

*Language Learning & Technology*  
Vol. 5, No. 1, January 2001, pp. 202-232

## THE EFFECT OF MULTIMEDIA ANNOTATION MODES ON L2 VOCABULARY ACQUISITION: A COMPARATIVE STUDY

Khalid Al-Seghayer  
University of Pittsburgh

### ABSTRACT

One aspect of second language teaching via multimedia to have received attention over the past few years is the impact of glossing individual vocabulary words through different modalities. This study examines which of the image modalities--dynamic video or still picture--is more effective in aiding vocabulary acquisition. The participants, 30 ESL students, were introduced to a hypermedia-learning program, designed by the researcher for reading comprehension. The program provides users reading a narrative English text with a variety of glosses or annotations for words in the form of printed text, graphics, video, and sound, all of which are intended to aid in the understanding and learning of unknown words. A within-subject design was used in this study with 30 participants being measured under three conditions: printed text definition alone, printed text definition coupled with still pictures, and printed text definition coupled with video clips. In order to assess the efficacy of each mode, a vocabulary test was designed and administered to participants after they had read the English narrative. Two types of tests were administered: recognition and production. In addition, a face-to-face interview was conducted, and questionnaires were distributed. Results of the both tests were analyzed using analysis of variance procedures. The investigation has yielded the conclusion that a video clip is more effective in teaching unknown vocabulary words than a still picture. Among the suggested factors that explain such a result are that video better builds a mental image, better creates curiosity leading to increased concentration, and embodies an advantageous combination of modalities (vivid or dynamic image, sound, and printed text).

### INTRODUCTION

In the realm of second language acquisition (SLA), the most recent effort to enhance the process of language learning has involved computer technology. In this regard, Garrett (1989) points out that the adoption of computer technology in foreign language education is part of a larger phenomenon known as the "new humanism" and "represents one of the most exciting developments coming out of the participation of advanced technology in education." Garrett contends that "new humanism" is an attitude whereby technology helps to integrate the efforts of researchers from different fields. Technology enables the humanists to investigate traditional concerns in novel approaches, exploiting technology potential to build on the values of a given sphere and to create "principled connections" among the discipline of the humanities (p. 104).

Since the initial introduction of computers into the field of second/foreign language education, a large number of practitioners have concurred that this technology holds great potential for language learning (Levy, 1997; Muyskens, 1997; Pennington, 1996; Warschauer, 1996; Warschauer &

Healey, 1998). This belief leads to what is known as Computer-Assisted Language Learning, more commonly referred to as CALL. Even though the field is still young, many language educators are endorsing its use as an essential component in language teaching. Embracing the use of computers seems to be due to the fact that computers are capable of performing multiple tasks and thus are more than simply text processors. The computer can organize, select, and present multiple sensory components.

Among the concerns often raised in the domain of CALL is how to use the potential of a computer to enhance the language learning process and how to use different media types in teaching and learning. The concern has been narrowed to the investigation of the efficacy of presenting information using multiple modalities, such as text, audio, still picture, and dynamic videos in the field of SLA. An area that has recently received attention is the impact of glossing individual vocabulary via annotations embodied by different modes and media. Researchers were inspired by the premise that a variety of glosses for words in various modalities, such as printed text, graphics, dynamic video, and sound, might have differing capacities to facilitate vocabulary acquisition and retention. (Chun & Plass, 1996; Davis & Lyman-Hager, 1997; Lyman-Hager, Davis, Burnett, & Chennault, 1993; Martinez-Lage, 1997).

In the context of second language vocabulary acquisition, this study investigates the relative efficacy of two different annotation modes in a multimedia environment: the printed text definition coupled with a still picture, and the printed text definition coupled with a dynamic video clip. The study's focal issue is to determine which mode of imagery--still picture or dynamic video--is more effective in aiding vocabulary acquisition.

This paper will first summarize research done over the years on the impact of multimedia annotations on second language vocabulary acquisition. Next, the comparison study will be discussed, together with its outcome and implications. Finally, suggestions for further research on the matter will be offered.

## **SURVEY OF RELATED LITERATURE**

A brief overview will be given of previous research on second language vocabulary acquisition enhanced by multimedia annotations. The overview will focus on three issues: the value and effect of multimedia annotations, the impact of electronic glossing, and the efficacy of dynamic videos and static pictures. In addition, related theories such as generative theory and dual-coding theory will be discussed. The results of the existing empirical studies and the assumptions and theories

they embody, along with the assumptions of the foregoing theories, constitute the theoretical basis for this study.

### **The Value and Effect of Multimedia Annotations**

The literature addressing multimedia annotations includes their presentation via printed text, as well as their presentation via the modes of audio delivery, of dynamic video imagery, and of still pictorial imagery. Of chief concern will be the modes that present imagery.

In discussing the utility of multimedia annotations, Davis (1989) notes that glossing through hypertext offers two features: (a) glosses are invisible and therefore do not interrupt the reading process; and (b) readers can obtain as much or as little information as desired regarding a particular concept or word mentioned in the text under study.

Jacobs (1994) also states that computerized glossing is an effective means that aids L2 vocabulary acquisition. Learners and researchers alike can benefit from glossing delivered through the computer. Learners are provided with access to glosses of unknown vocabulary items and a list of the accessed items is created automatically. Researchers can examine the created list to obtain information about individual learners such as their reading strategies and behaviors.

Chun and Plass (1996) emphasize that associating lexical items with different types of media fosters richness of recall cues and increases the likelihood of retention. The rationale is that because words are coded dually in two modes, they are learned better than those coded only in one mode. Dual coding provides more paths for retrieval, and as such, helps learners build two types of recall cues in memory.

Unlike traditional marginal glossing in the printed form, the computerized gloss is appealing, as pointed out by Davis and Lyman-Hager (1997), because the computer's capacity permits us to store more extensive glossing than a printed format does. Furthermore, a computerized gloss does not interrupt the reading process because the glossed item is invisible until the reader clicks on the target word.

Martinez-Lage (1997) contends that the value of multimedia annotations is attributable to different factors. First, multimedia annotations provide immediate access to the available annotated information in a program including textual, audio, and visual annotations. This access provides readers with the desired meaning immediately and allows them to pursue reading without

disturbing the reading process, a problem usually caused by stopping to look up words in a dictionary. Second, multimedia annotations give access to images. Visual information attached to a word enables readers to “confirm or reject hypotheses made about the meaning of a word” (p. 150). Finally, multimedia annotations provide an opportunity for readers to interact with the text making them actively involved in the reading process. Martinez-Lage and Herren (1998) added two aspects to the above merits of multimedia annotations. Multimedia annotations provide educators with novel ways through which they can (a) offer their students a means that might help in developing good reading strategies, and (b) help their students to move beyond the mechanical aspects of the reading process.

Finally, Mayer and Sims (1994) state that one of the functions of a multimedia program is to help learners construct referential connections between two forms of mental representation systems: the verbal and the visual one. These referential connections are more easily built when both verbal and visual materials are presented simultaneously.

Two prominent theories are believed to explain the value and effect of multimedia presentations in language learning environments. Mayer (1997) proposes the generative theory of multimedia learning, drawing on Wittrock’s generative theory and Paivio’s dual-coding theory. The application of the generative theory<sup>1</sup> of multimedia learning to vocabulary learning is based upon the assumption that learners of a second/foreign language have two separate verbal systems (L1 and L2) and a common imagery system. There is a suggestion that translations of words via simultaneous verbal and visual presentations would not only link the two verbal systems, but that this storage in the second verbal system would also have an additive effect on learning (Paivio & Desrochers, 1980). Dual-coding theory (Paivio, 1971) is based on the assumption that memory and cognition are served by two separate systems, one specialized for dealing with verbal information, such as printed words, and the other for nonverbal information, such as pictures or objects. Paivio and Begg (1981) assumed that the two systems are related to the sensory modality as explained in Table 1.

Table 1. The Two Systems That Serve Memory

Sensory Modality	Symbolic Systems	
	Verbal	Nonverbal
Visual	Printed words	Picture or objects
Auditory	Speech sounds	Environmental sounds
Tactual	Braille	Feelable objects
Kinesthetic	Motor feedback from writing	Motor feedback from haptic exploration of objects

From Pavio & Begg (1981, p. 68)

The two systems are presumed to be interconnected but capable of functioning independently. Interconnectedness is the characteristic by which representations in one system can activate those in the other. For instance, pictures can be named and words can evoke images. Independence is the characteristic by which nonverbal and verbal memory codes are aroused either directly by pictures and words, or indirectly by imagery and verbal encoding tasks. The relationship of the two systems has been shown to have some positive effects on recall.

