EXPLORING THE POTENTIAL OF HYPERMEDIA ANNOTATIONS FOR SECOND LANGUAGE READING

Robert Ariew
University of Arizona, Tucson

Gulcan Ercetin
Bogazici University, Istanbul, Turkey

This study explores whether there is a relationship between reading comprehension and the use of different types of hypermedia annotations. Data came from 84 intermediate and advanced level adult ESL learners using a tracking tool that recorded the amount of time readers spent on a given annotation. Results indicate that annotation use did not contribute to reading comprehension for the advanced group whereas over-reliance on certain annotations had a negative impact on the intermediate group. Moreover, prior knowledge was found to be an important variable related to reading comprehension. Finally, the qualitative data collected through a questionnaire and interviews revealed that hypermedia reading had a positive impact on the participants' attitudes towards reading on the computer.

INTRODUCTION

The use of computers affects many areas of daily life, including how we read and interact with texts. While we continue to read some material in a linear fashion, turning the pages of a book in sequence (or going from screen to screen one at a time), much of what we read is non-linear, as when we use a table of content to locate a specific piece of information in an instruction manual. It is also this kind of reading that we engage in when we use computers to access Internet sources. When we "surf" the Internet, information appears as a hypertext. A hypertext presents information in nodes and links integrating multiple forms of media. Reading becomes an exercise in acquiring information from nodes and then "jumping" to other nodes by means of links. Thus, reading is transformed from a linear process into a non-linear one.

Just as hypertexts came into widespread use on the Internet, hypermedia emerged. Hypermedia constitutes a system that combines hypertext and multimedia. While hypertext is characterized by its use of nodes and links to present textual information, hypermedia presents information through nodes and links integrating multiple forms of media. These nodes and links may provide textual, audio, or graphic annotations to aid reading comprehension. These added dimensions are important tools in promoting understanding of texts, especially texts in a second or foreign language. For this reason, many language teachers are incorporating hypermedia in their curricula.

Reading Glosses in Second Language (L2) Reading

Glosses have been used in second language reading materials for many years to help students understand texts. Glosses (1) can be prepared by the learners themselves or by professionals; they can be presented before or during reading; they may function to highlight or clarify important points or simply provide lexical or syntactic information; their focus may be textual or extratextual; they may appear in a first language or a target language; they may be verbal or visual in a variety of media.

The issue of effectiveness of glosses has been investigated through the use of tracking technology. Noblitt and Bland (1991) define tracking as "a [software] device for recording requests for information" (p. 121). Blake (1992) emphasizes that tracking helps researchers trace and explore learners' reading strategies. Hulstijn (1993) also considers tracking as a useful and valid means of gaining "a deeper understanding" (p.39) of reading strategies. Similarly, Collentine (2000) suggests that this technology provides insights into both the product and the process of learning. Hence, tracking allows the researchers or teachers to follow what the learners are actually doing.

Aust, Kelley and Roby (1993) and Roby (1999) compared the use of text-based hyper-dictionaries and conventional dictionaries. Their hyper-dictionaries were text-based and did not include any multimedia features. The variables investigated were the frequency of consultation of the dictionary, reading time and reading comprehension. Both studies revealed that learners tended to access computer dictionaries more frequently than conventional dictionaries. Moreover, access to computer dictionaries was found to reduce the amount of time used to read the text. However, neither study showed any differences in reading comprehension between the groups. Although this kind of research may provide useful information, tracking reading time and frequency of dictionary access on paper is unlikely to be as precise as on the computer, where a tracking tool may be used.

Tracking look-up behavior and reading time, Knight (1994) investigated the effectiveness of accessing a dictionary for reading comprehension. Participants were one hundred and twelve intermediate-level Spanish students categorized as low and high verbal ability groups based on their verbal scores on the American College Test. They were randomly assigned to "dictionary access" and "no dictionary access" conditions on the computer. Analyses of recall protocols revealed that the relationship between reading comprehension and word lookup was substantial (r=.68) for low verbal ability group whereas it was negligible for high verbal ability learners (r=.17). Low ability groups ostensibly benefit from dictionary access, but its effectiveness is questionable for high ability learners. The results also showed that dictionary access lead to more time on task regardless of the ability groups. However, only in the low ability group did an increase in reading time lead to improved comprehension.

