EFFECTIVENESS OF HYPERMEDIA ANNOTATIONS FOR FOREIGN LANGUAGE READING

Asim Sakar

ABSTRACT

This study first explores intermediate-level EFL learners' preferences for hypermedia annotations while they are engaged in reading a hypermedia text for general comprehension. Second, it examines whether there is a relationship between reading comprehension and the amount of annotation use. Finally, it investigates the effects of specific types of annotations on reading comprehension. The participants were 44 EFL adult learners studying English for Academic Purposes. Data were collected through a tracking tool, a reading comprehension test, a questionnaire, and interviews. Results indicate that learners preferred visual annotations significantly more than textual and audio annotations. On the other hand, a negative relationship was found between annotation use and reading comprehension. Especially, pronunciations, pictures, and videos were found to affect reading comprehension negatively. However, the qualitative data revealed that the participants had positive attitudes towards annotations and hypermedia reading in general.

INTRODUCTION

Hypermedia refers to computer-based applications that provide information in a nonlinear way through multiple types of resources such as text, graphics, sound, video, and animation (Kommers, et. al., 1996). It is suggested that hypermedia is potentially useful for second language (L2) reading because a given text can be made comprehensible for L2 readers by annotating it with multiple types of media in the form of glossaries (Davis, 1989; Martinez-Lage, 1997; Roby, 1999). In order to help L2 learners cope with an authentic text, glossaries have been considered more effective than simplifying the text (Widdowson, 1984). Thus, glossaries may be utilized to facilitate reading comprehension by providing information both at the word or sentence level and at the topic level.

While annotating L2 texts may be useful in helping learners cope with an authentic text, it also makes text processing more complicated. Current models of L2 reading that emphasize interaction of bottom-up and top-down processes for successful reading comprehension (Bernhardt, 1991) fall short in explaining text processing in a hypermedia

environment. In this case, the reader is not only engaged in processing the text, but also the verbal and visual input which are provided through the annotations.

Mayer (2001) proposes a cognitive theory of multimedia learning to explain how learning from verbal and visual input takes place. The theory is based on three main assumptions. The *dual channels assumption* is derived from Dual Coding Theory (Paivio, 1986) and suggests that visual and verbal information is processed in *separate channels*, however, one type of information may be transferred to the other through the interaction between the channels. The *limited capacity assumption* derived from Cognitive Load Theory (Sweller, 1994) suggests that each channel has a limited capacity. Finally, *active processing assumption* suggests that humans are actively involved in the construction of knowledge using cognitive processes such as selecting relevant information, organizing and integrating it with existing knowledge. Thus, when verbal and visual input is provided through a multimedia presentation, the learner selects relevant words and images and organizes them separately into verbal and pictorial models where connections among selected words and images are established to form a coherent mental structure. These two models are qualitatively different in that pictures provide holistic and nonlinear information while words provide discrete information in a linear way. Thus, the two models are supposed to complement each other. Learning takes place when the learner establishes connections between the corresponding portions of the verbal and visual model with the help of prior knowledge, hence integrating these two models in the working memory. Mayer provides empirical evidence supporting his theory and contends that learning in multimedia environments is facilitated when the information is presented through the verbal and visual channels in a way which doesn't overload the working memory such as presenting information by accompanying words and with pictures instead of only in words, placing words and pictures near rather than far from each other, presenting them simultaneously instead of successively, and so forth.

While evidence presented by Mayer is based on studies conducted with native speakers of English, few studies have investigated the effectiveness of multimodal information in L2 learning. The research below provides a review of the studies investigating the impact of multimodal electronic glosses on L2 reading comprehension.

Literature Review

Early studies on the effectiveness of electronic annotations compared hyper-dictionaries with conventional paper dictionaries (Roby, 1991 cited in Roby, 1999; Aust et al., 1993). These studies revealed that students tend to consult the hyper-dictionary more frequently than traditional dictionaries, while taking less time to read the text. However, neither study found significant differences between the groups in terms of reading comprehension. Such comparisons between traditional and electronic dictionaries may not be meaningful anymore since reading on the computer is becoming widespread. Therefore, we need to investigate the effectiveness of hypermedia dictionaries, addressing issues such as for whom and under what conditions they are effective. Later studies were confined to hypermedia dictionaries and provided insightful information by using tracking technology that recorded participants' interactions with a given text.

Several studies investigated learners' preferences regarding the types of annotations (Davis and Lyman-Hager, 1997; Lomicka, 1998; Chun, 2001, Ercetin, 2003). Davis and Lyman-Hager found that the forty-two intermediate level French learners had a strong preference for annotations providing L1 translations of words. This finding was confirmed by Lomicka's small-scale study with twelve French learners who were also provided with other types of glosses such as images, references, L2 definitions, and questions. Chun and Ercetin investigated learners' annotation preferences according to proficiency levels. Chun found that low verbal ability learners used both internal and external dictionaries more than high verbal ability learners. This finding was confirmed by Ercetin, who found that intermediate

proficiency students accessed annotations significantly more than advanced students. She also found that both groups accessed word definitions and video annotations which provided further information about the topic more frequently than the other types of annotations such as pronunciations of words, audio recordings and graphics providing extra information about the topic.

