

目 錄

CONTENTS

析辯近人對章學誠「浙東學術」一說所提之相關論點·····黃勇中 1 The Study of Zhang Xue Cheng's "Zhe Dong Xue Shu" ·····Huang, Yong-Zhong	1
IC 封裝導線架剪斷彎曲成形分析與模具設計·····許昭和、林蘇廣 25 Analysis of forming for cutting and bending leadframe of IC package and die design····· ·····Chao-Ho Hsu、Su-Guang Lin	25
多媒體註記系統對台灣大學生英語單字學習之影響·····謝怡娟、陳威廷 45 The Effects of Multimedia Annotations on Taiwanese EFL Learners' Vocabulary Learning····· ·····Yi-Chuan Hsieh、Waiting Chen	45
周原戰爭、祭祀類甲骨研究·····彭慧賢 73 The Research of Zhou Yuan Warfare and Sacrificed Ceremonies Oracle Bone Inscriptions····· ·····Hui-Hsien Peng	73
應用改良式基因演算法於供應鏈之模糊多目標決策問題·····李郁賢、黃士滔 99 Applying the fuzzy theory and modified genetic algorithm for multi-objective problem in supply chain····· Yu-Shian Li、Shih-Tao Huang	99
跨文化訓練、駐外適應與駐外績效之路徑分析模型之研究—以台商派駐大陸人員為例····· ·····程永明 125 A Path Analytical Model of Cross-Cultural Training, Expatriate Adjustment, and Expatriate Performance:A Survey of Taiwanese Expatriates in China·····Yung-Ming Cheng	125
探討在 QFN 製程中含丙烯酸熱塑型膠帶之殘膠現象····· ·····劉佳龍、黃千晏、洪秀婷、王易軒、劉耿廷、曾清桂、歐陽文忠 157 The Investigation of Residual Glue Phenomena of Thermoplastic Tapes of QFN Package in the Semiconductor Device Processing·····Jia-Long Liu、Chen-Yan Huang、 Xiu-Ting Hong、Yi-Xuan Wang、Kent-Ting Liu、Ching-Guey Tseng、Wen-Chung Ou-Yang	157
以 P7 及約略集合分析企劃 3C 連鎖量販店新型服務·····李靜芳、鄭琇文、陳耀茂 175 Planning New-type 3C Hypermarket Services by P7 and Rough Set Analysis····· ·····Ching-Fang Lee、Shiou-Wen Jeng、Yow-Mow Chen	175

科技大學學生對大一英文課程規劃的看法：以一所應用科技大學為例·····	
·····王佩玲、羅雅芬、許春滿、徐儷娜	197
Curriculum Design for Freshman English—An Analysis of Technological University Freshmen’s Perceptions·····	
·····Pei-Ling Wang, Ya-Fen Lo, Chuen-Maan Sheu, Lina Hsu	
關係傾向、轉換成本和產品類別對顧客保留意願的影響·····	
·····李曉青、周凱傑	223
The Effect of Relation Proneness, Switching cost and Product Category on Customer Retentions	
·····Hsiao-Ching Lee、Kai-Jie Zhou	
市電併聯型太陽能發電系統動態分析·····	
·····廖立智、陳盟仁、吳有基	243
Transient Behavior of Grid-Connected Photovoltaic Power Systems·····	
·····L.-C. Liao、 M.-J. Chen、 Y.-C. Wu	
風力與太陽能複合發電系統之製作·····	
·····郭文豐、李順晴	253
Fabrication of Wind Turbine and Solar Hybrid Power Conversion System·····	
·····Wen-Feng Kuo、Shun-Ching Lee	
Novel Sufficient LMI Conditions for Static Output Feedback Stabilization of Proper Linear Systems	
·····李慶祥 Ching-Hsiang Lee	263
濺鍍操作條件對摻鋁氧化鋅透明導電膜特性之影響·····	
·····王昇裕、蔡政賢	281
The Effects of Operating Conditions of Sputtering on the characteristics of Al ₂ O ₃ -doped ZnO Transparent Conducting Film·····	
·····Sheng-Yu Wang、Cheng-Hsien Tsai	

多媒體註記系統對台灣大學生英語單字學習之影響

The Effects of Multimedia Annotations on Taiwanese EFL Learners'

Vocabulary Learning

謝怡娟¹

清雲科技大學

應用外語系

助理教授

Yi-Chuan Hsieh

Ching-Yun University

Department of Applied Foreign Languages

Assistant Professor

E-mail: ychsieh@cyu.edu.tw

Tel: (03) 3481196 ext. 7903

陳威廷²

中原大學

應用外國語文學系研究所碩士

Waiting Chen

Department of Applied Linguistics

and Language Studies

M.A.

E-mail: waiting1213@yahoo.com.tw

摘要

本研究在探討傳統紙本式註記, 多媒體靜態圖片註記, 與多媒體動態影片註記對台灣大學生英語單字學習的影響。研究對象為一百四十二名台灣桃園地區的某校大學生, 他們被任意安排到三種不同的學習組別, 進行兩週的實驗。這三組學生分別學習所選的英文文章中之未知字: (1) 第一組以紙本式呈現所閱讀的文章中英文字彙的定義與例句 (2) 第二組以多媒體系統呈現所閱讀的文章中英文字彙的文字定義與例句, 加上靜態圖片闡釋字義 (3) 第三組以多媒體系統呈現所閱讀的文章中英文字彙的文字定義與例句, 加上動態影片闡釋字義。兩種不同類型的詞彙測驗 (確認測驗和造句測驗) 作為後測, 用以評估不同學習條件下各組學生的英語字彙學習成效。研究結果顯示第二組與第三組的學生的學習成效顯著地高過第一組的學生。第三組的學生的學習成效顯著地高過第二組與第一組的學生。研究結果顯示多媒體註記系統, 不管是將英文字彙的文字定義以靜態圖片或是以動態影片闡釋, 對於學生的英語單字學習成效均有正面的影響。

關鍵字: 多媒體註記, 電腦輔助語言教學, 英語單字學習

Introduction

Conventional English vocabulary instruction has been mainly based on the use of a dictionary or vocabulary lists with a focus on abstract, standard definitions and example sentences. However, such instruction for second language vocabulary has been criticized because the types of word knowledge embodied in a definition are too abstract and provide insufficient information about a word's actual usage (Anderson & Nagy, 1991).