Davis and Lyman-Hager (1997) investigated the relationship between the use of computerized glosses and reading comprehension and attitudes towards the use of glosses. Participants were forty-two randomly selected

intermediate-level French students, who were asked to read a literary text annotated with English definitions, French definitions, grammar explanations, pronunciation, cultural background information and pictures. Their reading comprehension was measured through a multiple-choice test and a recall protocol in English. The annotations were categorized as definitional glosses (English definitions) and nondefinitional glosses (all other types) after an examination of the participants' log data, which revealed that 85 percent of the information accessed was English definitions of words. Multiple regression analysis revealed that as the grade level increased, the score for the two comprehension measures also increased, whereas the frequency of access to the glosses decreased. The type of gloss was not significant in relation to reading comprehension.

Chun (2001) conducted a study with 23 students enrolled in a second year German course. Participants were given two different texts with comparable length and topics to read and then write a summmary of the texts. While one of the texts included an internal and external bilingual dictionary, the other had only an external bilingual dictionary. Results indicated that learners looked up more words when both an internal and an external dictionary were available and performed better on the recall task. However, no relationship was found between the number of words looked up and the number of propositions recalled. On the other hand, a positive significant relationship was found between the amount of time spent on task and the number of propositions recalled for one of the two texts. Participants were categorized as high and low ability based on their grades in German. High ability participants looked up fewer words and spent less time on task. However, no differences were found between the two groups in the number of propositions recalled, which can lead to a conclusion that look up behavior was more beneficial for low ability students.

The studies reviewed above indicate that easy access to annotations increases frequency of consultation and reduces study time. However, there may not be a direct relationship between annotation use and reading comprehension. Other variables, such as verbal ability and proficiency, may interact with the effects of annotations. Annotations, for instance, may be more beneficial for low verbal ability and low proficiency learners.

The purpose of the present study was to explore the relationship between the use of hypermedia annotations by intermediate and advanced L2 learners and their reading comprehension. Specifically, the study concerned the types of hypermedia annotations that are the best predictors of reading comprehension for intermediate and advanced ESL learners.

METHODS AND PROCEDURES

Participants

A total of 103 adult ESL students enrolled at the Center for English as a Second Language (CESL) at the University of Arizona during the Spring 2000 semester took part in this study. CESL students typically study English

for academic purposes. To keep the participants as homogeneous as possible, data for 19 participants were eliminated from the study because they did not study English for academic purposes. The remaining 84 participants reflected a variety of language and cultural backgrounds: Arabic (29), Spanish (16), Japanese (13), Korean (10), Chinese (5), Indonesian (2), Portuguese (2), Thai (2), Bulgarian (1), Turkish (1), German (1), Pular (1), Vietnamese (1). There were 58 male and 22 female participants, four did not indicate their gender. The average age was 24.06, ranging from 17 to 40. Students were placed at proficiency levels based on their performance on a standardized placement test. Each class consisted of 10 to 15 students. In all, 9 classes, with 34 intermediate and 50 advanced students served as participants.

Data Collection Instruments
CESL Placement Test

CESL uses a standardized test called the "Comprehensive English Language Test for Learners of English (CELT)" to place students at proficiency levels. CELT is designed to measure English language proficiency of intermediate and advanced ESL students. It has three sections: Listening, Structure and Vocabulary. Only the Listening and Structure sections are used for placement purposes by CESL. The reliability coefficients ranged from .82 to .94 for the listening section, and from .88 to .96 for the structure section.

Hypermedia Reading Software

A hypermedia reading software titled "Stephen Hawking's Universe" was developed based on a series of video programs produced by the Public Broadcasting Service (PBS). The text that accompanies these programs was obtained from the PBS website (http://www.pbs.org/wnet/hawking/html/home.html). The source text was much longer than necessary for purposes of this research and was shortened. No semantic adjustments such as changing vocabulary or simplifying the text were made, so as to avoid affecting the authenticity of the text.