As for the effectiveness of annotations, Davis and Lyman-Hager (1997) did not find any meaningful relationship between annotation use and reading comprehension although students reported positive attitudes toward the annotations. On the other hand, Lomicka (1998) found that full glossing (L1 translation and L2 definitions plus pronunciations, images, references, and questions) was more effective than limited glossing (L1 translation and L2 definitions) or no glossing, in facilitating L2 reading comprehension. However, Lomicka's findings should be taken with caution due to its small subject size. Knight (1994) investigated the extent to which low and high-verbal ability learners benefited from annotations by randomly assigning 112 Intermediate-level students of Spanish to dictionary access and no dictionary access conditions. Knight's study suggested that using a dictionary was more beneficial for low verbal ability students because the correlation between frequency of word lookup and reading comprehension was much higher for this group (r=.68) than the high verbal ability group (r=.17). This finding was also confirmed by Chun (2001), who found no difference in the number of propositions recalled between the high and low ability groups, and thus concluded that the latter group benefited more from an external and internal bilingual dictionary.

Studies by Chun and Plass (1996), and Ariew and Ercetin (2003) investigated the effectiveness of particular types of annotations on reading comprehension. Chun and Plass in a series of three studies conducted with 160 university students of German, found that presenting words with both visual and verbal annotations facilitated reading comprehension

more than words with no annotations or with verbal annotations only. On the other hand, Ariew and Ercetin found that video and graphics annotations had a negative impact on reading comprehension for intermediate learners of English while no relationship was found between annotation use and reading comprehension for advanced learners.

To sum up, the studies which are available to us have revealed insufficient and inconclusive results about what types of hypermedia annotations learners prefer to use and whether hypermedia annotations facilitate reading comprehension. Thus, this study aims to address these issues. More specifically, the research questions are:

a) What are the annotation preferences of intermediate level EFL learners when they are engaged in reading a hypermedia text?
b) Is there a relationship between the overall amount of annotation use and reading comprehension for intermediate level EFL learners?
c) In what way do particular types of annotations influence reading comprehension for intermediate level EFL learners?

It was hypothesized that learners would prefer videos and pictures to access more information about the topic (Ercetin, 2003), while verbal annotations providing word definitions would be preferred to comprehend the meanings of unknown words (Davis and
Lyman Hager, 1997; Lomicka, 1998). However, no direct relationship was expected between overall amount of annotation use and reading comprehension (Davis and Lyman Hager, 1997). Finally, visual annotations were expected to distract learners' attention and, therefore, hinder reading comprehension as found by Ariew & Ercetin (2003).

METHODOLOGY

Participants

A total of 125 Turkish students studying English for Academic Purposes at a Turkish university were given the Oxford Placement Test (Allan, 1992) to determine their proficiency levels in English. This placement test comprises two main sections, a listening test and grammar test, with 100 items in each section. The test provides percentage scores out of 200. For the purposes of this study, the score range between 135-145 was considered as intermediate level, which is roughly equal to level 5 in the IELTS, FCE and CCSE, level 2 in Cambridge exams, AP according to ARELS exams, and OP according to Oxford exams. The data for the 44 intermediate level participants were retained for analysis. 26 of the participants were male and 18 were female. The average age was 20, ranging between 18 and 25. The participants all had considerable experience with computers and were familiar with language learning CD-ROMs, word processing software, Internet and email.

Materials

The data were collected through a hypermedia reading text, a reading comprehension test, a questionnaire and interviews.

