



---

## The More, the Better? Effects of Multiple Modalities on EFL Listening and Reading Comprehension \*

Kim, Na-Young<sup>1</sup>

---

### Abstract

This study explores the effectiveness of multimedia modalities by comparing the effects of the use of different modality modes on EFL comprehension. The experiment was conducted during the 2020 academic year. The participants were 186 college EFL students who were selected from six intact classes at a university in Korea. They were divided into six groups at random and received the treatment depending on their experimental conditions: text-only (TO), audio-only (AO), text-picture (TP), audio-picture (AP), text-audio (TA), and text-audio-picture (TAP). For five weeks, all participants read, listened to, or both read and listened to the fairy tales with or without pictures in class. To test the modality effects, the TOEIC-based listening and reading pre- and post-tests were carried out. The whole data collected were analyzed with SPSS version 20 software. The study found that using different modalities in classes plays a beneficial role in increasing comprehension skills. TP, AP, and TA groups improved their listening skills while TO, TP, and TA groups increased their reading skills at a significance level of .05. Statistically significant group differences were also found based on experimental conditions in both listening and reading comprehension skills. Pedagogical implications and suggestions are discussed at the end.

**Keywords:** multimedia technology, modality mode, listening comprehension, reading comprehension, EFL learning

**Applicable level:** tertiary

---

\* This paper was supported by the Sehan University Research Fund in 2021.

<sup>1</sup> Corresponding author, Assistant Professor, Department of General Education, Sehan University, 33, Sehandae-gil, Sinpyeong-myeon, Dangjin-si, Chungcheongnam-do, Korea (E-mail: [nykim@sehan.ac.kr](mailto:nykim@sehan.ac.kr))

Received: June 30, 2021  
Revised: August 16, 2021  
Accepted: August 23, 2021

Copyright: © 2021 The Society for Teaching English through Media (STEM)  
This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

---

## I. INTRODUCTION

Multimedia can present information in multiple modalities including visual, auditory, and tactile modalities. According to Kalyuga (2000), one of the common sense beliefs surrounding multimedia is about the effectiveness of multiple modalities in instruction. Since the limited working memory can be expanded by using multiple sensory modalities, the amount of information processed through two or more channels can exceed the processing capacity of one single channel while reducing the cognitive load (Kalyuga, 2000). Therefore, using more than two multimedia modalities (e.g., both visual and auditory channels) are believed to be better than using one modality (e.g., either visual or auditory channel) in multimedia learning.

Mayer (2002) added that we learn best when utilizing multiple sensory channels as the use of multiple modalities extends the capacity of working memory to handle the information. It is likely that the more modalities we use, the better we learn. However, according to Kalyuga (2000), adding spoken words (e.g., audio explanations—auditory channel) to written texts (e.g., visual instructions—visual channel) can have negative rather than positive or neutral effects on learning. As both are verbally presented information, they are solely processed by the verbal working memory system while the nonverbal system left unused. Since the verbal working memory is overloaded, there are negative effects on learning. This can be explained by Paivio's (1991) dual coding theory, suggesting that we have two working memory systems, verbal and nonverbal, and meaningful learning occurs by using both processing systems. Therefore, even the use of both sensory modalities—auditory and visual channels—may not be effective in learning due to the working memory load (Kalyuga, 2000).

In this light, Mayer (2002) argued that the best way to practice multimedia instruction is using voice narration and visual graphics together, as this takes advantage of using both verbal and nonverbal working memories with two sensory channels—auditory and visual—without overloading one or the other. Sorden (2005) also claimed that the best learning occurs when pictures are combined with audio. This is based on the cognitive theory of multimedia learning (Mayer, 2002), drawing on Paivio's (1991) dual coding theory, Baddeley's (1992) working memory theory, and Sweller's (1988) cognitive load theory. To sum up, combining sound (verbal and auditory) with image (nonverbal and visual) is highly beneficial in learning by reducing the cognitive load while extending the working memory capacity.

In foreign language learning, the use of multimedia has become important. They have provided students with information in a non-linear way with different modalities. Thus, research studies have largely been conducted to investigate whether multimedia-assisted learning is effective for foreign language instruction. Many of them have revealed that the use of multimedia technology significantly improves students' foreign language comprehension skills (Chen, 2006; Diao & Sweller, 2007; Hsieh & Huang, 2020; Karbalaei & Zare, 2019; Kartal & Simsek, 2017; Mohamed, 2018; Rahman & Hajar, 2020; Turker, 2010; Yeh & Wang, 2003).

Nonetheless, Soruc (2015) pointed out that much research on multimedia learning is still needed especially in EFL settings. Particularly, there is a lack of empirical studies directly evaluating the effectiveness of different modality modes, and even so, previous studies have yielded conflicting results (Rogowsky, Calhoun, & Tallal, 2016). For example, Liu, Cao, and Wu (2019) claimed that the use of auditory modality (audio) is more effective than that of visual one (text), but Moyer (2011) found no significant difference between the two modalities. Studies comparing the effects of single (either text or audio) with dual (both text and audio) modality modes have also generated inconclusive results (Chang, 2009; Diao & Sweller, 2007; Kartal & Simsek, 2017; Liu et al., 2019; Mohamed, 2018; Rahman & Hajar, 2020; Rogowsky et al., 2016; Sweller, 2005b). In terms of the effectiveness of adding the visual input (image), some proved its positive effects on comprehension (Erfani, 2012; Hsieh & Huang, 2020) while others failed (Schnotz & Bannert, 2003; Suvorov, 2009). Accordingly, there have been contradictory findings regarding the effects of using text-picture and audio-picture modalities on comprehension skills (Chen, 2006; Karbalaei & Zare, 2019). Last but not least, although there has been a common belief that using multiple modalities is better than single one (Kalyuga, 2000; Mayer, 2002), scholars have challenged this 'the more, the better' belief in multimedia instruction (Mayer, Heiser, & Lonn, 2001; Yeh & Wang, 2003).

Taking all this into consideration, it is necessary to examine the modality effects on EFL learning in Korea. In light of this, this study aims to investigate whether different modalities all play a significant role in EFL listening and

reading comprehension skills. In particular, the current study tries to see if the use of more modalities results in more comprehension in Korean EFL context by comparing different modality modes including text-only, audio-only, text-picture, audio-picture, text-audio, and text-audio-picture. Research questions are as follows:

1. Can six modality modes (text-only, audio-only, text-picture, audio-picture, text-audio, and text-audio-picture) enhance listening and reading comprehension skills among Korean EFL students?

2. Are there any statistically significant differences in six modality modes regarding the effects on listening and reading comprehension skills among Korean EFL students?

