

The Effects of Multimedia-Enhanced Instruction on Knowledge Gain and Retention of ESL Learners: An Assessment of Mayer's Redundancy Principle

Araya Ramsin

Chulalongkorn University

Email Address: araya.r@chula.ac.th

Abstract

This quasi-experimental quantitative study investigated the effects of different formats of multimedia-enhanced instruction on knowledge gain and retention of second language learners. The study focused specifically on assessing Mayer's redundancy principle. Second language learners were randomly assigned into two groups and were presented with two formats of a multimedia-enhanced lesson: (1) images with audio and (2) images with audio and on-screen text. A repeated-measures analysis of covariance (ANCOVA) was carried out to evaluate change between pretest and immediate posttest, mean differences in knowledge gain between the two groups, group differences in growth from pretest to immediate posttest, change between immediate posttest and delayed posttest, and group differences in knowledge retention as indicated by difference scores between immediate posttest and delayed posttest, using gender and English proficiency as control variables. Additionally, an independent samples t-test was performed to assess the effect of the two formats of multimedia-enhanced

instruction on the learners' knowledge retention. Results showed that participants engaging with both formats of multimedia-enhanced instruction improved significantly from pretest to posttest. Moreover, participants who engaged with the multimedia-enhanced instruction where on-screen text was included were also found to perform significantly better than participants who were in the group where on-screen text was not present. Second language learners in both groups were also able to retain their knowledge at the same level. These results appear to contradict the redundancy principle as having on-screen text, which is considered redundant, presented together with images and audio in a multimedia-enhanced lesson was not detrimental to learning.

Keywords: Multimedia Learning, Redundancy Principle, Knowledge Gain and Retention, Second Language Learners

Introduction

With today's world being more visual than ever before, it cannot be denied that multimedia has an essential role to play in how people interact, communicate, and learn. Multimedia can be simply defined as presentations that integrate both words, which can be printed text or spoken text, and pictures, which can be static or dynamic (Issa et al., 2011; Mayer, 1997; Mayer & Moreno, 2003; van Merriënboer & Kirschner, 2018). A large body of research has been conducted on multimedia learning and repeatedly reported the incorporation of multimedia into the classroom to be more successful in enhancing learning compared to the traditional one in which lessons were taught mainly through spoken and written text. With multimedia-enhanced lessons, students across disciplines were found to be more motivated and engaged in their learning, which often resulted in observed favorable learning outcomes (Abdul Samat & Abdul Aziz, 2020; Ahmad & Yamat, 2020; Chuang & Liu, 2012; Erizar et al., 2018; Ilhan & Oruç,

2016; Moussa-Inaty & Atallah, 2012; Ramsin & Mayall, 2019; Schilling, 2009; Trevisan et al., 2010; Wang & Li, 2019; Winke et al., 2010). However, the focus of research on multimedia learning has recently proceeded from solely assessing the effectiveness of multimedia use in the classroom to exploring how multimedia can appropriately be incorporated into lessons to better benefit students with different needs in various educational settings (Chuang & Liu, 2012; Lusk et al., 2009). As multimedia integration into the classroom is now becoming a more widespread practice across learning disciplines, understanding how to employ multimedia in the process of teaching and learning in the classroom to best facilitate learners is crucial.

Problem Statement

As the effectiveness of multimedia technology in promoting classroom teaching and learning is apparent and has widely been accepted, the current audiences of multimedia instruction include diverse groups of learners in various academic disciplines (Abdul Samat & Abdul Aziz, 2020; Ahmad & Yamat, 2020); Chuang & Liu, 2012; Erizar et al., 2018; Ilhan & Oruç, 2016; Moussa-Inaty & Atallah, 2012; Ramsin & Mayall, 2019; Schilling, 2009; Trevisan et al., 2010; Wang & Li, 2019; Winke et al., 2010). There are times, however, when learners are not native speakers of the language multimedia instructional materials are made in, and among them are second language learners. Based on the researcher's experience as an English language university lecturer, it is often found that second language learners, especially those with limited English proficiency, tend to have difficulties comprehending the content presented when participating in a multimedia learning environment where the language of instruction is English, which is not their mother tongue. Previous studies have found that incorporating on-screen text, or caption, into multimedia instructional materials designed for second language learners can aid learning (Brasel & Gips, 2014; Chai & Erlam, 2008; Danan, 2004; Jae, 2019; Winke et al., 2013). However, according to the redundancy principle proposed by Richard Mayer as part of the cognitive theory of

multimedia learning, describing graphics or pictures (static or dynamic) using both on-screen text and narration repeats the text, and this redundant information, especially when presented simultaneously, can overwhelm the brain in processing information and result in cognitive overload (Clark & Mayer, 2008; Mayer, 2002).

In the researcher's previous work (Ramsin & Mayall, 2019), two multimedia learning environments developed for second language learners were investigated. Results suggested that second language learners participating in the multimedia learning environment that included images, audio, and on-screen text were able to achieve better learning outcomes than those in a multimedia learning environment where on-screen text was not present, contradicting the redundancy principle. Learning, however, also involves a relatively lasting change in behaviour (Cherry, 2022; Domjan, 2015; Flaherty, 1985; Gordon, 2000). Previously constructed knowledge can be used only if it is retained; therefore, knowledge retention is considered an integral part of learning and should not be thought of as separate from the learning process (Houston, 2001; Kosar & Bedir, 2018). This is especially significant in English language learning, in which the ability to retain basic language skills and knowledge, such as vocabulary and grammatical structures, is essential in order to have a good command of English and to be successful in proceeding to more advanced levels of the language (Kosar & Bedir, 2018; Niedfeldt, 2001).

Theoretical Framework

There are two theories underpinning this research study: the cognitive theory of multimedia learning of Richard Mayer (2002) and the second language acquisition theory of Stephen Krashen (1982).

Cognitive Theory of Multimedia Learning

Multimedia is a term used to refer to a wide range of different things and can be defined in a number of ways. It can sometimes be used to refer to the technologies used for information display, such as computers, networks, or devices; the representational format, such as text, graphics, or animation; or the sensory modalities in perceiving

information (Horz & Schnotz, 2008). According to Mayer (2002, 2014), multimedia is simply defined as the presentation in which words and pictures are integrated. Some examples of multimedia may include watching a recorded lesson on a tablet, playing a language game on a smartphone, or watching a weather forecast on a TV screen while listening to the corresponding words, music, and sounds. Multimedia learning, on the other hand, is defined as the learner's construction of knowledge from words and pictures (Mayer, 2014), and the process by which the learner builds mental representations from words and pictures is the focus of the cognitive theory of multimedia learning (Mayer, 2009).