Advances of computer-assisted language learning in educational applications have offered a new resource of authentic language input for language classrooms (Brucklacher & Gimbert, 1999). Sun and Dong (2004) also indicated that the rapid growth of multimedia computers as learning tools brings new possibilities to the design of L2 learning activities. Information presented in audio, video, graphical, animation, and verbal formats can be integrated to create an authentic and vivid multi-sensory language context. Numerous studies have demonstrated

that the provisions of both still images and dynamic videos in instruction are more effective than instruction without imagery in facilitating L2 learner's vocabulary development (Jacobs, et al., 1994; Kost, Foss & Lenzini, 1999). However, few studies have been conducted with the purpose of comparing the effect of still images with that of dynamic videos on L2 vocabulary acquisition.

The present study is significant in that it extends the knowledge of how multimedia annotation can facilitate the L2 readers' vocabulary learning. While it has been evident that multimedia annotation factors aided vocabulary learning in second language reading (Chun& Plass, 1996a, 1996b; Kost, Foss & Lenzini, 1999; Al-Seghayer, 2001), few researchers have explored which kind of multimedia modalities are more effective in aiding L2 vocabulary learning. In addition, the studies that directly probed the differential effect of still images and dynamic video clips on learning annotated words yielded inconsistent results (Chun& Plass, 1996a,

1996b; Al-Seghayer, 2001); hence, this study may shed more light in research on L2 vocabulary learning involving the multimedia learning systems and enlighten more educators to successfully integrate multimedia into L2 vocabulary instruction.

This study aimed to answer the following questions: (1) Did the EFL learners who used multimedia annotations in experimental groups (textual definition coupled with images or videos) outperform those using no multimedia annotation (control group) on the immediate post-test (a recognition section and a production section)? (2) Did the EFL learners who used multimedia annotations in experimental groups (textual definition coupled with images or videos) outperform those using no multimedia annotation (control group) on the delayed post-test (a recognition section and a production section)? (3) Did the experimental group which received dynamic video annotations score significantly higher than the experimental group which received still image annotations on the immediate post-test

(a recognition section and a production section)? (4) Did the experimental group which received dynamic video annotations score significantly higher than the experimental group which received still image annotations on the delayed post-test (a recognition section and a production section)?

Theoretical Background

Empirical studies on multimedia annotation indicated that computerized annotation had a beneficial effect on vocabulary acquisition and reading comprehension (Lyman-Hager, Davis, Burnett & Chennault, 1993; Tozcu & Coady, 2004). Learners who studied with computerized multimedia texts showed greater retention of vocabulary than those who learned with non-computerized texts (Al-Seghayer, 2001; Chun & Plass, 1996a, 1996b; Davis & Lyman-Hager, 1997). Likewise, Davis and Lyman-Hager (1997) pointed out that unlike traditional marginal glossing in the printed form, the computerized glossing allows us to store more extensive glossing than a printed format did.

Additionally, a computerized gloss does not interrupt the reading process in that the glossed word is not visible until the reader clicks on it. Unlike time-consuming dictionaries, multimedia annotations cause minimal interruption of the reading process (Davis & Lyman-Hager, 1997; Martinez-Lage, 1997; Nation, 2001). Using different methodologies, numerous studies have found that the effect of multimedia annotation on word retention is superior to conventional printed pages (Blake, 1992; Lyman-Hager, Davis, Burnett, & Chennault, 1993; Reinking & Rickman, 1990).

Multimedia annotations and vocabulary learning

Mayer (1997) proposed the “Generative Theory of Multimedia Learning” to explain the value of multimedia presentations for language learning. The application of the generative theory of multimedia learning to vocabulary learning assumed that L2 learners have two separate verbal systems (the first language and the second language) and a common imagery system. Translation of

words through simultaneous verbal and visual presentations would link the two verbal systems, and this storage in the second verbal system would also aid in learning (Paivio & Desrochers, 1980). Furthermore, according to Paivio (1971)’s dual-coding theory, the assumption is made that memory and cognition are served by two different systems, one to deal with verbal information, and the other to handle the nonverbal information. More specifically, information is stored in long term memory in either the verbal or nonverbal system. The dual coding theory suggests that learning can be facilitated when a person is exposed to more than one mode (Paivio, 1971). Supported by the Dual Coding Theory, multimedia language learning systems, with its ability to present verbal and nonverbal codes to learners, are believed to facilitate and enrich language learning (Najjar, 1996).

In Chun and Plass’s (1997) study, a multimedia program *CyberBuch* was designed to offer three kinds of multimedia annotations for learners to click on: text, text-picture, or

text-video. The research question was to probe how American university students learned German as a second language using the program in class (Chun & Plass, 1996a, 1996b; 1997). The effects of multimedia annotation on reading comprehension and incidental vocabulary acquisition were investigated. Before reading a story on the computer, students watched a video outlining the story's background. After reading that annotated story, students did a summary recall protocol and took a vocabulary test. The results showed that words annotated with text and picture (or text and video) were learned more effectively than words annotated with text only. In addition, students recalled more from words annotated with visual mode than words annotated with text or picture mode (Chun & Plass, 1996b). They attributed the effects to dual coding, which helped participants to perform better under the text- picture (and text-video) conditions. However, the researchers mentioned there was a need to compare influence of individual annotation type, for example, picture-only and video-only annotations on vocabulary learning.

