

## Effects of Text, Audio, and Graphic Aids in Multimedia Instruction for Vocabulary Learning

**Daesang Kim**

Office of Academic Affairs, Saint Mary-of-the-Woods College, USA // Tel: +1 812 535 5273 // [dkim@smwc.edu](mailto:dkim@smwc.edu)

**David A. Gilman**

Department of Curriculum, Instruction, and Media Technology, Indiana State University, USA // Tel: +1 812 237 2925 // [dgilman1@isugw.indstate.edu](mailto:dgilman1@isugw.indstate.edu)

### ABSTRACT

This study is an investigation of the use of multimedia components such as visual text, spoken text, and graphics in a Web-based self-instruction program to increase learners' English vocabulary learning at Myungin Middle School in Seoul, South Korea. A total of 172 middle school students (14 years of age) in five classes participated in the study. Each individual was required to complete several testing instruments such as a pretest, posttest, retention test, and attitude inventory. Participants learned better when they received "visual text and added graphics" or "visual text, added spoken text, and added graphics" instruction. Although the added multimedia components required learners to spend more time on the instruction, the extra time was not significant. The results lead one to conclude that an effective way to improve learning of English vocabulary is to offer graphics that illustrate what the vocabulary means.

### Keywords

Multimedia Learning, Web-based Self-Instruction, Admissible Probability Measurement Procedures, English as a Foreign Language Vocabulary Learning

### Introduction

With computer technology, Web-based learning has become a common choice in education institutions (Bauer, 2002, p. 31). Furthermore, the variety of media such as text, graphics, audio, and video for delivering content has attracted many instructors and students to use the Internet for distance education (Ali, 2003). These multimedia components get and hold learners' interest, which many researchers believe is important when teaching the video generation (Jonassen, 2000, p. 208). Visual text and graphics are some of the most popular tools in on-line learning. In many cases, graphics can be used to represent important information and are often used for supporting text (Newby, Stepich, Lehman, & Russell, 1996, p. 103). Using these techniques, the most widely used asynchronous on-line learning tool is courses primarily posted in visual text and static graphics (Liles, 2004).

English as a Foreign Language (EFL) learners often adopt various strategies to memorize vocabulary words. For instance, vocabulary learning is often used with strategies such as word lists or paired associations in which new words are presented with their translations. These strategies often require learners to memorize pair associations directly (Sun & Dong, 2004). Meara (1996) found that many researchers in vocabulary learning studies have explored various methods of vocabulary presentation and their corresponding effectiveness in retention. Some earlier studies found the following:

- There is a limit to the number of words that can be learned at one time (Crothers & Suppes, 1967, chap. 4).
- Rote repetition appears less efficient than using spaced recall and structured review (Atkinson, 1972).
- Silent repetition and silent writing are less effective than repeating the words aloud (Meara, 1996).

Brown and Payne (Hatch & Brown, 1995, p. 383) have identified five steps to vocabulary acquisition: (a) having sources for encountering new words; (b) getting a clear image, either visual or auditory or both, of the forms of the new words; (c) learning the meaning of the words; (d) making a strong memory connection between the forms and the meanings of the words; and (e) using the words.

Recently, a number of researchers have discussed the benefits of presenting information using multimedia components such as visual text, spoken text, graphics, and videos on language learning (Al-Seghayer, 2001; Chun & Plass, 1997; Duquette & Painchaud, 1996; Ehsani & Knodt, 1998). In their studies, information presented in text,

spoken words, graphics, and video formats can be integrated to create an authentic, attractive, and multi-sensory language context for EFL learners (Sun & Dong, 2004). Kost, Foss, and Lenzini (1999) found that EFL learners performed better on both production and recognition vocabulary tests when they were allowed to use a combination of visual text and graphics.

Designing pedagogically effective multimedia instruction in language learning based on theories has been an important issue (Chapelle, 1998; Hoven, 1999; Liu, Moore, Graham, & Lee, 2002; Watts, 1997). Mayer and Moreno (2002) focused on a cognitive theory of multimedia learning which combines dual coding theory (Paivio, 1986, chap. 4; Sadoski & Paivio, 2001, chap. 3), cognitive load theory (Sweller, Van Merriënboer, & Paas, 1998), and constructivist learning theory (Novak, 1998, chap. 3; Vygotsky, 1978, chap. 6). From dual coding theory they adopted the idea that verbal stimuli and nonverbal stimuli detected by our sensory systems are processed in different systems of the brain (verbal system and nonverbal system). From cognitive load theory they adopted the idea that “humans are limited in the amount of information that they can process in each channel at one time” (Mayer, 2001, p. 44). Sweller et al. explained that redundant memory load is caused by “the presentation format of instructions extraneous load” (Tabbers, Martens, & Merriënboer, 2004, p. 72). Mayer and Moreno (2002) finally concluded that “presenting too many elements to be processed in visual or verbal working can lead to overload” (p. 111). They also took the idea from constructivist learning theory that “meaningful learning occurs when learners actively select relevant information, organize it into coherent representations, and integrate it with other knowledge” (p. 111).

Mayer and Moreno (2002) found the following interesting results:

- Providing words with narration and animation helped learners’ performance more than words alone.
- Reducing the number of unneeded words and sounds helped learners’ performance.
- Providing words with narration helped learners’ performance more than on-screen text.
- Providing words as narration and animation helped learners’ performance more than narration, animation, and on-screen text.

