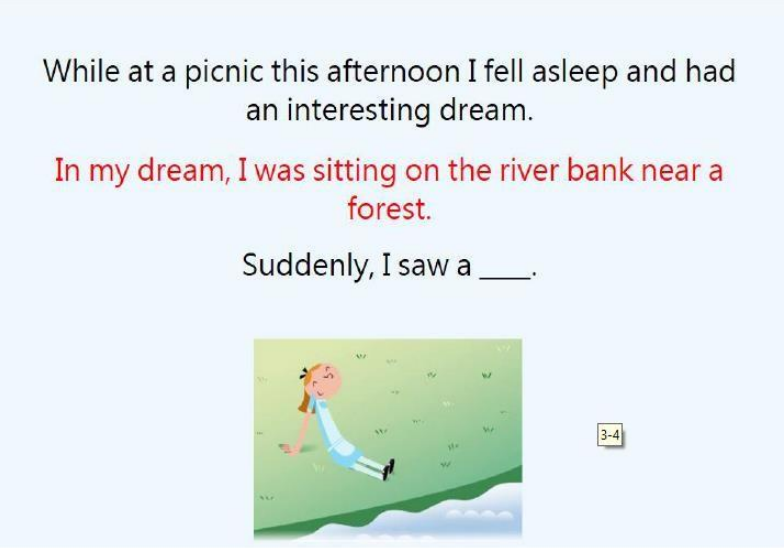

Figure 2. Reading section.

2.3.3 Predicting section

During the prediction section, the application prompts readers to make predictions about what would happen next. Participants make their selections after the questions are read and the light on the right side turns green. For the gestural group, participants make their prediction by either moving the main character to the option or making the same gestures (see figure 3) as they play a role as the story character in the story. Participants in the textual group (see figure 4(a)) and the visual group (see figure 4(b)) make prediction by clicking on one of the options.

Figure 3. Prediction section for the gestural group

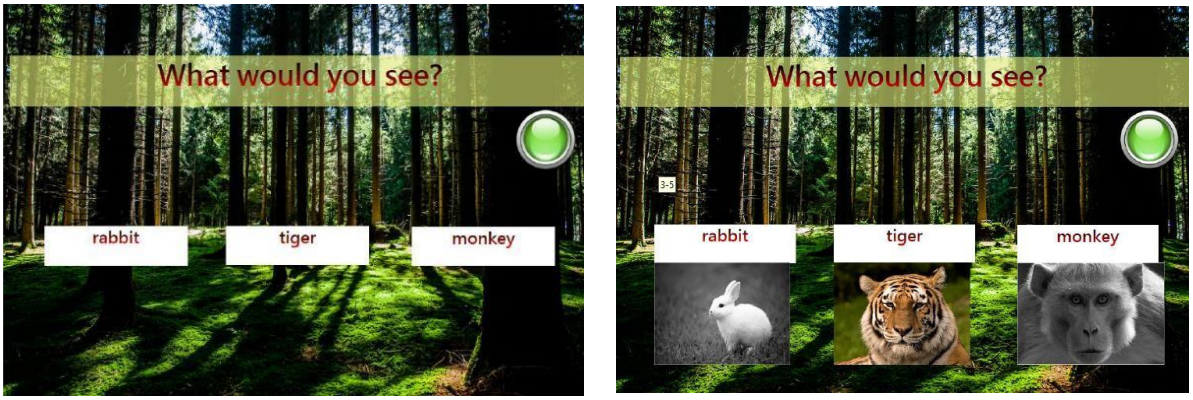

Figure 4. (a) Prediction section for the textual group (b) Prediction section for the visual group

2.4 Instruments

2.4.1 Pre-test

The pretests included demographic questionnaires and reading tests. To measure participants' reading ability, a short story titled Grandma's Mangoes was adapted from an eighth-grade English textbook in Taiwan. There were 10 comprehension questions and the possible scores were 10.

2.4.2 Post-test

To measure reading comprehension on I Entered a Wonderland, 10 questions tapping into their text-explicit, text-implicit and scripturally implicit knowledge and false knowledge, or concept that can be easily misunderstood, were designed. There were 10 comprehension questions, and the possible scores were 10. Since the language of the reading comprehension test can be a factor affecting reading comprehension (Brantmeier, 2005), the posttest was in participants' L1, Mandarin Chinese.

2.5 Procedure

Upon arrival, participants spent 5-10 minutes filling in a demographic questionnaire and taking a reading comprehension pre-test on Grandma's Mangoes. Each participant was then given a 5-minute introduction to the experimental procedure and content.

Figure 5. Participants engaged in different modes of reading during the reading task.

