The Effects of Multimedia Annotation and Summary Writing on Taiwanese EFL Students' Reading Comprehension

Wen-Chi Huang
National Taiwan Normal University

ABSTRACT

The present study investigates the effects of multimedia annotation through the discourse scheme and summary writing through the grounding theory (Chang, 1997) on text comprehension. Specifically, the study focuses on examining the influences of multimedia annotation from a special perspective, namely, the use of modified discourse scheme to trigger EFL learners' active reading process, rather than a passive way of receiving different types of annotations. Summary writing, a newly applied means to test learners' reading comprehension (Trites & McGroarty, 2005), is probed to see whether there exist wash back effects in assisting learners' reading performance. One hundred Taiwanese EFL university freshmen were divided into four groups, each group receiving a different treatment within a pretest-posttest-delay post-test design. Results of ANOVA indicated that both multimedia annotation and summary writing had significant positive effects on learners' reading performance, with obvious variances in the post and delayed post-test. Moreover, the repeated measure was applied to examine the lasting effects of the treatments. In addition, qualitative interview data relating to the participants' ideas and attitudes are also discussed.

INTRODUCTION

During the past twenty years, research has focused not only on examining the effects of teaching approaches in the classroom but also on how to engage learners in reading an abundance of texts. Beginning with the theoretical considerations and didactic concerns of advanced technology, Liontas (2002) has specifically provided ideas about how computer-assisted language learning and digitized multimedia-based learning materials may be beneficial to language instruction, learners, and teachers. In fact, Liontas has emphasized the important roles that "interaction" and "integration" assume, and proposed that making use of multimedia computer technology can have a significant impact on language teaching and learning.

To this end, a considerable number of studies have already investigated the effectiveness of different types of multimedia glosses and annotations under different premises (Akbulut, 2007; Ariew & Ercetin, 2004; Bowles, 2004; Chun & Plass, 1996a, 1996b; Jones & Plass, 2002; Lomicka, 1998; Wolfe, 2002), and these studies examine the validity of applied tools for second

language acquisition (SLA), based on the support of various theoretical frameworks (Hulstijn & Laufer, 2001; Mayer, 2001, 2005; Schmidt, 2001).

Multimedia annotation is one of the applications of computer-mediated communication (CMC), which provides interpersonal communication and offers access to authentic language input (Lo, Tsao, & Yeh, 2005). According to Sakar and Ercetin (2005), annotated L2 texts may be useful in helping learners cope with an authentic text; however, it is still possible that such annotations can make the reading process more complicated. In this way, the usefulness and effectiveness of annotation will depend on the types of annotation readers use and how they make use of it.

According to Yanguas (2009), most researchers to date have mainly investigated two issues related to annotations in general: (1) whether they do in fact assist students' overall comprehension of the text, and which exact type of annotation is more beneficial; and (2) if these types of annotations might promote incidental vocabulary learning. However, few studies have been done to highlight the importance of discourse-level while reading through annotation; namely, the effects of discourse analysis and multimedia annotation. Accordingly, the first research focus of the current study would emphasize the effects of multimedia annotations through discourse-level.

The second research focus of this study centers on the concept of "Reading to Learn" (Trites & McGroarty, 2005). Reading to Learn (RL) refers to summary, a shortened version of the original text, and the main purpose of such a simplification is to highlight the major points. A written summary has a clearly arranged structure and is written in a logical, chronological, and traceable manner (Trites & McGroarty, 2005). Because summaries should be significantly shorter than the original, minor facts have to be left out; however, all major conclusions should remain. According to Trites and McGroarty, asking readers to write a summary after reading the text is a better way to assess their comprehension. Nevertheless, seldom has research been done to include summary writing as part of the instruction to ascertain whether it can indeed lead to better reading comprehension.

Generally speaking, reading is viewed as an interactive process between the reader and the writer; however, seldom has research been done to investigate the interaction between online annotation (a medium of input) and summary writing (output). The purposes of this study are therefore twofold. Firstly, the study aims to investigate whether multimedia annotation through discourse-level and summary writing influence readers' text comprehension. EFL learners' reading and writing behaviors through the online environment have not yet been fully explored (Huang, Chern, & Lin, 2009), and having the access to know these specific behaviors may provide wash back effects to classroom instruction. Secondly, the study aims to examine to what extent and in what manner multimedia annotation and summary writing affects readers' comprehension. Furthermore, this study also focuses on the potential efficiency of the combination of annotation and summary on reading comprehension. Said another way, the study also tries to probe the perceptions and attitudes of the EFL learners about the application of online annotation and summary writing to uncover the strategy use during the reading process.

LITERATURE REVIEW

Multimedia Annotations and Second Language Reading

Annotations refer to the notes or glosses that readers makes for themselves, such as what the students make when reading texts or researchers create when noting references they plan to pursue (Wolfe, 2002). In this way, many instructors have recently advocated the benefits this practice might have for both readers and writers (Brantmeier, 2003). To be more specific, not only can annotations provide a good way for instructors to discuss different reading strategies, they can also help students see their reading as it takes place within the rules and conventions of a larger discourse community (Wolfe, 2002). For example, Salvatori (1996) asked students to photocopy and distribute their annotations on literary texts as a means of discussing the arguments that different readers construct from the text. Lunsford and Ruszkiewicz (1999) published essays accompanied through the uses of annotations to help students visualize the social nature of reading and to suggest how the backgrounds and biases of different readers affect their understandings and interpretations of a text.