Research on multimedia learning is centered around the belief that multimedia instruction should be designed based on how the human mind works in order to be effective and lead to meaningful learning. One theory that is well known and widely accepted in the world of multimedia instructional design is the cognitive theory of multimedia learning by Richard Mayer (2002). The theory is derived from several cognitive theories including Paivio's dual coding theory (Paivio, 1986), Baddeley's model of working memory (Baddeley, 1992), and Sweller's cognitive load theory (Chandler & Sweller, 1991). The fundamental proposition of the cognitive theory of multimedia learning is that people learn more deeply from words and pictures than from words alone, and that is because when people have the opportunity to build connections between words and pictures, learning becomes more meaningful (Mayer, 2002).

Mayer's cognitive theory of multimedia learning is based on three cognitive principles of learning: (1) humans process visual and auditory information in two different cognitive channels, (2) each cognitive channel has a limited capability, and (3) humans actively process this visual and auditory information as they learn. According to the theory, the human information-processing system includes three memory stores: sensory memory, working memory, and long-term memory (Mayer, 2002).

The cognitive theory of multimedia learning specifies five cognitive processes required for meaningful learning to occur in a

multimedia environment: (1) selecting relevant words from the presented text or narration, (2) selecting relevant images from the presented graphics, (3) organizing the selected words into a coherent verbal representation, (4) organizing selected images into a coherent pictorial representation, and (5) integrating the verbal and pictorial representations with prior knowledge (Mayer, 2014).

Based on his extensive research, Mayer suggests 12 principles that can provide guidance on how to design effective multimedia instruction, such as PowerPoint presentations, online learning courses, training videos, etc. These principles are the coherence principle, the signaling principle, the redundancy principle, the spatial contiguity principle, the temporal contiguity principle, the segmenting principle, the pre-training principle, the modality principle, the multimedia principle, the personalization principle, the voice principle, and the image principle (Mayer, 2009). As for this study, the emphasis was placed only on the redundancy principle, a principle that has largely been employed in the design of multimedia learning environments.

According to the redundancy principle, people learn best with just images and narration, not with images, narration, and written text all present together. This is because when images are already explained through narration, the written text here becomes unnecessary and is considered redundant information. Including this extra information, i.e., the written text, is, therefore, not recommended as it can overwhelm the brain in processing information, which can impede learning. It is hard for people to focus their attention on all three sources of information at the same time (Hoffman, 2006; Mayer, 2002).

Second Language Acquisition Theory

The United Nations Educational, Scientific, and Cultural Organization (UNESCO) defines the term second language as “a language acquired by a person in addition to his mother tongue” (Cook, 2001, p. 13). Second language is a broad term used to refer to any language learned in addition to a person’s first language, regardless of the purpose, the type of learning environment (formal or informal), or

the number of other non-native languages previously learned or acquired by the learner (Cook, 2001; Smith, 1994). This includes foreign languages, e.g., English as a foreign language for Thais, and languages which are not the learner's mother tongue but are widely used in his own community, e.g., English for Spanish-speaking Mexicans. The term second language is often abbreviated to L2 as opposed to L1—the mother tongue (Smith, 1994).

One very influential theory in the field of second language teaching is the second language acquisition theory developed by Stephen Krashen (1982). The theory is composed of five main hypotheses: (a) the acquisition-learning distinction, (b) the natural order hypothesis, (c) the monitor hypothesis, (d) the input hypothesis, and (e) the affective filter hypothesis. Of all these hypotheses, the one that is most often referred to when it comes to second language teaching and learning and is considered central to the theory of second language acquisition is the input hypothesis (Krashen, 1985a). The input hypothesis model claims that humans acquire language in only one way—by understanding messages or by receiving comprehensible input (Krashen, 1985a, 1985b), and in order for comprehensible input to be even more effective, it should also be compelling, or highly interesting to the learner (Krashen & Bland, 2014). Humans naturally focus on the meaning, rather than on the form or grammar, of the message as they proceed in their second language development (Cook, 2001; Neuman & Koskinen, 1992). Language learners need to be exposed to the target language, or input, to make progress in the language, and according to Krashen (1982), language acquisition happens best when the input is just slightly more advanced than the learners' level. The input hypothesis has already been employed in most language classes, at least partially. For example, grammar and vocabulary are taught step by step, and listening as well as reading activities contain mostly words familiar to learners. This allows learners to focus on new concepts without being overwhelmed, and they can build on what they have mastered (Niedzielski, 2020).

Comprehensible Input and Second Language Learning

Language acquisition occurs through an unconscious process and only with the necessary ingredient—comprehensible input (Lewis, 2020). In English language classes, comprehensible input simply refers to English language that learners can understand, which can be things they hear or read such as podcasts and recorded conversations, books, research articles, etc. One of the most crucial tasks for language teachers is to provide appropriate and meaningful messages for learners to understand the parts that are beyond their language knowledge (Cook, 2001). According to the the input hypothesis, the use of images and other realia is extremely helpful in second language learning, especially in the early stages, because they provide context helpful for acquiring the target language and help make input comprehensible (Krashen, 1985b).

Research conducted by Neuman and Koskinen (1992) found that second language learners were able to develop word meanings and language through comprehensible input. Results suggested that comprehensible input, which was captioned television in this study, was shown to provide a rich language environment enabling students to acquire new words through context while developing concepts in science. Students who viewed captioned television were reported to outperform those who viewed non-captioned television on all measures of word knowledge. In addition, they were also able to remember more science content than their counterparts. This demonstrated, therefore, that communicating messages to second language learners through different modalities appeared to enhance their learning, both the content knowledge and language, rather than negatively affecting their attention capacity (Neuman & Koskinen, 1992). Similar results were found in a number of studies conducted to investigate the effects of including text on screen to provide a rich learning environment and make language input more comprehensible (Adegoke, 2010; Linebarger, 2001; Ozdemir et al., 2016; Perez et al., 2013; Winke et al., 2010).

Purpose of the Study

The present study investigates the effects of different formats of multimedia-enhanced instruction on content knowledge gain and retention of second language learners. The study focuses specifically on assessing Mayer's redundancy principle by comparing the learning outcomes of two groups of students presented with two formats of a multimedia lesson: (1) images with audio and (2) images with audio and on-screen text.

Research Questions

The research questions guiding this study were as follows:

1. Is there a significant difference in knowledge gain between two groups of second language learners presented with different formats of a multimedia-enhanced lesson: (1) images with audio and (2) images with audio and on-screen text?
2. What is the effect of the two formats of multimedia-enhanced instruction on second language learners' knowledge retention?

Participants

The sample for this study consisted of 214 undergraduate students enrolled in an English course at a university in Thailand. The participants were recruited through nonprobability convenience sampling and were a sample of the entire undergraduate student population at this particular university of approximately 35,000. Permission was requested and granted from the director of the university language institute, who was in charge of all English courses offered at the university, to conduct the study. Permission was also granted from the English language instructors of all the potential participants to access their classes. Only students who provided consent to participate were included in the study.

A total of 11 classes of students, with approximately 25–35 students each, were recruited. Students within each class remained together as intact groups, but these 11 classes were randomly assigned

into two different groups defined by the combination of multimedia components within the multimedia-enhanced lesson the students attended. There were 110 students participating in Group 1 (images with audio only) and 104 students participating in Group 2 (images with audio and on-screen text). Relevant demographic information of the participants, which included gender, age, and English proficiency, was also collected (Tables 1 and 2). In this study, the participants' English proficiency was measured using the university's official English proficiency test. The total score was 120.

