Chih-Ming Chen*, I-Chun Chen "The Effects of Video-Annotated Listening Review Mechanism on Promoting EFL Listening Comprehension", *Interactive Learning Environments*, 2018. (Accepted) **[SSCI 收錄] (post print version)**

The Effects of Video-Annotated Listening Review Mechanism on Promoting EFL Listening Comprehension

Chih-Ming Chen[1] and I-Chun Chen[2]

[1]Graduate Institute of Library, Information and Archival Studies, National Chengchi University, No. 64, Section 2, ZhiNan Road, Wenshan District, Taipei City 116, Taiwan, R.O.C.

[2]E-learning Master Program of Library, Information Studies, National Chengchi University, No. 64, Section 2, ZhiNan Road, Wenshan District, Taipei City 116, Taiwan, R.O.C.

Abstract

In recent years, English instruction has emphasized the importance of listening comprehension; however, playing CDs in class is still the most widely used approach to practicing English listening skills in Taiwan. It is obvious that this training approach cannot satisfy individual needs, thus reducing the effect of promoting English listening comprehension performance. Thus, the study proposes a novel video-annotated learning review mechanism (VALRM) with two review modes—the delayed and immediate review modes, which can effectively assist individual learners to mark the English video sections that cannot listen clearly for individual review, to enhance learners' listening comprehension performance. This study adopts a mixed research methodology combining quantitative and qualitative approaches to examine the effects of learners in the experimental group using the VALRM and those in the control group using Youtube with self-determined learning review mechanism (SDLRM) on English listening comprehension performance. Analytical results show that the learners with the VALRM support for individual review achieved significantly better listening comprehension performance than those of using the SDLRM for individual review. More importantly, the experimental results confirm that the learners in the experimental group had significantly better listening comprehension abilities even though removing the VALRM support for individual review than those of the control group removing the SDLRM support. Additionally, the interview results show most of the interviewees agreed that using the VALRM is more effective in promoting English listening learning performance than using the SDLRM because they could find out the unclear parts and replay them easily and efficiently. The main contribution of this study is to present an effective computer-assisted English listening learning tool so that learners can efficiently improve their listening comprehension abilities based on the two review modes provided by the proposed VALRM in an autonomous learning environment.

Keywords: Video-annotated learning review mechanism (VALRM), Self-determined learning review mechanism (SDLRM), computer-assisted listening instruction, English listening learning, English listening comprehension

1. Introduction

Globalization has become a trend, and it is inevitable that modern people nowadays have to learn foreign languages to meet the developing trend. Since English probably is the most widely used language around the world, enhancing English skills is becoming more and more important to modern people. Undoubtedly, listening is the most dominant skill among English acquisition that we have to acquire (Mohsen, 2016) as well as has gradually become a prime concern to many language teachers (Brett, 1997; Chung, 1999). A number of language experts claimed that listening skills are very important to language communication (Nunan, 2002). Rost (1994) pointed out that listening provides input in language communication and is thus essential and fundamental to

speaking. However, listening is the most difficult skill for beginning English as a Foreign Language (EFL) learners because understanding a listening task requires learners to discriminate between sounds, understand the vocabulary and grammatical structure, gain familiarity with the intonation and stress, and contextualize the speech in terms of sociocultural utterances (Vandergrift, 1999).

In Taiwan's formal education, students start to learn English from elementary school until university; however, the importance of listening has long been neglected in Taiwan's English education so that most students only accept little listening training (Chung, 1999). Although listening practice has been included in English class in Taiwan's schools, the primary method of training listening skills mainly adopts CD player, which is not an efficient way, to support the listening comprehension and review process. To improve students' English listening comprehension performance, developing effective computer-assisted language learning (CALL) approaches to aid students' listening practices is definitely needed. Many previous studies (Chen, Wang, &·Lin, 2017; Lin & Chen, 2018; Chen & Chung, 2008) confirmed that the language review mechanisms can efficiently assist EFL learners to promote their language acquisition performance. For example, Chen, Wang and Lin (2017) presented an attention-based diagnosing and review mechanism (ADRM) based on brainwave detection to help learners identify the passages with low attention level in a lesson as review targets in order to perform efficiently and accurately review processes while reading paper-based English texts with digital pen support in autonomous learning environments. Lin and Chen (2018) presented an attention-based video lecture review mechanism (AVLRM) that can generate video segments for review based on students' sustained attention status, as determined using brainwave signal detection technology. Their study confirmed that AVLRM based on brainwave signal detection technology can precisely identify video segments that are more useful for effective review than those picked by student themselves. Also, Chen and Chung (2008) proposed a personalized mobile English vocabulary learning system based on Item Response Theory and learning memory cycle, which recommends appropriate English vocabulary for individual learner's vocabulary learning and review according to individual learner vocabulary ability and memory cycle.