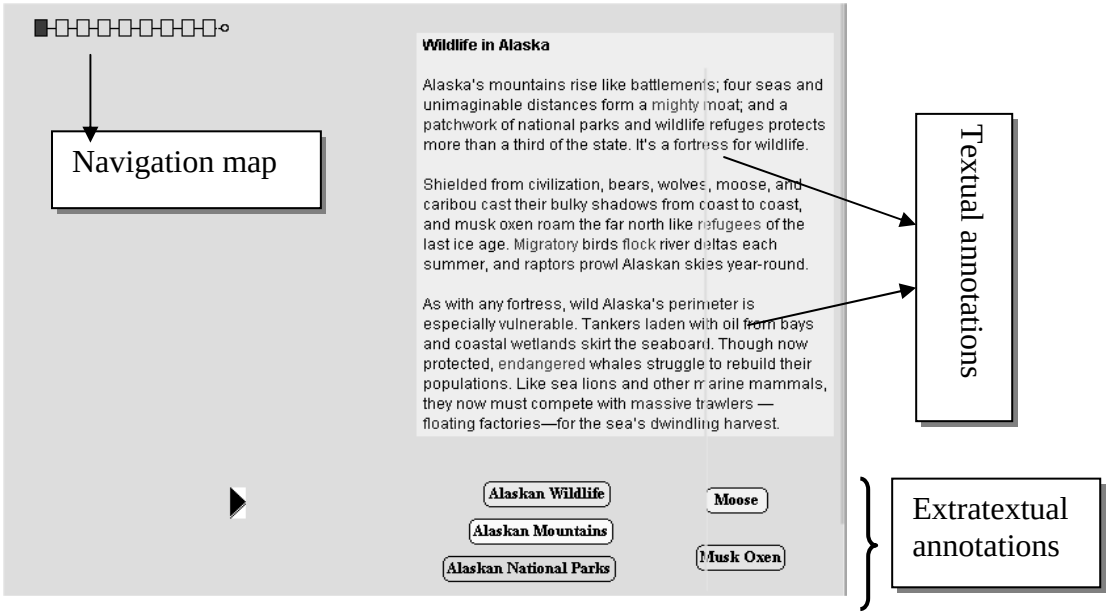
Hypermedia Reading Text

An authentic online reading text was selected from The National Geographic Journal website (http://www.nationalgeographic.com/greatland/) and was annotated with multiple types of media such as text, graphics, audio, and video by using a software designed by Ariew (1999). The text consisted of 900 words and had a linear organization, where the information was presented in 9 consecutive pages successively. A navigation map was provided for the readers showing their location in the document so that they would not get lost (see Figure 1). The annotations were conceptually categorized as (a) *textual annotations*: those providing

information about the text in the form of an internal gloss. The information consisted of a definition of a word, its pronunciation, and sometimes a picture to help understand the meaning of a given word (see Figure 2), (b*) extratextual annotations*: those providing background information about the topic (see Figure 3).

Figure 1

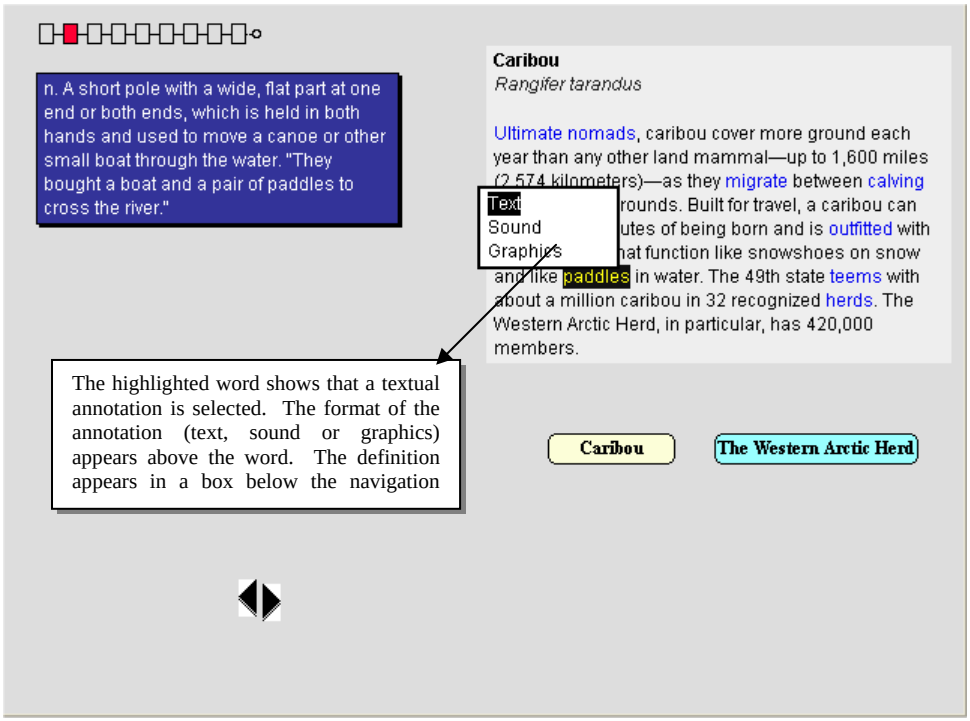
A Screen Shot Showing Textual and Extratextual Annotations



A total of 104 textual annotations were provided within the body of the text: An annotated word or phrase was a part of the text and was indicated by its blue color.