Reducing the amount of on-screen text makes more area available for graphics and labeled illustrations, which are necessary tools for teaching certain types of concepts. Some studies indicated that including the visual text in the illustration and labeling the illustrations improved learning (Koroghlanian & Klein, 2004).

However, as Mayer (2001) states, “all multimedia messages are not equally effective” (p. 79). For example, Mayer concluded “Schnotz, Bannert, and Seufert (in press) reported situations in which some learners reduced the amount of attention they paid to text when pictures were added” (p. 79). Tabbers et al. (2004) concluded that replacing visual text with spoken text and added graphics to the visual text both do not easily generalize to non-laboratory settings.

By better understanding the effect of individual components of multimedia, language educators will be able to design effective instruction for EFL learners. This study is an investigation of the effect of multimedia components such as visual text, spoken text, and graphics on increasing learning or decreasing redundant memory load in English vocabulary learning.

## **Method**

The primary objective of this research is to study the effects of six methods of instruction in a Web-based self-instruction program: visual text (Group A), visual text and added spoken text (Group B), visual text, and added graphics (Group C), visual text, added graphics, and added spoken text (Group D), reduced visual text and added spoken text (Group E), and reduced visual text, added graphics, and added spoken text (Group F). The researchers investigated the use of multimedia components such as visual text, spoken text, and graphics in a Web-based self-instruction program to increase English vocabulary learning at Myungin Middle School (MMS) in Seoul, South Korea.

Multiple choice questions are usually used to test a student's ability to recall information, to interpret data or diagrams, and to analyze and evaluate material. In this study, Shuford's Admissible Probability Measurement (APM) procedure was used to reduce guessing scores on the multiple choice tests (Gilman, 1967, p. 27). Shuford claimed that APM procedures operate scoring systems with a very unique property that guarantees that any student can maximize his expected score if and only if the student honestly reflects his or her degree of belief probabilities. The

formula depends only on the probability assigned to the correct answer and not on probabilities assigned to the other incorrect alternatives (Gilman, 1967, p. 27). The score obtained from an expressed probability to a correct response ( $r_k$ ) is expressed as a function  $g_k(r_k)$  such that

$$g_k(r_k) = \begin{array}{ll} 1 + \log r_k & \text{for } 0.01 < r_k \leq 1 \\ -1 & \text{for } 0.00 \leq r_k \leq 0.01 \end{array}$$

The possible scores range from minus one to plus one and are fairly related to the practice of giving the student one point for each correct answer and subtracting one point for each incorrect answer in order to discourage guessing in multiple choice testing. In other words, a value of 1 indicates that the student's response is correct for a question and a value of -1 indicates that the student's response is incorrect for a question. For instance, a student whose responses are all correct on a 30-item multiple choice test will receive 30 points. A student whose responses are all wrong for a 30-item multiple choice test will receive -30 points for the test.

There were two forms of response for each multiple-choice question. The first was for the response to each question (a, b, c, or d). The second required the learner to write a number from 0 to 100 to indicate how sure the student was that his or her response was correct. For instance, a student who was 100% sure that the response was correct would put "100" in the second space. If he or she was not completely sure, the second space should contain a smaller number. This number constitutes the degree of certainty score. From there, the admissible probability score is obtained by applying the formula listed above. For instance, if the student was 100% sure and his or her response was correct for a question, the admissible probability score would be 1 for the question; if the student was 0% sure that his or her response was correct for a question, the admissible probability score would be -1 for the question. If the student's response was wrong for a question, the admissible probability score would be -1 for the question.

## Research Questions

The research questions for this study were:

1. What are the differences in original learning among students who are taught under the six methods of instruction conditions as measured by raw score, mean degree of certainty estimate, and an admissible probability scoring procedure?
2. What are the differences in time to complete instruction among students who are taught under the six methods of instruction?
3. What are the differences in students' attitude toward instruction among students who are taught under the six methods of instruction conditions as measured by their scores on the attitude inventory?

## Participants

A total of 172 middle school students (14 years of age) in five classes participated in the study. The students had no previous experience with computer-assisted instruction in English vocabulary learning. All sampling procedures were random. All participants were separated into six groups, 43 students (Group A), 22 students (Group B), 34 students (Group C), 24 students (Group D), 24 students (Group E), and 25 students (Group F). The groups that used spoken text format (Group B, D, E, and F) were smaller than the non-spoken text groups (Group A and C) because many computers in the classroom did not have sound systems with headsets or speakers.

## Materials

The topic of the Web-based self-instruction was English vocabulary learning. The design of the English vocabulary instruction was based on the following criteria:

1. The items of English vocabulary were of appropriate difficulty level for Korean middle school students.
2. Graphics supported visual text.
3. Graphics were available for cueing meaning of the word from static or animated images.
4. Spoken text was used as narration to support visual text.
5. Reduced the amount of visual text on a screen left more area available for graphics and spoken text.

The illustration of the criterion 3, 4, and 5 are shown in Figure 1 with one example.