Later, Al-Seghayer (2001) conducted a study investigating which of the image modalities—dynamic video or still picture—is more effective in facilitating L2 vocabulary acquisition. The study used a program that asked the participants to read a narrative English text with a variety of annotations for words in the form of printed text, graphics, video, and sound. A within-subject design was used in this study with 30 participants being measured under three learning conditions: printed text definition alone, printed text definition coupled with still pictures, and printed text definition coupled with video clips. The result revealed that a video clip in combination with a text definition was more effective in teaching unknown vocabulary words than a still picture in combination with a text definition.

The differential results among the effect of pictorial vs. video annotations in Al-Seghayer's study (2001) and the study conducted by Chun and Plass (1996) can be attributed to several issues, including the mother tongue of the participants, the target language, the type of visual aids, and the type of tests through which

the learning outcome was assessed.

Given the fact that few studies are conducted to directly compare the effects of still picture annotation and dynamic video annotation on vocabulary learning, and that the mixed results yielded by previous studies, the present study attempted to investigate which multimedia annotation modalities is more helpful for L2 students to learn the new words.

Methodology

Pilot study and instrument design

A pilot study was conducted to help the researchers select the appropriate reading text and ascertain the level of difficulty of selected vocabulary. Based on the results of the pilot study, the instruments were modified. In addition, three professional Taiwanese university English instructors were involved in the pilot study to assist in the selection of target words and appropriate materials for learning. They also provided feedback and suggestions in designing the tests and the multimedia program. Additionally, thirty non-English major freshmen from a given university in Taiwan participated, and they

were not included in the latter's formal study.

The participating students went through two steps to help the researcher determine the target words for the later formal study.

Step 1: Selecting the Reading Texts

Ten reading texts were selected by the researchers based on their text length, the difficulty of the vocabulary, and presented topics. The participating students were expected to finish each reading text within 30 minutes, and vocabulary or grammar complexity should not overwhelm them. Additionally, the topic of each reading text should be related to the students' daily life. Five reading articles were narrative texts in English, and the other five articles were expository texts in English. Following that, three English instructors helped choose one narrative text and one expository text from all the reading texts based on the syntactic complexity, the text length, and the subject matter of the text selections. Finally, they reached a consensus to select two selections: "So Far, So Good: Recollections of a First-time Father" and "The Power of

Nature.”

Step 2: Selecting the Target Words

Three English instructors and 30 participants in the pilot study helped circle words from the two selected reading texts mentioned above. These words were assumed

to be the new words for the targeted subjects.

Finally, in the narrative text, “So Far, So Good: Recollections of a First-time Father,” 12 words were selected; while in the expository text, “The Power of Nature,” 16 words were selected as shown in Table 1.

Table 1. The Two Reading Texts And The Final Target Words

Reading text	Vocabulary	Number of words selected
So Far, So Good: Recollections of a First-time Father (Narrative text)	Invariably, ominous, trek, Lamaze, delivery room, privileged, snooze, diaper, fade, crawl, swell, efficient	12 words: Adverb = 1 Adjective = 3 Noun = 3 Verb = 5
The Power of Nature (Expository text)	Hurricane, massive, roar, overflow, hillside, debris, spin, tornado, flatten, property, collapse, tsunami, devastate, lava, volcano, landscape	16 words: Adverb = 0 Adjective = 1 Noun = 9 Verb = 6

Step 3: Designing the Multimedia Program

The design of the multimedia program was divided into three steps. First, still images and video clips were designed and developed to illustrate each of the 28 target words. Namely, each target word was attached to the explanation of its meaning with an image or a video clip. Some of these images or video clips were developed by the researchers. Others were downloaded from the Internet for fair use. After the image and video annotations for each target word were

developed, they were seriously examined by an expert with an educational technology background and with experience in developing multimedia programs for English instruction. Also three graduate students with English majors examined the annotations for the appropriateness and suitability in illustrating the selected target words. Later, the visual annotations were revised according to the suggestions of the expert and the graduate school students.

Secondly, the researcher designed the features of the programs and developed the multimedia program with the assistance of a computer programmer. Two reading articles with 28 annotations were the content in this multimedia program. The multimedia

program was designed using web design software, Adobe Dreamweaver CS 3™. The multimedia video group would be provided with textual definition, and dynamic videos on the right screen. The examples were shown in Figures 1 and 2.



Figure 1: Example of the multimedia program for the experimental image group



Figure 2: Example of the multimedia program for the experimental video group

Step4: Designing the Test Instrument

Knowing a lot of words is not enough for being proficient in a second or foreign language (Read, 2000). Learners need to be able to use the vocabulary knowledge effectively for various communicative purposes. Hence, we designed the vocabulary test to assess the effect of three types of treatments on the participants' vocabulary acquisition. The vocabulary test required the test-takers to show their understandings of the words' meanings, and to further demonstrate their abilities to use the words appropriately in their writing.

The participants had 40 minutes to finish the vocabulary test, which divided into two parts. The first part of the test was a recognition test containing 20 multiple-choice questions (testing learners' receptive competence). The second part of the test was a production test in which the test-takers were required to use the given 10 target words in sentences in order to test learners' productive competence.

As far as the production task was concerned, Al-Seghayer (2001) adopted a similar vocabulary assessment in which the test consisted of a recognition section for multiple-choice questions and a production section where the test-taker needed to define the given target words. However, Brown (2004) indicates that using a word in a sentence is more

authentic than defining the target word in vocabulary assessment tasks. Hence, the production task of the vocabulary tests in this study required the test-takers to use the target words in sentences.