Paivio's (1990) Dual Coding Theory (DCT)—the theoretical framework of multimedia-based learning—addresses the idea of pictorial-verbal system on knowledge construction. The DCT involves the activity of two distinct subsystems: a verbal system specialized for dealing directly with language and a nonverbal (imagery) system specialized for dealing with nonlinguistic objects and events. In addition, Mayer's (2001) Generative Theory of multimedia learning provides a similar rationale for learning and suggests that both verbal and visual information are accessed respectively in short-term memory, and then the verbal and visual representations may process the information coherently as a whole in the working memory, leading to a comprehensive understanding of the information. Usually, language learners will select useful information and organize them correspondingly into different models when they are provided the verbal and visual input through multimedia. Therefore, the connections can be established to form a mental and meaningful structure. Indeed, words in verbal models and pictures in other models provide different types of information, with the former discrete and linear, the latter holistic and nonlinear. Learners' comprehension, therefore, can be achieved when the models are integrated into related knowledge structures (Ariew, 2006).

Based on the Dual Coding Theory (Paivio, 1990) and Generative Theory of multimedia learning (Mayer, 2001), three main assumptions can be made to support the idea of the current study. The first is *dual channels*, referring to the idea that the verbal and visual information are processed in different channels, and the information of one channel can be transformed to another one through interaction. The second one is *limited capacity*, suggesting that each channel has limited capacity, and the last one is *active processing*, that is, learners are actively involved in the process of knowledge construction. This is why Yeh and Lo (2004) claim that multimedia annotations work as supportive tools helping readers learn how to move from reading to writing, such as highlight, questions, comments, examples and sharing which can scaffold various note-taking and reading strategy training. Therefore, in the current study, the multimedia annotation is treated as the visual information, and the summary writing as the verbal information.

Martinez-Lage (1997) gives a very detailed description of specific ways to use an authoring system to develop annotations and word glosses that students may use during the reading process. The use of word-glossing and annotations in computer-assisted language learning (CALL) is an area that has been researched extensively, and computerized reading with

glosses may enhance deep levels of text comprehension. Lomicka (1998) conducted a pilot study to explore how multimedia annotations influence the level of reading comprehension. Twelve participants joined the study and they were divided into three groups with full glossing, limited glossing, and no glossing to compare the effects of annotation. By applying the think aloud approach, the results showed that the participants benefited more from computerized reading with full glossing because it promoted a deeper level of text comprehension. Although it is noticeable that the number of participants and the evaluation of reading comprehension tests are problematic, this pilot study still provides some empirical evidence to support the usefulness of multimedia annotation. Wolfe (2002) has not only published a review article to cover the different models for annotation technologies currently available or under development, but also provided some characteristics about annotation, such as the advantages and impacts on readers (Table 1).

Table 1. Characteristics of Annotation (Wolfe, 2002)

Advantages of Annotation	1. Improving comprehension of source material
	2. Quote for later review
	3. Facilitate critical thinking
	4. Interpret and comment
	5. Record intermediate and unselfconscious reaction to text
Impacts of Annotation	1. Improving recall of emphasized items
	2. Influence perception of specific arguments
	3. Decrease tendencies to unnecessarily summarize

Ariew and Ercetin (2004) researched whether the relationship between reading comprehension and the use of different types of hypermedia annotations (texts, pictures, and audios) really exists. The participants were all English as a second language (ESL) learners from two different levels of proficiency, including advanced and intermediate. The data collection applied a tracking tool to record the amount of time participants used the annotation. Results demonstrated that the annotation uses did not lead to better reading comprehension for advanced ESL learners. Moreover, the application of annotation even contributed to negative effects for the intermediate learners. In addition, the variable identified that may help the participants positively was their prior knowledge. Finally, the results from questionnaire and interview showed hypermedia reading had a positive impact on the participants' attitudes toward reading on the computer.

Similarly, Sakar and Ercetin's (2005) investigated adult learners' English learning preferences for hypermedia annotations and the effects of multimedia annotations (texts, pictures, and audios) on reading comprehension. The participants were intermediate-level adults who were learners of English as a foreign language (EFL) and studied English for academic purposes (EAP). According to the results of the questionnaire and the interviews, participants preferred visual annotations significantly more than textual and audio annotations. However, the results of the reading comprehension test also revealed that annotations had negative effects on participants' reading comprehension, especially the annotations with audio-recordings and videos. In fact, it is possible that the data collection procedures themselves might be suspect here because these participants were asked to finish the reading and the test in only 45 minutes; in other words, they had to read the test with the annotations and then finish the test, and thus, the participants probably did not have enough time to accomplish all of these tasks successfully.

Multimedia Annotations and Second Language Vocabulary Learning

The research related to annotation (Hulstijn, Hollander, & Greidanus, 1996; Yanguas, 2009) mainly focuses on two issues: (1) to investigate the effects on reading, (2) to examine if annotation may promote vocabulary learning (Alessi & Dwyer, 2008; Yanguas, 2009), especially for incidental vocabulary learning (i.e., the picking up of new words during listening or reading tasks when the general goal of the task is comprehension rather than specifically learning those new words). Many studies confirmed that the use of multimedia annotation is effective in promoting different aspects of SLA, including vocabulary, listening, and general reading comprehension (Chun & Plass, 1996a; Lomicka, 1998; Ridder, 2002; Ercetin, 2003; Sakar & Ercetin, 2005; Yanguas, 2009). Ariew (2006) argues that the use of hypertext and hypermedia glosses may promote incidental vocabulary learning. In his view, hypermedia glosses play a role in assisting language learners to understand texts better, but the glosses also require effort in selecting the appropriate information in ways that learners are able to notice the information provided.