The attributes of multimedia annotation presented in the literature center around the dimensions of reader control, the multiplicity and combination of modes, the immediacy of access, and the absence of interruptions. The computer's capacity to control and orchestrate various modalities/media (e.g., static pictures, sound, animation, video) while affording critical human interaction features such as reader control, immediacy of access, and absence of interruption, has not been attainable through any other conventional instructional medium.

### **The Impact of Electronic Glossing on L2 Vocabulary Acquisition**

The impact of electronic glossing on L2 vocabulary acquisition has also been investigated in the literature. Lyman-Hager et al. (1993) conducted a study to examine the impact of a multimedia program on vocabulary acquisition. The researchers studied two groups: one composed of participants who read an excerpt from a story by F. Oyono (1956), *Une Vie de Boy*, via computer, and another composed of participants who read the story from a conventional printed text. Both groups had access to glosses. The computer group had access to multimedia annotations, while the traditional text group had access to a traditional book with gloss bearing definitions identical to those available to the computer group. Immediately after reading the story, participants were asked to write a recall protocol and, a week later, to take a vocabulary test. The results demonstrated that students who worked with the multimedia program showed a better ability to retain vocabulary than those who worked with the non-computerized text.

In order to examine the efficacy of multimedia annotations, to study incidental vocabulary learning, and to explore the relationship between look-up behavior and vocabulary test performance, Chun and Plass (1996) ran three studies with their second-year students of German using a multimedia program called *CyberBuch*. This program provides annotations through pictures, printed texts, and video. After reading the story, the students took a vocabulary test and wrote a recall protocol. The results indicated that recall for visual annotations was higher than for words annotated with text

alone, with a higher rate for incidental learning of vocabulary, and no correlation between the type of annotation looked up and performance on the vocabulary test. The authors postulated that dual coding is the key factor that helped participants to perform better under the text-picture and text-video conditions. They maintained that words annotated with both verbal and visual modes were learned better than words annotated only with the verbal ones because learners were able to construct referential connections between two forms of mental representation, the verbal and the visual.

Davis and Lyman-Hager (1997) examined participants' performance and attitudes pertaining to computerized L2 reading glosses. Forty-two intermediate level students of French read a glossed excerpt of *Une Vie de Boy* from a computer screen. The researchers administered a multiple-choice task, a written recall protocol, and an exit interview in which they asked participants to indicate their reaction to the program. Students showed positive attitudes toward the computerized glosses. However, the researchers did not find evidence of a relationship between computer use and comprehension. Although different types of glosses were available for consultation, they found that students used primarily the English definition of individual words and expressions. Davis and Lyman-Hager speculated that the positive attitude toward the computer-glossed format was based upon three factors: (a) it provided a coherent understanding of a text due to a lessening of the disruption of the reading process caused by conventional dictionary look-ups; (b) it made the participants more independent since they could find definitions by themselves without asking others to help them; and (c) it contained more material than a traditional dictionary.

Lomicka (1998) investigated the way multimedia annotations influence the level of comprehension. The participants were 12 college students in a second semester French course who read a text under one of three conditions: full glossing, limited glossing, or no glossing. The results indicated that computerized reading with full glossing promoted a deeper level of text comprehension. Lomicka proposed that the multimedia annotations were the key to text comprehension. More specifically, the multimedia annotations affected the generation of causal inferences and the construction of a situation model. That is, the computerized glosses helped learners in the construction of a situation model and led to the generation of causal inferences.

Thus, the existing multimedia annotation literature shows that computerized glossing does indeed have an impact on vocabulary acquisition in particular, and reading comprehension in general. The influence appears to stem from the availability of different types of information, the absence of interruptions during reading, the generation of causal inferences, and the construction of a situation model.

Two other points should also be noted. First, the available studies focused upon vocabulary learning without considering in-depth research on multimedia annotations and factors associated with them. In-depth research on multimedia annotations entails getting beyond retention to include issues like designing multimedia annotation in an individualized fashion to compensate for differences in the learners' verbal and visual abilities. The instructional treatments for selected verbal and visual modules should be designed on the basis of current theory and empirical findings to compensate for specific individual differences. In this respect, a number of areas need to be considered such as how to annotate electronic information, what issues need to be taken into account to make effective electronic annotations, what techniques can be embedded in an electronic environment to assure the absence of interruption and insure that readers will choose to use available aids efficiently in their reading. Second, it is interesting to note that the most common measures in glossing studies were recall protocols and post-reading tests. As pointed out by Myers (1990), both recall protocols and other post-reading tests represent an off-line measure of the product of comprehension. An exception was Lomicka's (1998) study, which used online think-aloud protocols.

## **L2 Vocabulary Learning via Video and Pictures**

The effects of video and pictures on L2 vocabulary acquisition particularly within multimedia environments have not been well explored. However, the effect of visual information on comprehending texts and learning other skills has received wide research attention.

Omaggio (1979) conducted the first empirical study on the impact of pictures on reading comprehension with beginning college French students. She found that pictures produced a significant improvement in reading comprehension in French. More recent research has narrowed the focus to include studies on the effects of visual imagery on L2 vocabulary acquisition. In a critical analysis of L2 vocabulary learning techniques, Oxford and Crookall (1990) stated that most learners are capable of associating new information to concepts in memory by means of meaningful visual images that make learning more efficient. Visual imagery is known to help learners package information more efficiently than they could if using just words alone. Moreover, the pictorial-verbal combination involves many parts of the brain, thus providing greater cognitive power (p. 17).

A study by Snyder and Colon (1988) investigated the influence of both audio and visual aids on facilitating second language acquisition. Two groups were taught for 7 weeks under two different

conditions. One was exposed to a standard curriculum with audio-visuals limited to overhead transparencies, audiotapes with accompanying fill-in pictures, and slides. The other group was exposed to a curriculum enriched with more audio-visual aids, such as additional overhead transparencies, audio tapes and slides, as well as one bulletin board for reviewing material, and another for students to bring in materials from home, pictures from magazines, cut-out dolls to identify and reinforce names for parts of the body, and picture flash cards. After testing both groups on the material covered, it was found that the group provided with additional audio-visual aids performed significantly better in vocabulary retention.

Kost, Foss, and Lenzini (1999) carried out a study comparing the effects of pictorial and textual glosses on incidental vocabulary growth for foreign language learners. Participants were asked to read a passage under one of three glossing conditions: textual gloss alone, pictorial gloss alone, and text combined with pictures. Performance on both production and recognition tests of 14 words was better for those who were allowed to use a combination of text and picture. The theoretical explanation for such results, the authors argue, is that processing information requires different degrees of cognitive effort. The two different representations allow plotting of the picture into one mental model and thereby provide a “stronger bond” than the plotting of the words (p. 94).

Little research exists on the efficacy of video in the domain of L2 vocabulary acquisition. Neuman and Koskinen (1992) state that captioned video with sound provides a semantically enriched context where the visual and the audio lend meaning to the printed words on the screen. Their study compares learning vocabulary through watching television, through reading and listening to a document, and through listening alone. Their results indicated that words were learned and retained best from watching television.

Brown (1993) used a videodisc to evaluate the effects of word frequency and saliency, or importance. In terms of frequency, the author examined two issues: whether the particular context in which the word appears makes a difference, and whether the general frequency of a word in the language permits accurate prediction if the word is going to be learned or not. With respect to word saliency, the author considered two issues: (a) whether giving the word to be learned an emphasis in the program would lead learners to acquire it, and (b) whether presenting the concept in the instructional material but delaying presenting the word associated with it until it appeared in an exercise or gloss would lead learners to acquire that word. Brown found that there was a relationship between word learning and the general frequency index of words learned, but that the specific frequency of words in the videodisc program was not as significant as their general



frequency. In terms of word saliency, she found that word learning was facilitated when concepts were activated by a visual image prior to the presentation of the corresponding word in written form. Brown ascribed this finding to the fact that evoking the concept in the form of an image before the word-form appeared made the vocabulary easier to remember.