Prior to the study, the researchers had asked three experienced English professors to check each multiple-choice item of the vocabulary test and to judge the validity and correctness of the item distracters. Then the testing items were revised and refined according to the suggestions. Two scorers were involved in the scoring process for the production section of the test. In order to avoid human error, subjectivity, and bias, the inter-rater reliability of the production test was 0.99. The reliability in the recognition section of the vocabulary test was 0.72 (Cronbach's α). Hence, overall the reliability of the vocabulary test was acceptable.

This vocabulary test was administered as the immediate post-test, and the delayed post-test. The content of testing items were the same, but the sequence of the items was dissimilar. In addition, only the annotated words shown in the multimedia program were evaluated in the test. The vocabulary test was shown in Appendix A.

Scoring of the vocabulary test

The first part of the test (the recognition task) involved discrete-point items that had only one correct answer. A correct answer received 5 point and an incorrect answer received 0 point. The total points of this section were 100. In

addition, the second part of the test (the production task) contained 10 items that required the test-takers to produce a sentence according to the word involving in each item. Each item was awarded 10 points, and the total points of this section were 100. Table 2 showed

the scoring criteria.

Each participant's final vocabulary test score was calculated as the mean score of their first part of the test score and their second part of the test score.

Table 2. Scoring Criteria for the Production Task on the Vocabulary Test

Categories	Grade	Evaluation criteria
Spelling	3 points	One mistake in spelling will deduct 1 point. Two mistakes in spelling will deduct 2 points. More than two mistakes in spelling will deduct 3 points.
Structure	3 points	One grammatical error will deduct 1 point. Two grammatical errors will deduct 2 points. More than two grammatical errors will deduct 3 points.
Meaning of the sentence	3 points	A sentence addressing the assigned vocabulary and expressing concrete ideas will acquire 3 points. A sentence addressing the assigned vocabulary but missing some points will acquire 2 points. Incomplete ideas of the sentence or somewhat off the assigned vocabulary will acquire 1 point.
Punctuation	1 points	One mistake or more than one will deduct 1 point.
A sentence is scored 0 if it contains no response, is written in the test-taker's native language, or consists only of keystroke characters.		
Total: 100 points (10 points for each item)		
Note: The scoring criteria was developed by referring to the <i>Brown, H. D. (2004). Language Assessment Principles and Classroom Practices. New York, USA: Longman. pp.236</i> , and revised according to the suggestions of two research experts and one native English-speaking teacher.		

Research Design

The purpose of this study was to investigate which multimedia modality—dynamic video clips or still images—were more effective in facilitating L2 vocabulary acquisition. An experimental design was used with three classes (142 participants) being randomly assigned to three different treatment conditions: (1) the control group (N=50): printed text definition alone, (2)

the experimental image group (N=47): textual definition coupled with images, and (3) the experimental video group (N=45): textual definition coupled with videos. Two of the classes as the experimental groups would be introduced to a multimedia-learning program, designed by the researchers to aid the learning of unknown words. The control group would read the same text with textual definition in paper-based form.

The Participants

One hundred and forty-two freshmen in a given university (located in northern Taiwan) participated in the formal study. They were from three classes with different majors—50 students studied in the Department of Business Administration, 47 studied in the Department of Accounting and 45 studied in the Department of Finance. The participants were selected for the present study because of their similar English proficiency level. Each class was randomly assigned to one of three treatment groups: the control group, the experimental-image group, or the experimental-video group.

In order to ensure that the English reading proficiency of these groups were equivalent, their scores attained in the Freshman English course last semester were compared. A one-way ANOVA was conducted to see whether there was a significant difference in the mean scores in the Freshmen English course last semester for the three groups. The results showed that no significant differences existed among the

three groups on the mean scores of the Freshmen English course last semester, [$F(2, 139) = .334, p > .05$]. Hence, these groups were equivalent in their English proficiency prior to the study.

Procedure

The treatments lasted for two weeks, one hour per week. In the beginning of the course, the experimental groups were introduced to the multimedia-learning program. Then, the participants of the experimental groups read a narrative text from the computer program for 25 minutes, while those in the control group read the same text in a paper-based format.

After participants had finished reading the text, they were asked to take an immediate vocabulary post-test for 25 minutes. The two experimental groups and the control group took a paper-and-pencil test. The procedure for the second week would be the same as that of the first week, except that the reading text was changed to be an expository text. Three weeks after the second week of this study, all the participants were asked to take two delayed post-tests (vocabulary tests of the

narrative text and expository text.).

Treatments

For the control group

The control group and the two experimental groups read the same reading articles and target words in this study. However, the annotation types for the target words were the textual definitions. The sentence examples of the words were also presented in a paper-based format. During the treatment, the roles of the teacher and researchers in the classroom were to monitor the participants' performance, and to control the time for the tests. They also made sure that all the participants paid full attention to the reading texts and finished the tests.

For the experimental-image group and the experimental-video group

The multimedia computer program for the two experimental groups (the image-group and the video group) was designed to facilitate the learners' vocabulary acquisition by providing the meanings of the annotated words through links to different types of information, such as textual

definition, still images or dynamic video clips.

The only difference between the image and the video groups was the presented annotation type. Strictly speaking, the image group can read the annotated word's textual definition and understand its meaning through still images, while the video group would be provided with textual definition and video annotation. The readers can click the annotated words as many times as they wished to see the meanings of the target words. During the treatments, the roles of the teachers and researchers in the computer classroom were to be the time keepers and the monitors to check the participants' performances. All participants of the experimental groups were advised to pay full attention to the content of the computer program. They read the text and finished the tests.