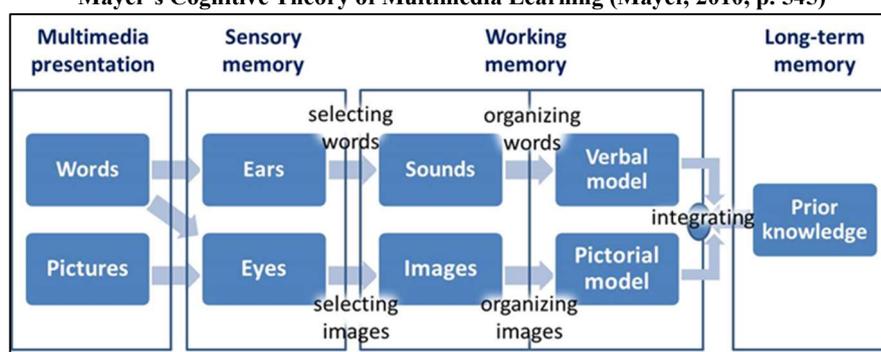
## II. LITERATURE REVIEW

### 1. Multimedia Learning

Multimedia learning happens when mental representations are built from words and pictures (Sorden, 2012). With the advancement of multimedia technologies, it has been easy for teachers to utilize multiple modalities in learning (Mayer, 2002). Now they can easily provide their students with visual information while also presenting related or supporting information through audio. According to Kalyuga (2000), using both visual and auditory channels is highly beneficial as the use of the two modalities extends the capacity of working memory to process the information. Furthermore, in multimedia learning, the information that cannot be understood in one channel can be integrated with the information in the other channel to be fully understood. Sorden (2005) noted that this is one of the strong points of multimedia instruction.

Mayer (2002) defined multimedia learning as the presentation of material using words and pictures. His cognitive theory of multimedia learning stated that auditory narration and graphical images produce both verbal and visual mental representations and integrate with prior knowledge to construct new knowledge (Mayer & Moreno, 2002). In multimedia learning, students attempt to make meaningful connections between words and pictures and learn more deeply than with words or pictures alone (Mayer, 2002). According to Mayer's Cognitive Theory of Multimedia Learning (Mayer, 2010), multimedia instruction aims to encourage students to build the coherent mental representations from presented materials. Finally, the theory tries to show how to structure multimedia instructional practices by employing cognitive strategies to help the students learn more efficiently (Sorden, 2012).

**FIGURE 1**  
Mayer's Cognitive Theory of Multimedia Learning (Mayer, 2010, p. 545)



There are three memory stores in the cognitive theory of multimedia learning (Mayer, 2010). They are sensory memory, working memory, and long-term memory. As a cognitive structure, sensory memory allows us to perceive new information (Sweller, 2005a). According to Mayer (2005), visual sensory memory briefly holds pictures and written texts as visual images while auditory sensory memory briefly holds sounds and spoken words as auditory images. Working memory is another cognitive structure in which information is consciously processed (Sweller, 2005a). It attends to or selects information from the sensory memory for processing and integration. The other cognitive structure is long-term memory that stores our knowledge base. It holds the entire store of knowledge for an

indefinite amount of time. According to Sorden (2012), we recognize the information in long-term memory when it's been transferred to working memory. Figure 1 above shows how memory works based on Mayer's cognitive theory of multimedia learning.

Paivio's (1991) dual coding theory, Baddeley's (1992) working memory theory, and Sweller's (1988) cognitive load theory are integral theories that support Mayer's cognitive theory of multimedia learning, and the theory is based on three assumptions: the dual-channel assumption, the limited capacity assumption, and the active processing assumption (Sorden, 2012). First, the dual-channel assumption is based on Baddeley's working memory theory and Paivio's dual coding theory and suggests that working memory has auditory and visual channels. The second assumption is based on Sweller's cognitive load theory stating that each subsystem of working memory has a limited capacity. According to Mayer (2010), working memory can only process a few pieces of material at one time. Thirdly, the active processing assumption suggests that students construct knowledge in a meaningful way when they give attention to the relevant materials, organize them into the coherent mental structure, and integrate them with their prior knowledge.

To summarize, providing sound with image can be highly beneficial in multimedia learning by reducing working memory load. Sorden (2005) also argued that better learning occurs when pictures (nonverbal and visual) are combined with narration or audio (verbal and auditory). Mayer (2002) explained that the use of voice narration and visual graphics takes advantage of both verbal and nonverbal working memories using two sensory channels—auditory and visual—without overloading one or the other. That is, the information can be processed by two channels when narration or audio is provided with pictures. As the use of multiple modalities extends the capacity of working memory to process the information (Kalyuga, 2000), students learn better when visual graphics are provided with words presented as spoken language as opposed to written language. Thus, the best way to practice multimedia instruction is using auditory explanations and visual graphics together (Mayer, 2002).

## 2. Previous Studies on Modality Effects on Comprehension Skills

With the advancement of multimedia technologies, it has been easy for teachers to use a variety of modalities (Mayer, 2002). Thus, related research has been carried out to investigate whether the use of multimedia technology significantly improves students' comprehension skills. First, by comparing the two modalities - visual modality (written texts) and auditory modality (spoken words) - some scholars have shown that the comprehension effect of the auditory channel is better than that of the visual channel. For example, Liu et al. (2019) investigated whether there are any differences in reading comprehension between the use of visual and auditory modalities and discovered that listening to audio is more effective than reading texts. However, Moyer (2011) found no statistical significance between the two modalities when it comes to reading comprehension. That is, comprehension levels are the same regardless of the delivery format—whether it's either visual or auditory.

Studies comparing the effects of single (either text or audio) with dual (both text and audio) modality modes on comprehension skills have also generated inconclusive results. For example, Chang (2009) found the superior effects of the dual modality mode (both reading and listening) on comprehension compared to the single mode (listening only). Liu et al. (2019) also revealed that the dual channel (visual and auditory) is better than any single channel in comprehension. Likewise, Miller (1937, as cited in Sweller, 2005b) showed that the presentation of both auditory and visual materials facilitates better reading and text comprehension than the presentation of either auditory or visual materials does. Rahman and Hajar (2020) added that reading while listening has certain advantages, especially to those who struggle to focus on reading activities. By reading and listening to text, students gain a deeper understanding of the language. Regarding listening comprehension, Kartal and Simsek (2017) revealed that reading while listening has a positive impact on listening comprehension skills. Mohamed (2018) also confirmed the usefulness of using both visual and auditory channels on developing listening comprehension.

Diao and Sweller (2007), however, showed statistically significant gains in the single modality mode (visual) rather than the dual mode (both auditory and visual). They compared reading-only group and reading plus concurrent listening group to investigate their effects on word decoding and text comprehension. According to the authors, this is an example of redundancy effect, and learning is inhibited if written texts and spoken words containing the same

information and presented simultaneously. Moreover, Rogowsky et al. (2016) witnessed no significant differences in comprehension between dual modality (reading and listening to texts simultaneously) and single modality of input (either reading or listening to texts). They performed equally in comprehension tests regardless of whether they read, listened to, or both read and listened to texts simultaneously.

With the encouraging evidence that comprehension can be enhanced by adding pictorial forms of presentation (e.g., pictures, graphics, maps, etc.) to verbal forms of presentation, research studies have also been performed to investigate the possible effectiveness of using pictures. Some have confirmed the positive effects of adding image on comprehension. For instance, Erfani (2012) investigated the probable effect of using pictorial contexts on reading comprehension ability and found that the experimental group provided with both reading passages and related pictures performed better than the control group without pictorial contexts. Hsieh and Huang (2020) reported that the audio input combined with visual aids also supports the improvement in listening comprehension skills by helping students visualize the auditory input and providing them with contextual information.