---

Added graphic



---

Added spoken text  
(Translation in English)

Narration1: "Word"  
Narration2: "Taut"  
Narration3: "Definition"  
Narration4: "pulled or stretched tight"  
Narration5: "Example sentence"  
Narration6: "The rope is stretched taut by the two men"


---

Reduced visual text

**Word:**  min  sec

**Taut**

**Definition:**



**Example sentence:**  
**The rope is stretched (taut) by the two men.**

Audio Player

◀ Play Stop ▶

Reduced text1: definition of the vocabulary

Reduced text2: description of the example sentence

---

*Figure 1: The example to describe criteria 3, 4, and 5*

The items of English vocabulary to be learned by students on the Web-based self-instruction program were: 1) tether, 2) wither, 3) tumble, 4) separate, 5) gorge, 6) fetter, 7) beacon, 8) crest, 9) awl, 10) ditch, 11) entice, 12) taut, 13) quench, 14) wizen, and 15) waylay. The length of each lesson was a maximum of 30 minutes and the time was controlled by computer. The six groups of Web-based self-instruction program are shown in Figure 2.

### Procedures

Each participant was required to complete a pretest, posttest, retention test, and attitude inventory for the study. A pretest was administered to the participants one week prior to the study. All students among the six different groups were required to respond to 30 questions regarding the vocabulary with a number from 0 to 100 to indicate their degree of certainty on the pretest. One week later, all students among the six different groups received multimedia instruction through a Web-based self-instruction program. Items of vocabulary were projected on the computer screen through the program. Then, all students among the six different groups were required to respond to 30

questions regarding the vocabulary with a number from 0 to 100 to indicate their degree of certainty on the posttest. Approximately one week later, all participants were required to again respond to 30 questions regarding the vocabulary with a number from 0 to 100 to indicate their degree of certainty on the retention test. In addition, all students were required to complete the 40 items of the attitude inventory.

Word:  min  sec

**Taut**

**Definition:**  
pulled or stretched tight

**Example sentence:**  
The rope is stretched (taut) by the two men.

⏪ ⏩

Group A (visual text)

Word:  min  sec

**Taut**

**Definition:**  
pulled or stretched tight

**Example sentence:**  
The rope is stretched (taut) by the two men.

Audio Player  
▶ Play ⏹ Stop ⏩


⏪ ⏩

Group B (visual text and added spoken text)

Word:  min  sec

**Taut**

**Definition:**  
pulled or stretched tight



**Example sentence:**  
The rope is stretched (taut) by the two men.


⏪ ⏩

Group C (visual text and added graphics)

Word:  min  sec

**Taut**

**Definition:**  
pulled or stretched tight



**Example sentence:**  
The rope is stretched (taut) by the two men.

Audio Player  
▶ Play ⏹ Stop ⏩

⏪ ⏩

Group D (visual text, added spoken text, and added graphics)

Word:  min  sec

**Taut**

**Definition:**

**Example sentence:**  
The rope is stretched (taut) by the two men.

Audio Player  
▶ Play ⏹ Stop ⏩


⏪ ⏩

Group E (reduced visual text and added spoken text)

Word:  min  sec

**Taut**

**Definition:**



**Example sentence:**  
The rope is stretched (taut) by the two men.

Audio Player  
▶ Play ⏹ Stop ⏩

⏪ ⏩

Group F (reduced visual text, added spoken text, and added graphics)

Figure 2: An example showing frames of Groups (Translation into English)

## Results

Data were analyzed using two analyses - mixed factorial design (the split-plot analysis of variance) and one-way ANOVA - to evaluate the results from the study with regard to the following variables:

1. Student's raw scores on
  - Pretest
  - Posttest
  - Retention test
2. Student's mean degree of certainty estimates on
  - Pretest
  - Posttest
  - Retention test
3. Shuford Admissible Probability Scores on
  - Pretest
  - Posttest
  - Retention test
4. Time required to complete multimedia instruction
5. Student attitude inventory

Table 1 presents the results of analyses of all variances with mean scores and standard deviation for each group on the pretest, posttest, and retention test.

Table 1. Mean scores and standard deviations for each group

		Group A (n = 43)	Group B (n = 22)	Group C (n = 34)	Group D (n = 24)	Group E (n = 24)	Group F (n = 25)
<i>Raw Scores</i>	Pretest	M = 10.67 D = 5.43	M = 9.77 D = 4.05	M = 9.65 D = 3.01	M = 10.96 D = 5.34	M = 10.71 D = 3.45	M = 11.40 D = 6.19
	Posttest	M = 25.21 D = 6.30	M = 21.95 D = 7.31	M = 27.44 D = 3.90	M = 26.17 D = 6.19	M = 21.08 D = 6.85	M = 25.48 D = 6.83
	Retention Test	M = 21.35 D = 8.37	M = 17.82 D = 9.53	M = 23.97 D = 7.47	M = 25.42 D = 4.85	M = 18.42 D = 8.77	M = 23.08 D = 8.95
<i>Degree of Certainty Estimates</i>	Pretest	M = .32 D = .29	M = .33 D = .28	M = .32 D = .28	M = .24 D = .17	M = .21 D = .21	M = .35 D = .21
	Posttest	M = .84 D = .26	M = .77 D = .29	M = .91 D = .17	M = .90 D = .18	M = .69 D = .31	M = .85 D = .27
	Retention Test	M = .65 D = .39	M = .56 D = .40	M = .73 D = .34	M = .81 D = .27	M = .60 D = .35	M = .77 D = .30
<i>Admissible Probability Scores</i>	Pretest	M = -16.11 D = 11.48	M = -16.58 D = 9.70	M = -17.75 D = 7.66	M = -16.68 D = 9.58	M = -20.10 D = 9.04	M = -13.15 D = 13.08
	Posttest	M = 17.34 D = 16.61	M = 10.43 D = 17.37	M = 23.72 D = 9.49	M = 21.28 D = 13.22	M = 6.34 D = 17.41	M = 18.39 D = 17.81
	Retention Test	M = 5.70 D = 22.71	M = -1.58 D = 24.03	M = 12.58 D = 19.34	M = 17.70 D = 13.31	M = 1.18 D = 20.39	M = 13.28 D = 19.59
<i>Time (sec)</i>		M = 676.21 D = 182.52	M = 740.00 D = 275.45	M = 675.97 D = 301.59	M = 725.39 D = 187.26	M = 786.22 D = 249.37	M = 758.71 D = 187.61
	<i>Attitude Inventory</i>	M = 91.29 D = 15.31	M = 92.52 D = 9.05	M = 92.12 D = 14.47	M = 92.00 D = 7.14	M = 93.46 D = 9.77	M = 90.64 D = 8.09