All were then taken to different research sites for receiving intervention in the textual, visual or gestural modes respectively (see figure 5). The application first guided them to imagine themselves in a story and act as the story characters. Next, they used the application to complete a reading task for their group. The task involved reading a series of passages from I Entered a Wonderland and predicting the story plots during reading. When making prediction about the story, they were prompted to select their answers using either the mouse or gestures according to the group to which they were assigned. Although different groups used different interactive modes, the passages and the prediction choices were identical. The whole reading task took around 10-15 minutes.

After completing the task, the participants were interviewed about their experience with the application with semi-structured questions and were given a post-test, by pen and paper, on their reading comprehension of the story. This part took between 5 to 10 minutes. All participants were debriefed and thanked for their participation.

4. Results

4.1.1 Pre-test results

To ensure that there were no differences between groups for reading ability, a one-way ANOVA was performed using modes (the textual, visual and gestural modes) as the between-subject variable and the reading scores on the pretest as the dependent variable. The results showed that there were no significant differences between groups, $F(2, 90) = .21$, $p > .05$, suggesting that the participants from the three groups were equivalent in terms of reading ability when they started the experiment (see table 1).

4.1.2 Post-test results by research question

The first research question concerns how different modes might affect reading comprehension when readers made predictions during reading. A one-way ANOVA was performed to examine the differences in reading comprehension between the three conditions, $F(2, 90) = 1.23$, $p > .05$ (see table 1). The results showed no significant differences between groups, suggesting that the text, picture and gesture modes led to comparable level of reading comprehension.

Table 1: Means and standard deviation of the pre- and posttest reading comprehension tests by condition.

	Reading pre-test			Reading post-test		
Condition	n	*M*	*SD*	n	*M*	*SD*
Textual mode	30	5.67	2.82	30	7.13	1.33
Visual mode	30	5.57	2.39	30	7.36	1.35
Gestural mode	31	5.97	2.40	31	6.74	1.95

The results indicate that the participants in the three modes achieved similar level of reading comprehension. This result echoes Homer et al.'s study (2014), which found that the Kinect-based storybook led to comparable comprehension as the physical storybook. However, the result does not support Glenberg et al.'s findings (2009), which show that the manipulation of the digital image can lead to better reading comprehension. One possible explanation is that the story adopted in this study is rather easy to comprehend for children who are between 3-6 grades.

The second question concerns whether the effects of making a prediction in the textual, visual and gestural modes on reading comprehension would vary as a function of reading ability. Reading levels, high vs. low, were determined based on two cutoff points. Those who scored below a score of 4 were categorized as low level and those who scored above a score of 7 were categorized as high level. A 3 (mode type) x 2 (high vs. low reading level) ANOVA was performed on reading scores on the posttests. Results showed no significant main effect of the mode type, $F(2, 64) = 0.34$, $p > .05$, a significant main effect of reading level, $F(1, 64) = 4.29$, $p = .04$, and a significant interaction, $F(2, 64) = 4.01$, $p = .02$. Further analysis of the simple main effect showed that high level reader scored significantly lower in the gesture mode (M = 6.5, MSE = 1.97) than in the text mode (M = 7.73, MSE = 1.22) and in the picture mode (M = 8.17, MSE = 0.83), $F(2, 40) = 5.01$, $p = .01$. Low lever reader did not score differently across the three modes.

Despite the finding that there is no difference in comprehension between the three modes, further analysis shows that readers' comprehension varies depending on their existing reading ability (see figure 6). It appears that readers with higher reading ability were interfered by gestures they performed during reading. It is possible that these readers might have their ways of comprehending the text such as building mental model through semantic activation. Activating the embodied experience might not benefit comprehension.

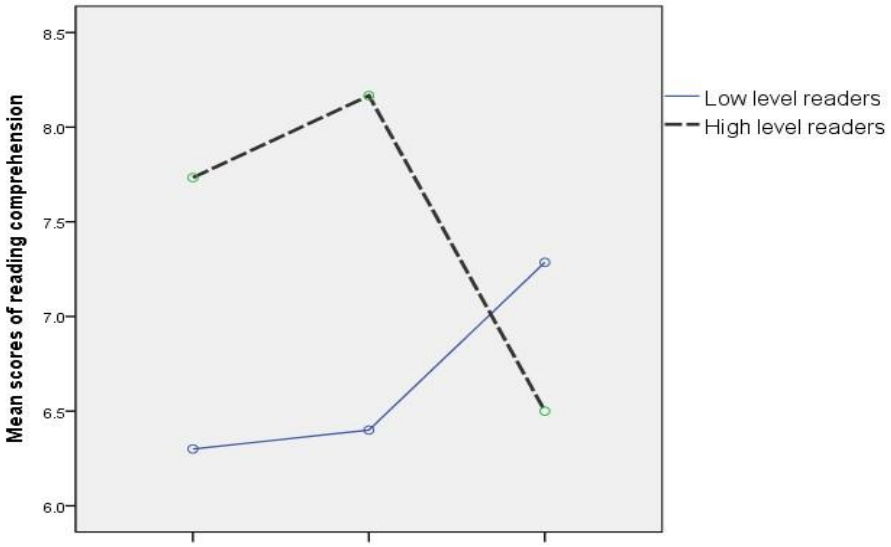

Figure 6. Mean scores of reading comprehenion following textual, visual and gestural modes for low and high level readers.