As shown in Table 1, the sample consisted of 214 participants, including 105 females (49.07%) and 109 males (50.93%). The Images with Audio Group consisted of 47 females (42.73%) and 63 males (57.27%) while the Images with Audio and On-Screen Text Group consisted of 58 females (55.77%) and 46 males (44.23%).

Table 1

Demographic Distribution for Gender of Participants

Characteristic	Images + Audio		Images + Audio + On-Screen Text		Combined	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Female	47	42.73%	58	55.77%	105	49.07 %
Male	63	57.27%	46	44.23%	109	50.93%
Total	110	100.00%	104	100.00%	214	100.0%

Table 2 provides descriptive statistics for the overall age and English proficiency of the participants. The mean age of all participants was 18.93 years ($SD = 0.58$). The mean age of the participants in the Images with Audio Group was 19.04 years ($SD = 0.62$) and the mean age of the participants in the Images with Audio and On-Screen Text Group was 18.83 years ($SD = 0.51$).

The average English proficiency score of all participants was 57.53 ($SD = 14.75$). The mean English proficiency score of the participants in the Images with Audio Group was 53.73 ($SD = 15.09$) and the mean English proficiency score of the participants in the Images with Audio and On-Screen Text Group was 61.55 ($SD = 13.32$).

Table 2

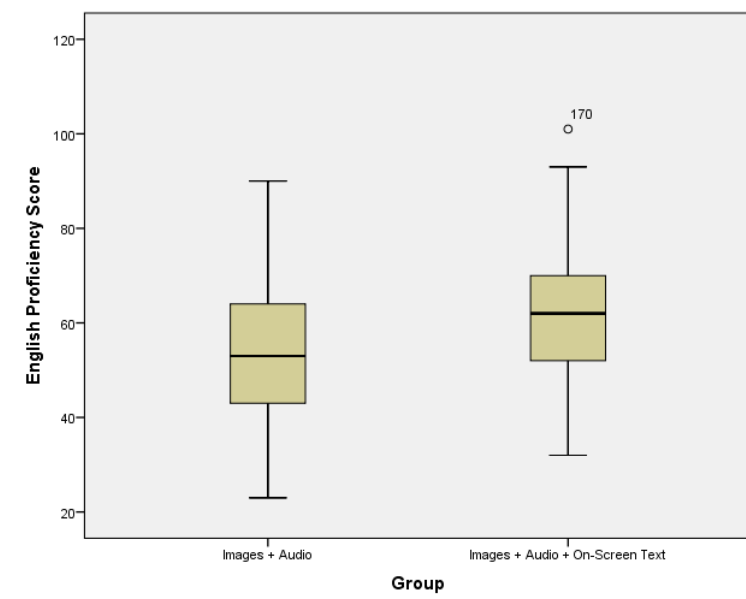
Descriptive Statistics for Age and English Proficiency of Participants

Variable	Images + Audio			Images + Audio + On-Screen Text			Combined		
	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>
Age	110	19.04	0.62	104	18.83	0.51	214	18.93	0.58
English Proficiency	110	53.73	15.09	104	61.55	13.32	214	57.53	14.75

The boxplots for the English proficiency scores of the two groups are presented in Figure 1. The median is approximately 54 for the Images and Audio Group and approximately 61 for the Images with Audio and On-Screen Text Group. The interquartile range and range are slightly larger for the Images and Audio Group, and there is one outlier in the Images with Audio and On-Screen Text Group. Because this outlier was not extreme in value, and the sample size was relatively large, this value was left intact for analysis.

Figure 1

Boxplots for English Proficiency Scores of Participants in Both Groups



Instrumentation and Materials

To achieve the research goal, there were two sets of instrumentation used in this study: two formats of a multimedia-enhanced lesson and tests.

Multimedia-Enhanced Lesson

Two different formats of a multimedia-enhanced lesson were created: (1) images with audio and (2) images with audio and on-screen text. The content of the lesson in both formats was exactly the same. The lesson was delivered in the form of a narrated video. In the first format, the video contained only images and audio while on-screen text was also included in addition to the images and audio in the second format of the lesson. The length of the lesson in both formats was 20 minutes. Since this was in an English class, the language used throughout the lesson was English. The lesson involved two related science topics, which were natural selection and evidence for evolution. This multimedia-enhanced lesson was a modified version of two combined videos originally created by Stated Clearly, a group of artists, scientists, and educators who have come together to create a series of short animations to teach science in a simple, friendly manner. The permission to use their videos as part of the research was granted by Jon Perry, the founder of Stated Clearly. The content presented in the lesson used in this study was validated by experts in the field. The lesson in both formats was installed into each of the computers in a computerized classroom in which the study was conducted. No internet access was required to complete the lesson.

Tests

Three tests were administered in this study: a pretest, an immediate posttest, and a delayed posttest. A total of 20 test items were developed by the researcher based on the content presented in the lesson. These 20 test items were used in the pretest to measure the participants' prior knowledge, in the immediate posttest to assess their knowledge gain, and in the delayed posttest to see how well the participants were able to retain the knowledge previously gained. The

same set of test items was used in order to ensure that the three tests are directly comparable. Although the three tests were identical in terms of their content, the sequence of the test items was arranged differently to minimize potential order effects that could occur. The tests were paper based, and they were mixed-format tests, consisting of different types of test items, which included multiple choice (5 items), short answer (4 items), fill-in-the-blank (1 item), true-false (3 items), and matching (7 items).

Content and Instrument Validation

The content of the tests was reviewed and validated by a team of experts, which included two second language instructors, two second language learners, and a scientist to ensure content accuracy and appropriateness of the tests. The team was asked to take the pretest before the lesson was given and take the posttest once the lesson was completed. A Wilcoxon signed-ranks non-parametric test (used because of the small sample size and lack of normality of the difference scores) was performed, and it was found that the participants, which were the team of experts, scored significantly higher on the posttest than the pretest ($z = 2.023$, $p = .043$). The effect size for the difference was large ($r = .90$). This provided supporting evidence for the validity of scores obtained from the test. Some of the test items, however, needed to be revised. There were items that appeared to be too complicated, as none of the participants were able to respond correctly to them even though the specific content knowledge needed was clearly presented in the video. Also, some of the test items seemed to measure extraneous knowledge as participants were able to get the correct answers simply by using prior, common knowledge. The tests used in this study were revised based on this process of content and instrument validation.

Data Collection Procedures

A list of participants of both groups—(1) images with audio and (2) images with audio and on-screen text—was prepared prior to data collection. The intervention took place in a computerized classroom where the researcher was available at all times. The arrangement was

made with each of the classes so that the instructors knew when they needed to bring their students into this computerized classroom. Students from the same classes stayed together and participated at the same time. All participants were given the same multimedia-enhanced lesson delivered in one of the two formats mentioned earlier. There were 11 classes of students participating in this study, and these students were randomly assigned into two groups. Approximately half of the students ($N = 110$) were in the images and audio group while the other half of them ($N = 104$) were in the images with audio and on-screen text group.