Data Analysis

A standard version of the Statistical Package for the Social Science 12.0 (SPSS) was used for all statistical analyses in this study. The vocabulary test was scored by the

students' responses to the test items. In the recognition section of the vocabulary test, five points were given to one correct answer in each multiple-choice question, and zero for an incorrect or missing answer. One scorer was involved in the scoring of the recognition section. In the production task, the participants were asked to make a sentence according to the given word. Each sentence contained 10 points, and the participants' answers were scored by two scorers according to the scoring criteria (See Table 2). To avoid subjectivity of choosing one of the two scorers as the final score for the production section, the mean of the two scores was calculated as the participants' final score for the production section. Each participant would receive total scores ranging from 0 to 100 points on each vocabulary test.

The students' scores on the immediate post-test and the delayed post-test were analyzed to investigate their short- and long-term vocabulary retention. Firstly, descriptive statistics were computed. A one-way analysis of variance (ANOVA) was

conducted to ascertain significant differences in short- and long-term vocabulary retention among the three groups. The independent variable of this study was the treatment (textual definitions in paper-based form, multimedia image annotation, and multimedia video annotation). The scores of the post-test were the dependent variables. Following the analysis of a one-way ANOVA, Post Hoc comparisons using the Turkey test were performed to identify differences between each pair of possible comparisons.

Additionally, the effect size and family of indices that measure the magnitude of a treatment effect, was estimated by the use of Cohen's *d*. Cohen (1988) defined effect size as the difference between two means divided by a standard deviation for the data. The effect size is regarded as small when $d \leq 0.2$, identified as medium when $d \leq 0.5$, and recognized as large when $d \geq 0.8$ (Cohen, 1988, p. 25). Next, the reliabilities of the immediate post-test and the delayed post-test were calculated. The inter-rater reliability in the production section was also

computed.

Results

The mean scores among the three groups on the immediate and the delayed post-tests

The immediate post-test and the delayed post-test results for the three groups are depicted in Table 3. For the control group, the mean score of the immediate post-test was 49.20 (SD=17.91) on the immediate post-test. For the experimental video group, the mean score of the immediate post-test was 73.17 (SD=14.48) on the immediate post-test. For the experimental image group, the mean score

of the immediate post-test was 63.88 (SD=15.41) on the immediate post-test.

Alternatively, on the delayed post-test, the control group's mean score dropped 5.77 points from the mean score of the immediate post-test. The experimental video group's mean score decreased 8.48 points on the delayed post-test when compared to the mean score of the immediate post-test. For the experimental image group, the mean score of the delayed post-test dropped 9.08 points from the mean score of the immediate post-test.

Table 3. Mean Scores On The Immediate Post-Test And the Delayed Post-Test Among The Three Groups

Dependent variables	Control group (N=50)		Experimental image group (N=47)		Experimental video group (N=45)	
	M	SD	M	SD	M	SD
Immediate post-test	49.20	17.91	63.88	15.41	73.17	14.48
Delayed post-test	43.43	21.16	54.80	17.38	64.69	14.72

ANOVA results of the immediate vocabulary post-test among the three groups

This section answered the first research question: “Did the EFL students who used

multimedia annotations in experimental groups (textual definition coupled with images or videos) outperform those using no multimedia annotation (control group) on the immediate post-tests (a recognition section

and a production section)?" It also helped to address the research question three: "Did the experimental group which received dynamic video annotations score significantly higher than the experimental group which received still image annotations on the immediate post-tests (a recognition section and a production section)?"

The one-way ANOVA results showed that there was a statistically significant difference among the three groups on the mean scores of the immediate post-test, [F (2, 139) = 27.58, $p < .001$]. The Turkey HSD post hoc test results further showed that the means of the control and the two experimental groups significantly differed from each other at the $p < .05$ level. The mean

score was 49.20 for the control group, 63.88 for the experimental image group, and 73.17 for the experimental video group on the immediate vocabulary post-test. The experimental video group and the experimental image group scored significantly higher than did the control group on the immediate post-test.

Furthermore, the Turkey HSD post hoc test results revealed that there was a statistically significant difference between the means of the experimental video group and the experimental image group on the immediate post-test. The experimental video group performed significantly higher than the experimental image group in the post-test (see Table 4).

Table 4 .A One-Way ANOVA Summary Table For The Effect Of Different Types Of Multimedia Annotations On The Participants' Immediate Post-test Scores Is Shown Below

Source	Sum of squares	Df	Mean squares	F-ratio	p
Between groups	14209.530	2	7104.765	27.580	.000 *
Within groups	35807.463	139	257.608		
Total	50016.993	141			

* $p < .05$

The effect size of the immediate post-test was medium ($d=.62$) for the comparison between the experimental video

group and the experimental image group. However, the effect size was large ($d=.88$) for the comparison between the experimental

image group and the control group on the immediate post-test. Also, the effect size was larger ($d=1.47$) for the comparison between the experimental video group and the control group on the immediate post-test. Effect size is a measure of the strength of the relationship between two variables (Cohen, 1988). In this case, the multimedia video annotations had a greater effect for the learners on vocabulary acquisition.

ANOVA results of the delayed vocabulary post-test among the three groups

This section answered the research question two: “Did the EFL students who used multimedia annotations in experimental groups (textual definition coupled with images or videos) outperform those students who did not use multimedia annotation (control group) on the delayed post-tests (a recognition section and a production section)?” It also addressed the research question four: “Did the experimental group which received dynamic video annotations score significantly higher than the experimental group which received still

image annotations on the delayed post-tests (a recognition section and a production section)?

The one-way ANOVA results revealed that there was significant difference among the three groups on the mean scores of the delayed post-test, $F(2, 139) = 16.899, p < .001$. The Turkey HSD post hoc test results showed that the means of the control group and the two experimental groups significantly differed from each other. The experimental video group and the experimental image group scored significantly higher than did the control group on the delayed post-test. The mean score was 43.43 for the control group, 54.80 for the experimental image group, and 64.69 for the experimental video group on the delayed post-test. Additionally, the Turkey HSD post hoc test results also demonstrated that there was a significant difference between the means of the experimental video group and the experimental image group. The experimental video group significantly outperformed the experimental image group on the delayed post-test (see Table 5).