However, there have been different findings regarding the effects of picture use on comprehension. Suvorov (2009) investigated the effects of context visuals on listening comprehension test scores, but no significant difference was found between the scores for audio-only and photo-added parts of the test. Schnotz and Bannert (2003) also conducted a study to analyze the effects of different kinds of multiple representations, verbal and pictorial, on comprehension. They compared text-only and text-graphic groups. The findings indicate that adding image is not always beneficial. Rather, the text-only condition performed better on comprehension quiz compared to conditions provided with text and illustrations together. Likewise, previous scholars have produced contradictory findings after comparing the effectiveness of text-picture and audio-picture multimedia inputs. For instance, Karbalaei and Zare (2019) discovered the superior effects of a combination of audio and picture on reading comprehension skills in comparison to a combination of text and picture. The authors confirmed that the visual information presented via audio-picture input acts as better retrieval cues compared to that via text-picture. In contrast, Chen (2006) reported that there is no significantly different effect between text-picture and audio-picture inputs on reading comprehension.

Several scholars have also investigated the effectiveness of a combination of text, audio, and picture and found that using pictorial context does not always improve learning. For example, Mayer et al. (2001) exposed students to pictures or animation (pictures with narration) with or without text and found that those without text comprehended better. According to them, when verbal information is presented in the auditory channel while pictures are in the visual channel, the attentional load is minimized. In contrast, when it is presented in both the auditory (narration) and visual (text) channels with pictures in the visual channel again, the load on this system is increased. That is, combining text with animation overloads the visual information-processing channel, making students split their visual attention. As the redundant visual information can interfere with learning, elimination of text paves the better way for comprehension in multimedia learning using pictorial context.

Yeh and Wang (2003) compared the three types of modality input—text, text-picture, and text-picture-audio—and found that the text-picture modality mode was the most effective while the text-picture-audio was the least effective. They offered some possible reasons for the ineffectiveness of a combination of text, picture, and audio. First, Asian EFL students who prefer visual learning style over auditory learning style, and therefore, they may not be skillful in using auditory information. Thus, the audio information cannot be processed effectively. Its fast speech rate can be another reason for the ineffectiveness. The students can be distracted if the speech rate exceeds their listening proficiency. Finally, too many elements of information can overwhelm the students' working memory. The combination of text, picture, and audio might not give the students enough time to process the information.

Studies have largely been conducted to confirm whether multimedia learning is effective in comprehension skills. However, much research on multimedia learning is still needed especially in EFL settings (Soruc, 2015). In addition, according to Rogowsky et al. (2016), there is a dearth of empirical studies directly evaluating the effects of different modalities on comprehension. Even so, previous modality studies have yielded conflicting results, and their findings cannot be compared directly due to the differences across studies in population characteristics and instructional conditions. Therefore, with a lack of consistency in literature regarding the effect of modality modes on comprehension, this study aims to investigate the extent to which modalities differently affect comprehension. By comparing six modality modes—text-only, audio-only, text-picture, audio-picture, text-audio, and text-audio-picture—this study examines the differential effects of modalities on EFL listening and reading comprehension skills.

### III. METHOD

#### 1. Participants

Participants in this study were composed of 186 EFL students who enrolled in a general English course at a university located in the southern part of Korea. There were 112 male students and 74 female students with a diverse group of majors including business administration, practical arts, social welfare studies, aeronautics and so on. All participants were Korean and had more than 10 years of English language learning experience in Korea. Their age ranged from 20 to 27.

In an effort to assess participants' English language proficiency, a simulated TOEIC test was carried out at the beginning of the course. The average test score of the participants was 15.97 out of 40, which can be interpreted as a low-intermediate proficiency level. This result revealed that the participants of the study were not proficient in English although they had studied English for many years.

For the current study, the participants were selected from six intact classes, taught by one of the researchers of this study. Each class was divided into six experimental groups at random and was randomly assigned to one of the six groups to find out the effectiveness of different modalities. The six modality modes used for class were text-only, audio-only, text-picture, audio-picture, text-audio, and text-audio-picture.

The text-only group (TO group hereafter) was made up of 32 students while the audio-only group (AO group hereafter) was 33. The text-picture group (TP group hereafter) was comprised of 29, and the audio-picture group (AP group hereafter) included 32. The text-audio group (TA group hereafter) was composed of 29 whereas the text-audio-picture group (TAP group hereafter) had 31. All participants in the study received the same input contents using fairy tales, but each in a different input modality.

**TABLE 1**  
**Homogeneity Test**

	TO (n = 32)		AO (n = 33)		TP (n = 29)		AP (n = 32)		TA (n = 29)		TAP (n = 31)		F	p
	M	SD	M	SD										
Listening Comprehension	7.66	3.83	7.70	3.48	7.28	3.99	7.78	4.47	7.76	4.84	8.84	3.57	.516	.764
Reading Comprehension	7.72	4.07	7.91	4.06	8.52	5.05	7.88	5.18	7.55	4.39	9.26	4.00	.625	.681

In order to assure the equivalence of the experimental groups before treatment, a homogeneity test was done with the pretest results. A one-way MANOVA was carried out to check whether the obtained data from the participants were homogeneous before intervention. As Table 1 shows, there were no statistically significant differences among the six groups in both listening and reading comprehension test score ( $p > .05$ ). The findings suggest that all participants in the present study were homogeneous in terms of English listening and reading comprehension skills at the beginning.

#### 2. Instruments

##### 1) Modality Modes

In order to figure out the effects of different modalities on English listening and reading comprehension, students were selected from six intact classes at a university in Korea. They were grouped into six experimental groups at random to find out the effectiveness of different modalities. All classes received the same input contents (fairy tales), but each in a different input modality depending on their experimental condition: text-only, audio-only, text-picture, audio-picture, text-audio, and text-audio-picture. The groups were named TO, AO, TP, AP, TA, and TAP. The intervention lasted for five weeks, excluding the time for pre- and post-tests as well as mid-term and final exams.

**FIGURE 2**  
Fairy Tale Stories



The fairy tale books were adopted from *Pinkfong Story* provided by Pinkfong. According to Arif, Triyono and Sahayu (2018), *Pinkfong* stories in extending utterances to young Indonesian EFL, they are known to be suitable for EFL learning by encouraging EFL students' English proficiency. They offer the same stories in video format as well as picturebook format. The picturebooks were from the collection of fairy tales, *The World's Best Fairy Tales in English*, published by Samsung Publishing Company (2020). Audio recordings are also available in compact disc (CD) format (See Figure 2 above). Based on the previous fairy tale study (Massi & Benvenuto, 2001), five were chosen for the in-class activities. The fairy tales used in this paper were as follows: *Beauty and the Beast*, *Snow White and the Seven Dwarfs*, *The Pied Piper of Hamelin*, *The Princess and the Pea*, and *Three Little Pigs*. Lee (2003) insisted that fairy tales are appropriate materials for EFL education. As a matter of fact, they have been commonly used as reading materials in EFL classrooms (Lazar, 1993). According to Massi and Benvenuto (2001), fairy tales are a great vehicle for enhancing foreign language development. They not only provide an authentic model of the target language (Kochiyama, 2013) but also motivate students to learn the language (Kosal, 2008).

In the TO group, fairy tales were provided solely in written texts. The participants in this group read the fairy tales in text format on a large white screen in front of a whole class. Regarding the AO group, the participants were provided with an audio file which was an audio representation of the written texts. This group listened to the same fairy tales in audio format through classroom computer speakers. In the TP group, the written texts were used together with some pictures. The participants in this group were given the fairy tales in text format as well as pictures that describe the fairy tale story. In the AP group, the students listened to the audio representation of the fairy tales, looking at the same pictures that depict the fairy tale story. The TA group was given the audio file concurrently with the identically written texts. In terms of the TAP group, the audio file was simultaneously accompanied with both the written texts and the relevant pictures. The participants in this group listened to the audio recording of the fairy tales and read the identically written texts while also looking at the pictures (See Figure 3 below).