For each of the classes, the researcher began by informing the participants of the purpose of the study as well as the steps they would have to go through. The researcher also ensured that the participants were aware that their participation was voluntary and may be withdrawn at any time without penalty or prejudice. Participants were then given an informed consent form to sign. Only those who provided consent to participate were included in the study. Each participant was then given a small card containing his/her name, English proficiency test score, and a blank space for participant code. The Participants' names were used here only to match the English proficiency scores with the right participants. Their names were then removed and not present anywhere else during or after the intervention. Then, a paper-based pretest with a participant code written on the front page was randomly distributed to each participant. Next, participants were instructed to copy the participant code from the test to their card so that they could put this same code on the posttest, which would be given to them afterward. Before taking the pretest, participants were required to provide some demographic information, which included their gender, age, and English proficiency test score. Once the pretest was completed, participants proceeded to the multimedia-enhanced lesson previously installed on the computers. Participants were not allowed to take notes. As soon as the lesson was complete, a paper-based posttest was administered to assess the participants' knowledge gain. Here, participants were required to write their participant code in the box located on the front page of the test. This allowed the researcher

to anonymously match participants' pretests to their posttests. To measure how well the participants in each group were able to retain their knowledge previously gained, a delayed posttest was administered one week after the intervention. The participant codes were also required here so that all three tests of each participant could be put together.

Data Analysis

The independent variable of interest in this study was the intervention, with two formats of a multimedia-enhanced lesson: (1) images with audio and (2) images with audio and on-screen text. English proficiency scores and gender were used as control variables. There were two dependent variables—knowledge gain as indicated by growth scores from the pretest to the immediate posttest and knowledge retention as indicated by difference scores between the immediate posttest and the delayed posttest. Means, standard deviations, and ranges for pretest scores, immediate posttest scores, delayed posttest scores, growth scores, difference scores between immediate posttest scores and delayed posttest scores, and English proficiency scores were computed using SPSS. A repeated-measures analysis of covariance (ANCOVA) was carried out to evaluate (1) change between pretest and immediate posttest, (2) mean differences in knowledge gain between the two groups, (3) group differences in growth from pretest to immediate posttest, (4) change between immediate posttest and delayed posttest, and (5) group differences in knowledge retention as indicated by difference scores between immediate posttest and delayed posttest, using gender and English proficiency as control variables. Additionally, an analysis of covariance was performed to assess group differences in immediate posttest scores, controlling for pretest scores, delayed posttest scores, gender, and English proficiency. Then, an independent samples t-test was performed to assess group differences in immediate posttest scores and delayed posttest scores. This was done in order to investigate the effect of the two formats of multimedia-enhanced instruction on second language learners' knowledge retention.

A power analysis was performed, and it was found that, assuming a moderate effect size for the difference between groups in the population, and using $\alpha = .05$, a total sample size of $N = 128$ participants would be required for 80% power. Power analysis also indicated that, assuming a large effect size for the difference between groups in the population, and using $\alpha = .05$, a total sample size of $N = 52$ participants would be required for 80% power. The details of these analyses are provided in Table 3. Eta-squared (η^2) was used to determine the effect size. After the data analysis was complete, the results of this study were discussed, and an interpretation of the data was written.

Table 3*Power Analysis Results*

F tests - ANCOVA: Fixed effects, main effects and interactions		
Analysis: A priori: Compute required sample size		
Input:	Effect size f	= 0.40
	α err prob	= 0.05
	Power (1- β err prob)	= 0.80
	Numerator df	= 1
	Number of groups	= 2
	Number of covariates	= 2
Output:	Noncentrality parameter λ	= 8.3200000
	Critical F	= 4.0426521
	Denominator df	= 48
	Total sample size	= 52
	Actual power	= 0.8068454
Analysis: A priori: Compute required sample size		
Input:	Effect size f	= 0.25
	α err prob	= 0.05
	Power (1- β err prob)	= 0.80
	Numerator df	= 1
	Number of groups	= 2
	Number of covariates	= 2
Output:	Noncentrality parameter λ	= 8.0000000
	Critical F	= 3.9175498
	Denominator df	= 124
	Total sample size	= 128

Actual power = 0.8013621

Statistical Analysis

Prior to analyses, internal consistency estimates of reliability (Cronbach's coefficient alpha) were calculated for the pretest, the immediate posttest, and the delayed posttest used in this study and descriptive statistics for the entire sample were computed. Cronbach's alpha was .63 for the pretest, .76 for the immediate posttest, and .71 for the delayed posttest. Table 4 reports the means, standard deviations, and skewness of pretest, immediate posttest (Posttest 1), delayed posttest (Posttest 2), growth from pretest to Posttest 1, and change from Posttest 1 to Posttest 2 of participants in the Images with Audio Group and the Images with Audio and On-Screen Text Group and of all participants combined.

Table 4

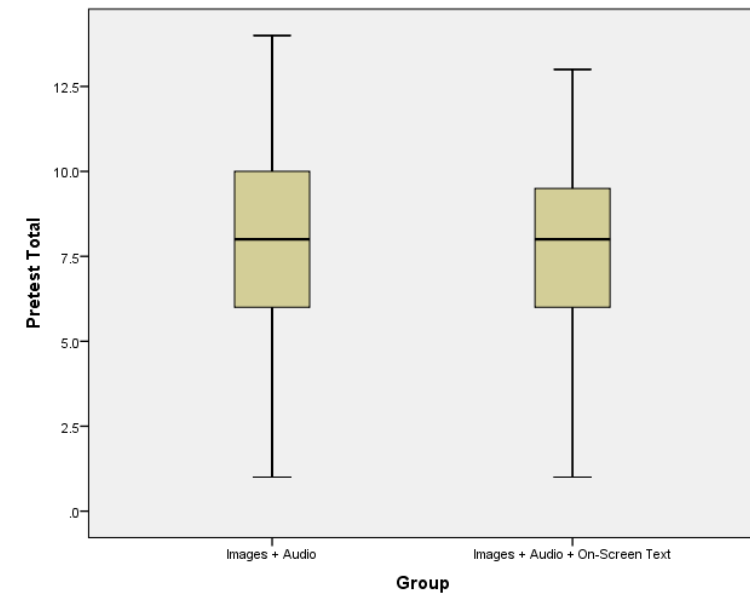
Descriptive Statistics for Pretest, Posttest 1, Posttest 2, Growth, and Change Scores

Variable	Images + Audio				Images + Audio + On-Screen Text				Combined			
	<i>N</i>	Mean	<i>SD</i>	Skew	<i>N</i>	Mean	<i>SD</i>	Skew	<i>N</i>	Mean	<i>SD</i>	Skew
Pretest	110	7.77	2.86	-0.32	104	7.41	2.91	-0.34	214	7.06	2.88	-0.33
Posttest 1	110	12.65	3.53	-0.52	104	15.04	2.89	-0.68	214	13.81	3.44	-0.66
Posttest 2	110	12.46	3.31	-0.49	104	14.43	2.96	-0.90	214	13.42	3.30	-0.66
Growth	110	4.87	3.13	-0.06	104	7.63	2.99	0.33	214	6.21	3.35	0.45
Change	110	0.18	2.35	-0.27	104	0.61	2.18	-0.02	214	0.39	2.27	-0.18

The boxplots for the pretest scores of the two groups as shown in Figure 2 appear to be quite symmetrical. The median is approximately 7.5. The interquartile range and range are slightly larger for the Images with Audio Group than for the Images with Audio Group and On-Screen Text Group.