Table 5. The One-way ANOVA Summary Table For The Effect Of Different Types Of Multimedia Annotations On the Participants' Delayed Post-Test Scores

Source	Sum of squares	Df	Mean squares	F-ratio	p
Between groups	10992.135	2	5496.068	16.899	.000 *
Within groups	45207.008	139	325.230		
Total	56199.143	141			

* $p < .05$

The effect size of the delayed post-test was medium ($d=.61$) for the comparison between the experimental video group and the experimental image group. The effect size was medium ($d=.59$) for the comparison between the experimental image group and the control group on the delayed post-test.

Furthermore, the effect size remained large ($d=1.17$) for the comparison between the experimental video group and the control group on the delayed post-test. Finally, Table 6 showed the effect size comparison between the immediate post-test and the delayed post-test.

Table 6. Comparing the Effect Sizes Between the Immediate Post-Test And The Delayed Post-test Between Groups

Comparison between groups	Immediate post-test effect size	Delayed post-test effect size
Experimental video group vs. Control group	$d= 1.47$	$d= 1.17$
Experimental image group vs. Control group	$d= 0.88$	$d= 0.59$
Experimental video group vs. Experimental image group	$d= 0.62$	$d= 0.61$

Reliability of the immediate post-test and delayed post-test

The reliability of the immediate post-test and the delayed post-test was conducted to ascertain whether the tests were trustworthy.

Furthermore, there were two raters who participated in the scoring process for the production section in the immediate post-test and delayed post-test. It was necessary to analyze inter-rater reliability to see whether

human error, subjectivity, and any biases existed in the scoring process. The average reliability was 0.70 for the immediate post-test and the delayed post-test, which was a satisfactory reliability level. In other words, reliability analysis showed that the reliability of the tests were acceptable. The average inter-rater reliability was 0.99 for the immediate post-test and the delayed post-test. The results showed that scoring was consistently conducted by the raters.

Discussion and Conclusions

The results of the present study confirmed the findings of the previous studies that multimedia annotations (image and video modes) are more beneficial to vocabulary acquisition as compared to applying textual annotation on a paper-based format (Chun and Plass, 1997; Al-Seghayer, 2001). Additionally, the results of this study also supported the study of Al-Seghayer (2001) which stated that dynamic video was a better mode of multimedia annotation on the vocabulary acquisition for L2 learners. The researchers found that students who had

access to multimedia video annotations performed better in vocabulary tests than the students receiving instruction with multimedia image annotations.

The results of the present study also revealed that annotations, no matter if they were multimedia image annotations with textual definitions, or multimedia video annotations with textual definitions, had positive effects on immediate and delayed post-tests. The findings in this study were consistent with the findings of previous studies that L2 learners can benefit from annotations on vocabulary learning (Hulstijn, 1992; Lyman-Hager and Davis, 1996; Chen, 2002; Davis & Lyman-Hager, 1997; Martinez-Lage, 1997; Nation, 2001; Chun and Plass, 1997; Al-Seghayer, 2001). These annotated words in the multimedia program were significant for the learners in that the words were highlighted with immediate meaning support. Therefore, the probability of wrong guesses for the meaning of the annotated word was greatly reduced. Furthermore, unlike time-consuming

dictionary look-ups, annotations caused minimal interruption of the reading process (Davis & Lyman-Hager, 1997; Martinez-Lage, 1997; Nation, 2001).

This study results indicated that the effect of multimedia annotation on word retention was superior to conventional printed pages (Blake, 1992; Lyman-Hager, Davis, Burnett, & Chennault, 1993; Reinking & Rickman, 1990; Najjar, 1996). There may be several factors contributing to the significantly superior performance in the short-and long-term vocabulary retention for the two experimental groups than for the control group in the present study.

The first factor may be that the computerized annotations provide more extensive and comprehensive annotation than a paper-based format. In other words, multimedia annotations provided immediate access to the annotated information in modalities including textual and visual annotations, affording the readers to access information immediately. The second factor was the possibility to offer images through

multimedia annotations.

Visual information, like an image or a video clip, was attached to a word, which helped readers confirm or reject guesses made about the meaning of a word. Thirdly, multimedia annotated words would not interfere with the readers' reading process, and annotated information met the readers' immediate need for vocabulary assistance. Therefore, the students' performance increased in this low-anxiety learning environment (Davis & Lyman-Hager, 1997; Martinez-Lage, 1997; Nation, 2001). Finally, the results of the study supported the Dual Coding theory (Paivio, 1971) that suggested that learning can be facilitated when a person is exposed to verbal and nonverbal information. Multimedia language learning systems, with its ability to present verbal and nonverbal codes to learners, facilitated and enriched language learning in the present study.

Similar to the results conducted in Al-Seghayer's (2001) study investigating effects between image and video annotations

on the effectiveness of facilitating L2 vocabulary learning, the results of the present study depicted that video annotations with textual definitions were more successful than image annotations with textual definitions in teaching unknown words. Nevertheless, the findings of the present study differed from those obtained by Chun and Plass (1996a). In their study, they investigated incidental vocabulary learning, the effectiveness of different types of annotations on vocabulary learning, and the relationship between look-up behavior and performance on vocabulary tests. They found that words with text plus image annotations were recalled significantly better than those with text plus video annotations. Chun and Plass contended that because images can be viewed for as many times as the readers wished, the mental model of the annotated information can be readily developed. On the contrary, video annotations were usually short in length, so the annotated pieces of information were difficult to be established in the long-term memory. Unlike these findings, this study

showed different results. The participants in the multimedia video group scored significantly higher than those in the multimedia image group on the immediate post-test and delayed post-test.