**FIGURE 3**  
Six Experimental Groups



A classroom computer was used to expose students to visual or auditory input.<sup>1</sup> The participants in the TO, TP, TAP groups read texts on the projector screen at the front of the classroom. Pictures were also provided on the screen for the TP, AP, and TAP groups. For the textual and pictorial input, *Pinkfong* picturebooks were converted to Portable

<sup>1</sup> In this scenario, a teacher manipulates the input, and students have no control over the input. That means, the participants in this study read, listened to, or both read and listened to the same fairy tales with or without pictures within the same time limit.

Document Format (PDF) files and shown to the students. Classroom projector speakers were used when providing the AO, AP, and TAP groups with audio files. They allowed the students to listen to the same sounds simultaneously. The speaker volume was adjusted to a comfortable level. The average length of audio recordings lasted 4 minutes. The average number of words was 304. That means, the written texts were read at a pace of 76 words per minute on average. Regardless of their experimental condition, each group had the same amount of time to read, listen to, or both read and listen to the fairy tales with or without pictures.

## 2) TOEIC Listening and Reading Test

In EFL settings, English language is one of the most crucial skills for employability (Hsu, 2014). To get a job or to get promoted, what is required is adequate English competence. The majority of EFL universities, therefore, have encouraged their students to improve their English skills for their job searches. College students have tried to make themselves proficient for the competitive job markets by strengthening their English skills (Kim, 2020). According to Kim (2019), job seekers in Korean EFL context need to prove their English skills by submitting their official test scores. Specifically, most of them are requested to submit their ETS TOEIC test scores.

The TOEIC tests in EFL fields have been considered to be an effective measure to hire and evaluate employees (Sirikanjanawong & Wasanasomsithi, 2018). The tests are considered to be a rapid, convenient, and cost effective means of evaluating students' English language proficiency on the same score scale, providing a common standard of measurement for students' English language skills (Kim, 2018). They have also been considered as a well-established test in EFL contexts (Daller & Phelan, 2006). As standardized proficiency tests, they allow relatively easy comparisons of EFL students (Weaver, 2016). For the current study, therefore, TOEIC-based pre and post tests were adopted. According to Kim (2019), the TOEIC listening and reading test scores represent how well the students listen and read in English.

In this vein, all participants took simulated TOEIC listening and reading tests as pre and posttests to investigate the impacts of modality modes on EFL listening and reading comprehension skills. The tests were paper-and-pencil, multiple-choice assessments. They consisted of two timed sections—listening and reading—with 20 questions in each section. Given the original test-taking time, 9 minutes were given for the listening section and 15 minutes for the reading.<sup>2</sup> All questions were adopted from the practice tests provided by official ETS website (<https://www.ets.org/s/toEIC/pdf/toEIC-listening-reading-sample-test-updated.pdf>). Test scores were confirmed by the number of correct answers. One question with a correct answer was worth a single point. The scores ranged from 0 to 20 for each section.

## 3. Procedure

The focus of this research study was to ascertain whether the different modality modes have a notable effect on EFL students' listening and reading comprehension skills. The whole experiment was conducted during the 2020 academic year at a university in South Korea. For the experiment, a total of 186 undergraduate students from six general English classes were recruited. In this university, it was mandatory for all students to take more than two general English courses before graduation. The course title was General English 1. The primary course objectives were to promote students' general English language skills including listening and reading skills. The textbooks used for this course were TOEIC official test preparation books for listening and reading (Educational Testing Service, 2018a, 2018b).

The whole experiment was carried out during the participants' regular class times. There were five sessions of two-hour English class during the five-week experimental period.<sup>3</sup> In order to strengthen their English listening and reading skills, all groups were taught the same curriculum provided by the same teacher. The same fairy tales were

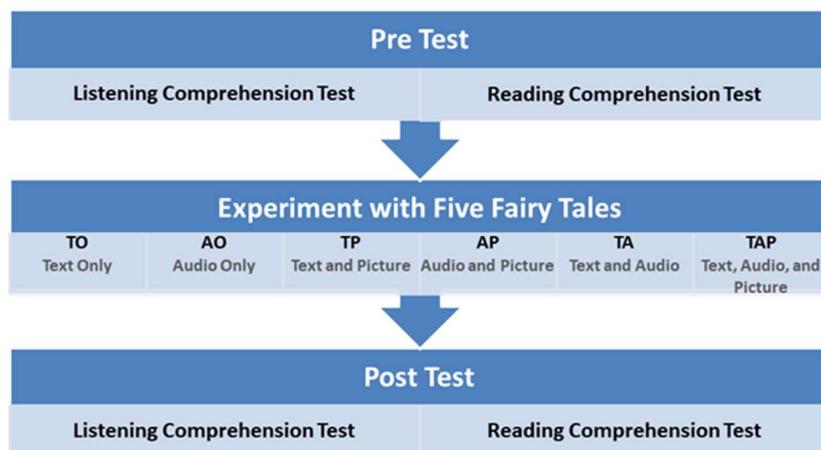
---

<sup>2</sup> The original TOEIC listening and reading tests have 100 questions in each section, and 45 minutes are given for the listening section and 75 minutes for the reading section.

<sup>3</sup> The whole experiment took place in a relatively short period of time. Due to the outbreak of coronavirus diseases (COVID-19), social distancing was enforced. In Korea, the Korean Ministry of Education decided to lock down school campuses, and it was not possible for students to attend their face-to-face class. Although some schools began to open in phases, they were closed again after reopening during the pandemic.

given for additional in-class activities. However, they were provided under six different conditions depending on multimedia modalities used for class. In the current study, the six modality conditions used were: text-only, audio-only, text-picture, audio-picture, text-audio, and text-audio-picture. Depending on their experimental conditions, the participants read, listened to, or both read and listened to the selected fairy tales with or without pictures (See Figure 4).

**FIGURE 4**  
**Research Procedure**



During the experimental period, the TO group read the fairy tales in text format on the screen in front of the classroom while the AO group listened to the same fairy tales in audio format through classroom computer speakers. In terms of the TP group, the participants read the written texts with some pictures that describe the fairy tale story. The AP group was given the same fairy tales in audio format with the same pictures on the projector screen. The TA group listened to the audio recording of the fairy tales while reading the identically written texts on the screen. For the TAP group, the audio file was simultaneously provided with both the written texts and pictures that depict the fairy tale story.

To expose students to multiple modality input, a classroom computer was used. For the visual input, a projector screen was used at the front of the classroom. The students were able to read texts or look at pictures on the white large screen in front of the class. For the auditory input, classroom projector speakers were used. The students listened to the same sounds simultaneously when provided with audio files. The speaker volume was adjusted to a comfortable level. Each group had the same amount of time to read, listen to, or both read and listen to the fairy tales with or without pictures. The classroom teacher pre-cued the different modalities (text, audio, or picture) and monitored each participant to make sure that there were no interruptions. The researcher also monitored participants' compliance.