### Students' Raw Scores

The mixed factorial design (the split-plot analysis of variance) was used to analyze data from mixed designs – one within-subjects factor (“test”) and one between-subjects factor (“group”). There were three levels of factor in the

within-subjects design in which subjects took all three tests (“pretest”, “posttest”, and “retention test”) and there were 6 levels of factor in the between-subjects design corresponding to the six groups of participants.

The repeated factor, “test,” was a statistically significant main effect for students’ raw score, Wilks’ Lambda = .174,  $F_{(2,165)} = 390.784$ ,  $p = .000$ , effect size = .826. The interaction between factors “test” and “group” was also a significant main effect for students’ raw scores, Wilks’ Lambda = .848,  $F_{(10,330)} = 2.831$ ,  $p = .002$ , effect size = .079. For a follow-up analysis on a significant interaction between the factors “test” and “group”, the ANOVA was used to test for the “group” factor on the levels of the “test” factor.

The ANOVA performed on the pretest data indicated that there was no significant difference in students’ raw scores between the groups,  $F_{(5, 166)} = 0.570$ ,  $p = .723$ , effect size = .017. However, the ANOVA performed on the posttest data revealed that there was significant difference between the groups,  $F_{(5,166)} = 4.220$ ,  $p = .001$ , effect size = .113. A follow-up post-hoc test, Tukey’s Honestly Significant Difference (HSD) test was conducted to evaluate pair-wise differences among the means. The results of the Tukey’s HSD test indicated that there was a significant difference between the groups (Group B & C and C & E). These results demonstrate that Group C students who received “visual text and added graphics” ( $M = 27.44$ ,  $SD = 3.90$ ) earned significantly higher raw scores than Group B students who received “visual text and added spoken text” ( $M = 21.95$ ,  $SD = 7.31$ ) and Group E students who received “reduced visual text and added spoken text”, ( $M = 21.08$ ,  $SD = 6.85$ ) on the posttest.

From the retention test, the ANOVA data also revealed that there was a statistically significant difference between the groups,  $F_{(5,166)} = 3.487$ ,  $p = .005$ , effect size = .095. Tukey’s HSD post-hoc test indicated that there was a significant difference between the groups (Group B & D and D & E). These results indicate that Group D students who received “visual text, added spoken text, and added graphics” ( $M = 25.42$ ,  $SD = 4.85$ ) earned significantly higher raw scores than Group B students who received “visual text and added spoken text” ( $M = 17.82$ ,  $SD = 9.53$ ) and Group E students who received “reduced visual text and added spoken text” ( $M = 18.42$ ,  $SD = 8.77$ ) on the retention test.

These results indicate that the students in Group C (“visual text and added graphics”) and Group D (“visual text, added graphics, and added spoken text”) learned and retained English vocabulary more effectively than students who received the other types of instruction as shown in Figure 3.

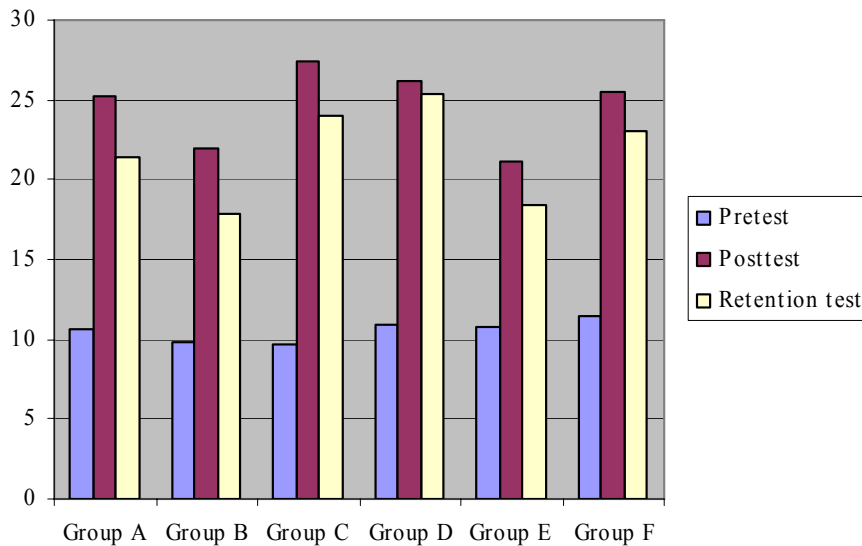


Figure 3. Group’s means for students’ raw scores

### Students’ Mean Degree of Certainty Estimates

The mixed factorial design was used to analyze the mean degree of certainty estimates. The repeated factor, “test,” was a statistically significant main effect for students’ mean degree of certainty estimates, Wilks’ Lambda = .236,

$F_{(2,165)} = 267.809, p = .000$ , effect size = .764. However, the interaction between the factors “test” and “group” was not a significant main effect for students’ mean degree of certainty estimates, Wilks’ Lambda = .908,  $F_{(10,330)} = 1.630, p = .098$ , effect size = .047.