Figure 2

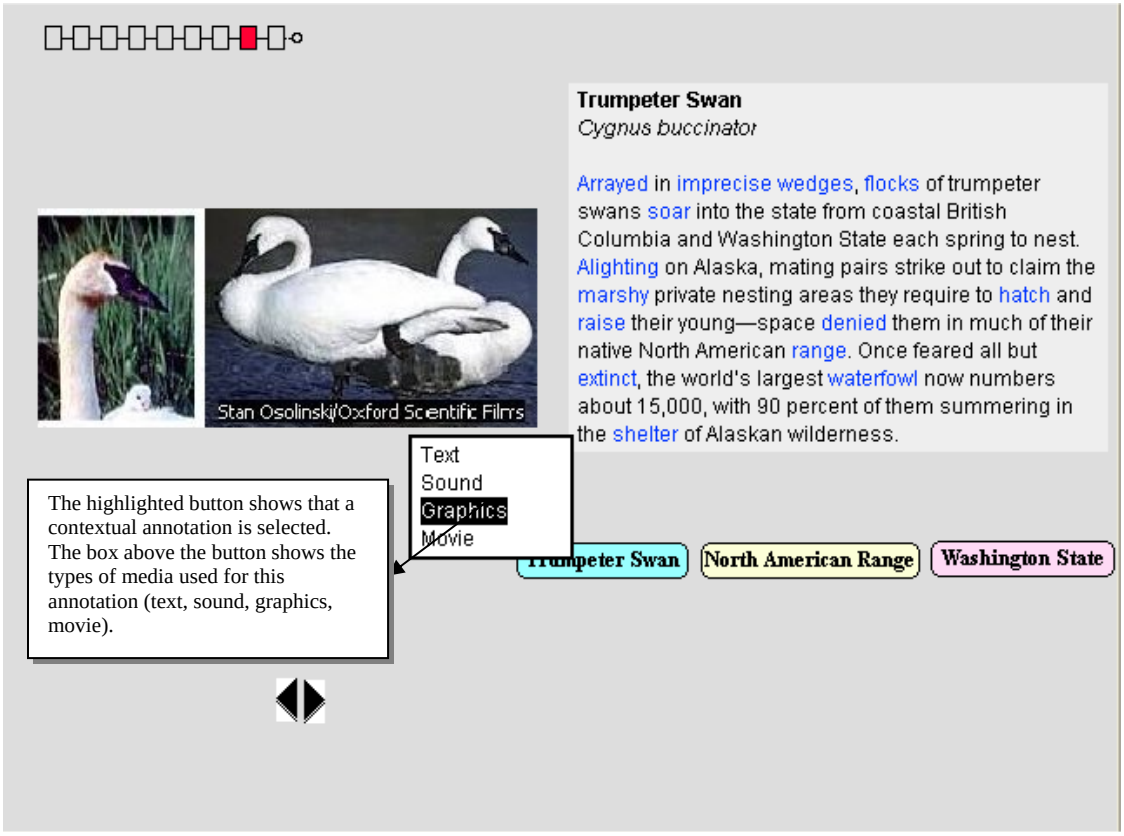
A Screen Shot Showing a Textual Annotation



A total of 26 extratextual annotations were provided below the text and were indicated by colored buttons placed on the page.

Figure 3

A Screen Shot Showing an Extratextual Annotation



When a participant clicked on a highlighted word, phrase, or a background information button, they could see in what forms of media the information was available (i.e. text, graphics, sound, or video). To ensure that the participants actually viewed the annotations, they were required to hold the left mouse button down after they clicked on a particular annotation. So, the amount of time the participants engaged the mouse was considered to be the time they viewed a given annotation.

Participants were able to choose and view as many annotations as provided. Thus, they were allowed to view the same annotation more than once. The software tracked every interaction of the reader with the text, including which annotations the reader chose to view,

how much time (in seconds) the reader spent on a particular annotation, the order in which the annotations were selected, and the total amount of time spent on reading the text. The data were saved as a log file. Since the tracking tool was hidden, the collection of the data regarding the participants' interaction with the text did not hinder the flow of reading.

Reading Task

There were two separate reading tasks, thus two reading sessions. During the first reading, the subjects were asked to read the text for general comprehension and use the annotations to help them understand the text. They were told they would be asked to write everything they remembered without looking at the text. This reading session including the recall test took approximately 45 minutes. During the second reading the subjects were given a reading comprehension test, of which they were not informed beforehand. This time they were allowed to access the text to answer the questions. The first reading was necessary to make sure that the participants made use of the annotations. It was thought that if they had been given the comprehension test during the first reading, learners might have used the annotations in a limited way.

Reading Comprehension Test

Although the participants completed a recall task after the first reading, their recall protocols were not used to measure their reading comprehension since this was a totally new task to them. As indicated above, the goal of the first reading was to ensure the use of glosses. Reading comprehension was measured with a test developed by the researchers. The reading comprehension test contained 20 questions consisting of 15 multiple-choice questions and 5 open-ended questions, with maximum possible score of 20. As the goal was to examine whether annotations facilitated comprehension of the text, the questions on the test assessed the important information provided in the text, not in the annotations. The subjects were not informed that the questions would involve information in the main text only. The questions

were based on the main ideas and details in the text as identified by a native speaking teacher of English. The multiple choice questions mostly elicited specific factual information, whereas the open-ended questions required analysis and synthesis of the information in the text. Both multiple-choice and open-ended questions had the same weight as 1 point for each correct answer. Scoring open-ended questions was done entirely based on the content of the answer; students were not penalized for errors of grammar, spelling, or punctuation. The test was given on paper and the reading text was available on the computer to the students during the test.

Questionnaire and Interviews

A questionnaire was given after the experiment, with the purpose of obtaining information about the participants' experience with hypermedia reading and their perceptions of the usefulness of the annotations.

The interviews, which took place the week after the treatment were conducted on with 12 volunteering participants and in their first language. The purpose of the interviews was to collect supplementary data about participants' use of annotations while reading.