Before and after the treatment, the TOEIC-based pre and post tests were performed in all groups to confirm the effects of different modality modes on EFL listening and reading comprehension skills. According to Kim (2019), the TOEIC listening and reading test scores represent how well the students listen and read in English. They were paper-and-pencil and multiple-choice assessments consisting of two timed sections with 20 questions each. For the listening section, 9 minutes were given while 15 minutes for reading. The number of correct answers indicated the test takers' scores. One credit was assigned to each item. The test score ranged from 0 to 20 for each section.

#### 4. Data Analysis

Data obtained from the study were analyzed with SPSS statistical software version 20. In order to evaluate participants' listening and reading comprehension skills, descriptive statistics were first computed. Paired samples *t*-tests were then administered to examine whether there were significant mean score changes between pre and posttests. A one-way multivariate analysis of variance (MANOVA) was used to analyze group differences on test scores for each test. The significance level was set at .05.

In order to check for possible differences among the six experimental groups in the pretest scores, the first one-way MANOVA was carried out. This homogeneity test was run to check whether the participants were homogeneous or not before they were treated. The findings showed no group differences in both listening and reading comprehension test scores ( $p > .05$ ), suggesting that the participants were homogeneous before the treatment in terms of English listening and reading comprehension skills. Thus, the second one-way MANOVA was conducted to determine whether the posttest scores vary and differ for modality mode. Scheffe post hoc test was also performed to verify which means significantly differ from which other ones.

## IV. RESULTS

### 1. Changes in EFL Listening and Reading Comprehension Skills

The current research study aims to ascertain whether the different modality modes have a notable effect on EFL students' listening and reading comprehension skills. For the experiment, all participants engaged in extra in-class activities using fairy tales and read, listened to, or both read and listened to the story with or without pictures. Five same fairy tales were provided in a different way depending on experimental conditions (modality modes). To confirm the effectiveness of multiple modalities, the pre and post tests used in the current study were the simulated TOEIC tests based on ETS TOEIC listening and reading tests.

For analyses, descriptive statistics were computed first. After that, paired samples *t*-tests were administered to determine the effectiveness of multiple modalities. Table 2 below shows the results for the effects of using different modalities on English listening comprehension skills.

**TABLE 2**  
**The Effects of Using Different Modalities on EFL Listening Comprehension Skills**

Group	Pre Test ( $n = 186$ )		Post Test ( $n = 186$ )		<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
TO	7.66	3.83	7.22	3.71	-1.149	.259
AO	7.70	3.48	8.58	4.48	1.727	.094
TP	7.28	3.99	10.83	3.96	5.701	.000*
AP	7.78	4.47	10.72	5.15	4.483	.000*
TA	7.76	4.84	12.55	5.51	5.917	.000*
TAP	8.84	3.57	10.16	4.69	1.140	.263

In terms of listening comprehension skills, findings in Table 2 reveal that there were significant changes in mean scores of some experimental groups between pre and posttests. To be specific, the TP group significantly increased their listening test scores from 7.28 on the pretest to 10.83 on the posttest ( $t = 5.701$ ,  $p = .000$ ). Likewise, the AP group significantly improved their listening comprehension skills ( $t = 4.483$ ,  $p = .000$ ). Their pretest score was 7.78 while their post test score was 10.72. The TA group also showed statistically significant mean score changes between pre and posttests ( $t = 5.917$ ,  $p = .000$ ). Their mean score on the pretest was 7.76 and it increased to 12.55 on the post test.

To sum up, statistically significant mean score changes were witnessed between the listening comprehension pre and posttests in the three experimental conditions: text-picture, audio-picture, and text-audio conditions. It was found that the three groups—the TP, AP, and TA groups—significantly improved their listening comprehension skills as a result of using multiple modalities. These findings suggest that EFL students can benefit from fairy tales in text format with pictures, in audio format with pictures, or in audio format with the identically written texts. In other words, using multimedia modalities in a combination of texts and pictures, audio and pictures, or texts and audio can lead to improvement in English listening skills.

**TABLE 3**  
**The Effects of Using Different Modalities on EFL Reading Comprehension Skills**

Group	Pre Test ( $n = 186$ )		Post Test ( $n = 186$ )		$t$	$p$
	$M$	$SD$	$M$	$SD$		
TO	7.72	4.07	9.72	4.07	3.819	.001*
AO	7.91	4.06	8.09	4.66	0.329	.744
TP	8.52	5.05	12.24	4.52	8.004	.000*
AP	7.88	5.18	7.09	5.22	-1.847	.074
TA	7.55	4.39	11.83	5.31	5.375	.000*
TAP	9.26	4.00	9.52	6.60	0.190	.851

Regarding reading comprehension skills, findings in Table 3 represents that there were also significant mean score changes in some experimental groups between pre and posttests. To be specific, the TO group significantly increased their reading test scores from 7.72 on the pretest to 9.72 on the posttest ( $t = 3.819, p = .001$ ). Likewise, the TP group significantly improved their reading comprehension skills ( $t = 8.004, p = .000$ ). Their pretest score was 8.52 while their post test score was 12.24. The TA group also showed statistically significant mean score changes between pre and posttests ( $t = 5.375, p = .000$ ). Their mean score on the pretest was 7.55 and it increased to 11.83 on the post test.

To summarize, statistically significant changes in mean scores were found between the reading comprehension pre and posttests in the three experimental conditions: text-only, text-picture, and text-audio conditions. It was found that the three groups—the TO, TP, and TA groups—significantly improved their reading comprehension skills as a result of using multimedia modalities in class. These findings suggest that EFL students can benefit from fairy tales in text format with or without pictures or audio recordings. That is, using multimedia modalities—texts only, text and pictures, or texts and audio—can bring about improvement in English reading skills.

## 2. Group Differences in EFL Listening and Reading Comprehension Skills

The current research study focuses on different modality effects on EFL students' listening and reading comprehension skills. In order to check for possible differences among the six experimental groups before experimental treatment, a one-way MANOVA was first run with the pretest scores. The obtained data were said to be homogeneous since the significant levels were greater than .05 in both listening and reading tests. This means that the participants were homogeneous before the treatment in terms of English listening and reading comprehension skills. Thus, the second one-way MANOVA was conducted to see if there are any group differences in posttest scores.

The MANOVA results showed a statistically significant group difference:  $F(10, 358) = 7.59, p = .000$ ;  $Wilks' \Lambda = 0.681$ ,  $partial \eta^2 = .18$ . In other words, the participants' posttest scores were significantly dependent on which modality mode they had been provided with ( $p = .000$ ). More specifically, it was found that the different modality modes had statistically significant effects on both listening ( $F(5, 180) = 5.06; p = .000$ ;  $partial \eta^2 = .12$ ) and reading ( $F(5, 180) = 4.75; p = .000$ ;  $partial \eta^2 = .12$ ), as shown in Table 4 below. That is, the participants benefited differently from the use of different multimedia modalities.

**TABLE 4**  
**Group Differences in English Listening and Reading Comprehension Skills**

	TO ( $n = 32$ )		AO ( $n = 33$ )		TP ( $n = 29$ )		AP ( $n = 32$ )		TA ( $n = 29$ )		TAP ( $n = 31$ )		$F$	$p$
	$M$	$SD$	$M$	$SD$	$M$	$SD$	$M$	$SD$	$M$	$SD$	$M$	$SD$		
	Listening Comprehension	7.22	3.71	8.58	4.48	10.83	3.96	10.72	5.15	12.55	5.51	10.16		
Reading Comprehension	9.72	4.07	8.09	4.66	12.24	4.52	7.09	5.22	11.83	5.31	9.52	6.60	4.747	.000

In order to verify which means were significantly different from which others, Scheffe post hoc test was then run. The post hoc test results showed that mean scores for listening were statistically significantly different between the TO group and the TA group ( $p = .002$ ), and the AO group and the TA group ( $p = .048$ ). Mean reading scores were also statistically significantly different between the TP group and the AP group ( $p = .011$ ), and the AP group and the TA group ( $p = .027$ ). Table 5 below indicates the main findings of the Scheffe test.