The text was annotated with the aid of multiple forms of digital media to aid comprehension. There were annotations in the form of text, graphics, audio and video. Based on the interactive theory of reading (Bernhardt, 1991; Kintsch & Van Dijk, 1978; Rumelhart, 1977), two categories of annotation emerged to facilitate both top-down and bottom-up processes; *textual annotations* provided information about the text, such as definitions of words (textual text annotations), their pronunciation (textual audio annotations) and graphics illustrating definitions (textual graphics annotations), whereas *contextual annotations* provided extra background information about the topic in the form of text (contextual text annotations), audio (contextual video annotations), graphics (contextual graphics annotations) and video (contextual video annotations). Textual annotations appeared within the body of text; annotated words or phrases were a part of the text and were indicated in blue. On the other hand, contextual annotations appeared outside the text and were indicated by green buttons placed on the margins of the screen. When

participants clicked on a word or phrase highlighted in blue (textual annotation), or a background information button (contextual annotation), a small information box appeared and provided a choice of the forms of media through which the information was available (i.e. text, graphics, sound, or video) (see Appendix I). They were then able to choose and view as many annotations as were available.

The reader's interaction with the text was tracked by keeping a record of the types of annotations the reader chose to view, the amount of time (in seconds) the reader spent on viewing the annotations and the number of times the annotations were accessed. The tracking feature provided an unobtrusive tool for collecting information about a participant's interaction with the text. This information was saved as a log file.

Reading Comprehension Test

A reading comprehension test was prepared to test students' comprehension of the text. To develop the questions, one native speaker and two nonnative speakers—all graduate students—read the text and took notes. They believed that their task was to take a comprehension test, during which they would not be able to see the text. The questions were based on the notes taken by these informants. The rationale behind this was that these experienced readers would take notes on the main and important ideas; therefore, questions based on these notes would test the main ideas of the text itself (but not of the annotations, which were not provided to the informants). Thus, the comprehension test did not include any information from the annotations.

The test had 14 items of three types: (a) six short-answer questions, (b) six multiple-choice questions and (c) two open-ended questions. Scoring of the test entailed consideration of different weights based on the question type. Correct multiple-choice responses received a score of one point, whereas each correct answer to a short-answer question or an open-ended question was worth two points. The highest possible score on the test was 22.

The reading text was not available to the participants during the test, but they were allowed to take notes while they were reading based on the argument that reading comprehension is more validly measured when students do not have the text available (Pederson, 1986). Hence, participants in this study were allowed to look at the notes they took during reading, but were not allowed to refer to the text during the comprehension test. The participants completed the test via computer. Each screen provided two questions with space below the questions to enter answers. Participants could go back and forth between the questions and review them as many times as they wanted. However, once they clicked the "quit" button, the program ended, and no further changes could be made.

Prior Knowledge Test

This test consisted of ten true and false questions. The purpose in using the test was to ascertain how much the participants already knew about the topic. Since the test was completed just before the reading, the questions

were different from those on the reading comprehension test to minimize sensitizing respondents to the test. The items in the prior knowledge test involved general concepts about the topic, whereas those for the reading comprehension test involved more specific questions about the text.

The participants' familiarity with reading hypertext documents was determined based on the frequency with which they accessed the Internet. They used the Internet for a variety of purposes including e-mail, chatrooms, searching for information. The group was rather heterogeneous with a mean of 8.60 hours per week (SD= 7.64, min= 1, max= 35). 46.7 percent of the participants reported that they used the internet for 1 to 5 hours per week, 23.4 reported using the internet for 6 to 10 hours per week, 18.2 percent used it for 11-15 hours per week while 11.7 percent used it for than 20 per week.

Background Questionnaire

Participants completed a background questionnaire after the experiment; the purpose for it was to find out more information about the participants and their experience with electronic reading. In addition to demographic information, items related to the participants' experience with technology, in general, as well as the usefulness or difficulty of use of certain features of the software.

Student Interviews

Twenty volunteers participated in interviews no more than 3 days after completing the comprehension test. These semi-structured interviews took place on a day other than the experiment and were approximately one half hour in length. The purpose of the interviews was to provide supplementary data concerning participants' use of annotations while reading the electronic text.