Chun and Plass (1996a) were early researchers who began work on the research of second language vocabulary acquisition through multimedia annotations (*CyberBuch*). The fundamental theory of their research focused on the advantage of multimedia application because, in addition to traditional definitions of words, it can provide various types of input for learners, such as pictures and videos. In this way, they investigated how vocabulary was learned and also the effectiveness of the multimedia annotation. The results indicated that participants received significantly higher scores on words that were annotated with pictures + text than for those with video + text or text only.

Later, Chun and Plass (1996b) also carried out a similar research based on the idea that reading should be an interactive process that might require both top-down and bottom-up processing. In order to examine how reading comprehension can be facilitated by the multimedia application, two different tools were used. On the macro level (top-down processing), the effects of a dynamic visual advanced organizer were investigated, while on the micro level (bottom-up processing), the effects of multimedia annotations for vocabulary items were studied. The data collection included a questionnaire, a vocabulary test, and a recall protocol. The results of the recall protocol showed that the multimedia advanced organizer facilitated reading comprehension, and multimedia annotations also helped the reading comprehension for individual vocabulary items. However, the interactive effects were not very significant. Moreover, regarding the research design itself, it would have been better for the author to have a pretest and then the scores of individual vocabulary items so that the relationship of reading comprehension could be more comparable.

Since 2000, more and more researchers joined to investigate the effects of multimedia annotations on vocabulary acquisition (Akbulut, 2007; Al-Seghayer, 2001; Ariew & Ercetin, 2004; Sakar & Ercetin, 2005; Yanguas, 2009; Yoshii & Flaitz, 2002). Al-Seghayer (2001) and Akbulut (2007) had similar research designs that included annotations with text, pictures, and video clips for three different groups of participants. In Al-Seghayer's (2001) study, thirty participants were divided into three groups with dynamic video, pictures, and text annotations. The data were collected through tests of recognition and production, face-to-face interview, and questionnaires. The results showed that the dynamic video clip is more effective in teaching unknown vocabulary than other modalities (pictures and texts.) Moreover, the results of face-to-face interview and questionnaires also indicated that video better built a mental image, better

created curiosity leading to increased concentration, and embodied an advantageous combination of modalities (video, pictures and texts). In the same way, Akbulut's (2007) data were collected through pretest, posttest, and delayed posttest. In terms of incidental vocabulary test, participants performed significantly better with the help of visual (pictures and videos) annotations. However, the results of reading comprehension tests did not show any significant difference among the groups.

Table 2. Brief Summaries of Recent Studies on Reading with Annotation

	Vocabulary Acquisition	Text +	Graphs +	Video +	Audio +
Chun and Plass (1996a)	●	●	●●	●	
Chun and Plass (1996b)	●	●	●●	●	
Lomicka (1998)		●	●		
Al-Seghayer (2001)	●	●	●	●●	
Yoshii and Flaitz (2002)*	●	●	●		
Ariew and Ercetin (2004)		●●	●		● (negative)
Sakar and Ercetin (2005)		●●	●		● (negative)
Ariew (2006)		●	●	●	●
Akbulut (2007)	●	●	●●	●●	
Yanguas (2009)*	●	●	●		

● Research design

●● Better effects

* The combination of text & picture has best effects

Yoshii and Flaitz (2002) studied the effects of annotation types on incidental vocabulary retention in multimedia reading setting. In their research, three different types of annotation were probed, including text-only, picture-only, and the combination of the two. Statistically, the results of an ANOVA analysis revealed that participants in the combination group (text and picture) significantly outperformed those in other groups. With a similar research design, Yanguas (2009) carried out a study based on the theoretical framework of "attention" to examine the effects of different types of multimedia glosses, including textual, pictorial, and textual + pictorial. Both the quantitative and qualitative results all showed that the three experimental groups (with the help of multimedia glosses) performed better than the control group (without any multimedia gloss). Furthermore, no significant results could be found from the production task. As for the reading comprehension test, the group with the help of textual and pictorial multimedia glosses performed significantly better than other groups. Table 2 above provides brief summaries of recent studies on annotations, including the research design and effects of the used annotations.

Second Language Reading and Discourse

Teufel, Carletta, and Moens (1999) introduced an annotation scheme with seven categories for discourse-level argumentation, arguing that the scheme is useful for readers to establish the summary. Moreover, the annotation scheme can be learned or trained by learners and this might be an alternative means for reading comprehension. However, the discourse scheme, according to Teufel et al. (1999), was used for helping readers to better understand the thesis, rather than general and shorter texts for EFL learners. Actually, almost no study has been done to investigate the role of annotation through discourse perspective, which is very crucial to

reading comprehension, because, as already mentioned, reading is an interactive process requiring learners to read beyond the lines to arrange the information of the texts well. It is, therefore, hypothesized that the annotations through discourse-level scheme may help learners to comprehend the texts better and more efficiently.

In sum, different researchers had various designs and applied several multimedia annotations to compare the effects on reading comprehension and vocabulary acquisition. However, the results varied because the types of multimedia annotation used were not similar despite positive effects on reading comprehension. In the current study, the application of multimedia annotations would be studied in a discourse-level perspective to see if learners can indeed benefit from it to enhance their reading comprehension, and this discourse scheme is modified based on the one by Teufel et al.