Duquette and Painchaud (1996) compared listening to a dialogue with and without visual aids in order to determine which would better allow learners to guess the meaning of new words. Participants listened to a dialogue on the subject of driving a car under two conditions. Under the first condition, they listened to a dialogue while a videotape played. Under the second condition, they listened to the same dialogue with only an audiotape. Results indicated that the learners in the video group made gains on 8 unfamiliar words out of a total of 40, whereas the audio- only group learned only 3. The authors speculated that the “match between prominent visual cues and linguistic ones allows for the inference of unfamiliar words” (p. 158).

The only study that undertook to compare the effect of video clips and static pictures on comprehension and retention of a written passage was carried out by Hanley, Herron, and Cole (1995). Their study showed that a video clip is a more effective organizer than a picture. Twenty-eight college students of French, divided into two groups, took part in the study. Group One was shown a short video clip with French narration. Group Two was presented the same narrative with the teacher reading it aloud while presenting four still pictures of the context. The students in Group One performed significantly better on a comprehension and retention test than those in Group Two. The authors contended that the video clip proved to be more effective in aiding comprehension and retention because video aids in “conceptualizing language,” that is, linking language form to meaning ( p. 63).

General observations can be made about the literature pertaining to the use of video and pictures to teach new vocabulary. It is important that selected pictures and video segments depict meanings precisely and do not contain too much information. In other words, they should be simple illustrations. The research did not compare the effects of videos and pictures on L2 vocabulary acquisition. Both seemed more effective than other modes of presentation, particularly audio. This is simply because students can visualize what the word means and relate the words to actual objects.

This survey of the literature thus makes a strong case for the use of multimedia annotations. The studies unanimously support the use of multimedia to enhance L2 vocabulary acquisition because it is a single interactive presentation environment for diverse instructional resources including printed

texts, photographs, slides, dynamic audio, and dynamic video. The varied and interactive nature of multimedia instruction makes reading, listening, and speaking both engaging and enjoyable. Additionally, the immediacy of access and student independence make learning more efficient and effective. Finally, the combination of media enhances learning because words are dually coded resulting in referential connections that are constructed between the verbal and visual systems.

Multiple studies demonstrate that both dynamic video and still pictures are more effective than instructional modes that lack imagery. However, no study has been conducted with the main aim of comparing the efficacy of dynamic video with that of still pictures. Because each mode is powerful and available, determining which is more effective is a compelling question. The following study attempts to address this issue.

## **THE STUDY**

A key issue in L2 vocabulary acquisition is whether or not learning would be improved if information about an individual lexical item were to be presented simultaneously via multiple modes. Although there are studies of the efficacy that multimodal glosses within a computer program have on vocabulary acquisition, to date there do not seem to be studies focusing, as the main objective, on comparing the efficacy of different types of pictorial annotations. More specifically, no investigation has been done on the difference between the impact of static pictures and dynamic video on learning annotated words. Therefore, this study was designed to investigate the effects of different types of annotations on vocabulary acquisition on the macro scale, and to compare the efficacy of printed text definition coupled with still pictures as opposed to video on the micro scale.

The central question which guided this study is Which is more effective for facilitating vocabulary acquisition: video mode or static picture? Due to the contextual richness, meaningfulness, facility of recall, and cultural authenticity of video, the researcher hypothesized that video is a more effective tool to foster the acquisition of new words in a foreign language. It was anticipated that video annotations would be more effective in teaching unknown vocabulary than still picture annotations. Consequently, it was anticipated that the participants' performance on the vocabulary test will be significantly higher for words that had been annotated with video clips than for words annotated with pictures. Nevertheless, in analyzing the data, a null hypothesis of no difference was tested. It was expected, however, that this null hypothesis would be rejected in favor of the directional hypothesis that video clip annotations are more effective than still picture annotations.

## METHOD

### Participant Selection

To test the aforementioned hypothesis, a convenience sample was selected consisting of 30 ESL participants (17 males, 13 females) who were enrolled in the English Language Institute (ELI) at the University of Pittsburgh. They can be grouped according to their native language as follows: 13 Arabic, 4 Japanese, 6 Korean, 3 Spanish, and 4 Thai speakers. The distribution according to country is shown in Table 2. They were selected with regard to their proficiency level and length of time studying English in the US. They were placed in level three, the intermediate level, based on their performance on the reading section of the institution's placement test. They all scored in the range of 70 (out of 100 points) on the reading section. Each of them had spent at least two semesters in the host environment and had attained intermediate-proficiency TOEFL scores of 450-500. The participants' instructor indicated that they had intermediate proficiency in reading regardless of their language backgrounds. The participants were asked to provide demographic information about their background, length of time studying English, length of stay in the United States, and their TOEFL score. Since this was a within-subject design, all participants were exposed to the same treatment conditions. Note that the participants had different language backgrounds. However, since a within-subject design was used in this study, each participant served as his/her own control, and these different language backgrounds were therefore controlled.

Table 2. Characteristics of Participants

Country	Male	Female	Total
Japan	1	3	4
Korea	2	4	6
Kuwait	5	-	5
Saudi Arabia	7	1	8
Spain	-	3	3
Thailand	1	3	4
Total	17	13	30

### Materials

**Software: Overview of Content and Organization.** The interactive multimedia computer program used in this study was designed by the researcher to enhance L2 vocabulary acquisition by providing readers with the meaning of a target word via hypermedia links to multiple modalities. The computer program provided students who were reading a narrative English text with annotations for target words in various modes--text, graphics, video and sound--all of which

were intended to aid in the understanding and learning of unknown words. Instead of a traditional gloss, the program provided various multimedia links. Users could read the target word's printed textual definition, hear its pronunciation, or view its meaning via a still picture or video.

The program was written in an HTML editor called Dreamweaver 2.0 (1998) and then published on CD-ROM. HTML was chosen as the programming language because it provides for easy integration of hypermedia, and it is suitable for both PC and Mac platforms. The audio component was processed with Cool Edit Pro 1.0 (1997), the pictures were processed with Adobe Photoshop 5.0 (Adobe, 1998), and the video clips were processed with Adobe Premiere 5.0 (Adobe, 1998). For further description concerning other aspects of the program see Appendix A.

***Target Word Criteria.*** The annotated words were controlled for frequency, grammatical category, morphological category, and visual complexity, and for whether they involved concrete or abstract concepts. With respect to conceptual quality or type, using the word frequency corpora of Francis and Kucera (1982), it was found that the 10 video words had a mean of 26 occurrences per million words, the picture words had a mean of 24 occurrences per million words, and the text words had a mean of 27 occurrences per million words. In terms of grammatical category, the selected words for video clips, pictures, and text shared the same characteristics. Each of them had one adjective, two verbs, and seven nouns. Regarding visual complexity, the selected words ranged from 4 to 10 letters in length. The average length of the selected words in each group was as follows: 8 for picture words, 7.5 for both video and for text words. The selected video clips, pictures, and text words all involved concrete concepts. A tracking device was built into the program to ensure that our English definition and annotated modes were exclusively consulted by the participants.

***Passage.*** A narrative passage from Mundahl (1993) was especially adapted selected for this study. The passage tells of a Native American boy who belongs to a tribe called the Lakota. The Lakota people lived on the land that is now a part of the United States. Their culture was closely connected to the Earth. They taught their children that the Earth was their mother, and that they were its caretakers. In this story, a young Lakota boy grows old and watches his world vanish. The text seemed appropriate for intermediate-level ESL learners after applying three criteria in addition to the cloze procedure: syntactic complexity, text length, and content.

The selected passage met the criterion of appropriate syntactic complexity for intermediate ESL learners because it consisted of short, uncomplicated sentences. The mean length of the sentences was seven words, and the percentage of simple sentences in the passage was 79%. By contrast,

the percentage of complex sentences was 1% while that of compound sentences was 20%. The high percentage of simple sentences thus made the story appropriate for intermediate ESL learners.

The adapted text was also selected based upon its length. It had about 1,300 words. Learners at this level are not trained to read long passages. In the English Language Institute (ELI) at the University of Pittsburgh, they are usually assigned passages of about this length. In addition, since learners were asked to perform multiple tasks, a decision was made not to engage them in a long reading.