The different results between Chun and Plass' study (1996a) and the present study can be attributed to the types of visual aids included in the annotations. In Chun and Plass' study, participants were native speakers of English learning German as a foreign language. The principles of the selection for the visual aids included still images that were familiar to the participants and used video scenes that were related to German culture. As a result, videos were unable to reinforce the mental images of the annotated words, while still images contributed a greater effect of retrieval cues for the participants. In contrast, for the present study, the principles of the visual aid selection included still images and dynamic videos that were familiar to the participants and avoided connecting to a particular culture.

Our findings confirmed that dynamic video annotations can better facilitate L2 vocabulary learning in short-and long-term retention than the still image annotations. Dynamic stimuli were more easily remembered and more effective in helping learners to build mental images between the annotated words and their meanings because dynamic stimuli depicted connections more readily (Al-Seghayer, 2001).

Educational Implications

Second language learners need to know a large percentage of the words in order to understand target language input. Studies had shown direct evidence in favor of explicit vocabulary instruction (Paribakht & Wesche, 1997; Zimmerman, 1997). Zimmerman (1997) also indicated that systematic vocabulary instruction may result in better acquisition than simply having students read or learn words from context. Additionally, Grabe and Stoller (1997) pointed out that reading advances vocabulary knowledge, and vocabulary knowledge enhances reading development. With multimedia programs or

software as a learning tool, students are better able to broaden their vocabulary base to improve their reading comprehension.

It should be noted that computers cannot solve all educational problems. They can not react to a student's frustrated look, and computers cannot determine the right kind of encouragement that a qualified teacher can offer when a student's self-esteem needs a boost (Rieber, 1994). In other words, multimedia programs should be regarded as a learning tool by the language teachers for capitalizing on the strengths of computer technology to support vocabulary learning and instructional goals. Jonassen et al. (1999) contended that multimedia programs do not teach students. But rather, students only learn when they construct knowledge and think and learn through experience.

Technology is merely a tool to enable learners to construct knowledge. That is, understanding cannot be conveyed to students through computer technology. Rather, students construct understanding themselves through computer technology (Jonassen,

2000). Therefore, teachers need to carefully select multimedia programs appropriate for learners according to the objectives, difficulty level and instructional design. Also, teachers are expected to take the learners' needs as well as individual differences into consideration. Instructors are suggested to provide sufficient guidance and support for those who learn with multimedia programs.

References

- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5(1), 202-232.
- Anderson, R. C., & Nagy, W. E. (1991). Word meanings. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 690-724). New York: Longman.
- Blake, R. J. (1992). Second language reading on the computer. *ADFL Bulletin*, 24(1), 17-22.
- Brown, H. D. (2004). *Language assessment principles and classroom practices*. (1st ed.). New York, USA: Longman.
- Brucklacher, B., & Gimbert, B. G. (1999). Role-playing software and WebQuests: What's possible with cooperative learning and computers. *Computer in the Schools*, 15(2), 37 - 48.
- Chen, H. H. J. (2002). Effects of L1 and L2 glosses on reading comprehension and vocabulary retention. In *Proceedings of the nineteenth international conference on English teaching & learning in the Republic of China* (pp. 117-129). Crane Publishing, Taipei, ROC.
- Chun, D. M., & Plass, J. L. (1996a). Effects of multimedia annotations on vocabulary acquisition. *Modern Language Journal*, 80, 183-198.
- Chun, D. M., & Plass, J. L. (1996b). Facilitating reading comprehension with multimedia. *System*, 24(4), 503-519.
- Chun, D. M., & Plass, J. L. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology*, 1(1), 60-81.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.).

- Hillsdale, NJ: Lawrence
- Davis, J. N., & Lyman-Hager, M. (1997). Computers and L2 reading: Student performance, student attitudes. *Foreign Language Annals*, 30, 58-72.
- Earlbaum Associates. Davis, N. (1989). Facilitating effects of marginal glosses on foreign language reading. *The Modern Language Journal*, 73(1), 41-48.
- Grabe, W., & Stoller, F. L. (1997). Content-based instruction: Research foundations. In M.A. Snow, & D. M. Brinton (Eds.), *The content-based classroom: Perspectives on integrating language and content* (pp.5-21). NY: Longman.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: experiments in incidental learning. In: P.J.L. Arnaud & H. Bejoint (eds.), *Vocabulary and applied linguistics* (pp.113-125). London: Macmillan.
- Jacobs, G., Dufon, P., & Hong, F. C. (1994). L1 and L2 glosses in reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*, 17(1), 19-28.
- Jonassen, D. H., Peck, K. L., & Wilson, E. G. (1999). *Learning with technology: A constructivist perspective*. Upper Saddle River, NJ: Prentice Hall.
- Jonassen, D. H. (2000). *Computers as mindtools for schools: Engaging critical thinking*. Upper Saddle River, N.J.: Merrill.
- Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals*, 32(1), 89-113.
- Lomicka, L. (1998). To gloss or not to gloss: An investigation of reading comprehension online. *Language Learning and Technology*, 1(2), 41-50.
- Lyman-Hager, M., & Davis, J. N. (1996). The case for computer-mediated reading: Une Vie de Boy. *The French Review*, 69(5), 775-790.