Procedures

The data collection was completed in two phases. The first phase took place in the computer lab and lasted two hours. After a ten-minute demonstration on how to operate the software, the participants were asked to read the text for general comprehension and completed a recall task. Next they were given a reading comprehension test to complete during which they were allowed to have access to the text. For each participant, two log files were saved on the hard disk: one for the participants' interaction with the text during the first reading, the other for their second interaction while answering comprehension questions. Immediately after they finished the comprehension test, they started filling in the questionnaire, which was attached to the test. The second phase of data collection involved

semi-structured interviews with 12 volunteer participants within one week after the experiment.

RESULTS

Learners' preferences regarding types of annotations

In order to find out what types of annotations the participants preferred, the frequency of access to annotations was determined by taking the ratio of the total number of times the participants accessed a given annotation to the total number of times it occurred in the software. Since there were two reading sessions, the frequency of access to annotations in the two readings were combined. Table 1 shows frequency of access to annotations in the first and second reading, the average of two readings, and the total number of time annotations occurred in the text.

Table 1

Frequency of Access to Textual & Contextual Annotations

	Textual Annotations			Contextual Annotations			
	Text	Audio	Graphics	Text	Audio	Graphics	Video
Reading I	45	30	38	10	10	25	28
Reading II	6	1	4	1	1	3	3
Average	50.91	30.55	42.52	10.56	11.23	28.43	31.02
Number of Annotations	102	102	44	25	25	25	12

As the table indicates participants tended to use the annotations during the first reading that involved general comprehension. The reason for infrequent annotation use during the second reading might be either because the subjects had already seen the annotations or because they felt no need to use them because the answer was found in the main text.

The average of two readings was divided by the total number of times a particular annotation occurred in the software since the frequency of occurrence for each annotation type

was different. Since the distribution was not normal, a square-root transformation was applied to the data. Table 2 provides the means and the standard deviations after the data is transformed.

Table 2

Means and Standard Deviations for the Frequency of Access to Textual & Contextual Annotations after the Square Root Transformation

	Textual Annotations			Contextual Annotations			
	Text	Audio	Graphics	Text	Audio	Graphics	Video
Mean	0.68	0.28	0.94	0.54	0.56	1.01	1.48
SD	0.19	0.18	0.28	0.37	0.37	0.36	0.63

As the table shows, the most frequently accessed annotations were video and graphic annotations which provided extra information about the topic and graphic annotations that provided the meanings of words. On the other hand, audio annotations that provided pronunciations of words and those that provided extra information about the topic were accessed the least. In order to determine whether differences among the means were significant, a one-way repeated measures ANOVA test was conducted (see Table 3). The sphericity assumption of the ANOVA test was checked with Mauchly's sphericity test. In case of the violation of the sphericity assumption, adjustments were made to the ANOVA results using the Geisser-Greenhouse epsilon, which provides an F-test using a much more stringent criterion.

Table 3

ANOVA Summary Table for the Frequency of Access to Annotations

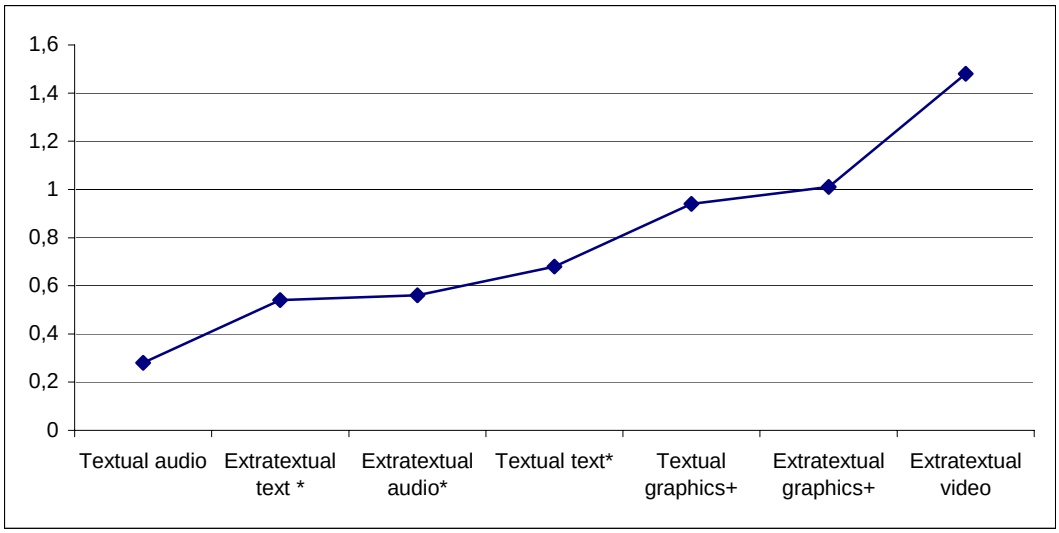
Source	**df**	**SS**	**MS**	**F**
Within Subjects	264	67.51	6.96	
Annotation type	6	41.13	6.86	67.04***
Residual	258	26.38	.10	
Total	307			

****p<.001*

This analysis revealed a significant difference in the frequency of access to different annotations, $F(6,258) = 67.04$, $p < .01$ (with the Geisser-Greenhouse correction), suggesting that certain types of annotations are preferred significantly more than others (see Figure 4).