With regard to the content, there was no reason to assume that any group of ESL learners knew more or less about Native American culture in general, and that of the Lakota tribe in particular than any other group. Therefore, it seemed reasonable to expect that ESL learners with similar residency experience in the US were roughly comparable in their background knowledge of the passage topic. Although the passage contained culture-specific content, it did not require that its readers have much background knowledge about the Lakota tribe. Moreover, the passage provided simple descriptions and explanations for culture-specific events and terminology.

A cloze procedure was also employed to estimate the difficulty of the text. Four intermediate levels students who did not take part in this study, were introduced to the cloze procedure and asked to supply every deleted fifth word in each sentence starting with the second sentence. After multiplying the total number of exact word replacement by two, the students' cloze score was 65%. This indicated that the selected passage met the intermediate-level competency.

## **Design**

It was decided to employ a within-subject design, allowing the same participants to be exposed to the same treatment conditions and then assessing them on the dependent variable after they participated in the experimental treatment. They were asked to read the story included in the multimedia program, to take a vocabulary test, to respond to a questionnaire, and participate in a short interview. In other words, the independent variables, that is, printed text definition alone, printed text definition coupled with still pictures, and printed text definition coupled with video clips, were experimentally manipulated by exposing all participants to the experimental treatment condition. The dependent variable, that is, acquiring new words in a foreign language, was measured by two types of vocabulary tests. The Friedman test, a non-parametric analog to a repeated-measures one-way ANOVA, was employed to see if there were significant differences between printed text definitions coupled with video clips, printed text definitions coupled with

pictures, and printed text definition alone. All the null hypotheses were tested at a .05 level of significance.

## **Procedures**

The experimental procedure consisted of five different phases.

1. First, participants met individually with the researcher at a computer lab where each was asked to fill out a background questionnaire and then given a brief introduction to the program, its objectives, and its methods. The investigator demonstrated to each participant how the program worked, that is, how to move from one section of the story to another, and how to click on an annotated word. The investigator also showed each participant that clicking on words allowed them to hear the selected word pronounced, read a definition, and see either a picture or a video clip. They were told that they could consult the annotated words whenever they wished and as many times as they wished.
2. Second, each participant read the story individually, using the multimedia program which contained multimodal annotations for the more difficult vocabulary items. Participants spent approximately 30 minutes reading the story.
3. Third, after reading the story, participants were asked to take a vocabulary test (described in the next section). Participants spent approximately 15 minutes completing the test.
4. Fourth, participants were asked to respond to a questionnaire asking them to indicate which type of annotation mode helped them learn the annotated words best.
5. Finally, participants engaged in a short interview in which they were asked to indicate whether the particular mode conveyed the meaning of the lexical item.

## **Instruments**

**Vocabulary Test.** A vocabulary test was designed in order to assess the effect of each mode. The participants were asked to take this test after reading the story. They were not informed in advance that they would be tested because it was assumed that if they knew, they would consciously try to learn the new words. It was hoped that attempting to prevent the participants from making such a conscious effort would create a more natural reading environment.

Two types of tests were administered: recognition and production.

- The recognition test required participants to answer multiple-choice items with four alternatives. The test consisted of 21 English words that were annotated in the story, for which the participants were asked to select the corresponding English equivalents. Out of the 21 words, 5 were annotated with a textual definition, 5 with text and image, and 5 with text and video (see Figure 1).

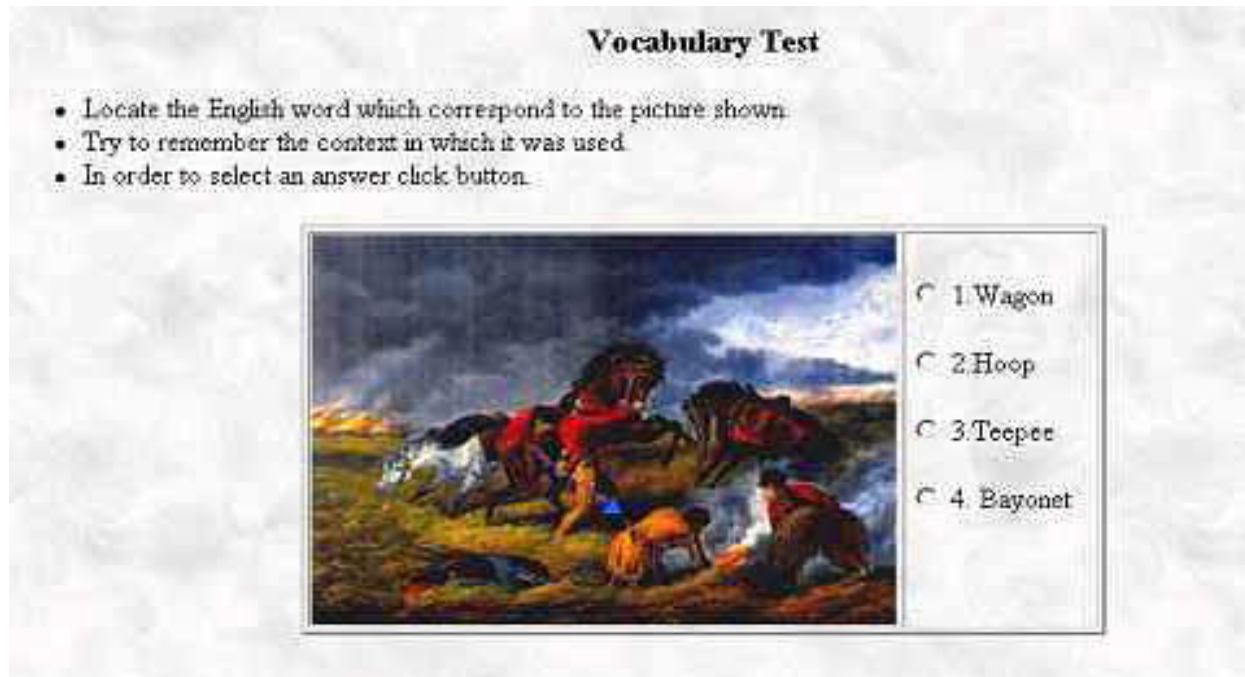


Figure 1. Example of the Recognition Test

- The production test, on the other hand, required the participants to briefly define in English, 6 selected words that were annotated in the story. Out of the six words, two were annotated with only text information (i.e., a definition), two with both text and image, and two with text and video (see Figure 2).

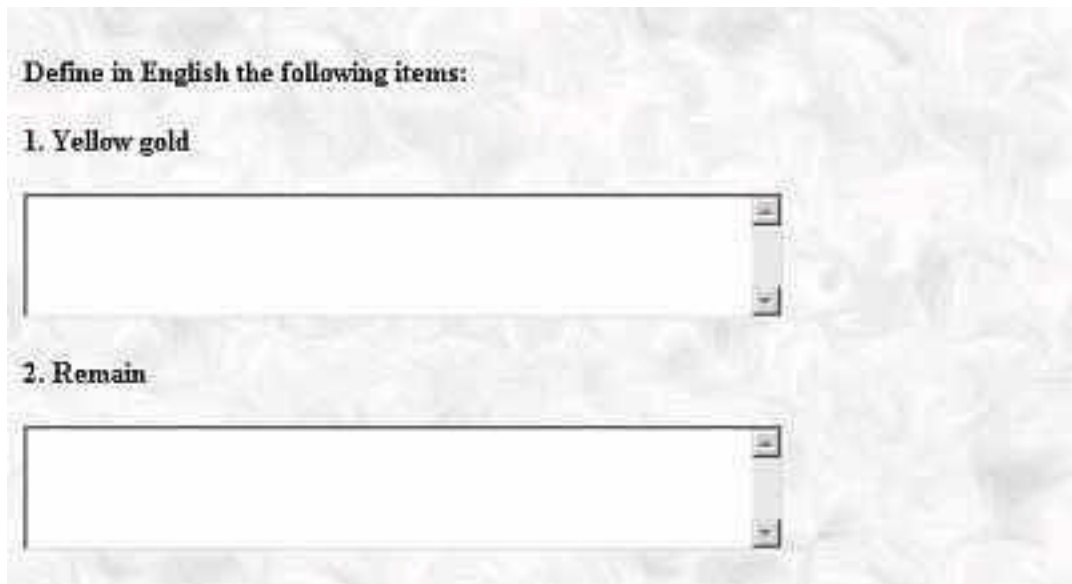


Figure 2. Example of the Production Test

The test design took several issues into account. A criterion was established for the words to be tested. Therefore, words were selected according to their importance to the story, level of difficulty, and part of speech (i.e., 6 verbs, 6 nouns, 9 adjectives/adverbs). In addition, the participants' instructor was consulted to see whether the students had prior familiarity with these words.