- Lyman-Hager, M., Davis, J. N., Burnett, J., & Chennault, R. (1993). *Une Vie de Boy: Interactive reading in French*. In F. L. Borchardt & E.M.T. Johnson (Eds.), *Proceedings of CALICO 1993 annual symposium on assessment* (pp. 93-97). Durham, NC: Duke University.
- 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, E. (1997). *Multimedia learning: Are we asking the right questions?* *Educational Psychologist*, 32(1), 1-19.
- Najjar, L. J. (1996). *Multimedia information and learning*. *Journal of Educational Multimedia and Hypermedia*, 5, 129-150.
- Nation, I, S, P. (2001). *Learning vocabulary in another language*. NY: Cambridge University Press.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.
- Paivio, A., & Desrochers, A. (1980). *A dual-coding approach to building memory*. *Canadian Journal of Psychology*, 34(4), 388-899.
- Paribakht, S. & Wesche, M. 1997. *Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition*. In Coady, J. & Huckin, T. (Eds.), *Second language vocabulary acquisition* (pp. 174-200). Cambridge: Cambridge University Press.
- Read, J. (2000). *Vocabulary and testing*. In Schmitt N. and McCarthy M. (Eds.), *Vocabulary: Description, acquisition and pedagogy*. (pp. 303-320). Cambridge: Cambridge University Press.
- Reinking, D. & Rickman, S. S. (1990). *The effects of computer-mediated texts on the vocabulary learning and comprehension of intermediate-grade readers*. *Journal of Reading Behavior*, 12(4), 395-411.
- Rieber, L. P. (1994). *Computers, Graphics, & Learning*. Madison, WI: Brown & Benchmark.
- Sun, Y, & Dong, Q. (2004). *An Experiment*

- on Supporting Children's English Vocabulary Learning in Multimedia Context. *Computer Assisted Language Learning*, 17(2), 131-147.
- Tozcu, A., & Coady, J. (2004). Successful learning of frequent vocabulary through CALL also benefits reading comprehension and speed. *Computer Assisted Language Learning*, 17(5), 473-495.
- Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (Eds.), *Second language acquisition: A rationale for pedagogy*. NY: Cambridge University Press.

Appendix A: The Vocabulary Test

A. The First Part (Recognition Task)

A. Please choose the appropriate answer to complete the sentences.

1. Because Claire _____ hangs on the bed every morning, she always arrives late.
 - a. invariably
 - b. variably
 - c. efficiently
 - d. inefficiently
2. Those black clouds are a bit _____.
 - a. grumpy
 - b. ominous
 - c. fortuitous
 - d. imperative
3. To go to Mount Kailash in Tibet, Tom needs to _____ for several days across the desert.
 - a. spin
 - b. trek
 - c. crawl
 - d. roar
4. One of the topics included in _____ classes is massage techniques to ease the pain of labor and to enhance relaxation.
 - a. Lamaze
 - b. Rehabilitation
 - c. Massage
 - d. Foreign language
5. Flowers soon _____ when cut.
 - a. fade
 - b. flourish
 - c. swell
 - d. roar
6. Jack's face was _____ with toothache.
 - a. faded
 - b. contracted
 - c. swelled
 - d. overflowed
7. Normally, a baby _____ before he or she can walk and talk.
 - a. creeps
 - b. crowds
 - c. climbs
 - d. crawls
8. _____ is a room in hospital where women give birth.
 - a. Recovery room
 - b. Delivery room
 - c. Intensive care unit
 - d. Pharmacy
9. The ice _____ when the sun shone on it.
 - a. faded away
 - b. disappeared
 - c. melted
 - d. dissolved
10. The ink _____ all over the desk yesterday.
 - a. evaporated
 - b. spilt
 - c. flowed
 - d. flawed
11. Gambling can have a _____ effect on families.
 - a. privileged
 - b. optimistic
 - c. devastating
 - d. invariable
12. When a washing machine _____, it squeezes water out of the clothes by turning the drum around and around quickly.
 - a. overflows
 - b. breaks down
 - c. swells
 - d. spins
13. When an explosion has destroyed something large, a lot of _____ are left.
 - a. diapers
 - b. property

- c. lava
d. debris
14. A _____ is a very strong wind that goes quickly around in a funnel
- whirly
 - torpedo
 - typhoon
 - tornado
15. This exercise helps to _____ a flabby stomach.
- swell
 - flatten
 - flatter
 - spin
16. There were fears that the roof would _____ because of the weight of snow.
- be flattened
 - collapse
 - be devastated
 - remain intact
17. A patent is a form of intellectual _____.
- poverty
 - property
 - prosperity
 - savings
18. Some _____ are not immediately dangerous because they are not active and have become dormant. Others will never be dangerous again because they are completely extinct.
- floods
 - volcanoes
 - lava
 - plate movement
19. _____ refers to an area that is beautiful to look at or has a particular type of appearance.
- Continent
 - Landlord
 - Basin
 - Landscape

20. When earthquakes happen, they may also cause _____, which damage coastal regions.
- tsunamis
 - continental drift
 - plate movement
 - the collapse of the buildings

B. The Second Part (Production Task)

Please write down the complete sentences using the provided words.

Example (witness) : He witnessed an accident in the park yesterday.

- efficient :
- privileged :
- snooze :
- diaper :
- admit :
- overflow :
- roar :
- massive :
- hillside :
- hurricane :

Abstract

The purpose of this study was to investigate the effects of two types of multimedia modalities: video and image annotations, on L2 vocabulary acquisition. An experimental design was used in this study. One hundred and forty-two participants from three classes were randomly assigned to three treatment conditions: printed text definition alone, printed text definition coupled with images, and printed text definition coupled with videos. The experimental groups were introduced to a multimedia-learning program designed by the researcher for reading comprehension. This program provided the readers with a narrative English text and an exposition English text with a total of 28 annotated words in the modalities of printed text, images, and video. To assess the learning outcomes of students in different treatment conditions, two types of vocabulary tests were administered to the participants after they read the text. The tests were administered as the immediate post-test and the delayed post-test. The results indicated that the experimental video group and the experimental image group significantly outperformed the control group on the immediate post-test and the delayed post-test. Moreover, the experimental video group performed significantly better than the experimental image group on the immediate post-test and the delayed post-test.

Keywords : Second language learning, English vocabulary learning, Multimedia annotations, Computer-assisted language learning