5. Conclusion

This study is among the few studies to include bodily interaction in a multimodal reading environment and to compare how textual, visual and gestural modes affect young readers' story comprehension. While most studies investigate vocabulary learning during reading and word-level decoding, or meaning-making during reading, this study examined the cohesive understanding of the text. During reading, young readers were guided to select text, select pictures or act out to make predictions about the story plots. Overall, the results showed that the textual, visual and gestural modes have similar effect on reading comprehension when they are combined with prediction strategy in the multimodal interactive environment. However, when reading ability was taken into account, the high level readers' comprehension was negatively affected by the gesture mode while the low level readers' comprehension was not affected. The findings suggest that pictures and gestures can be integrated into a multimodal reading environment to make reading an active learning experience. However, in designing an interactive digital storybook involving gestures, it is recommended that individual differences in reading ability be addressed to maximize its effect. An adaptive mechanism can be designed to diagnose reading ability at the beginning and then recommend the appropriate learning modes for subsequent reading activity.

Some limitations should be noted. First, the present study did not include delayed comprehension test due to the arrangement of the summer camp where the participants were recruited. Several studies have shown that bodily involvement can slow forgetting rate (Kao et al., 2014; Chao et al., 2013) as measured by the delayed tests. Future studies should look into the effect of longer test interval on reading comprehension. Second, the reading level defined in this study is unique to its population and thus should be interpreted with cautions. Future studies should continue to examine how technology-enabled embodiment affects reading comprehension.

Acknowledgements

This study was supported by the Ministry of Science and Technology, Taiwan, under the project numbers of MOST104-2511-S-110-009-MY3, MOST104-2511-S-110-007-MY3, and MOST103-2511-S-110-002-MY3. We would like to thank Adam Henze for his help with the audio recording.

References

Ajideh, P. (2003). Schema theory-based pre-reading tasks: A neglected essential in the ESL reading class. *The reading matrix, 3*(1), 1-14.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(1), 617-645. doi: doi:10.1146/annurev.psych.59.103006.093639

Brantmeier, C. (2005). Effects of reader's knowledge, text type, and test type on L1 and L2 reading comprehension in Spanish. *The Modern Language Journal, 89*(1), 37-53.

Coiro, J. (2003). Reading comprehension on the Internet: Expanding our understanding of reading comprehension to encompass new literacies. *The Reading Teacher, 56*(5), 458-464.

Chao, K. J., Huang, H. W., Fang, W. C., & Chen, N. S. (2013). Embodied play to learn: Exploring Kinect-facilitated memory performance. *British Journal of Educational Technology, 44*(5), E151-E155.

Chen, N. S., & Fang, W. C. (2014). Gesture-Based technologies for enhancing learning. In Huang et al. (Eds), *The New Development of Technology Enhanced Learning* (pp. 95-112). doi: 10.1007/978-3-642-38291-8_6

Dalton, B., Proctor, C. P., Uccelli, P., Mo, E., & Snow, C. E. (2011). Designing for diversity the role of reading strategies and interactive vocabulary in a digital reading environment for fifth-grade monolingual English and bilingual students. *Journal of Literacy Research, 43*(1), 68-100.

Glenberg, A. M., Brown, M., & Levin, J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology, 32*(3), 389-399.

Glenberg, A. M. (2011). How reading comprehension is embodied and why that matters. *International Electronic Journal of Elementary Education, 4*(1), 5.

Homer, B. D., Kinzer, C. K., Plass, J. L., Letourneau, S. M., Hoffman, D., Bromley, M., ... & Kornak, Y. (2014). Moved to learn: The effects of interactivity in a Kinect-based literacy game for beginning readers. *Computers & Education, 74*, 37-49.

Kuo, F. R., Hsu, C. C., Fang, W. C., & Chen, N. S. (2014). The effects of embodiment-based TPR approach on student English vocabulary learning achievement, retention and acceptance. *Journal of King Saud University-Computer and Information Sciences, 26*(1), 63-70.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and cognitive processes, 18*(4), 443-468.

Walsh, M. (2006). The'textual shift': Examining the reading process with print, visual and multimodal texts. *Australian Journal of Language and Literacy, 29*(1), 24.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625-636.

Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology, 13*(2), 48-67.