The testing was parallel to the modality in which information was presented to the users in the program. Hence, the vocabulary test displayed images, video clips, and definitions, as did the program. Students were asked to indicate whether they knew the word before they read the story. They were asked to select the best corresponding English word from the list after having read the story. They were also informed that once they had selected an answer they could not change it and that they would not be told whether or not their answer was correct. Students' answers were automatically sent to the investigator's e-mail account. The reliability of the recognition and production tests using Kuder-Richardson KR-20 was 0.70.

Because the recognition test involved discrete-point items that have only one correct answer, a correct answer received 1 point and an incorrect answer received 0 points. The same procedure was applied to the production test. For the production text, 1 point was given to a correct or partially correct answer.



## **Questionnaire**

A questionnaire was employed to determine the most effective annotation modes. Participants were asked to rate the helpfulness of each of the three links on a scale of 1-3, with 1 indicating a very helpful link and 3 indicating the least helpful link. Responses from the questionnaires were analyzed and converted into percentages to indicate the type of annotation that the participants first remembered when reading the word on the test and to indicate the type of definition used as the retrieval cue to recall the word.

## **Interview**

In order to determine the most effective annotation modes from the participants point of view, informal individual 15-minute face-to face interviews were conducted. During the interviews, they were asked one closed-ended question and three open-ended questions. The closed-ended question was asked to query participants as to whether or not the three types of annotations conveyed the meaning of a selected lexical item. They were given the choice of agreeing, disagreeing, or taking the position of undecided. The participants were also asked three open-ended questions:

- (a) Which of the three modes helped you to learn the target word best?
- (b) Which of the three modes conveyed the meaning of a selected item?
- (c) What are the technical features that computerized glossing has and that helps you, as a language learner, to learn unfamiliar words?

The open-ended questions were used to allow for more freedom of response, to permit follow-up by the interviewer, to elicit more information from the interviewees, and to check the accuracy of the answers to the closed-ended question.

## **RESULTS AND DISCUSSION**

As stated earlier, this study was carried out to determine which is the more effective annotation mode: text with dynamic video or text with still pictures. In order to reach a conclusion, three instruments were employed. A vocabulary test was administered, questionnaires were distributed, and a face-to-face interview was conducted. The results obtained for each of the foregoing will be considered in turn.

The learning outcome as measured by the vocabulary test was analyzed for each of the 21 words tested. Table 3 shows the overall results on both the recognition and production vocabulary tests. The figures show that words presented under the printed text definition coupled with video clips produced the best results among the three. The mean and percentage of correct answers for words with video and text annotations were 4.3 (87%), compared to 3.3 (67%) for words with picture and text, and 2.7 (53%) for text alone. Table 3 also shows the standard deviations for each of the three modes. For video it was .759 and for picture it was .952. By contrast, for text alone it was 1.586, indicating that these scores were more heterogeneous than the scores for the other two conditions.

Table 3. Overall Results on Both Recognition and Production Vocabulary Test

	Video & Text	Picture & Text	Text
Participants	30	30	30
# Items	7	7	7
% correct answers	87%	67%	53%
Mean	6.1	4.7	4.03
SD	.759	.952	1.586
Minimum	5	3	2
Maximum	7	7	7

By looking at the frequency scores of each mode, one can conclude that most of them were moderate for the video annotation. The range of the scores for text-annotated words was greater than that for the scores on the picture annotated words.

Table 4. Frequencies

Video			Picture			Text		
Score	Freq.	Percent	Score	Freq.	Percent	Score	Freq.	Percent
5	7	23.3	3	2	06.7	2	5	16.7
6	13	43.3	4	12	40.0	3	8	26.7
7	10	33.3	5	10	33.3	4	8	26.7
			6	5	16.7	5	2	06.7
			7	1	03.3	6	4	13.3
						7	3	100.0

In order to see whether there was a difference on vocabulary scores generated from the three modes of annotation, the Friedman test, a non-parametric analog to a repeated-measures one-way ANOVA was utilized. The non-parametric test was used for two reasons. First, the distribution of scores across the annotation modes did not meet the normality assumptions required for ANOVA. Second, the number of score levels obtained within each annotation mode was not sufficiently

continuous for an analysis-of-variance approach; however, they were ordinal in nature and so were appropriate for a non-parametric rank test.

Using the Friedman test, it was found that the vocabulary scores were not all equal across the three annotation conditions,  $X^2 (df= 2, N= 30) = 28.876, p < 0.001$ . To determine specific differences among the three annotation conditions, three pair-wise comparisons were performed using Wilcoxon matched-pair signed-rank tests, adjusting the Type I error rate of 0.05 for the multiple comparisons using the Bonferroni multiple-comparison procedure. Thus, the nominal alpha level for all three tests was 0.0167. It was found that performance on the text-plus-video was significantly better than on the text-plus-picture condition ( $z = 4.018, p < 0.001$ ), indicating that the video condition is more effective than the picture condition. It was also found that the performance on the text only condition, and the text-plus-video condition was significantly different ( $z = 4.219, p < 0.001$ ), showing that the former is more effective than the latter. However, there was no significant difference between performance on the text-plus-picture condition and text alone treatment in the printed text definition coupled with picture condition and the text definition alone condition ( $z = 2.038, p = 0.042$ ).

The results of the study are consistent with the conclusion proposed by the dual coding theory, as suggested by Paivio (1986). The dual coding theory states that when both verbal and visual materials are presented, learners can construct referential connections between these two forms of mental representation, and thus learn more effectively. As expected, all the words that were dually coded (i.e., those with picture plus text and those with video plus text annotations) were learned better than words with text only definitions. However, words with text and video annotations were remembered better than words with text and picture.

The results of the present study differ from those obtained by Chun and Plass (1996). Chun and Plass examined incidental vocabulary learning, the effectiveness of different types of annotations on vocabulary acquisition, and the relationship between look-up behavior and performance on vocabulary tests. Their results showed that across the three studies annotations including printed text with still imagery were remembered better than annotations including printed text with video. An analysis of the mean scores for correct answers on the vocabulary test showed a significant difference between scores for words associated with a picture and printed text, and those associated with a text definition and video. Words with still pictures and definitions were recalled significantly better than words with video and definitions. Chun and Plass postulated the following theoretical explanations for their findings. Since pictures can be viewed for as long as the learner wishes, they allow for the development of a mental model of the information. By contrast,

because video presentations are usually short, they do not allow the information to be established in long-term memory. Chun and Plass state that “the pace of the presentation of information is not sensitive to the cognitive constraints of the learner, and its transient character, therefore, does not allow the student to reflect and to refresh short-term memory” (p. 193). The current study, on the other hand, reported the opposite finding, that is, that participants performed better in both recognition and production vocabulary tests on the words that had textual definitions and video. The analysis of the mean scores for correct answers on the vocabulary test showed a significant difference between scores for words annotated with text-plus-video and text-plus-picture. The suggested theoretical explanations will be discussed below.

The difference in results between the effect of pictorial versus video glosses obtained in this study and those obtained by Chun and Plass (1996) can be attributed to a number of 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. The participants in the Chun and Plass study were native speakers of English studying German as a foreign language. The textual definitions were presented in English. By contrast, the participants in this study came from different linguistic backgrounds and were studying English as a second language. The textual definition was presented in the target language, (i.e., English). In addition, the selection of visual aids could have also had an effect. It could be that Chun and Plass selected still pictures that were familiar to the participants and included video scenes that were related to German culture. As a result, pictures were more helpful as retrieval cues for these participants. In the current study, however, an attempt was made to select visual aids (pictures and video clips) that were not associated with a particular culture. The difference in results could be also traced back to the nature of the assessment used in the two studies. Chun and Plass utilized a vocabulary recognition test that required learners to provide an English equivalent. In the first two studies, they asked the participants to indicate the appropriate English equivalent without accompanying pictures or video clips. In the third study, they presented either a picture or a video clip and asked the participants to choose the German word (from a list of six) that corresponded to the picture, video, or definition. By contrast, the current study employed both production and recognition tests. The recognition test required the participants to select an answer out of a list of four items. The production test, on the other hand, required the participants to briefly define in the target language (i.e., English), a selected word used in the story. The test paralleled the modality in which information was presented to the user in the program. Hence, the vocabulary test, like the program, displayed images, video clips, and definitions.