Figure 2

Boxplots for the Pretest Scores of Participants Both Groups



Research Questions and Results

Research Question #1: Is there a significant difference in knowledge gain between two groups of second language learners presented with different formats of a multimedia-enhanced lesson: (1) images with audio and (2) images with audio and on-screen text? A repeated-measures analysis of covariance (ANCOVA) was performed to evaluate change between pretest and immediate posttest, mean differences in knowledge gain between the two groups, and group differences in growth from pretest to immediate posttest. Gender was employed as a factor and English proficiency score as a covariate. Additionally, analysis of covariance was also conducted to assess group differences in immediate posttest scores, controlling for pretest scores, gender, and English proficiency. Eta-squared (η^2) and Cohen's d were used to compute the effect size.

For the repeated-measures ANCOVA analyses, which used the pretest and posttest comprehension scores as the paired outcomes, Box's test of equality of covariance matrices indicated that covariances

were equal across groups, $p = .429$. Table 5 shows the results for ANCOVA. The results indicated that, across groups, a statistically significant change between pretest to immediate posttest occurred, $F(1, 209) = 14.26$, $p < .001$. Computation of eta-squared suggested a small effect size ($\eta^2 = 0.05$); however, this statistic was likely influenced by sample size and reflected the large amount of unexplained error variation in the data. Cohen's d , which reflects a standardized index of mean growth as opposed to variance accounted for, was $d = 1.85$, indicating a very large increase in knowledge gained. The observed growth from pretest to immediate posttest also differed significantly between the two treatment groups, $F(1, 209) = 29.14$, $p < .001$. The effect size was moderate ($\eta^2 = 0.11$). Figure 3 shows that participants in the Images with Audio and On-Screen Text Group appear to display a significantly higher amount of growth than those in the Images with Audio Group. However, there was no significant difference in growth between males and females, $F(1, 209) = 0.53$, $p = .469$. The results also suggest that there was no significant three-way group \times gender \times time interaction effect, $F(1,209) = 0.09$, $p = .762$.

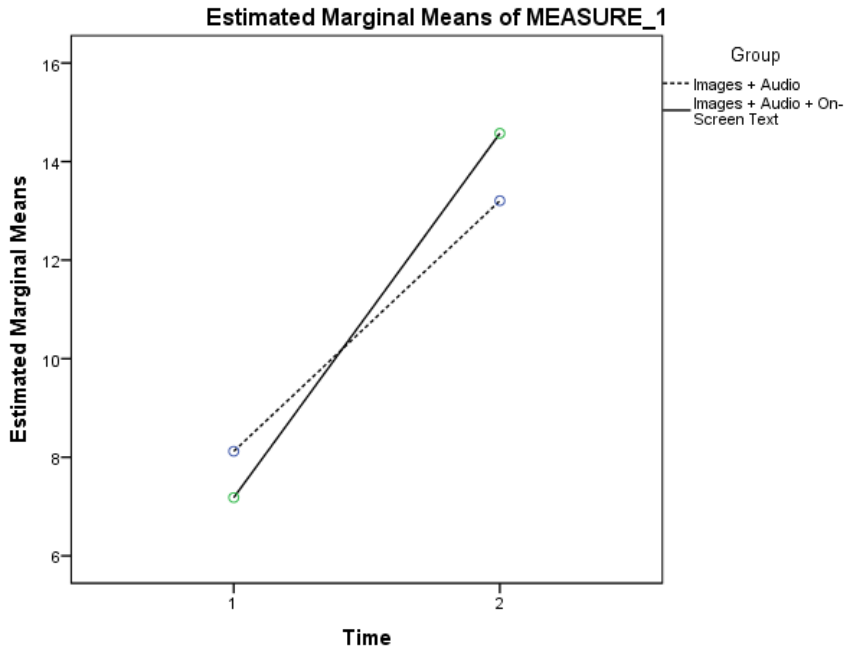
Table 5

Results for the Repeated-Measures Analysis of Covariance (ANCOVA) of Knowledge Gain Across Time by Group and Gender

Source	Type III Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>
Time	63.478	1	63.478	14.261	<.001
Time \times English Proficiency	57.094	1	57.094	12.827	<.001
Time \times Group	129.713	1	129.713	29.141	<.001
Time \times Gender	2.345	1	2.345	.527	.469
Time \times Group \times Gender	.408	1	.408	.092	.762
Error(Time)	930.292	209	4.451		

Figure 3

Mean Growth Scores (Knowledge Gain) Across Time by Group



Covariates appearing in the model are evaluated at the following values: English Proficiency Score = 57.53

The second analysis employed ANCOVA, with the immediate posttest scores as the outcome. Levene's test was not statistically significant, $p = .721$, indicating equal variances across groups. As displayed in Table 6, a test of between-subject effects indicated that group had a statistically significant effect on the immediate posttest scores, $F(1,208) = 22.78$, $p < .001$. The effect size was $\eta^2 = 0.06$, suggesting a moderate effect. Analysis of the data shows that participants in the Images with Audio and On-Screen Text Group appear to perform significantly better than those in the Images with Audio Group. However, gender and group \times gender interaction were not found to significantly affect the immediate posttest scores, $F(1,208) = 0.58$, $p = .447$ and $F(1,208) = 1.91$, $p = .169$ respectively.

Table 6*Results for the Analysis of Covariance (ANCOVA)*

Source	Type III Sum of		Mean		
	Squares	<i>df</i>	Square	<i>F</i>	<i>p</i>
Corrected Model	1215.172 ^a	5	243.034	38.531	<.001
Intercept	311.658	1	311.658	49.410	<.001
English Proficiency	339.038	1	339.038	53.751	<.001
Pretest Total	238.746	1	238.746	37.851	<.001
Group	143.671	1	143.671	22.778	<.001
Gender	3.662	1	3.662	.581	.447
Group × Gender	12.028	1	12.028	1.907	.169
Error	1311.973	208	6.308		
Total	43331.000	214			
Corrected Total	2527.145	213			

a. R Squared = .481 (Adjusted R Squared = .468)

Research Question #2: What is the effect of the two formats of multimedia-enhanced instruction on second language learners' knowledge retention? As displayed in Table 7, the difference mean score between immediate posttest and delayed posttest was 0.18 for the Images and Audio Group and 0.61 for the Images with Audio and On-Screen Text Group.