**TABLE 5**  
**Scheffe Post Hoc Test**

		<i>MD</i>	<i>SE</i>	<i>SE</i>
Listening	TO < TA	5.33	1.18	.002
	AO < TA	3.98	1.18	.048
Reading	AP < TP	5.15	1.31	.011
	AP < TA	4.73	1.31	.027

As can be seen from Table 5 above, in the listening comprehension test, the TA group scored significantly higher in comparison with the TO and AO groups. The mean score of the TA group was 12.55 for the listening posttest. The TO group received 7.22 for the same test while the AO group obtained 8.58. While the TA group both read and listened to the fairy tales simultaneously, the TO group just read and the AO group just listened to the fairy tale story. The findings revealed that providing auditory input with the identically written texts had more positive effects on listening comprehension compared to providing either texts or audio only. It can be said that this study proved the superior effects of the dual modality modes (using both text and audio) on listening comprehension compared to any single mode (using either text or audio only).

Regarding the reading comprehension test, the AP group scored significantly lower than the TP and TA groups. The mean score of the AP group was 7.09. The TP group obtained 12.24 on the same test while the TA group did 11.83. While the AP group was given auditory input with pictures, the TP group received texts with pictures. The TA group was provided with fairy tales in both text and audio format. The findings of this study showed that using multimedia modalities in a combination of text and picture led to better performance in English reading comprehension skills compared to that of audio and picture. Furthermore, the current study also proved the superior effects of combining text and audio on comprehension compared to integrating audio and picture.

To sum up, EFL students can benefit differently from multimedia modalities depending on which modality mode they use. To be specific, students who received both text and audio input will benefit more in listening comprehension than those given either text or audio. Providing multimedia modalities in a combination of audio and picture is less effective in reading comprehension compared to providing them in that of text and picture or text and audio.

## V. DISCUSSION

The results of the pre and post tests show that using some different modalities in class plays a beneficial role in increasing EFL comprehension skills. Specifically, it can be said that students increase their reading skills when provided with fairy tales even in text format solely. When given in audio format with pictures, fairy tales seem also effective in facilitating listening skills. In the case of using texts with additional modalities (with pictures or audio), it is likely that students promote both their listening and reading comprehension skills. However, it does not seem that EFL students benefit from using the auditory modality solely or using both visual and auditory modalities together with pictures.

These findings corroborate previous modality studies. For example, in the current study, increased reading comprehension skills were observed when fairy tales were given only in text format. This is in accordance with previous findings proposing that students can improve their reading skills when exposed to written texts solely (Diao & Sweller, 2007). Regarding a combination of audio and picture, Hsieh and Huang (2020) reported that the audio input combined with visual aids supports the improvement in listening comprehension by helping the students to visualize the auditory input, providing them with contextual information. This study also proved the positive effects of the use of both auditory and visual modalities on listening comprehension skills.

When given text with picture, the participants in this study promoted their listening as well as reading comprehension. Previous scholars have also acknowledged that comprehension can be facilitated by adding pictures to verbal forms of presentation. In particular, the current study supports Erfani (2012) showing the beneficial effects of using text and picture on comprehension. The findings of the present study also provide evidence to support previous scholars indicating the positive effects of the combined use of text and audio on the improvement of comprehension skills. Many of them proved the usefulness of their simultaneous use on enhancing listening (Kartal

& Simsek, 2017; Mohamed, 2018) and reading comprehension (Chang, 2009; Liu et al., 2019; Rahman & Hajar, 2020; Sweller, 2005b).

In order to test for the significance of group differences, a one-way MANOVA with group as an independent factor was also computed. The findings revealed that there were statistically significant group differences in both listening and reading comprehension skills. Specifically, providing both audio and text had more positive effects on EFL listening comprehension skills compared to providing either text or audio only. That means this study proved the superior effects of dual modality mode on listening over single mode. Regarding reading comprehension, there were also significant differences among the different dual modality modes (text and picture, audio and picture, and text and audio). A combination of text and picture resulted in better performance compared to that of audio and picture. A concurrent use of text and audio also generated better results in reading skills compared to that of audio and picture did.

Previous studies comparing the effects of dual (both text and audio) and single (either text or audio) modality modes on comprehension skills have also yielded similar results with the current findings. For example, Liu et al. (2019) found that the dual channel is more effective than any single channel in comprehension. Accordingly, Kartal and Simsek (2017) revealed that the presentation of both auditory (audio) and visual (text) materials facilitates listening comprehension. Mohamed (2018) also confirmed that reading while listening has a positive impact on listening comprehension skills. The findings of the current study support these previous findings, showing the superior effects of a combination of text and audio on EFL listening comprehension skills.

Interestingly, this study found that the TAP group where the audio file was simultaneously accompanied with both the written texts and the relevant pictures did not perform better than did the other experimental groups—the TO, AO, TP, AP, and TA groups—on both listening and reading comprehension tests. That is, providing auditory input concurrently with text and picture did not yield better performance compared to providing text only, audio only, text with picture, audio with picture, or text with audio. These findings challenge one of the common beliefs about the effectiveness of multiple modalities in instruction. Previous scholars (Kalyuga, 2000; Mayer, 2002) proposed that using multiple modalities is better than single one as the limited working memory can be expanded with the increased information processing capacity and the reduced cognitive load. According to them, the best learning occurs when multiple sensory channels are utilized. However, in this empirical study, using more modalities did not result in more comprehension in both listening and reading.

This result can also be explained by the previous research. Mayer et al. (2001) also found the inferior effects of the concurrent use of text, picture, and audio. The authors claimed that combining text with picture overloads the visual information-processing channel, making students split their visual attention. Yeh and Wang (2003) also proved the ineffectiveness of text-picture-audio mode. They cited the additional auditory input as reasons for this. According to them, due to the EFL students' preference for visual learning over auditory learning, the audio information cannot be processed effectively. That is, the students may not be skillful in processing the auditory information, and even be distracted by its fast speech rate. In addition, according to Yeh and Wang (2003), the combination of text, picture, and audio provides the students with too many elements of information, which can overwhelm their working memory. Thus, the students may be lack of time to handle the given information. In line with the previous studies, this study challenges 'the more, the better' beliefs in multimedia instruction.

## VI. CONCLUSION

Multimedia has played an important role in foreign language teaching and learning, and research has largely been conducted to prove its effectiveness in foreign language education. Although many of them have shown the significant effects on learning, there is still a dearth of research in EFL contexts (Soruc, 2015). Specifically, empirical studies comparing the effectiveness of different modality modes are needed, given that previous relevant studies have produced conflicting results (Rogowsky et al., 2016). In this regard, the current study investigated the effects of different modality modes on EFL learning. By comparing six modality modes—text-only, audio-only, text-picture, audio-picture, text-audio, and text-audio-picture—this study examined the modality effects on English listening and

reading comprehension skills.