The second instrument employed to determine the most effective annotation modes was a questionnaire. Participants were asked to rate the helpfulness of each of the three links on a scale of 1-3, with 1 indicating the very helpful link and 3 indicating the least helpful link. Responses from the questionnaires were analyzed and converted into percentages to indicate the type of annotation that the participants first remembered when reading the word on the test and to indicate the type of definition used as the retrieval cue to recall the word. As Figure 3 shows, in the case of the video condition, 86.6% of the participants rated video as very helpful while 70% rated pictures as a very helpful mode. By contrast, only 10% of the participants said the textual definitions were very helpful. In fact, 60% said the textual condition was the least helpful as compared to 10% and 3.3% for the picture and video respectively.

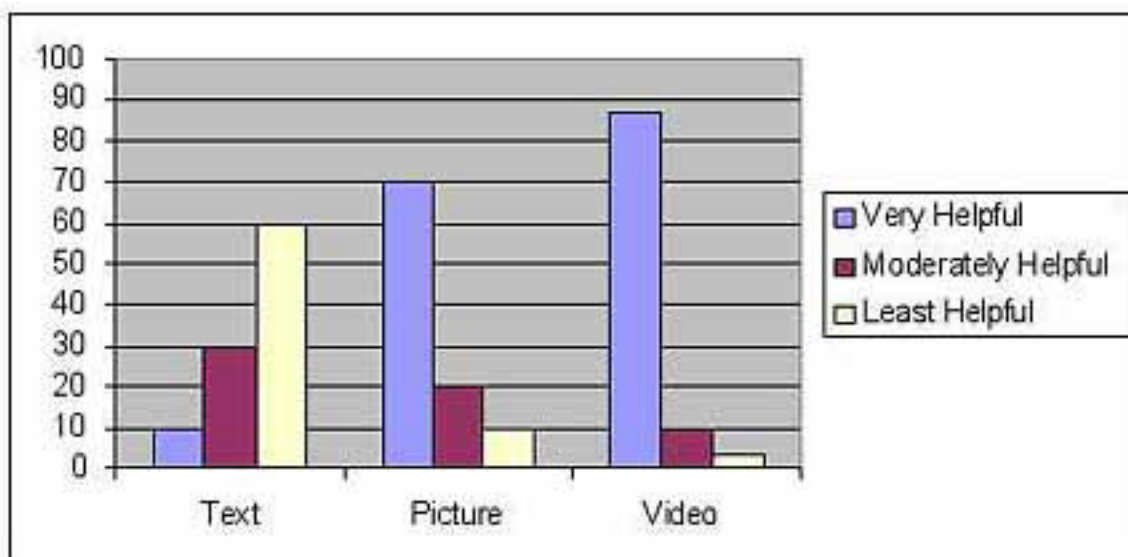


Figure 3. Participants' perspective regarding which of the three media links helped them learn words best

The qualitative ratings or results correspond to the performance on the vocabulary test. Words annotated with video were reported to be more helpful than other types of annotation. Furthermore, video clips were used as the retrieval cue for remembering vocabulary words. The overwhelming majority of students' responses indicated that video annotations provided the best retrieval cue for remembering the meaning of a word.

The last instrument employed to determine the most effective annotation modes was a face-to face interview. Participants were asked a question to determine whether or not the type of annotations conveyed the meaning of a selected lexical item. They were given the choice of agreeing, disagreeing, or taking the position of undecided. As shown in Figure 4, participants agreed (90%) that the video clips showed the meaning of a word more clearly than other links. Almost all participants agreed that both the video (90%) and the still picture (76%) links showed what the

words mean. However, for the textual definitions, they gave different opinions. About 46.6% agreed, 30% disagreed, and 23.33% were undecided.

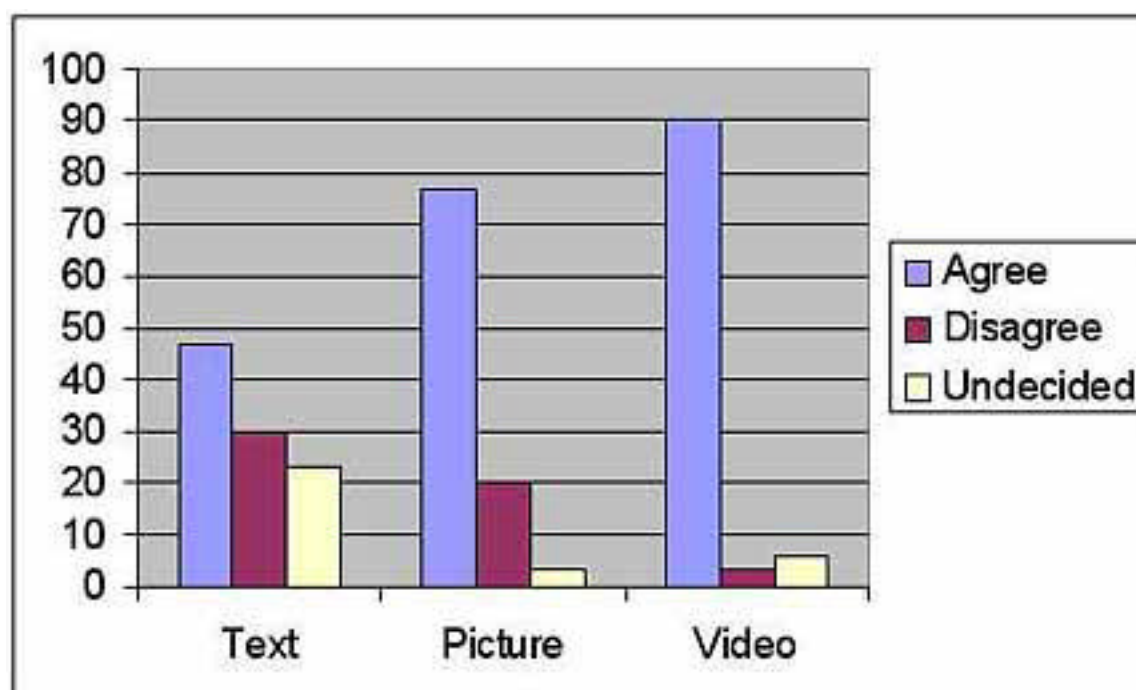


Figure 4. Participants' perspective regarding which type of annotations conveyed the meaning of a selected lexical item

The participants were also asked three open-ended questions. In their responses to the first question, participants acknowledged the helpful contextual support offered by still pictures, and even more so, by video clips. The following are examples of their comments indicating which of the three media helped them to learn the target word best. "Watching live action helped me remember the meaning of these words." "My curiosity pushes me to watch the video to the end because I wanted to know what is going to happen." "Pictures enable me to visualize the meaning of the word."

In their answers to question two, participants indicated that video clips reflected more closely the meaning of the target word than the other two modes. The following statements are examples made by the participants indicating which of the three media depicted the meaning of the target word. "Video explains to me exactly what the word means." "Video clips matches with the meaning of the word better than pictures." "Pictures are good, but I like the videos because they show me the meaning of the word."

In their replies to question three, participants enumerated a number of technical features of computerized glosses. The following are examples of the participants' answers as to the technical features of computerized glosses that helped them learn unfamiliar words. "I can get the meaning of the word right away. I don't have to stop reading and then look up the word in the dictionary." "When studying vocabulary in a program like this, I can hear the pronunciation of the word and see a picture explaining the meaning to me." "I can check the meaning as many times as I want and that does not take me a long time to do."