Table 7*Descriptive Statistics for Group Differences in Immediate and Delayed Posttest Scores*

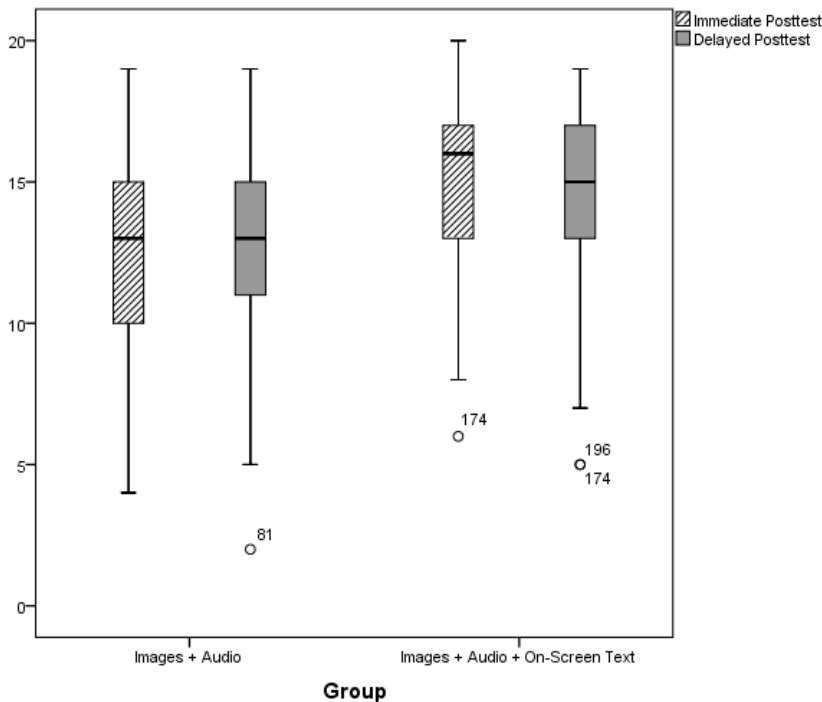
	Group Statistics (Knowledge Retention)			
	N	Mean	SD	Std. Error Mean
Images + Audio	110	0.18	2.35	0.22
Images + Audio + On-Screen Text	104	0.61	2.18	0.21

The boxplots for the immediate posttest scores and the delayed posttest scores of the two groups are displayed in Figure 4. For the

Images with Audio Group, the medians of the immediate and the delayed posttests appear to be at the same level, which is approximately 13. However, the interquartile range and range are slightly larger for the immediate posttest than for the delayed posttest. For the Images with Audio and On-Screen Text Group, the median of the immediate posttest is slightly higher than that of the delayed posttest, which are 16 and 15 respectively. The interquartile range and range of the immediate and delayed posttests, however, appear to be similar.

Figure 4

Boxplots for the Immediate and the Delayed Posttest Scores of Participants in Both Groups



An independent samples t-test was carried out to assess group differences in immediate posttest scores and delayed posttest scores in order to investigate the effect of the two formats of multimedia-enhanced instruction on second language learners' knowledge

retention. There was no statistically significant difference in immediate posttest scores and delayed posttest scores between the Images with Audio Group and the Images with Audio and On-Screen Text Group, $t(212) = -1.37, p = .17$ as displayed in Table 8.

Table 8

Independent Sample Test for Learners' Knowledge Retention

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Retention	Equal variances assumed	.051	.822	-1.366	212	.173	-.42395	.31030
	Equal variances not assumed			-1.369	211.917	.172	-.42395	.30963

Discussion of Findings

This quasi-experimental study was conducted to investigate the effects of different formats of multimedia-enhanced instruction on knowledge gain and retention of second language learners. The focus of the study was placed specifically on Mayer's redundancy principle, which is part of the cognitive theory of multimedia learning. The second language learners' prior knowledge of the learning content was assessed by a pretest before they engaged with one of the two formats of a multimedia-enhanced lesson. After that, a posttest was administered to measure their learning outcomes. To answer the research question of whether second language learners participating in a multimedia-enhanced lesson that included only images and audio were different in terms of knowledge gain from those who participated in a multimedia-enhanced lesson containing images, audio, and on-screen text, their growth scores, which were the differences between pretest and posttest scores, were compared. Then, to investigate the effect of the two formats of multimedia-enhanced instruction on the

learners' knowledge retention, an independent samples t-test was carried out to assess group differences in immediate posttest scores and delayed posttest scores.

According to the cognitive theory of multimedia learning, humans have two separate cognitive channels for processing information: the visual channel and the auditory channel, and each channel has limited capability for processing (Mayer 2002). The redundancy principle, which is based on the cognitive theory of multimedia learning, suggests that humans learn better and more deeply with images and narration as opposed to images, narration, and on-screen text.

When images are already explained through narration, visual text, or on-screen text, becomes redundant and should be eliminated. Mayer (2002) explains that cognitive overload could occur in the visual channel when both images and text are presented visually at the same time. Learning is, therefore, hindered under such an instructional condition as it is hard for learners to focus their attention on all three types of information presented to them at the same time (Clark & Mayer, 2008; Hoffman, 2006; Mayer, 2002; Mayer et al., 2001).

The psychological reason for this is that when all three sources of information are presented together in a multimedia learning environment, images enter the learner's cognitive system through the eyes and are processed in the visual channel, while the narration enters through the ears and is processed in the auditory channel. However, the on-screen text also enters the learner's cognitive system through the eyes and is processed in the visual channel. This means the limited cognitive resources in the visual channel must be shared in processing both the images and the on-screen text, and especially, if the pace of the lesson is fast, it is very likely that cognitive overload will occur. As a result, some information or learning content presented may not be selected and transferred into a mental representation (Clark & Mayer, 2008). On the other hand, when only images and narration are presented, the images, which is in the form of pictorial information, enters through the eyes and is processed in the visual channel while the narration, which is in the form of verbal information, enters

through the ears and is processed in the auditory channel. With redundant information (i.e., on-screen text) eliminated, the chances for cognitive overload is minimized. Learners are, therefore, able to engage in appropriate cognitive processing and learn more effectively (Clark & Mayer, 2008; Mayer, 2002). This was supported by a number of research studies such as those conducted by Acha (2009) and Moussa-Inaty and Atallah (2012) who found that including redundant text in a multimedia lesson appeared to impair learning as their participants' working memory could have been overloaded due to redundancy when images and on-screen text were presented simultaneously. DeLeeuw et al. (2010) employed eye tracking to test how the addition of written text (as opposed to narration) could affect how a learner processed a diagram in multimedia learning. Their results suggested that accompanying diagrams with written text could lead to visual overload.