Summary Writing (Reading to Learn)

The TOEFL 2000 reading construct paper (Enright et al., 1998) suggested that a "Reading to Learn" task would require students to recognize the larger rhetorical frame organizing the information in a given text and carry out a task demonstrating awareness of this larger organizing frame. Moreover, they proposed the idea that in 'reading to learn' readers must integrate and connect information presented by the author with what they already know. In the research by Trites and McGroarty (2005), the participants received two different reading comprehension tests: one was a basic comprehension test with multiple choices and the other was a "reading to learn" that required them to write summaries of the texts they read to demonstrate their understanding. The results indicated that summary writing better assesses readers' comprehension than multiple choice questions. However, Trites and McGroarty (2005) did not conclude that asking readers to write summaries positively enhances their understanding of the texts. That is to say, enhancing reading comprehension on the one hand and being better alternatives for reading assessment on the other hand seem to be different ideas in themselves; it cannot be overstated that summary writing (Reading to Learn) can not only assess readers' understanding accurately but also enhance their comprehension positively. Additionally, it should be noted that, in the current study, summary writing is considered a medium to help learners while reading, instead of simply a way to assess learners' reading performance.

Given the above explanations, the study focuses on wash back effects to examine if teaching learners how to write summaries may help them understand the texts better, and this instruction may concentrate on the discursive concept of grounding. Information units in discourse are broadly classified into two major types based on the functions they serve: those that push forward these story line and those that provide supporting materials for the advancement of the story line (Givon, 1979; Chang, 1997). The former is generally referred to as "foregrounding" and the latter as "backgrounding" (Givon, 1979). Foregrounding, according to Chang (1997), tends to be realized syntactically as main clauses, while backgrounding tends to be realized as subordinate clauses. Robinson (1984) points out that each well-constructed paragraph should consist of a topic sentence, followed by sentences that coordinate with or are subordinate to the main point. The second hypothesis of this study is, through the instruction of the discursive concept of grounding, learners may know how to write summaries that can lead to better comprehension of the original texts.

The Present Study

Reading comprehension does not only occur during the reading process, it will equally occur after the process because the readers will most likely try to combine and integrate the main ideas of the texts to understand the central meaning. As already noted, most studies pay attention to the connections between multimedia annotation and vocabulary learning, while few studies discuss the role of discourse. Moreover, in line with Ariew (2006), "the design and implementation of hypermedia or hypertext materials are important factors in reading comprehension, and, consequently, the means that annotations are accessed and presented may have a critical impact on readers' comprehension" (p. 207). Therefore, the current study will explore the effects of discursive multimedia annotation and summary writing on reading comprehension. Again, although Trites and McGroarty (2005) mentioned that asking readers to write summary after reading can specifically and accurately evaluate the reading ability, they did not mention that such means can indeed enhance readers' comprehension. Thus, this study investigates whether (a) multimedia annotation and summary writing effects reading comprehension, and (b) whether the influences and effects of multimedia annotation and summary writing can interact positively to help readers have better comprehension toward the texts. This study is guided by the following three research questions:

1. Do Taiwanese university EFL students perform differently and benefit from the treatments (multimedia annotation and summary writing) they receive? To what extent do Taiwanese university EFL students perform differently according to the treatments they receive?
2. If Taiwanese university EFL students perform differently according to the treatments (multimedia annotation and summary writing) they receive, what are the lasting effects of the treatments over time?
3. What are the Taiwanese university EFL students' attitudes toward the application of the treatments (multimedia annotation and summary writing)?

METHODS

Participants

The pool of participants consisted of approximately 105 freshmen non-English majors, ages 19 to 20. Enrolled in the General English course, these participants included 60 females and 45 males; however, 5 of them did not participate in all three tests and are therefore not included in the analysis of data. During the research experiment, all participants were separated into four groups in order to receive different treatments to examine the effects on their reading comprehension.

Discourse Scheme

As mentioned above, the discourse scheme by Teufel et al. (1999) was used for reading the thesis in scientific domain, not for general reading comprehension for EFL learners. In this way, the discourse scheme (Table 3) for multimedia annotation of current study is modified based on the one by Teufel et al.

Table 3. Discourse Scheme for Annotation

Reading for Gist	Readers need to get a general sense of what a reading passage is basically about. In other words, they read to understand the main topic, or theme of the passage.
Understanding the Main Idea	Once readers have determined the text type of a passage, and what it is generally about, they then read on to understand the main idea of the passage. Understanding the main idea of a text means being able to identify the most important point or information in the passage.
Paraphrasing	Paraphrasing involves identifying a restatement of a section in a passage that retains the basic meaning while changing the words, often explained in a more simplified form. A paraphrase often clarifies a more ambiguous original statement in the text by putting it into alternative words that are often more easily understood.
Identifying Details	Identifying details in a text answer specific questions (*e.g., who, what, when, where, why*).
Making Inferences	Readers must be able to guess some things and make clear assumptions from the information, facts, opinions and author's feelings presented in the passage.
Vocabulary	The definitions of the vocabulary in the text may be given through different clues to help readers identify how the author explains his or her meaning.
Understanding References	This involves focusing on specific meaning of 'pronoun references' used throughout a passage (*e.g., this, those, their, it*).

During the instruction, participants were asked to make use of the multimedia annotation based on the discourse scheme to annotate the article they read. Various annotations were applied to mark different sources of information. For example, in order to get the main ideas, participants might use annotations to mark the topic sentences, and moreover, when encountering some unknown words, they could use annotation to derive some possible definition from the context. In general, the discourse scheme is used as a guideline for participants to annotate the reading passage.

Instruments

The present study employs two main instruments:

1. *Annotation Tool*: According to Wolfe (2002), new technologies can improve and enhance annotation practices through multimedia features unique to hypermedia environments. In this study, users can create coding schemas for their annotations by selecting from multiple colors, fonts, and presentation styles. This multimedia annotation is a tool that can create highlights in multiple colors on texts from any web page or PDF file. Readers can distill important facts just as if the texts had been printed out. By doing so, the participants can make use of this annotation tool to highlight the ideas based on the importance level to comprehend the texts. Readers can also add sticky notes or colorful boxes with their own comments while making inferences or understanding references. If needed, readers can equally use this tool to organize, rearrange, and delete the annotations of the web page or PDF file to gain a better understanding of the texts or reading materials in question. In the current study, the participants will apply the annotation tool during the reading process.