The ANOVA performed on the pretest ( $F_{(5, 166)} = 1.286, p = .272$ , effect size = .037) and the retention test ( $F_{(5,166)} = 1.953, p = .088$ , effect size = .056) data indicated that there were not statistically significant differences in mean degree of certainty estimates between the groups. However, from the posttest, the results of the ANOVA were statistically significant,  $F_{(5, 166)} = 2.910, p = .015$ , effect size = .081. The results of Tukey’s HSD test also indicated that there was a significant difference in means between the groups (Group C & E). In other words, Group C students who received “visual text and added graphics” ( $M = .91, SD = .17$ ) earned significantly higher mean degree of certainty scores than Group E students who received “reduced visual text and added spoken text” ( $M = .69, SD = .31$ ).

Therefore, the results indicate that students in all groups earned higher scores in general, indicating that their degree of belief probabilities increased when they received multimedia instruction as shown in Figure 4.

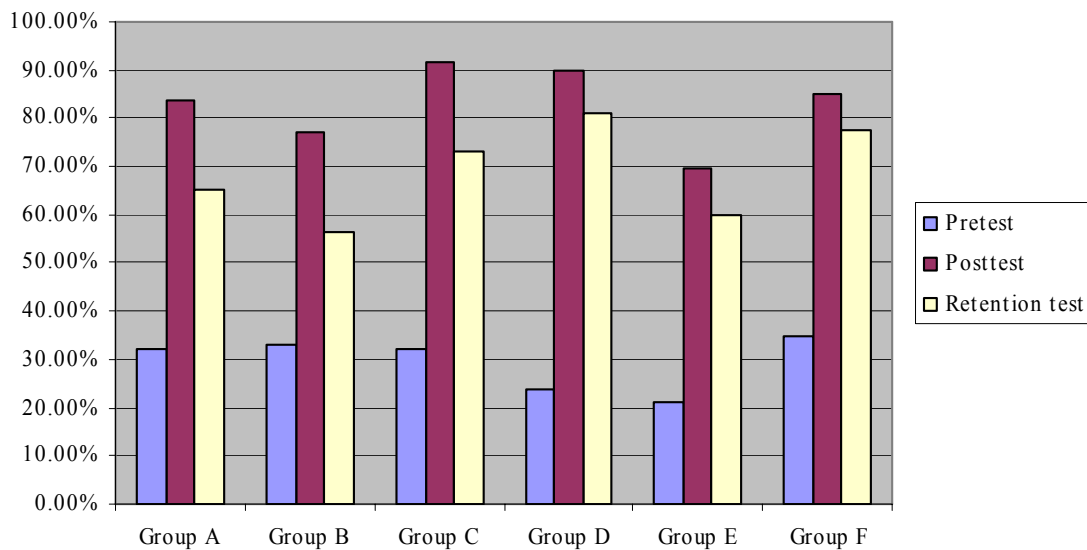


Figure 4. Group’s means for students’ mean degree of certainty estimates

### Shuford Admissible Probability Scores

The mixed factorial design was also used to analyze the students’ admissible probability scores. The repeated factor, “test,” was a statistically significant main effect for students’ admissible probability scores, Wilks’ Lambda = .186,  $F_{(2,165)} = 360.841, p = .000$ , effect size = .814. The interaction between the factors “test” and “group” was also a statistically significant main effect, Wilks’ Lambda = .845,  $F_{(10,330)} = 2.9.4, p = .002$ , effect size = .081.

The ANOVA performed on the pretest data indicated that there was no significant difference in admissible probability scores between the groups,  $F_{(5,166)} = 1.222, p = .301$ , effect size = .036. However, from the posttest, the ANOVA was statistically significant,  $F_{(5,166)} = 4.789, p = .000$ , effect size = .126. Tukey’s HSD test indicated that there was a significant difference in the means between the groups (Group B & C, C & E, and D & E). These results indicate that Group C students who received “visual text and added graphics” ( $M = 23.72, SD = 9.49$ ) earned significantly higher scores than Group B students who received “visual text and added spoken text” ( $M = 10.43, SD = 17.37$ ) and Group E students who received “reduced visual text and added spoken text” ( $M = 6.34, SD = 17.41$ ). It also shows that Group D students who received “visual text, added graphics, and added spoken text” ( $M = 21.28, SD = 13.22$ ) earned significantly higher scores than Group E students who received “reduced visual text and added spoken text” ( $M = 6.34, SD = 17.41$ ).

For the retention test, the ANOVA was also statistically significant,  $F_{(5,166)} = 3.378$ ,  $p = .006$ , effect size = .092. Tukey's HSD test indicated that there was a significant difference in means between groups (Group B & D) on the retention test. The result indicates that Group D students who received "visual text, added graphics, and added spoken text" ( $M = 17.70$ ,  $SD = 13.31$ ) earned significantly higher scores than Group B students who received "visual text and added spoken test" ( $M = -1.58$ ,  $SD = 24.03$ ).