Data Collection Procedures

The participants took part in this study at the Instructional Technology Facility of the College of Education at the University of Arizona during regular class periods. The data collection consisted of two main phases. The first phase, which lasted two class periods, took place in the lab, where the participants completed a prior knowledge test, received a demonstration on how the software worked, read the text for its content using the annotations to help them understand it and completed the comprehension test. Participants were allowed to take notes while reading the text because they did not have access to the text during the comprehension test. Moreover, the participants were not told that the questions did not concern the information in the annotations to avoid the possibility of opting to read the text only instead of using the annotations. Two log files were saved on the hard disk for each participant upon exiting the program: one showing their annotation use, and the other showing their answers to the comprehension test questions. Finally, some of the participants completed the background questionnaire the same day, and others the next day. During the second phase of the study 20 volunteering participants were interviewed within three days after the experiment.

Data Analysis

This study was exploratory in nature and focused on the behavior of a certain group of learners in a given environment with an unobtrusive data collection tool. The data provided by the log files and used in the analyses were: (a) the amount of time spent on textual annotations and (b) the amount of time spent on contextual annotations. Other variables of interest were participants' prior knowledge of the subject and their reading comprehension scores. A multiple regression analysis with reading comprehension as the criterion and the amount of time spent on annotations and prior knowledge as the predictors served to reveal whether a relationship existed between the independent and dependent variables. Qualitative data from questionnaires and interviews permitted investigation of the question in more detail.

RESULTS

An independent samples *t-test* indicated the advanced group ($\underline{M}$=11.72, SD=4.55) performed significantly better on the reading comprehension test than did the intermediate group ($\underline{M}$=9.16, SD=4.21), $\underline{t}$(82)=-2.5, $\underline{p}$=.01. To determine whether there was a relationship between annotation use and the learners' performance, correlations between dependent and independent variables were investigated. Table 2 shows the intercorrelations among the variables for the intermediate group.

Table 2: <u>Intercorrelations among the Variables for the Intermediate Group</u>

	2	3	4	5	6	7	8	9
1. Reading Comprehension	0.53**	-0.15	-0.13	-0.07	-0.38*	-0.29	0.08	-0.43**
2.Prior Knowledge	--	0.11	-0.01	-0.13	-0.12	-0.16	0.04	-0.25
3. Textual Graphics		--	0.60***	0.32	0.50**	0.19	0.06	0.18
4. Textual Audio			--	0.31	0.36*	0.30	0.05	0.01
5. Textual Text				--	-0.16	-0.26	-0.06	-0.21
6. Contextual Graphics					--	0.69***	0.38*	0.60***
7.Contextual Audio						--	0.53***	0.47**
8.Contextual Text							--	0.51**
9.Contextual Video								--

*p<.05, **p<.01, ***p<.0001

For the intermediate group, there were three variables that were moderately or substantially associated with the dependent variable. A significant positive relationship between reading comprehension and prior knowledge ($\underline{r}$=.53) emerged, which suggested that the higher the prior knowledge, the higher the reading comprehension. There was a significant

negative relationship between reading comprehension and the time spent on
video ($\underline{r}$=-.43) and graphics ($\underline{r}$=-.38) annotations, which may suggest that
video-based aids distract attention and interfere with comprehension. The
correlations between reading comprehension and the time spent on the other
types of annotations were either low or negligible.

Table 3: <u>Intercorrelations among the Variables for the Advanced Group</u>

	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1. Reading Compre-hension	0.29*	-0.22	-0.21	-0.08	0.05	0.04	0.26	0.09
2. Prior Knowledge	--	-0.13	0.03	-0.20	-0.11	0.03	0.05	-0.02
3. Textual Graphics		--	0.25	0.58***	0.12	-0.04	-0.14	-0.05
4. Textual Audio			--	0.37*	0.30*	0.25	0.10	0.12
5. Textual Text				--	0.14	-0.04	0.03	-0.12
6. Contextual Graphics					--	61***	.52***	0.67***
7. Contextual Audio						--	.59***	0.63***
8. Contextual Text							--	0.45**
9. Contextual Video								--

*p<.05, **p<.01, ***p<.0001

For the advanced group, correlations of the independent variables with
the dependent variable were negligible except for prior knowledge. There was a
significant positive relationship between reading comprehension and prior
knowledge for this group ($\underline{r}$=.29), yet this is a weak relationship.