2. *Summary Writing/Reading to Learn*: Here participants are required to write a summary of

what they have read. According to Trites and McGroarty (2005), such summary writing after reading can be considered as "Reading to Learn" because readers have to demonstrate their understanding of the texts by providing a summary. Since the summary writing is considered as another means for reading comprehension test, we hope to see the wash back effects it may bring to learners while reading. In this way, summary writing is the second treatment and is carried out to investigate the effects on participants' reading comprehension. Specifically, this treatment is applied for participants through the instruction of foregrounding and backgrounding.

In this study, all participants were divided into four groups randomly, and Table 4 shows the arrangement of treatments and the research design of the current study.

Table 4. The Treatments of Each Group

Group 1 (G1)	No Treatment
Group 2 (G2)	Treatment 1: Multimedia Annotation
Group 3 (G3)	Treatment 2: Summary Writing
Group 4 (G4)	Combination: Treatment 1 + Treatment 2

Again, the effects of writing a summary cannot be definitive, especially when it is combined with Treatment 1 (multimedia annotation). The design is done this way to examine the effects of different treatments. The results from Group 1 can be the baseline for comparison with the results from groups with different treatments.

Reading Materials

1. Reading materials from the Internet (i.e., *The Student Post*) were selected for instruction to train participants on how to use the multimedia annotation.
2. For the experiment itself, the reading texts/materials were chosen from the retired version of General English Proficiency Test (GEPT), high-intermediate level, and these reading texts/materials were then transformed into PDF files for them to read. As concerns the pre-test, the practice version of GEPT was used to examine if the reading abilities of the participants exhibited similar levels.
3. As for post- and delayed post-test, two different retired versions of GEPT were also applied to assess the participants' reading performance. All of the reading tests were multiple-choice questions, including several reading texts and 20 questions in total. The total score is 100, with 5 points for each item.
4. The readability (Flesch/Flesch-Kincaid Readability Tests) was employed to ensure similar difficulty level for all selected texts/articles.

Data Collection Procedures

The total experiment lasted for two months, including the instruction of the treatments and all three different tests; each meeting would last for about two hours. Specifically, Table 5 refers to the data collection procedures, which begin with the pre-test, the instruction and practice of the treatments, and then followed by the post and delayed post-test. At the final stage, five volunteer participants from each group were invited to carry out the interview to share their points of view about the use of different treatments.

Table 5. Six Main Phases of Data Collection

1st Stage (1 hour)	The first phase was about the pre-test of the participants' reading performance. The reading comprehension test is selected from the retired version of GEPT (high-intermediate level).
2nd Stage (4 hours)	The second phase was about the instruction of how to use the annotation tool (1 hour). Participants will be given another hour to practice and get familiar with the use of the annotation tool.
3rd Stage (4 hours)	The third phase was about the participants' formal reading with the use of different treatments (2 different versions of reading materials, 1 version for 1 hour).
4th Stage (1 hour)	The fourth phase was about the post-test of the participants' reading performance. The reading comprehension test is selected from the retired version of GEPT (high-intermediate level), but the version for the post-test is different from the pre-test with the same level.
5th Stage (1 hour)	The fifth phase (one month after the fourth stage) was about the delayed post-test of the participants' reading performance.
6th Stage (1 hour)	The last phase was about the participants' interview (1 hour for each experimental group respectively).

Data Analysis

Following a quantitative approach, two different analyses were applied to answer the study's three research questions. First, one-way analysis of variance (one-way ANOVA) was applied to calculate and compare the differences of participants' reading performance among the groups (between-group effects). Second, a repeated-measure design of ANOVA was applied to examine the participants' reading performance among different stages of tests (within-group effects). In addition, five volunteer participants from each experimental group respectively were interviewed about their perceptions and attitudes toward the application of the treatments (annotation tool and writing summary) they received. These qualitative results were useful to explain the reading comprehension test performance by using the video-recording and open-ended questionnaire.

In general, the first two research questions are answered by applying the quantitative approaches to examine if there is any difference among the groups, while the third research question is answered by the qualitative approach to see the participants' viewpoints about the treatments they received.

RESULTS AND DISCUSSION

Before the experiments, all 100 participants received the pre-test to make sure their reading proficiency was similar. The mean scores of pre-test (Table 6) among the four groups ranged from 50.8 to 53, and the one-way ANOVA analysis revealed that all participants had similar reading proficiencies ($F = .493, p < .688$) before the experiment.

Table 6. Results of Pre-test

	G1	G2	G3	G4		
Mean (SD)	52 (9.242)	51.8 (8.765)	50.8 (8.573)	53 (8.977)	$F = .493$	$P < .688$

The first research question addresses whether Taiwanese university EFL students perform differently and benefit from the treatments (multimedia annotation and summary writing) they receive, and to what extent Taiwanese university EFL students perform differently according to the treatments they receive. After the participants joined the experiments, they received the post-test, and the results showed that the mean scores of the four groups ranged from 55.4 to 82.8. As shown in Table 7, the one-way ANOVA analysis showed that all experimental groups presented significantly different reading performance ($F = 53.344$, $p < .0001$) after they received the treatments, with G2 the best, followed by G3 and G4. G1 participants showed similar reading performance with the results of pre-test, while other participants of the experimental groups (G2-G4) performed better than the results in the pre-test. Therefore, participants who used the treatments performed better than those who did not employ any treatment. Among the three experimental groups, the *Scheffé Post Hoc* comparison test disclosed that G2 and G4 were significantly different, that is, the scores of participants who used multimedia annotation were significantly different from the score of participants who used both multimedia annotation and summary writing.