The results of the present study, however, showed that participants, who were second language learners of English, engaging with both formats of multimedia-enhanced instruction: Images with Audio and Images with Audio and On-Screen Text, improved significantly from pretest to posttest, $p < .001$. Moreover, participants who engaged with the multimedia-enhanced instruction where on-screen text was included were also found to perform significantly better than participants who were in the group where on-screen text was not present, $p < .001$. It might not be clear, though, whether or not cognitive overload in the visual channel occurred as predicted by the redundancy principle, but if it did, then it would have been so little that it did not impede learning in this study. As for knowledge retention, learners in both groups appeared to be able to retain their knowledge at the same level as demonstrated by their change scores from the immediate posttest to the delayed posttest, which were not significantly different, $p = .17$.

These results appear to contradict the redundancy principle as having on-screen text, which is considered redundant, presented together with images and audio in a multimedia-enhanced lesson was not detrimental to learning. This is also consistent with the results obtained from the researcher's previous work in which learners who

participated in a multimedia learning environment where on-screen text was provided were able to achieve higher posttest and growth scores than those in a group where on-screen text was not provided. However, the difference in scores between the two groups at the time was not statistically significant (Ramsin & Mayall, 2019). These results supporting the use of on-screen text to facilitate learning were in line with what was found in previous studies on multimedia learning. Participants' learning outcomes were found to significantly improve whether or not visual text was included while participating in a multimedia lesson (Karakas & Saricoban, 2012; Yuksel & Tanriverdi, 2009). Their results, as well, suggests that providing on-screen text in addition to images and audio did not appear to adversely impact learners. Another study was conducted by Winke et al. (2010) to examine the effects of captioning during video-based listening activities. Their results revealed that having both images and written text together on screen appeared to enhance the learning outcomes of foreign language learners. The written text, presented in the form of captions in their videos, appeared to be a crucial element and was shown to be beneficial to students as it enabled them to understand the videos better. This is consistent with what was found in a study conducted by Munassar et al. (2010). In their study, the researchers compared two groups of EFL students participating in two instructional modes: redundancy mode and modality mode. Results revealed that learners who were exposed to the redundancy mode performed significantly better and were significantly more motivated than their counterparts in the modality mode. In the same way, McCrudden et al. (2014) carried out an experiment to clarify the boundary conditions for determining when redundancy is beneficial in multimedia learning. Specifically, they wanted to find out if embedding text segments within a visual display redundant with the text passage would affect memory and transfer. The redundant text segments were found to enhance performance on measures of memory, and the statistically significant effect sizes on the memory tests were large, suggesting that embedding redundant text within visual displays may be of practical importance.

As suggested by the results of this study, providing all three types of information—pictorial information, verbal information, and visual text information—was not found to negatively affect learning. It, instead, appeared to be helpful to learners as demonstrated by their learning outcomes, which were significantly better than when only two types of information—pictorial information and verbal information—were available. The positive effect of including on-screen text found here could possibly be backed up by the theory of second language acquisition, particularly the input hypothesis (Krashen, 1982). According to the input hypothesis, receiving comprehensible input is necessary in language learning (Ipek, 2009; Johnson, 2004; Krashen, 1985a). However, language learning materials should be slightly above learners' abilities. That is because when learning materials are too easy, learners may get bored and stop paying attention. In contrast, learners may struggle and give up if learning materials are too difficult. According to Danan (2004), language learners often find audiovisual input in a foreign language challenging and difficult to comprehend, and so the use of on-screen text to supplement audio narration in this study can, therefore, be considered as comprehensible input because, with on-screen text made available, learners are able to visualize what they hear and can be more certain of ambiguous input. When written text is provided on screen, learners are also able to analyze words and phrases, which can lead to a better understanding of the learning content being presented (Chai & Erlam, 2008; Danan, 2004; Winke et al., 2013).

Conclusion

Multimedia provides richer and more effective learning environments for different types of learners as evidenced in extensive research previously conducted. However, designing effective multimedia learning environments can be a challenging task and they can come in a wide variety of different formats. Mayer (2002) suggests very helpful and widely accepted multimedia design principles as part of his cognitive theory of multimedia learning to provide guidelines for instructional designers and educators who wish to incorporate

multimedia into their teaching, which may include designing a PowerPoint presentation, developing an online course, or creating a video to be used in a flipped classroom. The multimedia design which was the focus of the present study was the redundancy principle, of which the fundamental proposition is that people learn better and more deeply from pictures and narration than from pictures, narration, and on-screen text. It is better to eliminate on-screen text because it is redundant and can overwhelm the visual channel as the learner will need to process both images and on-screen text at the same time (Clark & Mayer, 2008; Mayer, 2002).

As implied by the results of this study, including on-screen text, although redundant, may be beneficial to some groups of people such as second language learners, who can make use of the written text provided to them in making sense of the learning content presented in a multimedia-enhanced lesson. Various factors such as the learner's needs, subject areas, and learning context should all be taken into consideration when designing a multimedia-enhanced lesson because even a carefully designed lesson that works well with one group of learners might not be successful when used in a different learning context. Therefore, it is important that multimedia-enhanced lessons are customized to fit the target groups of learners and to better address their needs. Further research might investigate how multimedia should be incorporated into lessons designed for different groups of learners in other academic disciplines to effectively facilitate learning.

References

- Abdul Samat, M. S., & Abdul Aziz, A. (2020). The effectiveness of multimedia learning in enhancing reading comprehension among indigenous pupils. *Arab World English Journal*, 11(2), 290–302. <https://dx.doi.org/10.24093/awej/vol11no2.20>
- Acha, J. (2009). The effectiveness of multimedia programmes in children's vocabulary learning. *British Journal of Educational Technology*, 40(1), 23–31.
- Adegoke, B. A. (2010). Integrating animations, narratives and textual information for improving Physics learning and listening skills.

- Electronic Journal of Research in Educational Psychology*, 8(2), 725–748.
- Ahmad, W. I. W., & Yamat, H. (2020). Students' perception on learning English language through conventional and digital storytelling. *International Journal of Academic Research in Business and Social Sciences*, 10(2), 484–504.
- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556–559. <https://doi.org/10.1126/science.1736359>
- Brasel, S. A., & Gips, J. (2014). Enhancing television advertising: Same-language subtitles can improve brand recall, verbal memory, and behavioral intent. *Journal of the Academy of Marketing Science*, 42(3), 322–336. <https://doi.org/10.1007/s11747-013-0358-1>
- Chai, J., & Erlam, R. (2008). The effect and the influence of the use of video and captions on second language learning. *New Zealand Studies in Applied Linguistics*, 14(2), 25–44.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2
- Cherry, K. (2022, March 24). *What is learning?* Very Well Mind. <https://www.verywellmind.com/what-is-learning-2795332>
- Chuang, H-H., & Liu, H-C. (2012). Effects of different multimedia presentations on viewers' information-processing activities measured by eye-tracking technology. *Journal of Science Education and Technology*, 21(2), 276–286.
- Clark, R. C., & Mayer, R. E. (2008). *E-learning and the science of instruction*. Pfeiffer.
- Cook, V. (2001). *Second language learning and language teaching*. Oxford University Press.
- Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta*, 49(1), 67–77. <https://doi.org/10.7202/009021ar>
- DeLeeuw, K. E., Mayer, R. E., & Giesbrecht, B. (2010). How does text affect the processing of diagrams in multimedia learning? In A. K. Goel, M. Jamnik & N. H. Narayanan (Eds.), *Diagrammatic*