In light of the correlations between the independent variables and
reading comprehension, a hierarchical multiple regression analysis was
conducted for the intermediate group. The decision regarding which variables
would enter the equation was made after examining the relationships among
the variables and reading comprehension. The variables with low and
negligible relationships did not enter the multiple regression analysis. Thus,
the variables that entered the equation for the intermediate group were prior
knowledge, the time spent on contextual video annotations and the time spent
on contextual graphics annotations. A check of the multicollinearity for
predictor variables revealed that multicollinearity was not a problem for this
analysis. Table 4 shows the results of the regression analysis.

At the first step of the analysis, prior knowledge entered the equation
and accounted for 28 percent of the variability (<u>Adj. R^2</u>=0.26) in reading
comprehension, $\underline{F}$(1, 33)= 12.50, $\underline{p}$<.001. At the second step in the regression
analysis, time spent on the contextual video annotations entered the model,
adding an incremental R^2 change of 10 percent to the model, $\underline{F}$(2, 33)=4.32,
$\underline{p}$<.001. At the third step, time spent on contextual graphics annotations
entered the model and added another two percent to the R^2. However, the
unique contribution of this variable to the model was not significant.

Moreover, the contribution of the time spent on video annotations became insignificant when this variable entered the model, probably due to the correlation between these two variables (r=.60). Therefore, this variable was taken out of the model and it was concluded that the variables that significantly explained the variability in the reading comprehension of the intermediate learners were prior knowledge and the time spent on the contextual video annotations. For the advanced group, the only variable that was significantly related to reading comprehension was prior knowledge (r=.29) and the coefficient of determination indicated that this variable accounted for 9 percent of the variability in reading comprehension.

Table 4: <u>Beta Weights Obtained in Multiple Regression Analysis Explaining the Variability in Reading Comprehension for the Intermediate Group</u>

Independent Variable	**Step 1** Beta	T	**Step 2** Beta	T	**Step 3** Beta	T
Prior knowledge	1.02	3.54***				
Contextual Video Annotations	0.87	3.07**	-4.36	-2.19*		
Contextual Graphics Annotations	0.88	3.14**	-2.61	-1.06	-2.38	-1.18

<u>Note.</u> N= 34
<u>Note.</u> Beta weights are unstandardized multiple regression coefficients.

Results of the regression analysis revealed different pictures for the intermediate and advanced learners. First, prior knowledge had a more important role for the intermediate learners than the advanced learners. Second, annotation use explained some of the variability in the reading comprehension for the intermediate group, whereas it did not explain any variability in the advanced learners' reading comprehension. Thus, the intermediate learners relied on prior knowledge and contextual video annotations to compensate for their lower level of proficiency. However, there was a negative relationship between contextual video annotations and reading comprehension, which may suggest that video annotations hindered comprehension instead of enhancing it. On the other hand, no relationship was found between the advanced learners' annotation use and reading comprehension. This suggests that the advanced group did not rely on the annotations to understand the text.

Although the quantitative analyses did not indicate a significant contribution of annotation use to reading comprehension, the data obtained from the questionnaire and interviews revealed that the participants perceived the annotations to be useful and reading in a hypermedia environment enjoyable, interesting and easier to understand.

In terms of the usefulness of specific types of annotations for reading comprehension, the groups showed agreement on the usefulness of most types of annotations. They considered definitions of words highly useful because they increased the "speed of reading" without "getting bored". However, the groups did not show an agreement on the usefulness of audio annotations, which were rated "not at all useful" for comprehension by the advanced group. The interviews revealed that the intermediate group perceived the pronunciations of words and audio recordings as an opportunity to improve their pronunciation. They, too, did not view these annotations contributing to reading comprehension.

As for the contextual annotations, the responses were mixed. Some participants did not consider these annotations useful because they thought they provided too much information or because there was not enough time to use them.

DISCUSSION

This study did not find a clear-cut relationship between annotation use and reading comprehension even though the participants considered annotations useful for text comprehension.