Table 7. Results of Post-test

	G1	G2	G3	G4		
Mean	55.4	82.8	77	74.6	$F = 53.344$	$p < .0001$
(SD)	(9.456)	(7.916)	(8.292)	(6.602)		

In accordance with the research design, all participants received the delayed post-test one month after the post-test. The results of the one-way ANOVA analysis of the delayed post-test (Table 8) also demonstrated that all experimental groups performed significantly different ($F = 18.338$, $p < .0001$) from the control group, with the mean scores ranging from 55.8 to 72.4. G1 participants also showed little change in reading test; on the contrary, among the three experimental groups (G2-G4), G3 performed best in reading comprehension test, followed by G2 and G4. In terms of the delayed post-test, participants with the help of multimedia annotation and summary writing still demonstrated better understanding than those without any treatment. However, different from the results of post-test, according to the *Sheffe Post Hoc* comparison test, G3 participants achieved significantly better performance than G4 participants in the results of delayed post-test. Said another way, participants with the help of summary writing did significantly better than those with the help of multimedia annotation and summary writing at the same time.

Table 8. Results of Delayed Post-test

	G1	G2	G3	G4		
Mean	55.8	68.2	72.4	65.6	$F = 18.338$	$p < .0001$
(SD)	(9.456)	(7.916)	(8.292)	(6.602)		

The second research question examined the lasting effects of the treatments over time, that is, do Taiwanese university EFL students perform differently according to the treatments (multimedia annotation and summary writing) they receive? Theoretically, the results of between-group effects can be answered and explained through the data analysis of ANOVA; nevertheless, the results of within-group effects should be examined and discussed through the approach of repeated measure of ANOVA. In this way, Table 9 shows the performance of the

three experimental groups among the various times of the tests. The results show that G2, G3, and G4 had significantly different reading performance in different tests respectively ($F =$ 316.343, $p < .0001$). In other words, the three experimental groups performed differently under the influence of the treatments (multimedia annotation and summary writing), while the effects of each group apparently varied with each other.

Table 9. Results of Within-group Effects

	SS	df	MS	F	Sig. of F
Within cells	6069.833	198	1101.639	316.343	.000
A	21399.500	2	10699.750		

In fact, Figure 1 might illustrate specific information of the within-group effects among the three tests. As seen, the controlled group (G1) maintained a similar reading performance, while G2-G4 participants in the three experimental groups showed that they were influenced by the various treatments (multimedia annotation and summary writing). G2 participants, with the help of multimedia annotation, seemed to do best in the post-test, followed by G3 participants who applied the summary writing for reading comprehension, and then G4 participants, with the help of the combination of the two treatments, seemed to perform not as well as the former two groups. As for the delayed post-test, on the contrary, each treatment had different lasting effects. G3 participants, with the aid of summary writing, did best in reading comprehension, followed by G2 participants, who were under the assistance of multimedia annotation, and lastly G4 participants, who applied both treatments.

Figure 1. Within-group Effects among the Three Tests
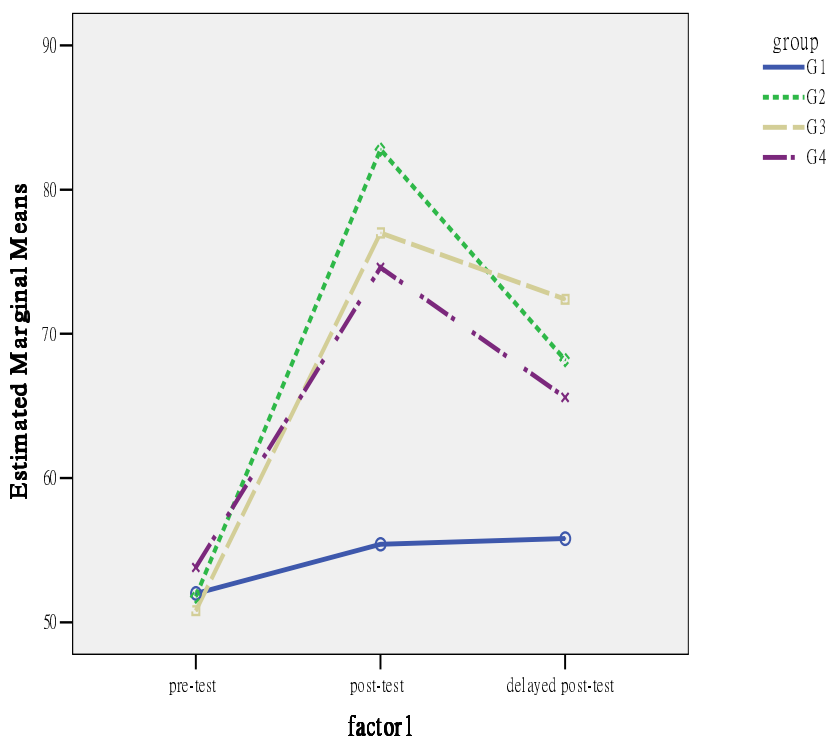
In accordance with quantitative analysis of one-way ANOVA, compared with the controlled group (G1), participants in all three experimental groups (G2-G4) performed significantly better in reading comprehension tests. That is to say, the treatments in this study, including multimedia annotation and summary writing, played a facilitative and positive role in helping readers to comprehend the reading texts better. Participants, under the influences of multimedia annotation and summary writing, apparently did better in reading comprehension of both post- and delayed post-test. As for the G1 participants, who did not employ multimedia annotation and summary writing, they did not improve the reading comprehension significantly among the tests. Since the quantitative results showed that the treatments facilitated learners reading comprehension and enhanced their performance, it is important to know to what extent the variances were among different groups.