Figure 4

Distribution of Means for the Frequency of Access to Annotations



The results of the pairwise comparisons with Tukey post-hoc tests indicated that pronunciations of words were preferred significantly less and videos were preferred significantly more than all other types of annotations. Insignificant differences are shown with the same sign in Figure 4. Thus, these results partly confirmed the first hypothesis, which stated that the participants would use graphics and video annotations to get extra information about the text. However, the hypothesis that word definitions would be used to retrieve the meanings of words was not confirmed since the participants accessed pictures more frequently than definitions.

The effect of overall annotation use on reading comprehension

Annotation use was analyzed in terms of the total number of times the participants accessed the annotations and the total amount of time they spent on the annotations. Even though there was a substantial positive relationship between the two variables ($r=.70$), it was thought that the total number of times the participants accessed annotations did not

necessarily reflect the total amount of time they spent utilizing the annotations. Thus, the participants were categorized into three groups according to z-scores on these variables: Low users (LU), moderate users (MU) and high users (HU), with roughly equal number of subjects in each group. Table 4 shows the ranges for the frequency of access to annotations based on raw frequencies and their z-score equivalents as well as the mean comprehension test scores for the groups.

Table 4

Reading Comprehension Scores Based on Frequency of Access to Annotations

Groups	Range	Range in Z scores	Number	Mean	SD
Low Users (LU)	55 – 152	-1.80 - -0.64	14	8.14	3.11
Moderate Users (MU)	161 – 247	-0.53 – 0.50	15	7.53	2.88
High Users (HU)	252 – 439	0.56 – 2.80	15	6.33	2.19

A one-way ANOVA was conducted to determine whether there was a significant difference among the groups. Results of the ANOVA test are provided in Table 5.

Table 5

ANOVA Summary Table for Reading Comprehension Scores Based on Frequency of Access to Annotations

Source	**Df**	**SS**	**MS**	**F**
Between Groups	2	24.77	12.38	1.64
Within Groups	41	308.78	7.53	
Total	43	333.55		

The analysis showed that comprehension test scores of the groups did not differ significantly $F(2,41)=1.64$, $p=0.206$, suggesting that more frequent access to annotations does not necessarily facilitate reading comprehension .

As for the relationship between the amount of time the participants spent on annotations and their comprehension test scores, the subjects were again divided into three groups based on the amount of time they spent on reading. Table 5 shows the ranges for each

group based on raw frequencies and their z-score equivalents as well as the means and SD's for the groups' reading comprehension scores.

Table 6

Reading Comprehension Scores Based on Access Time to Annotations

Groups	Range	Range in Z scores	Number	Mean	SD
Low Users (LU)	179 – 1079	-2.23 - -0.3	14	8.71	3.02
Moderate Users (MU)	1089 – 1444	-0.36 – 0.37	15	7.07	2.66
High Users (HU)	1454 – 2259	0.39 – 2.05	15	6.27	2.25

A one-way ANOVA test was conducted to determine whether there was a significant difference among the groups (see Table 7)

Table 7

ANOVA Summary Table for Reading Comprehension Scores Based on Access Time to Annotations

Source	**Df**	**SS**	**MS**	**F**
Between Groups	2	44.82	22.41	3.18
Within Groups	41	288.72	7.04	
Total	43	333.55		

The analysis suggests that comprehension test scores of the groups are somewhat different because the F test is only marginally significant $F(2,41)=3.18$, $p=0.052$. Since the F test result was larger than that of the frequency of access, the difference between the groups was considered worthy of further investigation through post-hoc comparisons using a Tukey procedure. The difference in the mean comprehension scores for LU and HU approached significance ($p=.045$), suggesting that spending more time on annotations may not facilitate reading comprehension.

These results partly confirmed the second hypothesis that there would not be a direct correlation between annotation use and reading comprehension. While this hypothesis was

confirmed in terms of frequency of access to annotations, it was not confirmed with regard to the amount of time spent on annotations.

The effect of particular annotations on reading comprehension

The findings above suggest little or no relationship between overall annotation use and reading comprehension. However, an examination of the correlations between specific types of annotations and reading comprehension revealed that the use of certain annotations did have some relationship to reading comprehension which was statistically significant. Therefore the relationship between reading comprehension and specific types of annotations were also examined. Table 9 provides the correlations between annotation types and reading comprehension.

Table 9

Intercorrelations among the Variables for Frequency of Access to Annotations

	2	**3**	**4**	**5**	**6**	**7**	**8**
1. Reading Comprehension	-0,10	-,24	-0,39**	-0,04	-0,18	-0,39**	-0,37*
2. Textual Text	--	0,49**	0,08	0,06	-0,14	0,08	-0,05
3. Textual Graphics		--	0,12	-0,09	0,08	0,12	0,00
4. Textual Audio			--	0,46**	0,39**	1,00***	0,48**
5. Extratextual Text				--	0,33*	0,46**	0,37*
6. Extratextual Graphics					--	0,39**	0,54**
7. Extratextual Audio						--	0,48**
8. Extratextual Video							--

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

Significant correlations were found between reading comprehension and pronunciations of words (r=-.39, p<.01), audio recordings (r=-.39, p<.01), and videos (r=-.37, p<.05) providing extra information about the topic. However, these correlations were negative. Other variables had low or negligible effect on reading comprehension.