- representation and inference* (pp. 304–306). Springer.
https://doi.org/10.1007/978-3-642-14600-8_36
- Domjan, M. (2015). *The principles of learning and behavior* (7th ed.). Cengage Learning.
- Erizar, Syahputra, A., & Hidayati, T. (2018). The effectiveness of use multimedia in English teaching at SMPN 1 Meulaboh. *Advances in Social Science, Education and Humanities Research*, 208, 47–48. <https://doi.org/10.2991/icssis-18.2019.10>
- Flaherty, C. (1985). *Animal learning and cognition*. McGraw-Hill College.
- Gordon, W. C. (2000). *Learning and memory*. Cengage Learning.
- Hoffman, B. (2006). *The encyclopedia of educational technology*. Montezuma Press.
- Horz, H. & Schnotz, W. (2008). Multimedia: How to combine language and visuals. *Language at work*, 4(4), 43–50.
- Houston, J. P. (2001). *Fundamentals of learning and memory* (4th ed.). Harcourt Brace Jovanovich College Publishers.
- Ilhan, G. O., & Oruç, S. (2016). Effect of the use of multimedia on students' performance: A case study of social studies class. *Educational Research and Reviews*, 11(8), 877–882.
- Ipek, H. (2009). Comparing and contrasting first and second language acquisition: Implications for language teachers. *English Language Teaching*, 2(2), 155–163.
- Issa, N., Schuller, M., Santacaterina, S., Shapiro, M., Wang, E., Mayer, R. E., & DaRosa, D. A. (2011). Applying multimedia design principles enhances learning in medical education. *Medical Education*, 45, 818–826. <https://doi.org/10.1111/j.1365-2923.2011.03988.x>
- Jae, H. (2019). The effectiveness of closed caption videos in classrooms: Objective versus subjective assessments. *Journal of Instructional Pedagogies*, 22.
- Johnson, M. (2004). *A philosophy of second language acquisition*. Yale University Press.

- Karakas, A., & Saricoban, A. (2012). The impact of watching subtitled animated cartoons on incidental vocabulary learning of ELT students. *Teaching English with Technology*, 12(4), 3–15.
- Kosar, G., & Bedir, H. (2018). Improving knowledge retention via establishing brain-based learning environment. *European Journal of Education Studies*, 4(9), 208–218.
<https://doi.org/10.5281/zenodo.1298918>
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Krashen, S. D. (1985a). *Inquiries and insights*. Alemany Press.
- Krashen, S. D. (1985b). *The input hypothesis: Issues and implications*. Longman.
- Krashen, S., & Bland, J. (2014). Compelling comprehensible input, academic language and school libraries. *CLELE Journal*, 2(2), 1–12.
- Lewis, R. (2020, November 4). *What is comprehensible input and why does it matter for language learning?* Leonardo English.
<https://www.leonardoenglish.com/blog/comprehensible-input>
- Linebarger, D. L. (2001). Learning to read from television: The effects of using captions and narration. *Journal of Educational Psychology*, 93(2), 288–298.
- Lusk, D. L., Evans, A. D., Jeffrey, T. R., Palmer, K. R., Wikstrom, C. S., & Doolittle, P. E. (2009). Multimedia learning and individual differences: Mediating the effects of working memory capacity with segmentation. *British Journal of Educational Technology*, 40(4), 636–651.
- Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, 32(1), 1–9.
- Mayer, R. E. (2002). *Multimedia learning*. Cambridge University Press.
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.
- Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 43–71). Cambridge University Press.

- Mayer, R. E. (2014). Introduction to multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 1–24). Cambridge University Press.
- Mayer, R. E., Heiser, H., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*(1), 187–198.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52.
- McCrudden, T. M., Hushman, C. J., & Marley, S. C. (2014). Exploring the boundary conditions of the redundancy principle. *The Journal of Experimental Education, 82*(4), 537–554.
- Moussa-Inaty, J., & Atallah, F. (2012). Multimedia use in higher education in the UAE: A cognitive load theory perspective. *Journal of Educational Multimedia and Hypermedia, 21*(2), 127–142.
- Munassar, W. A. S., Yahaya, W. A. J. W., & Chong, T. S. (2010). When is the redundancy principle effective in the multimedia learning of a foreign language? *Journal of Institutional Research South East Asia, 8*(2), 16–36.
- Neuman, S. 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.
- Niedfeldt, K. S. (2001). *A study of the impact of time lapse on language retention in beginning level Spanish classes*. [Master's thesis, University of Wisconsin-Stout].
<https://www.semanticscholar.org/paper/A-STUDY-OF-THE-IMPACT-OF-TIME-LAPSE-ON-LANGUAGE-IN-Niedfeldt/388ec86380187621cf085e44be8a9d45e10276a2>
- Niedzielski, B. (2020, January 14). *Input hypothesis*. UCLA Humanities Technology (HumTech).
<https://humtech.ucla.edu/instruction/input-hypothesis/>
- Ozdemir, M., Izmirlı, S., & Sahin-Izmirlı, O. (2016). The effects of captioning videos on academic achievement and motivation:

- Reconsideration of redundancy principle in instructional videos. *Educational Technology & Society*, 19(4), 1–10.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Perez, M. M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41(3), 720–739.
- Ramsin, A., & Mayall, H. (2019). Examining the effects of different multimedia learning environments on the learning outcomes of second language learners. *Journal of Educational Multimedia and Hypermedia*, 28(3), 307–330.
- Schilling, K. (2009). The impact of multimedia course enhancements on student learning outcomes. *Journal of Education for Library and Information Science*, 50(4), 214–225.
- Smith, M. S. (1994). *Second language learning: Theoretical foundations*. Longman.
- Trevisan, M. S., Oki, A. C., & Senger, P. L. (2010). An exploratory study of the effects of time compressed animated delivery multimedia technology on student learning in reproductive physiology. *Journal of Science Education and Technology*, 19(3), 293–302. <https://doi.org/10.1007/s10956-009-9200-4>
- van Merriënboer, J. J. G., & Kirschner, P. A. (2018). *Ten steps to complex learning: A systematic approach to four-component instructional design* (3rd ed.). Routledge.
- Wang, L., & Li, J. (2019). Development of an innovative dual-coded multimedia application to improve reading comprehension of students with imagery deficit. *Journal of Educational Computing Research*, 57(1), 170–200.
- Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14(1), 65–86.
- Winke, P., Gass, S., & Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *The Modern Language Journal*, 97(1), 254–275.

Yuksel, D., & Tanriverdi, B. (2009). Effects of watching captioned movie clip on vocabulary development of EFL learners. *The Turkish Online Journal of Educational Technology*, 8(2), 48–54.