These results lead to the conclusion that there was a significant difference in students' admissible probability scores between the groups when students received multimedia instruction as shown in Figure 5.

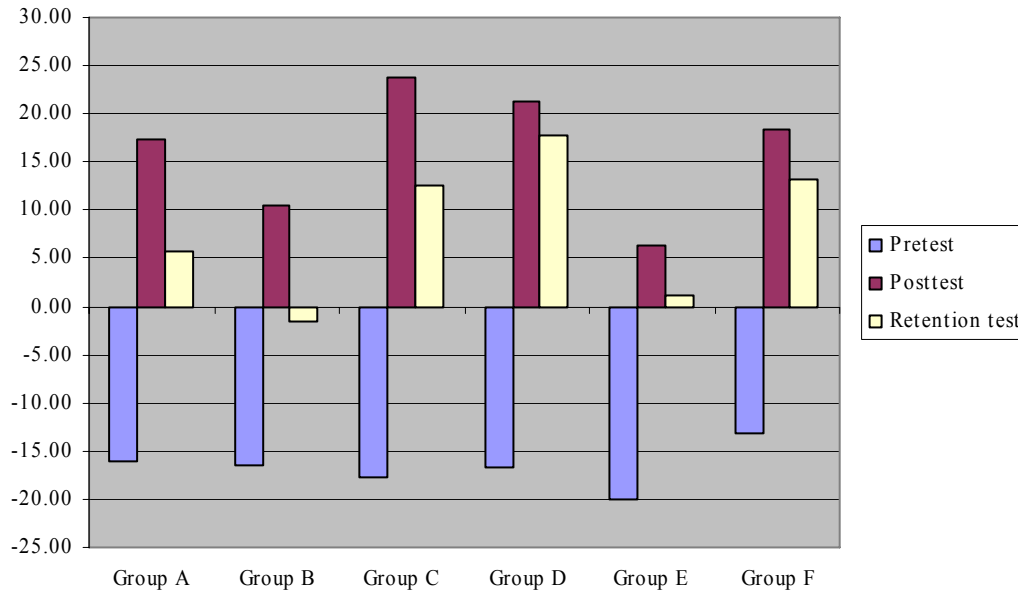


Figure 5. Group's means for students' Shuford Admissible Probability Scores

### Time Required to Complete Instruction & Student Attitude Inventory

The ANOVA showed that there was no significant difference among students who were taught under the six methods of multimedia instruction conditions either in the time required to complete the instruction ( $F_{(5, 161)} = 1.070$ ,  $p = .379$ , effect size = .032) or in the student attitude inventory ( $F_{(5, 163)} = .175$ ,  $p = .972$ , effect size = .005).

### Discussion

*Hypothesis 1: There are no differences in original learning among students who are taught under the six methods of multimedia instruction conditions as measured by raw score, mean degree of certainty estimate, and the Shuford Admissible Probability Scoring Procedure.*

The mixed factorial design (the split-plot analysis of variance) on data obtained from the scores of students who were taught under the six methods of multimedia instruction conditions as measured by the student's raw scores and the Shuford Admissible Probability Scores on the posttest and retention test showed differences between the six methods of multimedia instruction at the .05 level of significance.

No significant differences were found among treatment groups on data of the students' mean degree of certainty estimates on the pretest and retention test. Thus, this result leads one to conclude that there were no significant differences among the treatment groups with respect to the degree of certainty of knowledge on the pretest and retention test. However, the results show that, in general, students earned a higher score which indicates that they increased their degrees of belief probabilities when they received multimedia instruction.

The students received higher scores in general, which indicates that they learned better when they received “visual text and added graphics” or “visual text, added graphics, and added spoken text” in their instruction than did students who received other types of instruction (“visual text”, “visual text and added spoken text”, “reduced text and added spoken text”, or “reduced text, added graphics, and added spoken text”). In other words, when visual text was presented with graphics, students may be more motivated to success and achievement in vocabulary learning.

*Hypothesis 2: There are no differences in time to complete instruction among students who are taught under the six methods of multimedia instruction conditions.*

Results of the ANOVA on data obtained from the students who were taught under the six methods of multimedia instruction conditions as measured by the time required to complete instruction showed no differences between the six methods of multimedia instruction at the .05 level of significance.

Although providing multimedia components required individuals to spend more time to complete the instruction, the amount of time spent was not significant. This result leads one to conclude that there were no significant differences among the six methods of instruction with respect to time needed to complete instruction.

*Hypothesis 3: There are no differences in students' attitude toward instruction among students who are taught under the six methods of multimedia instruction conditions as measured by their score on the student attitude inventory.*

Results of the ANOVA on data obtained from the students who were taught under the six methods of multimedia instruction conditions as measured by score on the student attitude inventory showed no differences in attitude toward instruction among students who were taught under the six methods of multimedia instruction conditions. This result leads one to conclude that there were no significant differences among the treatment groups with respect to student attitude.

*Based on visual text.* The instruction based on visual text (Group A), in general, helped students to learn and retain English vocabulary more effectively than the instruction based on visual text and added spoken text (Group B) and the instruction based on reduced visual text and added spoken text (Group E).