A multiple regression analysis was conducted on reading comprehension scores as the dependent variable with sound and video annotations as the independent variables. Because there was a high correlation between textual audio and extratextual audio annotation use, the

former was not included in the analysis to avoid multicolliniarity among the independent variables. Results indicated that the model was significant ($F_{2, 41}$=5.00, p<.01). The two variables together explained 20% of the variability in reading comprehension ($\underline{R}^2$=.20). However, the negative relationship between these variables and reading comprehension may suggest that these annotations hindered reading comprehension instead of enhancing it. Thus, these findings partially confirmed the third hypothesis that videos would hinder reading comprehension.

To summarize the quantitative findings, learners preferred visual annotations to help them understand the text. However, these annotations did not necessarily facilitate reading comprehension. On the contrary, certain types of annotations, i.e., audio and video annotations, had a negative impact on reading comprehension. Moreover, the more the time the participants spent on annotations, the worse they performed on the reading comprehension test.

Findings from the Questionnaire and Interviews

Qualitative data were obtained from the questionnaire and interviews to cross-analyze the quantitative findings. The participants were asked to rate the usefulness of annotations for reading comprehension in the questionnaire (see Table 8).

Table 8

Usefulness of Annotations Rated by the Participants

		Not at all useful	Slightly useful	Useful	Very useful	Essential
Annotations that provided information about words	Text (definition)	0%	18%	41%	26%	15%
	Audio (pronunciation)	26%	15%	26%	22%	11%
	Graphics	0%	0%	18%	37%	45%
Annotations that provided extra information about the topic	Text	0%	22%	41%	30%	7%
	Audio	7%	30%	30%	18%	15%
	Graphics	0%	0%	18%	41%	41%
	Video	0%	7%	11%	22%	60%

There seems to be a consensus among the participants regarding the usefulness of video and graphics annotations since the majority of them rated these annotations either ‘very useful’ or ‘essential’. Although verbal annotations providing word definitions and extra information about the topic were not rated as highly as video and graphics annotations, they were still considered to be ‘useful’. On the other hand, there does not seem to be a consensus on the usefulness of audio annotations. Thus, the ratings suggest that participants perceived visual annotations to be most vital for text comprehension.

Participants who were interviewed also stated that they preferred the visual annotations to retrieve background information about the topic as suggested by one participant: “Because the topic was new to me, the annotations about the topic were very helpful, especially the visual annotations.” Another participant provided a similar response: “Visuals and movies themselves told everything about the topic. How am I supposed to know about an animal I’ve never seen and heard; however, pictures and movies showed me the animal directly much better than the words.” Moreover, the participants especially emphasized that the videos ‘made the text more interesting’ as stated by a participant: “I liked and used movies a lot. I watched them to understand the topics without reading the main text. I did the comprehension test using the information from the movies.”

As for the usefulness of annotations providing word definitions, participants indicated that definitions of words allowed them find the meaning of words easily without slowing the reading pace as stated by one participant: “This is very different from traditional reading for language learners. I usually skip the words I don’t understand and never check for their meanings from a dictionary; however, with this text, I can easily check their meaning at a single click. Furthermore, this is more permanent due to double information (pictures and text together).” On the other hand, another participant stated using such annotations when he could not guess the meaning from context: “Pictures helped to understand the meanings in case I

couldn’t do that using the context.” Thus, for all of the participants who were interviewed, graphical cues for words were "necessary" or "essential. However, pronunciation of words was not considered very important for understanding the text.

One participant summarized this groups’ preferences succinctly: “I first checked the pictures, secondly the video, then the audio finally the text. I mostly used the annotations that provided information about the topic.”

To sum up, the participants’ perception of the usefulness of annotations was consistent with their actual annotation preferences because the log data also showed that they accessed the visual annotations significantly more frequently and spent significantly more time on these annotations than other types. Interviews revealed that most of the subjects considered visual annotations essential because they were interesting and motivating.

Discussion

An examination of the specific types of annotations used by the learners revealed that all learners consistently preferred visual information (i.e. graphics and/or videos). These annotations were rated as highly useful by the participants. This finding is similar to Ercetin (2003) who also found that both advanced and intermediate learners preferred visual annotations.

As for the relationship between preferences for annotations and reading performance, no relationship between the two variables were found in terms of the frequency of access to annotations, however, a negative relationship was observed when the question was investigated in terms of the amount of time spent on annotations. These findings partly confirm the findings of Davis and Lyman-Hager (1997) who also found no relationship but are contrary to the findings of Lomicka (1998), who found full glossing to be more useful than partial or no glossing.