The above results show that words that had only printed textual annotations received the lowest scores across the three sources of data. By contrast, words annotated both visually and textually had high scores. This result confirms the value of multimedia learning and of dual coding which suggests that words annotated with both textual and visual modes of information are learned better than those annotated with only the textual mode. In this regard, Underwood (1989) suggested that "a commonplace principle of human learning is visual memory. We remember images better than words, hence we remember words better if they are strongly associated with images" (p. 19). The results concur with those of previous studies, particularly Kost et al. (1999), who showed that performance on both production and recognition tests of fourteen words was better for those who were allowed to view pictures. This is also true of the work of Duquette and Painchaud (1996), whose results showed that participants were able to learn the meaning of new words with the help of visual aids. Listening to a dialogue with a videotape enabled participants to acquire 8 unfamiliar words out of a total of 40. By contrast, the other group, which was given another presentation, learned just three unfamiliar words.

What needs to be studied carefully, however, are the results obtained from video annotations as opposed to those from pictures. The learning outcome measured by the vocabulary test and the participants' reports favors video over picture annotations. According to the available data, the dynamic mode (video) is more effective than pictures. This could be attributed to the fact that dynamic stimuli are more easily remembered and more effective in helping learners to build mental images because they more readily depict connections or provide a gestalt. The results of the current study agree with those of Hanley et al. (1995) who attempted to compare the effect of video clip and pictures on comprehension and retention of a written passage. Their study showed that a video clip is a more effective organizer than a picture. The experimental group that was shown a short video clip with French narration scored significantly higher than the control group that was exposed to still pictures with the narrative read by a teacher. The authors contended that the video clips proved to be more effective in aiding comprehension and retention because they facilitated "conceptualizing language," that is, linking language form to meaning.

Another explanation for the present findings is that the participants' attention and concentration were higher when watching the video presentation because they wanted to know what was going to happen in the next segment of the video clips. This could have made presenting the target vocabulary in the form of a video a more memorable experience and, in the long run, a better retrieval cue. Research evidence suggests that the stimulation of curiosity enhances the acquisition of information (Day, 1982; McCombs, 1972).

A further explanation as to why video facilitates vocabulary learning can be attributed to the redundancy hypothesis proposed by Sherwood, Kinzer, Hasselbring, & Bransford (1987) who suggested that video coupled with a related text helped students retain information because they received the same information twice. This between-channel redundancy might facilitate information processing, reduce error and information loss, and increase recall. In addition, video's contextual richness and cultural authenticity might make the information more meaningful, and hence more memorable.

## **CONCLUSIONS**

This study explored the efficacy of multimedia annotations for learning unknown lexical items. The results of this investigation suggest that a video clip in combination with a text definition is more effective in teaching unknown vocabulary than a picture in combination with a text definition. Participants learned and recalled more words when video clips were provided than when pictures were made available. The variety of modality cues can reinforce each other and are linked together in meaningful ways to provide an in-depth experience. Among the possible factors that may explain these results were the following: video better helps learners build a mental image, curiosity increases concentration, and video's combination of modalities (dynamic image and sound) facilitate recall. These results are viewed as a starting point for further exploration into the use of video-based technology within multimedia environments.

## **Pedagogical and Theoretical Implications**

The aforementioned findings and discussion have pedagogical and theoretical implications for language learning and for learning with multimedia in general. Furthermore, they direct our attention to some important designing principles that need to be taken into consideration when developing instructional multimedia materials.



As far as pedagogical implications are concerned, what has been mentioned above constitutes evidence addressing the design of multimedia instruction for second-language learning. Stated more specifically, what has been presented above demonstrates that exposing learners to multiple modalities of presentation (i.e., printed text, sound, picture, or video) produces a language-learning environment which can have a real impact on learning. Another pedagogical implication that can be drawn on the basis of the above findings is that “organizing information in working memory seems to be aided by learners making connections between the verbal and visual system, and this helps in linking information to components of the mental model in long-term memory” (Chun & Plass, 1996, p. 517).

Assuming that students are accustomed to traditional glosses and might not know how other types of glosses foster vocabulary learning, instructors need to spend some time training students to use software in the most beneficial way. Instructors need to explain to the students the efficacy of multimedia annotations in order to ensure that all informational categories available will be consulted by them. Thus, teachers should encourage their learners to exploit the different modalities when looking up the meaning of unfamiliar words. Instructors and program developers should consider including interesting and relevant visual material in their programs in order to increase learners’ motivation to allocate the required mental effort to learn the unknown words.

As far as theoretical implications are concerned, the study’s results support the generative theory of multimedia learning (Mayer, 1997). The basic theme of this theory is that the design of multimedia instruction affects the degree to which learners engage in the cognitive processes required for meaningful learning within the visual and verbal information processing system (Mayer, 1997). This theory suggests that presenting the explanation in words and corresponding illustrations is effective because it helps guide learners’ cognitive processes. In vocabulary learning, learners tend to build visual and verbal cues for retrieving stored information from memory. Storing information in memory is not supposed to be a difficult task, but retrieving it is expected to be difficult. In order to make the task easy for learners, we can provide multiple retrieval cues by integrating two different forms of mental representations.

In order to create effective multimedia instructional materials, two principles need to be considered. The first principle is that instructional materials designed to accommodate individual differences should combine the use of integrated media. Pusack and Otto (1997) contend that it is important to consider that students may have personal modes or combinations of modes that work best for them as individuals; thus, we must never assume that specific media will be put to the same use or have the same effect on all students. The second principle that needs to be taken into consideration is

that, as indicated by Chun and Plass (1997), the selection of the mode of presentation should be based on how it best supports a particular cognitive process. Cognitive processes are said to be supported by the characteristics of the particular mode. Therefore, an instructional designer should make a sound judgment regarding which mode of presentation is more suitable to a given learning situation.

It seems appropriate to conclude this section on implications and design principles by stressing the importance of media combinations in language learning. Information is easy to conceptualize, and the rich learning environment that multimedia instruction creates captures the learners' attention. Thus, the study of multimedia annotations should at least partially provide curriculum designers, program developers, and teachers with a better understanding of what accounts for students' success in acquiring target language vocabulary.

### **Limitations**

There were some limitations to this study. First, the small size of the sample population (N=30) sheds doubt on the validity of the observed significance. A replication study with a greater number of subjects is needed in order to obtain reliable and generalizable results. Second, assessment of the learning outcome was measured only with multiple-choice and production tests. The problem lies with the need to devise alternative assessment techniques that tap various aspects of vocabulary knowledge. Third, the study does not consider analyzing the online individual performance data such as the user's study path and reaction time when interacting with a given item. Such an examination would provide qualitative information about the cognitive process underlying the participant's learning activity. Finally, only short-term retention was studied. It would be interesting to examine the effects of different modes of annotation on long-term retention.

### **Suggestions for Future Research**

This study represents a preliminary effort to empirically examine the efficacy of multimedia annotations on L2 vocabulary acquisition by comparing two modes of annotation embedded in a multimedia program: video clip versus still picture. Further research is needed for a thorough understanding of this issue and for confirmation of the findings. This is especially true when considering that there may be additional variables that would add different intrapersonal effects based on learning style preferences which were excluded from this study. It is recommended that this study be replicated with a larger number of participants from the same background. It would be interesting to compare results across levels of proficiency. Another aspect to be investigated is the method by which non-concrete items can be visualized or represented by non-textual

annotations. These suggested avenues of research might shed more light on L2 vocabulary acquisition involving the computer. They should enlighten us as to which combinations of media will enhance L2 vocabulary learning the most. Finally, it is hoped that the outcome of this study will be of some use to future research studies.

## APPENDIX A: Further Description Concerning Other Aspects of the Program

Screen layout design principles were considered. The screen was divided into four frames. The top frame was reserved for the title of the story, the left frame was reserved for the text, the right frame was used for the visual aids, and the bottom frame was used for navigation. Several aesthetic features were employed in the screen display such as colors, stars, boldface, and graphics. Different colors were deliberately assigned to different sections to maintain the users' interest and attention. Bright colors that could annoy and distract the users were avoided.

The adapted text had six sections presented on the left side of the screen. Each section was accompanied by a picture on the right side representing the theme of the section. For example, the first section in the narrative portrayed the ways through which a boy's father taught him the history of his nation. Hence, a picture was selected to reflect this theme. While going over the story, users could look up any of the annotated words. Thirty words were annotated and each annotated word was marked with an asterisk (\*). All the textual definitions for the annotated words were taken from *The Longman Dictionary of Contemporary English* (1987). Ten words had additional glosses in the form of color pictures, and 10 had additional glosses in the form of short videos. In order to look up a gloss, students clicked on the word, causing the annotations (definition, picture, video) to appear on the right side of the screen (see Figure 5). All annotations were accompanied by an audio component, that is, a native speaker pronouncing each individual word. The following are some examples of different types of annotations.