The findings of the study revealed that EFL students can benefit from the use of different modalities in comprehension. To be specific, using multimedia modalities in a combination of text and picture, audio and picture, or text and audio appeared to improve English listening comprehension. Providing text with or without picture or audio was beneficial for increasing EFL reading skills. There were also significant group differences in listening and reading comprehension. It was found that the students benefited differently from the use of different modality modes. More specifically, providing both audio and text yielded better performance in listening compared to providing either text or audio. A combination of audio and picture, however, showed the inferior effects on reading compared to that of text and picture or text and audio.

Some of these findings support previous studies whereas others yield different results. First, this study corroborates previous findings proposing that students can improve reading comprehension when exposed to text itself (Diao & Sweller, 2007). In the present study, a combination of text and picture also promoted students' listening and reading comprehension, supporting previous studies (Erfani, 2012; Mayer, 1999). When provided with audio and picture, the participants in this study also enhanced their listening skills, as Hsieh and Huang (2020) reported. The current study also provides evidence to support previous research showing the positive effects of the combined use of text and audio on the improvement of listening (Kartal & Simsek, 2017; Mohamed, 2018) and reading comprehension skills (Chang, 2009; Liu et al., 2019; Rahman & Hajar, 2020; Sweller, 2005b).

When it comes to a comparison of different modality modes, the findings of the study agreed with previous studies suggesting the superior effects of dual mode (both text and audio) over single one (either text or audio) on listening comprehension skills (Kartal & Simsek, 2017; Mohamed, 2018). Likewise, the current study confirmed previous findings showing the ineffectiveness of text-picture-audio mode (Mayer et al., 2001; Yeh & Wang, 2003). However, given the common beliefs about the effectiveness of multiple modalities in multimedia instruction (Kalyuga, 2000; Mayer, 2001), this result can be interpreted as a challenge to 'the more, the better' beliefs in multimedia instruction. Thus, further study is needed in this area.

It is also necessary to consider some other findings that yielded different results from the previous ones. First, regarding the findings about the superior effects of a combined use of text and picture on reading comprehension compared to that of audio and picture, previous scholars have produced opposite results (Kalyuga, Chandler, & Sweller, 1999; Karbalaei & Zare, 2019). In addition, regarding the effects of the concurrent use of text and audio on reading comprehension, the current study disagreed with previous studies suggesting the ineffectiveness of their combination (Kalyuga, 2000). Moreover, the findings of this study showing the inferior effects of a combination of audio and picture compared to that of text and audio should be further investigated considering that previous scholars including Mayer (2002) and Sorden (2005) recommended the combined use of audio and picture as the best way to practice multimedia instruction. Clearly, these inconclusive findings need more investigation.

As can be seen above, inconsistent findings between the previous studies and the present one raise questions while requiring more research on the effects of different modality modes on English comprehension skills. Nevertheless, results of the current study still yield some practical and pedagogical implications. Language instructors in EFL settings can use multimedia modalities in their class to increase students' comprehension skills. To do so, EFL teachers should have knowledge about different characteristics of modality modes and be able to select the effective mode for their pedagogical goals. For example, to improve English listening comprehension, students should be exposed to dual modality modes such as text and picture, audio and picture, or text and audio. They can make use of text with or without picture or audio to enhance reading comprehension skills. In particular, a dual use of text and audio is more recommended for listening compared to a single use of text or audio. A combination of text and audio or text and picture is suggested for better performance in reading comprehension rather than that of audio and picture.

Limitations and suggestions for future research are as follows: First, findings of the study cannot be generalized to other settings as the participants in this study were restricted to 186 intermediate-low level students in Korean EFL field. If it were applied to a higher number of participants in different contexts, different results could be obtained. Another limitation is that it was not a longitudinal study. The whole experiment took place in a short period of time (5 weeks) due to the outbreak of coronavirus diseases (COVID-19). Further research that will look into the same dimension with the current study may use longitudinal studies. Finally, in this study, five selected fairy tales were given as input contents in a different modality. However, the modality effects can depend on the input contents and

therefore may yield different results with different ones.

## REFERENCES

- Arif, M., Triyono, S., & Sahayu, W. (2018). Pinkfong stories in extending utterances to young Indonesian EFL learners: A case study. *Applied Linguistics Research Journal*, 4(2), 15-24.
- Baddeley, A. (1992). Working memory: The interface between memory and cognition. *Journal of Cognitive Neuroscience*, 4(3), 281-288. <https://doi.org/10.1162/jocn.1992.4.3.281>
- Chang, A. (2009). Gains to L2 listeners from reading while listening vs. listening only in comprehending short stories. *System*, 37, 652-663. <https://doi.org/10.1016/j.system.2009.09.009>
- Chen, Z. (2006). *The effects of multimedia annotations on L2 vocabulary immediate recall and reading comprehension: A comparative study of text-picture and audio-picture annotations under incidental and intentional learning conditions* (Unpublished doctoral dissertation). University of South Florida, Florida. Retrieved from <https://digitalcommons.usf.edu/cgi/viewcontent.cgi?article=3477&context=etd>
- Daller, H., & Phelan, D. (2006). The C-test and TOEIC as measures of students' progress in intensive short courses in EFL. In R. Grotjahn (Ed.), *The C-test: Theory, empirical research, applications* (pp. 101-119). Bern: Peter Lang. Retrieved from [https://www.swansea.ac.uk/images/artsandhumanities/4\)%202006%20Daller%20and%20Phelan%20The%20C-test%20and%20TOEIC.pdf](https://www.swansea.ac.uk/images/artsandhumanities/4)%202006%20Daller%20and%20Phelan%20The%20C-test%20and%20TOEIC.pdf)
- Diao, Y., & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations. *Learning and Instruction*, 17(1), 78-88. <https://doi.org/10.1016/j.learninstruc.2006.11.007>
- Educational Testing Service [ETS]. (2018a). *TOEIC official test preparation book LC*. Seoul: YBM.
- Educational Testing Service [ETS]. (2018b). *TOEIC official test preparation book RC*. Seoul: YBM.
- Erfani, S. M. (2012). Pictures speak louder than words in ESP, too! *English Language Teaching*, 5(8), 164-169. <https://doi.org/10.5539/elt.v5n8p164>
- Hsieh, Y., & Huang, S. (2020). Using an e-book in the secondary English classroom: Effects on EFL reading and listening. *Education and Information Technologies*, 25(2), 1285-1301. <https://doi.org/10.1007/s10639-019-10036-y>
- Hsu, L.-Y. (2014). Integrating culture with project-based instruction in an EFL classroom. *English Teaching and Learning*, 38(1), 61-90. <https://doi.org/10.6330/ETL.2014.38.1.03>
- Kalyuga, S. (2000). When using sound with a text or picture is not beneficial for learning. *Australian Journal of Educational Technology*, 16(2), 161-172. <https://doi.org/10.14742/ajet.1829>
- Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 13(4), 351-371. [https://doi.org/10.1002/\(sici\)1099-0720\(199908\)13:4<351::aid-acp589>3.0.co;2-6](https://doi.org/10.1002/(sici)1099-0720(199908)13:4<351::aid-acp589>3.0.co;2-6)
- Karbalaei, A., & Zare, A. (2019). A comparison of the effect of textual, audio and textual-pictorial and audio-pictorial annotations on enhancing reading comprehension among Iranian EFL learners. *Teaching English with Technology*, 19(3), 40-67. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1224608.pdf>
- Kartal, G., & Simsek, H. (2017). The effects of audiobooks on EFL students' listening comprehension. *The Reading Matrix: An International Online Journal*, 17(1), 112-123. Retrieved from <https://www.readingmatrix.com/files/16-7w4b733r.pdf>
- Kim, N.-Y. (2018). A study on chatbots for developing Korean college students' English listening and reading skills. *Journal of Digital Convergence*, 16(8), 19-26. <https://doi.org/10.14400/JDC.2018.16.8.019>
- Kim, N.-Y. (2019). Effects of captioning order on EFL listening and reading: Should captions be first or second? *English Language and Literature Teaching*, 25(4), 1-22. <https://doi.org/10.35828/etak.2019.25.4.1>
- Kim, N.-Y. (2020). Languages on the screen: A study on reversed subtitling and EFL receptive language skills. *Multimedia Assisted Language Learning*, 23(1), 37-66. <https://doi.org/10.15702/mall.2020.23.1.37>