Factors related to reading comprehension were different for the intermediate and advanced groups. The intermediate learners relied on top-down processes (prior knowledge and contextual annotations) to compensate for their lower level of proficiency. This may provide evidence for the "short-circuit hypothesis" (Clarke, 1988), which suggests that limited proficiency may lead to the use of poor reading strategies. In this case, limited proficiency might have led to context-biasedness (Carrell, 1988) and reliance on contextual annotations might have hindered comprehension instead of enhancing it, since there was a negative correlation between access to the contextual annotations and reading comprehension. On the other hand, relying on prior knowledge enhanced reading comprehension for this group.

Advanced learners seemed to have spent more time on understanding the text and consulting the annotations when necessary. It should be noted that prior knowledge explained less variability in reading comprehension for the advanced group than for the intermediate group while annotation use was not related to reading comprehension at all for this group. Similarly, Knight (1994) did not find any effect of access to glosses on the reading comprehension of high verbal ability subjects. Thus, there must be other factors related to the advanced learners' text comprehension. These factors have not yet been identified. There are some candidates, however. For instance, it is possible that high linguistic competence might have enabled advanced learners to utilize good reading strategies for successful comprehension in a hypermedia environment. This may provide evidence for the theory that good reading skills are activated at a certain threshold level of linguistic ability (Anderson, 1991; Clarke, 1988; Devine, 1988; Taillefer, 1996).

Although the participants considered annotations useful for reading comprehension, the contribution of annotation use to reading comprehension could not be detected through the reading comprehension test used in this study. A more open-ended test, such as a recall protocol, may have revealed more about the quality of learning from annotations.

Even though learning from annotations might have provided a richer knowledge of the topic, the availability of annotations might have hindered comprehension as well. Aust et al. (1993) found that when annotations were readily accessible, readers used them excessively to look up even familiar information. The interviews in this study provided some evidence for Aust et al.'s finding; some participants indicated that they tended to use almost all of the annotations because they were easily accessible.

In summary, this study found no relationship between annotation use and reading comprehension for the advanced group. Advanced learners must have relied on their cognitive resources for text comprehension. The study also suggested that annotation use might hinder reading comprehension for the intermediate group. This group might have had difficulty in comprehending the information in the annotations as well.

Prior knowledge contributed more to reading comprehension than annotation use for both the intermediate and advanced learners. Given the nature of the reading comprehension test, it is not surprising that there was a moderate relationship between prior knowledge and reading comprehension. In other words, the test tapped the knowledge that was in the text, not in the annotations. Even though the role of the contextual annotations was to provide schematic knowledge about the topic, it was probably difficult for the learners to integrate the information obtained from the contextual annotations with the information embedded in the text in such a short time. Thus, prior knowledge, which was already acquired and resided in the long-term memory led to better reading comprehension. The interviews also revealed that prior knowledge apparently had an impact on the way readers interacted with the text. Participants with prior knowledge utilized the contextual annotations based on their interest or curiosity or to fill in the gaps in their already existing schemata and to get a more complete understanding of the topic. On the other hand, participants with no prior knowledge used the annotations to achieve a better understanding of the concepts.

Pedagogical Implications

This study has shown that L2 learners interact with texts differently according to their proficiency level and prior knowledge. It has also shown that video annotations had a negative impact on reading comprehension for intermediate learners, which suggests that video may be distracting. Practitioners should be cautious about the use of video annotations especially with lower proficiency learners.

One may be tempted to suggest not using hypermedia annotations at all to aid reading comprehension based on these findings. However, the interviews revealed that learners considered annotations highly useful for

reading comprehension. How can we, then, explain the lack of quantitative evidence for the usefulness of annotations? One plausible explanation is that the participants in this study did not have much experience with reading hypermedia texts. Reading in a hypermedia environment does not only require effective use of top-down and bottom-up reading skills, but also other skills such as establishing the relations between graphics and text (Bolter 1998; Hedley et al., 1994), constructing and deconstructing visual and verbal images (Beavis, 1998), knowing when to read a definition or an explanation (Venezky, 1994), skimming, note taking, analyzing charts and diagrams (Venezky, 1994), reading and interpreting any one text in the presence of others (Lemke, 1998), distinguishing relevant and reliable information (Landow, 1992; Synder, 1998), navigating through complex and continually changing systems of information (Leu, 1999) and understanding how links function (Bolter, 1998). All this implies that these skills should be taught as a part of the reading curriculum when hypermedia texts are used for instruction.