From Mayer and Moreno's research (2002), the results should have indicated that providing words as narration (spoken text) helped students' performance more than on-screen text (visual text). However, the results show that written words alone may help students to learn and retain English vocabulary. A possible reason for this is that EFL learners often adopt various strategies to memorize vocabulary words such as word lists or paired associations in which new words are presented with their translations. These strategies often require learners to memorize pair associations directly. Vocabulary learning studies in South Korea have often focused on learning based on visual text alone in printed materials. In other words, written words alone may be better than text and added spoken text because Korean students are more accustomed to memorizing lists.

*Added spoken text.* The instruction based on visual text and added spoken text (Group B) and based on reduced visual text and added graphics (Group E) reduced students' ability to learn and retain English vocabulary. According to cognitive load theory (Mayer, 2001; Sweller et al., 1998; Tabbers et al., 2004), the results should have indicated that students could reduce memory load with instruction based on reduced text with added spoken text (Group E). However, in fact, this aid did not help students to learn and retain English vocabulary. This result indicates that neither replacing visual text with spoken text nor adding graphics to the visual text easily generalize to all educational settings.

The lower scores for Groups B (visual text and added spoken text) and E (reduced visual text and added spoken text) may indicate problems in phonic learning. Students may have difficulty in knowing exactly how the words are pronounced. Because many EFL learners in Korea are accustomed to memorizing new words without knowing exactly how they are pronounced, spoken text in the instruction created an unnecessary distraction.

Another possible reason for this is that the test was designed to measure only students' understanding of a word's meaning and did not measure their knowledge of the word's pronunciation. The spoken text seemed to be a

distraction to students who are accustomed to learning a foreign language mainly in the current written test format. Thus, this result leads one to conclude that an effective way to improve learning of new English vocabulary is to avoid the addition of spoken text when explaining what the vocabulary means in Korean.

*Added graphics.* The instruction based on visual text and added graphics (Group C), based on visual text, added graphics, and added spoken text (Group D), and based on reduced text, added spoken text, and added graphics (Group F) helped students to learn and retain English vocabulary more effectively than the other types of instruction as demonstrated on the pretest and retention test. The results show that in general, students earned a higher score, which indicates that they learned better when they received graphics or graphics and spoken text in their instruction than did students who received other types of instruction (Group A, B, and E). In other words, when visual text is presented with graphics, students may be motivated to success and achievement in vocabulary learning.

The likely reason is that text does not usually translate in a manner that is meaningful to the student, while a graphic allows the student to visualize the definition in a more meaningful way. Some words cannot be translated directly and retain meaning for middle school students. When students received the instruction based on visual text only, for example, showing that “taut” means “pulled or stretched tight” may not allow learners to explain what that definition means in ways that make sense to them. The results supported the conclusion that students performed better on vocabulary tests when they were allowed to use a combination of visual text and graphics (Kost et al., 1999).

The results support the concept from dual coding theory (Paivio, 1986, chap. 4; Sadoski & Paivio, 2001, chap 3) that students are likely to build connections between verbal (visual text) and nonverbal (graphics) representations. As Mayer (2001) concluded, we assume that processing of visual text takes place initially in the nonverbal channel and then moves to the verbal channel of the brain. The results appear to indicate that providing both visual text and graphics helped students to “select relevant information, organize it into coherent representations, and integrate it with other knowledge” (Mayer & Moreno, 2002, p. 111) as meaningful learning.

The results appear to indicate that replacing visual text with spoken text and adding graphics to the visual text both do not easily generalize to all educational settings (Tabbers et al., 2004). A possible reason for this is that reduced text was not sufficient to explain the definitions in Korean.

In addition, learning new vocabulary within the context of the instruction based on visual text and added graphics may indicate that

“...students have to make informed guesses as to the meaning of a new word in light of available linguistic cues in the context as well as the relevant knowledge in the learner’s mind, including general knowledge of the world, awareness of the situation, and relevant linguistic knowledge.” (Sun & Dong, 2004, p. 132)

## **Conclusion**

This study was an investigation of the use of multimedia components such as visual text, spoken text, and graphics in a Web-based self-instruction program to increase students’ English vocabulary learning at Myungin Middle School in Seoul, South Korea. The findings of the study support the idea that the use of visual media supports vocabulary acquisition and helps increase achievement scores. In particular the results lead one to conclude that an effective way to improve the learning of English vocabulary is to offer graphics to illustrate the definition. Students were likely motivated to success and achievement in vocabulary learning when visual text was presented with graphics because text alone did not usually translate in a manner that is meaningful to the learners, while graphics allowed them to visualize the definition in a more meaningful way.

Vocabulary learning is often used with strategies such as word lists or paired associations in which new words are presented with their translations. These strategies with visual text alone may be outdated and irrelevant to students who are accustomed to visual stimuli and have shorter attention spans. The findings of this study indicate that developers of vocabulary learning instruction and curriculum should reconsider their use of multimedia within their presentations. For example, because presenting too many elements in visual or verbal form can lead to reduced ability to learn and retain vocabulary, visual text, spoken text and graphics must be carefully planned and utilized in the instruction. Replacing text and graphics with spoken text can create an unnecessary distraction for EFL learners

who are accustomed to memorizing new words without knowing exactly how they are pronounced. Reduced text or graphics sometimes does not sufficiently explain the definitions of new words, while spoken text helps developers to save space or time to present messages in the instruction. Developers should select relevant graphics to illustrate the meanings of words with appropriately sized images in multimedia instruction. Integrating text and graphics can allow learners to visualize definitions of the words in a way that fosters meaningful learning outcomes. It can help students to have meaningful learning through the cognitive process, or in other words, “selecting relevant words and images, organizing them into coherent verbal and visual representations, and integrating corresponding verbal and visual representations” (Mayer & Moreno, 2002, p. 111). This has ramifications beyond the world of computer-based learning. Printed instructions and in-class lessons should also be designed to improve learner retention through the implementation of graphics alongside current uses of media.