- Kochiyama, A. (2013). Teaching gender awareness through fairy tales in the EFL classroom. *English Review: Journal of English Education*, 2(1), 1-7. Retrieved from <https://journal.uniku.ac.id/index.php/ERJEE/article/view/86/47>
- Kosal, I. (2008). *Enhancing writing and improvisational speaking skills through fairy tales in EFL preparatory classrooms* (Unpublished doctoral dissertation). Selçuk University, Konya, Turkey.
- Lazar, G. (1993). *Literature and language teaching*. Cambridge, MA: Cambridge University Press. <https://doi.org/10.3138/cmlr.52.3.507>
- Lee, C. (2003). The uses of enchantment: Fairy tales as instructional materials to facilitate primary English language education. *Journal of National Taipei Teachers College*, 16(1), 29-48. Retrieved from <https://academic.ntue.edu.tw/ezfiles/7/1007/img/41/16-1-2.pdf>
- Liu, H., Cao, S., & Wu, S. (2019). An experimental comparison on reading comprehension effect of visual, audio and dual channels. *Proceedings of the Association for Information Science and Technology*, 56(1), 716-718. <https://doi.org/10.1002/pra2.148>
- Massi, M. A., & Benvenuto, A. M. (2001). Using fairy tales to develop reading and writing skills. *The Catesol Journal*, 13(1), 161-171. Retrieved from <https://files.eric.ed.gov/fulltext/ED474366.pdf#page=157>
- Mayer, R. E. (1999). Multimedia aids to problem solving transfer. *International Journal of Educational Research*, 31, 611-623. [https://doi.org/10.1016/s0883-0355\(99\)00027-0](https://doi.org/10.1016/s0883-0355(99)00027-0)
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/cbo9781139547369.002>
- Mayer, R. E. (2002). Multimedia learning. *The Annual Report of Educational Psychology in Japan*, 41, 27-29. [https://doi.org/10.5926/arepj1962.41.0\\_27](https://doi.org/10.5926/arepj1962.41.0_27)
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). New York, NY: Cambridge University Press. <https://doi.org/10.1017/cbo9780511816819>
- Mayer, R. E. (2010). Applying the science of learning to medical education. *Medical Education*, 44, 543-549. <https://doi.org/10.1111/j.1365-2923.2010.03624.x>
- Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, 93, 187-198. <https://doi.org/10.1037/0022-0663.93.1.187>
- Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, 12(1), 107-119. [https://doi.org/10.1016/s0959-4752\(01\)00018-4](https://doi.org/10.1016/s0959-4752(01)00018-4)
- Mohamed, M. M. K. (2018). Using audiobooks for developing listening comprehension among Saudi EFL preparatory year students. *Journal of Language Teaching and Research*, 9(1), 64-73. <https://doi.org/10.17507/jltr.0901.08>
- Moyer, J. E. (2011). *Teens today don't read books anymore: A study of differences in comprehension and interest across formats* (Unpublished doctoral dissertation). University of Minnesota, MN. Retrieved from <https://hdl.handle.net/11299/116037>
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45(3), 255-287. <https://doi.org/10.1037/h0084295>
- Rahman, A., & Hajar, I. (2020). The effect of audiobook on reading comprehension of the eleventh grade students of SMA Negeri 2 Buru. *ELT Worldwide*, 7(2), 104-115. <https://doi.org/10.26858/eltww.v7i2.14931>
- Rogowsky, B. A., Calhoun, B. M., & Tallal, P. (2016). Does modality matter? The effects of reading, listening, and dual modality on comprehension. *SAGE Open*, 6(3), 1-9. <https://doi.org/10.1177/2158244016669550>
- Samsung Publishing Company. (2020). *The world's best fairy tales in English* (5th ed.). Seoul: Samsung.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13(2), 141-156. [https://doi.org/10.1016/s0959-4752\(02\)00017-8](https://doi.org/10.1016/s0959-4752(02)00017-8)
- Sirikanjanawong, N., & Wasanasomsithi, P. (2018). Relationship between the ICAO language proficiency requirements (LPRs) and test of English for international communication (TOEIC) scores of flight attendants in Thailand. *LEARN Journal*, 11(1), 64-86. Retrieved from <https://so04.tci-thaijo.org/index.php/LEARN/article/view/135870/101472>
- Sorden, S. D. (2005). A cognitive approach to instructional design for multimedia learning. *Informing Science*, 8, 263-279. <https://doi.org/10.28945/498>

- Sorden, S. D. (2012). The cognitive theory of multimedia learning. *Handbook of Educational Theories, 1*, 1-22. Retrieved from [https://sorden.com/portfolio/sorden\\_draft\\_multimedia2012.pdf](https://sorden.com/portfolio/sorden_draft_multimedia2012.pdf)
- Soruc, A. (2015). What makes redundant presentation of multimedia learning difficult? *Procedia-Social and Behavioral Sciences, 191*, 2844-2848. <https://doi.org/10.1016/j.sbspro.2015.04.402>
- Suvorov, R. (2009). *Context visuals in L2 listening tests: The effects of photographs and video vs. audio-only format* (Unpublished doctoral dissertation). Iowa State University, Ames, Iowa. Retrieved from <https://core.ac.uk/download/pdf/38911438.pdf>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12* (2), 257-285. [https://doi.org/10.1207/s15516709cog1202\\_4](https://doi.org/10.1207/s15516709cog1202_4)
- Sweller, J. (2005a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19-30). New York, NY: Cambridge University Press. <https://doi.org/10.1017/cbo9780511816819>
- Sweller, J. (2005b). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159-168). New York, NY: Cambridge University Press. <https://doi.org/10.1017/cbo9780511816819.011>
- Turker, S. (2010). *The effectiveness of audio books on the reading comprehension of selected texts by university EFL students at different proficiency levels* (Unpublished doctoral dissertation). Bilkent University, Ankara, Turkey. Retrieved from <https://www.thesis.bilkent.edu.tr/0003964.pdf>
- Weaver, C. (2016). The TOEIC IP test as a placement test: Its potential formative value. *JALT Journal, 38*(1), 5-26. <https://doi.org/10.37546/jaltj38.1-1>
- Yeh, Y., & Wang, C. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *CALICO Journal, 21*(1), 131-144. <https://doi.org/10.1558/cj.v21i1.131-144>