Specifically, as indicated in Figure 1, the experimental groups (G2-G4) outperformed significantly the controlled group (G1) in post- and delayed post- reading comprehension tests. Participants in G2, with the help of multimedia annotation, achieved the best performance in post-test. The positive results of applying annotation in the present study are quite similar to the results obtained in such studies as Al-Seghayer (2001), Yoshii and Flaitz (2002), Ariew and Ercetin (2004), Sakar and Ercetin (2005), Akbulut (2007), and Yanguas (2009). However, the differences center on the value of this study, that is, the positive effects of discourse schemes through multimedia annotation. As already discussed, most researches to date focused on the "different means" of the information provided (i.e., text, pictures, and/or video), wherein learners apply the annotation in a passive manner, instead of annotating and processing the information in an active way to read more comprehensively. In the present study, participants did take full advantage of annotation based on discourse schemes to vigorously annotate different levels of information.

Expressed more precisely, G3 participants, with the help of summary writing, achieved the second best performance in post-test, while the effects were not as significant via the multimedia annotation. Indeed, asking learners to engage in summary write based on the grounding theory does facilitate understanding in that it helps learners how best to process information and how best to arrange ideas in a coherent manner. G4 participants, with the help of both multimedia annotation and summary writing, performed only better than G1 participants. It was assumed that, if both treatments had positive effects on reading comprehension, G4 participants would perform much better than the other groups. However, in the post-test and delayed post-test, the G4 results showed that the combination of treatments was not the best, although participants still demonstrated better performance than the G1 participants. One possible explanation for this outcome might be the working load itself, that is, the learners' cognition load is not enough to apply two different treatments at the same time. In fact, the use of multimedia annotation and summary writing might hinder the effects of reading comprehension with each other. This assumption was confirmed by the participants' attitudes and deserves further discussion.

Concerning the reading performance in delayed post-test, G2 participants still performed better than G1 or G4 participants. The lasting effects of multimedia annotation were not as strong as the G3. Asking learners to engage in summary writing based on the grounding theory required more cognitive processing, resulting in better lasting effects of reading performance (Jonassen, 1985). Although the application of multimedia annotation was still beneficial, learners' retention of this treatment did not function as well as with the summary writing. Without the help derived from multimedia annotation, learners might not know how to annotate different information both

appropriately and systematically. Similar to the results obtained in the post-test, G4 participants only performed better than G1 participants, supporting the conclusion that learners' cognition cannot process two treatments simultaneously. Thus, annotation and summary writing might only be confounding factors, rather than facilitative ones.

The third research question sought to uncover the Taiwanese university EFL readers' attitudes toward the application of the different treatments. For each group, five volunteers were invited to join the interview to share their thoughts. G2 volunteer participants agreed on the positive use of the discourse scheme because "the multimedia annotation was useful and clear for them to underline or put marks on important information or key points." These volunteers considered the discourse scheme to be most crucial because "the scheme could lead them to understand the reading coherently, and with the help of multimedia annotation, they might leave some notes to identify the details and the pronominal references." Further, one volunteer mentioned that, when encountering some unknown words, he could make use of the discourse scheme to annotate the related meaning from the context. One volunteer remarked that the combination of discourse scheme and multimedia annotation might help her outline the logical organization of the text, and as a result, she found it easier to distinguish the implicitly stated ideas from the explicitly stated ones. Four volunteers also expressed their difficulties in the delayed post-test. Although they still remembered some principles of the discourse scheme, without the application of multimedia annotation, "they had difficulties in comprehending and identifying the texts." Moreover, three volunteers even mentioned that they forgot the rules of the discourse scheme, thus, they somehow had to answer the questions by their basic reading abilities. In general, it could be stated that these G2 volunteers only expressed their short retention of the principles of discourse scheme and placed too much reliance on the multimedia annotation.

Similarly, G3 volunteer participants also agreed on the positive effects summary writing had on their reading performance by saying that "the grounding theory allowed them to identify the topic and the subordinate sentences, and consequently, they could write a brief summary based on the topic sentences or ideas." Interestingly, two volunteers expressed their anxiety as well when asked to write a summary, mostly because they felt unprepared to write a well-crafted summary due to their limiting writing abilities, even though they knew their summary writing would not be evaluated. It is important to note here that in the current study the summary writing was considered as a mediated treatment to assist learners to comprehend the text, instead of a test to be assessed. Thus, the results showed its positive effects, especially in the delayed post-test. When asked how the summary writing really helped them, the volunteers expressed that "they needed to totally understand the text in order to write a summary." Summary writing indeed requires more cognitive processing, and, because of this, learners may have better understanding and longer retention on the text.

Finally, the G4 volunteer participants performed better than the G1 participants in the post- and delayed post-test. From the outset, it was assumed that if the two treatments (multimedia annotation and summary writing) have positive effects on learners' reading performance, then G4 participants, having access to both treatments, would outperform all other groups. Results obtained did not confirm this assumption. During the interview, the volunteers revealed that, although they received both instructions of the treatments, they encountered the problem of how best to use the treatments. In other words, they did not know which one they were supposed to apply first during the test. Indeed, they considered the two treatments as confounding factors which required too much cognitive demanding and the "should-be positive

effects" diminished when they were answering the questions in both tests. In sum, the assumed interactive effects between input (reading) and output (writing) appeared not to exist.