Future research could focus on measuring students’ knowledge of a word’s pronunciation as well as its meaning. In a test that measures students’ knowledge of the word’s pronunciation, the spoken text may not be a distraction to English vocabulary learning. Other future studies could replicate our methods to discover whether the results would be similar for vocabulary learning for languages other than English and with students from countries other than Korea. Although “all multimedia messages are not equally effective” for EFL students (Mayer, 2001, p. 79), it is hoped these findings can be replicated and expanded for the use of vocabulary learning in other languages and cultures.

## References

- Ali, A. (2003). Instructional design and online instruction: Practices and perception. *TechTrends*, 47 (5), 42-45.
- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning and Technology*, 5 (1), 201-232.
- Atkinson, R. C. (1972). Optimizing the learning of a second-language vocabulary. *Journal of Experimental Psychology*, 96, 124-129.
- Bauer, J. F. (2002). Assessing student work from chatrooms and bulletin boards. In R. S. Anderson, J. F. Bauer, & B. W. Speck (Eds.), *Assessment strategies for the online class: From theory to practice* (pp. 31-36). San Francisco: Jossey-Bass.
- Brown, C., & Payne, M. E. (1994). Five essential steps of processes in vocabulary learning. *Paper presented at the TESOL Convention*, March 8-12, 1994, Baltimore, MD, USA.
- Chapelle, C. (1998). Multimedia call: Lessons to be learned from research on instructed SLA. *Language Learning and Technology*, 2 (1), 22-34.
- Chun, D., & Plass, J. (1997). Research on text comprehension in multimedia environments. *Language Learning and Technology*, 1 (1), 60-81.
- Crothers, E., & Suppes, P. (1967). *Experiments in second-language learning*, New York: Academic Press.
- Duquette, L., & Painchaud, G. (1996). A comparison of vocabulary acquisition in audio and video contexts. *The Canadian Modern Language Review*, 53 (1), 143-172.
- Ehsani, F., & Knodt, E. (1998). Speech technology in computer-aided language learning: Strengths and limitations of a new CALL paradigm. *Language Learning and Technology*, 2 (1), 45-60.
- Gilman, D. A. (1967). A comparison of the effectiveness of feedback modes for teaching science concepts by means of a computer-assisted adjunct auto-instruction program (Doctoral dissertation, Pennsylvania State University). *Dissertation Abstracts International*, 29, 182.
- Hatch, E., & Brown, C. (1995). *Vocabulary, semantics and language education*, Cambridge, UK: Cambridge University Press.
- Hoven, D. (1999). A model for listening and viewing comprehension in multimedia environments. *Language Learning and Technology*, 3 (1), 88-103.

- Jonassen, D. (2000). *Computers as mindtools for schools engaging critical thinking* (2<sup>nd</sup> Ed.), Upper Saddle River, NJ: Prentice-Hall.
- Koroghlanian, C., & Klein, D. J. (2004). The effect of audio and animation in multimedia instruction. *Journal of Educational Multimedia and Hypermedia*, 13 (1), 23-46.
- 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.
- Liles, B. (2004). Going the distance. *Sound & Video Contractor*, 22 (3), 48.
- Liu, M., Moore, Z., Graham, L., & Lee, S. (2002). A look at the research on computer-based technology use in second language learning: A review of the literature from 1990-2000. *Journal of Research on Technology in Education*, 34 (3), 250-273.
- Mayer, R. E. (2001). *Multimedia learning*, Cambridge, UK: Cambridge University Press.
- Mayer, R. & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, 12, 107-119.
- Meara, P. (1996). The classical research in L2 vocabulary acquisition. In G. Anderman & M. Rogers (Eds.), *Words, words, words: The translator and the language learner* (pp. 27-37), Clevedon, UK: Multilingual Matters.
- Newby, J., Stepich, A., Lehman, D., & Russell, D. (1996). *Instructional technology for teaching and learning: Designing instruction, integrating computers, and using media*, Englewood Cliffs, NJ: Prentice Hall.
- Novak, J. D. (1998). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*, Mahwah, NJ: Lawrence Erlbaum.
- Paivio, A. (1986). *Mental representations: A dual coding approach*, New York, NY: Oxford University Press.
- Sadoski, M. & Paivio, A. (2001). *Imagery and text: a dual coding theory of reading and writing*, New York, NY: Lawrence Erlbaum.
- Schnotz, W., Bannert, M., & Seufert, T. (2002). Towards an integrative view of text and picture comprehension: Visualization effects on the construction of mental models. In J. Otero, A. Graesser & J. A. Leon (Eds.), *The psychology of science text comprehension* (pp. 385-416), Mahwah, NJ: Erlbaum.
- 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.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10 (3), 251-296.
- Tabbers, K. H., Martens, L.R., & Merriënboer, J. J. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology*, 74 (1), 71-82.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*, Cambridge, MA: Harvard University Press.
- Watts, N. (1997). A learner-based design model for interactive multimedia language learning packages. *System*, 25 (1), 1-8.