Regarding the effects of specific types of annotations, it was found that there were low correlations between most of the annotations, and negative correlations between reading comprehension and pronunciations, pictures, and videos. These annotations might have distracted readers and interfered with reading comprehension. These findings comply with the findings of Ariew & Ercetin who found a negative impact on visual annotations for intermediate learners whereas no impact was found for advanced learners. Since the task (reading for general comprehension) and the reading environment (similar types of annotations, similar text organization) are comparable in these two studies, direct comparisons between the studies are possible. The findings of both studies provide evidence supporting Cognitive Load Theory when learners are at intermediate level. In other words, provision of additional input using a variety of presentation modes might have lead to 'redundancy' and 'split attention' effects (Mayer, 2002) and thus, hampered reading comprehension at this level.

To conclude, learners perceive easy access to annotations highly useful in helping them cope with an L2 text. However, easy access might have lead them to use the annotations excessively as emphasized by Aust et al. (1993), who found that when annotations were readily accessible, readers used them excessively to look up even familiar information. This may, in turn, have resulted in 'deleterious effects' due to 'task interference between modes of presentation,' which would occur 'when the visual and verbal processes are not automated but require executive resources' and 'directing of attention to the type of information perceived as more important or more interesting, away from the other mode which may in fact contain more important information' (Chun & Plass' 1997).

Pedagogical Implications

The results have implications both for teachers and materials developers. Intermediate level EFL learners seem to prefer visual information in order to understand a text. They have

positive attitudes toward reading on the computer because of the capability of the computer to provide such information. However, the presentation of visual information along with additional information through different channels successively may not facilitate reading comprehension. As suggested by Mayer (2002), simultaneous presentation of visual and verbal information may be more effective. Moreover, while annotations may be useful in helping learners build schemata about a given topic, learners at lower proficiency levels may not have the resources to process such additional information. Thus, it is recommended that hypermedia activities be integrated into the reading curriculum and learners be trained in effective reading strategies for such an environment such as knowing when to read a definition or an explanation (Venezky, 1994) or establishing the relations between graphics and text (Bolter 1998; Hedley et al., 1994).

Limitations of the Study

Several limitations have been identified in this study; therefore, study findings should be taken into consideration accordingly.

First, an experimental study which better controls confounding variables should be conducted for more accurate inferences about the effect of hypermedia annotations on reading comprehension. Moreover, qualitative data such as the kind obtained from think aloud protocols would provide deeper insights into the process of utilizing annotations while reading the hypermedia text. Second, inferences in this study are made based on a reading task that required both reading for general comprehension for specific information. Thus, the findings apply only to those situations where a similar task is involved. Moreover, the reading comprehension test was limited in detecting the quality of learning from annotations. A different test such as a recall protocol might have revealed different results. Third, the findings of the study cannot be generalized to other contexts and proficiency levels unless the study is replicated in the different contexts and with different proficiency levels Finally, there

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

REFERENCES

Ariew, R. (1999). Reading Toolbox (Version 2.0). [Computer Software]. Tucson, AZ.

Ariew, R. & Ercetin, G. (In Press). Exploring the potential of hypermedia annotations for second language reading. Computer-Assisted Language Learning Journal, 17(2).

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.

Bernhardt, E. B. (1991) Reading development in a second language: Theoretical, empirical and classroom perspectives. New Jersey: Ablex Publishing.

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.

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

Chun, D. M. & Plass, J. L. (1997) Research on text comprehension in multimedia environments [online]. Language Learning & Technology, 1(1), 60-81. Available: http://llt.msu.edu

Chun, D., & Plass, J. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal, 80*(2), 183-198.

Chun, D. M. & Plass, J. L. (1996b) Facilitating reading comprehension with multimedia. System, 24 (4), 503-519.

Davis, J. N. (1989) Facilitating effects of marginal glosses on foreign language reading. The Modern Language Reading, 73 (1), 41-52.

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

Ercetin, G. (2003) Exploring ESL learners' use of hypermedia reading glosses. CALICO Journal, 20(2), 261-283.

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.

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

Kommers, P. A. M., Grabinger, S., & Dunlap, J. C. (1996) Hypermedia learning environments: Instructional design and integration. New Jersey: Lawrence Erlbaum.

Lomicka, L. (1998) To gloss or not to gloss: An investigation of reading comprehension online. Language Learning and Technology, 1(2), 41-50. Available: http://polyglot.cal.msu.edu/llt/vol1num2/roby/index.html

Martinez-Lage, A. (1997) Hypermedia technology for teaching reading. In M. D. Bush, R. Terry (Eds.). Technology Enhanced Language Learning (pp.121-163). Lincolnwood, IL: National Textbook Company.

Mayer, R. E. & Moreno, R. (2002) Animation as an aid to multimedia learning. Educational Psychology Review, 14(1), 87-99.

Paivio, A. (1986) Mental Representations: A Dual Coding Approach. Oxford: Oxford University Press.

Roby, W. B. (1999) What's in a gloss? [online] Language Learning & Technology, 2 (2), 94-101. Available: http://polyglot.cal.msu.edu/llt/vol2num2/roby/index.html.

Sweller, J. (1994) Cognitive load theory, learning difficulty and instructional design. Learning and Instruction, 4, 295-312.

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

Widdowson, H. G. (1984) Teaching Language as Communication. Oxford: Oxford University Press