Provided that learners acquire such skills, reading in a hypermedia environment may promote learner autonomy (Healey, 1999) with its provision of individualized learning (Soo, 1999). The reader can determine the pace and path of reading. Thus, hypermedia reading can accommodate those learners who cannot keep up with the learning pace in a classroom environment.

Limitations of the Study

This study poses several limitations that suggest a need for caution concerning the results obtained. First, the use of a reading comprehension test and imposing a time constraint might have affected the participants' use of annotations. As indicated by some of the participants, had there not been a test at the end of the reading and had there been more time to read, they might have interacted with the text differently, and this would have better simulated a real-life task.

Second, the traditional nature of the reading comprehension test did not match the nontraditional nature of the text. The participants were tested linearly and verbally, which may not have detected the quality of learning that occurred from hypermedia reading. For instance, a goal-oriented use of the hypermedia text might have provided different results in annotation use and learning. Moreover, had the responses to the questions been made in the participants' own language, the quality of the responses might have been different from their responses in English, especially for intermediate learners. A different test such as a recall protocol, written in the participants' own language, might have revealed different results and might have been more reflective of the learning achieved from interacting with the text.

Third, there may be other factors that are closely related to learners' interaction with a hypermedia text, such as reading goals, learner styles, reading strategies, experience with computers and reader's interest in the topic. These factors were not investigated in this study and they may be more related to performance and annotation use.

Finally, this study does not permit to form any conclusions about the effects of hypermedia on reading comprehension. Such conclusions would require a study that is truly experimental in nature. In other words, we would need to isolate the annotation types, randomly assign participants into groups and control all the other variables that would be related to their performance.

REFERENCES

Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *Modern Language Journal, 75*(4), 460-472.

Aust, R., Kelley, M. J., & Roby, W. (1993). The use of hyper-reference and conventional dictionaries. *Educational Technology Research and Development, 41*(4), 63-73.

Beavis, C. (1998). Computer games, culture and curriculum. In I. Synder (Ed.), *Page to screen: Taking literacy into the electronic era* (pp. 234-255). London: Routledge.

Bernhardt, E. B. (1991). *Reading development in a second language: Theoretical, empirical & classroom perspectives*. Norwood, NJ: Ablex.

Blake, R. J. (1992). Second language reading on the computer. *ADFL Bulletin, 24*(1), 16-22.

Bolter, J. D. (1998). Hypertext and the question of visual literacy. In D. Reinking, M. C. McKenna, L. D. Labbo, & R. D. Kieffer (Eds.), *Handbook of literacy and technology: Transformations in a post-typographic world* (pp. 3-14). Mahwah, NJ: Lawrence Erlbaum.

Carrell, P. (1988). Some causes of text-boundness and schema interference in ESL reading. In P. Carrell, J. Devine, & D. Eskey. *Interactive approaches to second language reading* (pp. 101-113). Cambridge, UK: Cambridge University Press.

Clarke, M. A. (1988). The short circuit hypothesis of ESL reading – or when language competence interferes with reading performance. In P. L. Carrell, J. Devine, & D. E. Eskey (Eds.), *Interactive approaches to second language reading* (pp.114-124). New York: Cambridge University Press.

Collentine, J. (2000). Insights into the construction of grammatical knowledge provided by user-behavior tracking technologies. *Language Learning & Technology, 3*(2), 44-57.

Chun, D. (2001). L2 reading on the Web: Strategies for accessing information in hypermedia. *Computer Assisted Language Learning, 14*(5), 367-403.

Davis, J. J., & Lyman-Hager, M. (1997). Computer and L2 reading: Student performance, student abilities. *Foreign Language Annals, 30*(1), 58-72.

Devine, J. (1988). The relationship between general language competence and second language reading proficiency: Implications for teaching. In P. Carrell, J. Devine, & D. Eskey, (Eds.), *Interactive approaches*

to second language reading (pp. 260-277). Cambridge, UK: Cambridge University Press.