CONCLUSION

The current study investigated the effects of multimedia annotation and summary writing on EFL university freshmen's reading comprehension. It focused on the role of multimedia annotation through discourse scheme and the derived wash back effects of summary writing. The results showed that both multimedia annotation and summary writing may have positive influences on assisting learners to understand the text, with lasting effects depicting few differences. G2 learners made use of the multimedia annotation based on the principles of discourse scheme, while G3 learners benefited from the summary writing. Their retention was longer due to more cognitive processing. Given the positive effects of summary writing, its use could potentially become an alternative approach for English teachers to consider when offering instruction in English. As for the summary combination of the two treatments, while helpful, such combination was not strong enough meriting further pedagogical attention here. Perhaps learners need more instruction how best to utilize the two treatments and not make them the confounding factors.

Wen-Chi Huang *is currently a Ph.D. student of TEFL at the National Taiwan Normal University in Taiwan. His main research areas and interests include second language reading and acquisition, learner classroom interaction, and language testing.*

Email: glavinehuang@ntnu.edu.tw

REFERENCES

Akbulut, Y. (2007). Effects of multimedia annotations on incidental vocabulary learning and reading comprehension of advanced learners of English as a foreign language. *Instructional Science, 35*, 499-517.

Alessi, S., & Dwyer, A. (2008). Vocabulary assistance before and during reading. *Reading in a Foreign Language, 20*(2), 246-263.

Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology, 5*(1), 202-232.

Ariew, R., & Ercetin, G. (2004). Exploring the potential of hypermedia annotations for second language reading. *Computer Assisted Language Learning Journal, 17*(2), 237-259.

Ariew, R. (2006). A template to generate hypertext and hypermedia reading materials: Its design and associated research findings. *The Reading Matrix, 6*(3), 195-209.

Bowles, M. A. (2004). L2 glossing: To CALL or not to CALL. *Hispania, 87*(3), 541-552.

Brantmeier, C. (2003). Technology and second language reading at the university level: Informed instructors' perceptions. *The Reading Matrix, 3*(3), 50-74.

Chang, W. C. (1997). *Freshman English composition: An error analysis from the discourse*

perspective. Taipei: The Crane Publishing.

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

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

Enright, M., Grabe, W., Mosenthal, P., Mulcahy-Ernt, P., & Schedl, M. (1998). *A TOEFL 2000 framework for testing reading comprehension: A working paper*. Princeton, NJ: Educational Testing Service.

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

Givon, T. (Ed.). (1979). *Syntax and semantics 12: Discourse and Syntax*. New York, NY: Academic Press.

Huang, H., Chern, C., & Lin, C. (2009). EFL learners' use of online reading strategies and comprehension of texts: An exploratory study. *Computers & Education, 52*, 13-26.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, *80*(3), 327-339.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*(3), 539-558.

Jonassen, D. H. (1985). Learning strategies: A new educational technology. *Programmed Learning and Educational Technology*, *22*(1), 26-34.

Jones, L., & Plass, J. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal, 86*(4), 546-561.

Liontas, J. I. (2002). CALL Media digital technology: Whither in the new millennium? *CALICO Journal, 19*(2), 315-330.

Lo, J. J., Tsao, C. H., & Yeh, S. W. (2005). Using online annotation to facilitate reading: System development and empirical studies. *Cgung Yuan Journal, 33*(2), 193-214

Lomicka, L. L. (1998). To gloss or not to gloss: An investigation of reading comprehension online. *Language Learning & Technology, 1*(2), 41-50.

Lunsford, A., & Ruszkiewicz, J. (1999). *The presence of others* (3rd ed.). New York, NY: St. Martin's Press.

Martinez-Lage, A. (1997). Hypermedia technology for teaching reading. In M. Bush & R. Terry (Eds.), *Technology enhanced language learning* (pp. 121-163). Lincolnwood, IL: National Textbook Company.

Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.

Paivio, A. (1990). *Mental representations: A dual coding approach*. New York, NY: Oxford University Press.

Ridder, I. (2002). Visible or invisible links: Does the highlighting of hyperlinks affect incidental vocabulary learning, text comprehension, and the reading process? *Language Learning & Technology, 6*(1), 123-146.

Robinson, S. E. (1984). *Coherence in student writing*. Unpublished doctoral dissertation, Harvard University.

Sakar, A., & Ercetin, G. (2005). Effectiveness of hypermedia annotation for foreign language reading. *Journal of Computer Assisted Learning, 21*, 28-38.

Salvatori, M. (1996). The argument of reading in the teaching of composition. In B. Emmel, P. Resch, & D. Tenney (Eds.), *Argument revisited, argument redefined: Negotiating meaning in the composition classroom* (pp. 181-197). Thousand Oaks: SAGE.

Schmidt, R. (2001). Attention. In P. Robinson, (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Teufel, S., Carletta, J., & Moens, M. (1999). An annotation scheme for discourse-level argumentation in research articles. *The Proceedings of the Ninth Conference on European Chapter of the Association for Computational Linguistics* (pp. 110-117). University of Bergen, Norway.

Trites, L., & McGroarty, M. (2005). Reading to learn and reading to integrate: New tasks for reading comprehension tests? *Language Testing, 22*(2), 174-210.

Wolfe, J. L. (2002). Annotation technologies: A software and research review. *Computers and Composition. 19*, 471-497.

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

Yeh, C. Y., & Lo, J. J. (2004). An investigation into vocabulary learning strategies used by senior high school students in Taiwan. *Journal of Applied Foreign Languages*, *2*, 1-22.

Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of text and annotation types. *CALICO Journal, 20*(1), 33-58.