If users decided to click on the word *ripe*, they would read only the definition (*ripe*: [fruit] ready to be eaten) and hear the word pronounced. If users chose to click on the word *melt*, they would read the definition (to cause a solid to become liquid: the sun melted the snow), see a color picture (illustrating melted snow), and hear the word pronounced. If users clicked on the word *lullaby*, they would read the definition (a pleasant song used for putting children to sleep), watch a 20-second video segment, showing a woman putting a baby on her lap and singing a lullaby to put it to sleep, and hear the word pronounced.



Figure 5. Example of an annotated word accompanied by a color picture

## **NOTE**

1. The generative theory of multimedia learning draws on two studies. The first is Wittrock's (1990) study of generative theory. The second is Paivio's (1986) study of dual-coding theory as well as extensions of studies that test the conditions under which the presentation of visual and verbal material promotes learning.

The generative theory of multimedia posits that learners engage in three major processes--selecting, organizing, and integrating--when they are presented with visual and verbal information such as illustrations and text. When presented with text, the learner must first select relevant words to be retained as a text base in verbal working memory. When presented with illustrations, the learner must select relevant images to be retained as an image base in visual working memory. Second, the learner must organize the text base into a coherent verbal representation and the image base into a coherent visual representation. Third, the learner must integrate the verbal and visual representations by making one-to-one connections between the features of the two. According to the generative theory, meaningful learning is enhanced when a learner can construct and coordinate visual and verbal representations of the same material.

## **ACKNOWLEDGEMENTS**

I am indebted to Dr. Richard Donato for his guidance, support, and encouragement. Heartfelt thanks are also expressed to Dr. Keiko Koda for her encouragement and insightful suggestions that shaped the first draft of this manuscript. My deepest gratitude is extended to the anonymous reviewers for their valuable suggestions.

## **ABOUT THE AUTHOR**

Khalid Al Seghayer is a PhD candidate in the Language and Literacy Program with a specialization in Foreign Language Education/ Applied Linguistics in the Department of Instruction and Learning at the University of Pittsburgh. His research interests include Computer-Assisted Language Learning and Second Language Reading Acquisition.

E-mail: [khast5+@pitt.edu](mailto:khast5+@pitt.edu)

## **REFERENCES**

Adobe Photoshop 5.0 [Computer Software]. (1998). Mountain View, CA: Adobe Systems.

Adobe Premiere 5.0 [Computer Software]. (1998). Mountain View, CA: Adobe Systems.

Ausubel, D. (1978). *Educational Psychology: A Cognitive View*. New York: Holt, Rinehart and Winston.

Brown, C. (1993). Factors affecting the acquisition of vocabulary: Frequency and saliency of words. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 263-286). Norwood, NJ: Ablex.

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

- Chun, D., & Plass, J. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology*, 1(1), 60-81. Retrieved November 30, 2000 from the World Wide Web: [http://llt.msu.edu/vol1num1/chun\\_plass/default.html](http://llt.msu.edu/vol1num1/chun_plass/default.html).
- Cool Edit Pro1.0 [Computer Software]. (1997). Phoenix, AZ: Syntrillium Software Corporation.
- Davis, N. (1989). Facilitating effects of marginal glosses on foreign language reading. *The Modern Language Journal*, 73(1), 41-48.
- Davis, N., & Lyman-Hager, M. (1997). Computers and L2 reading: Student performance, student attitudes. *Foreign Language Annals*, 30(1), 58-72.
- Day, I. (1982). Curiosity and the interested explorer. *Performance and Instruction*, 21(4), 19-22.
- Dreamweaver 2.0 [Computer Software]. (1998). San Francisco, CA: Macromedia, Inc.
- Duquette, L., & Painchaud, G. (1996). A Comparison of vocabulary acquisition in audio and video contexts. *The Canadian Modern Language Review*, 54(1) 143-172.
- Francis, W., & Kucera, A. (1982). *Frequency analysis of English usage: lexicon and grammar*. Boston: Houghton Mufflin Company
- Garrett, N. (1989). Computers in foreign language teaching and research: A new humanism. In William H. Graves (Ed.), *Computing across the curriculum: Academic perspectives* (pp. 101-139). McKinney, TX: Academic Computing Publications, Inc.
- Hanley, J., Herron, C., & Cole, S. (1995). Using video as advance organizer to a written passage in the FLES classroom. *The Modern Language Journal*, 79(1), 57-66.
- Jacobs, G. (1994). What lurks in the margin: Use of vocabulary glosses as a strategy in second language reading. *Issues in Applied Linguistics*, 4(1), 115-137.
- Kost, C., Foss, P., & Lenzini, J. (1999). Textual and pictorial gloss: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals*, 32(1), 89-113.
- Levy, M. (1997). *Computer-assisted language learning: context and conceptualization*. Oxford, UK: Oxford University Press.
- Lomicka, L. (1998). To gloss or not to gloss: An investigation of reading comprehension online. *Language Learning and Technology*, 1(2), 41-50.
- Longman dictionary of contemporary English. (1987). London: Longman.
- Lyman-Hager, M., Davis, N., Burnett, J., & Chennault, R. (1993). *Us 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.
- Martinez-Lage, A., & Herren, D. (1998). Challenges and opportunities: Curriculum pressures in the technological present. In J. Harper, M. Lively, & M. Williams (Eds.), *The coming of age of the*

profession: Issues and emerging ideas for teaching of foreign languages (pp. 141-167). Boston: Heinle & Heinle Publishers.

Mayer, E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, 32(1), 1-19.

Mayer, E., & Sims, K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, 86(3), 389-401.

McCombs, L. (1972). The validation of a measure of state epistemic curiosity in a computer-assisted learning situation. Paper presented at the annual meeting of the American Psychological Association, Honolulu, HI.

Mundahl, J. (1993). *Tales of dreams, A multicultural reader*. New York: Addison Wesley Publishing Company, Inc.

Muyskens, J. (Ed.). (1997). *New ways of learning and teaching: focus on technology and foreign language education*. Boston: Heinle & Heinle Publishers.

Myers, L. (1990). Causal relatedness and text comprehension. In D. A. Balota, G. F. d'Arcais, & K. Rayner (Eds.), *Comprehension process in reading* (pp. 361-375). Hillsdale, NJ: Erlbaum.

Neuman, B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, 27(1), 95-106.

Omaggio, C. (1979). Picture and second language comprehension: Do they help? *Foreign Language Annals*, 12(2), 107-116.

Oxford, R., & Crookall, D. (1990). Vocabulary learning: "A critical Analysis of techniques." *TESL Canada Journal* 7(2), 9-30.

Oyono, F. (1956). *Une Vie de Boy*. Paris: Julliard.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.

Paivio, A. (1986). *Mental representation: A dual-coding approach*. New York: Oxford University Press.

Paivio, A., & Begg, I. (1981). *Psychology of language*. Englewood Cliffs, NJ: Prentice Hall.

Paivio, A., & Desrochers, A. (1980). A dual -coding approach to building memory. *Canadian Journal of Psychology*, 34(4), 388-899.

Pennington, M. (Ed.). (1996). *The power of CALL*. Houston, TX: Athelstan.

Pusack, J., & Otto, S. (1997). Taking control of media. In M. Bush & R. Terry (Eds), *Technology enhanced language learning* (pp. 1-64). Lincolnwood, IL: National Textbook Company.

Sherwood, R. D., Kinzer, C. K., Hasselbring, T. S., & Bransford, J. D. (1987). Macro-contexts for learning: Initial Findings and issues. *Applied Cognitive Psychology*, 1(2), 93-108.

Snyder, H., & Colon, I. (1988). Foreign language acquisition and audio-visual aids. *Foreign Language Annals*, 21(4), 343-384.

Underwood, J. (1989). HyperCard and interactive video. *CALICO*, 6(3),7-20.

Warschauer, M. (1996). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp. 3-20). Tokyo: Logos International.

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching*, 31, 57-71.

Wittrock, C. (1990). Generative process of comprehension. *Education Psychologist*, 24(4), 345-376.

[About LLT](#) | [Subscribe](#) | [Information for Contributors](#) | **Masthead** | [Archives](#)