Healey, D. (1999). Theory and research: Autonomy in language learning. In J. Egbert, & E. Hanson-Smith (Eds.), *CALL environments: Research, practice, and critical Issues* (pp. 391-402). Alexandria, VA: TESOL, Inc.

Hedley, C. N., Hedley, W. E., & Baratta, A. N. (1994). Visual thinking and literacy. In N. J. Ellsworth, C. N. Hedley, & A. N. Baratta (Eds.), *Literacy: A redefinition* (pp.109-126). Hillsdale, NJ: Lawrence Erlbaum.

Hulstijn, J. H. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The effect of task and learner variables. *Modern Language Journal, 77* (2), 139-147.

Kintsch, W., &Van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.

Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *Modern Language Journal, 78*(3), 285-299.

Landow, G. P. (1992). *Hyper Text: The convergence of contemporary critical theory and technology*. Baltimore, MD: John Hopkins University Press.

Lemke, J. L. (1998). Metamedia Literacy: Transforming Meanings and Media. In D. Reinking, M. C. McKenna, L. D. Labbo, & R. D. Kieffer (Eds.), *Handbook of literacy and technology: Transformations in a post-typographic world* (pp. 283-302). Mahwah, NJ: Lawrence Erlbaum

Leu, D. J. (1999). Literacy and technology: Deictic consequences for literacy education in an Information Age [On-line]. Available: http://web.syr.edu/~djleu/Handbook.html.

Noblitt, J. S., & Bland, S K. (1991). Tracking the learner in computer-aided language learning. In F. B. Freed (Ed.), *Foreign language acquisition research and the classroom* (pp. 120-131). Lexington, MA: Heath and Company.

Pederson, K. M. (1986). An Experiment in Computer-Assisted Second Language Reading. *Modern Language Journal, 70*(1), 36-41

Roby, W. B. (1999). What's in a gloss? *Language Learning and Technology, 2*(2), 94-101.

Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI* (pp.573-603). Hillsdale, NJ: Lawrence Erlbaum.

Soo, K. (1999). Theory and research: Learning styles, motivation, and the CALL classroom. In J. Egbert, & E. Hanson-Smith (Eds.), *CALL environments: Research, practice, and critical issues* (pp. 289-301). Alexandria, VA: TESOL, Inc.

Synder, I. (1998). Beyond the hyper: Reassessing hypertext. In I. Synder (Ed*.), Page to screen: Taking literacy into the electronic era* (pp. 125-143). London: Routledge.

Taillefer, G. F. (1996). L2 reading ability: Further insight into the short-circuit hypothesis. *Modern Language Journal, 80*(4), 461-477.

Venezky, R. (1994). Literacy and the textbook of the future. In N. J. Ellsworth, C. N. Hedley, & A. N. Baratta (Eds.), *Literacy: A redefinition* (p. 39-54). Hillsdale, NJ: Lawrence Erlbaum.

NOTES

(1) The terms "glosses" and "annotations" are used interchangeably in this article.

APPENDIX I

Figure 1: A Screen Shot Showing the Layout of the Screen

The highlighted word shows that a textual annotation is selected. The format of the annotation (text, sound or graphics) appears above the word. The definition appears in a box below the navigation map.

Navigation map

Textual annotation

n. a release of mechanical, chemical, or nuclear energy in a sudden and often violent manner with the generation of high temperature and usually with the release of gases

The Big Bang
Many scientists of the early 20th century, including Albert Einstein, viewed the universe as static and eternal. However, in 1927 Edwin Hubble showed that the entire universe was expanding. The Big Bang Theory is a broadly accepted way for the origin and evolution of our universe. According to the Big Bang Theory, the universe began to expand after a super powerful explosion of concentrated matter and energy ten to twenty billion years ago. Since then, the universe has continued to expand, gradually increasing the distance between our galaxy and other galaxies. Gravity slows the expansion of the universe. If the universe is dense enough, the expansion of the universe will eventually reverse and the universe will collapse. If the density is not high enough, then the expansion will continue forever. Thus, the density of the universe will determine its ultimate fate.

text
Sound
Graphics

Big Bang Theory

Stephen Hawking

Graphic and video annotations appear in this space.

Textual annotations are available when word is colored blue.

Contextual annotations