Video annotation tools have been successfully applied in the CALL field or other disciplines (Chiu et al., 2016; Zarzour & Sellami, 2018). Fu, Schaefer, Marchionini and Mu (2006) indicated that the development of video annotation tools must be considered together with the embedding video navigation and manipulation tools. A video annotation tool generally provides the functions — selecting the segment of the annotated information object, adding the notes and creating the link, although the way each function is implemented in specific systems differs considerably. Chiu et al. (2016) proposed a video annotation learning system that learners can annotate and highlight the interesting or important contents to make them more memorable. Zarzour and Sellami (2018) proposed a linked data-based annotation system that includes an annotating function, a linked data enrichment function, a sharing function and faceted search function to reduce the students' cognitive load and to improve their learning achievement. However, to the best of our knowledge, there has been little research to propose effective computer-assisted listening review mechanisms based on video annotation to improve English listening comprehension performance. Although the VoiceTube provides an immediate review mode that can support a learner to immediately replay any video section not comprehending well to support listening skills' training, this review mode may be not suitable for all learners. This study thus presents a video-annotated listening review mechanism (VALRM) with two review modes— delayed and immediate review to promote EFL learners' listening comprehension abilities. This study aims to assess whether or not the effects of using the VALRM to improve English listening learning on listening comprehension performance are significantly superior to that of using the self-determined listening review mechanism (SDLRM) provided by the frequently used video player platform like Youtube. Results of this study can be regarded as references to further study on computer-assisted listening designs in order to improve teaching qualities and learning effectiveness of English listening acquisition and to decrease time

spent using when learners practice listening.

2. Literature Review

2.1 English Listening Learning

Listening comprehension has long been neglected in language research and generally considered as an ability that can be developed without effective assistance (Osada, 2004). Some researchers claimed that listening comprehension was an innate ability that people were born to have, and therefore there was no need to learn it. Some even considered listening is equal to hearing. However, many researchers have challenged this perspective and pointed out that the ability to listen is different from the ability to hear. Schnell (1995) further indicated that hearing is merely a physiological process, whereas listening is an interpretive process. That is, hearing is an innate ability that people possess; in contrast, listening requires concentration to catch the meaning of the context. Taylor (1964) indicated that hearing involves physiological process of sound while listening is a complex process in which listeners accumulate sounds first, identify "short sound sequences as words, and then translate larger word sequences into meaning."

From 1940s, the founders of the listening skill, James Brown, Ralph Nichols, and Carl Weaver, started to pay attention to the recognition of listening (Bozorgian, 2012). After that, more and more researchers carried on studies associated with listening practice and training (Feyten, 1991). Listening comprehension skills thereby started to receive more attention (Wu, 2004). Meanwhile, a variety of theories and instructional designs which aimed at assisting teachers and learners to develop effective listening strategies were proposed since then, and listening has finally been viewed as a distinctive skill (Brown, 1990), a fundamental role played in language acquisition, rather than a secondary one.

Undoubtedly, listening plays a pivotal role in language acquisition. Hence, it is essential to find out efficient ways to facilitate it. In recent years, language researchers have not only emphasized the importance of listening but also stressed that of learners' motivation (Field, 1998). Many language researchers have explored effective ways to help learners improve their listening comprehension performance. To understand the listening problems that Taiwanese students encounter during listening practices, several previous studies were conducted (Chao & Chien, 2005; Sun, 2002; Chen, 2013). Chao and Chien's study (2005) revealed that Taiwanese students had English listening difficulties caused by the factors of texts, listeners, processes, and speakers. Sun's study (2002) indicated that the most difficulty in English listening for Taiwan's students was "forget the meaning of the word." Chen's study (2013) showed that unfamiliar vocabulary, rapid speech rate and linking sounds between words are three major factors affecting Taiwanese students' English listening. Moreover, listening practice has a property and progress of continuity, which may cause learners' anxiety and frustration when they cannot catch up with others, and the unsatisfied performance therefore leads to lower motivation to learn.

Underwood (1989) indicated that performing a listening practice by repeatedly listening some passages was usually decided by teachers in a traditional classroom instruction so that individual learner had no chance to choose the reviewing sections by himself/herself. The frustration of making decision on their learning may thereby decrease learners' motivation to practice listening. Lui (2008) examined the relationship between the use of listening strategy and listening ability of Taiwanese university students, indicating a significant positive relationship between strategy use and listening proficiency exists. Also, Wang and & Treffers-Daller confirmed that vocabulary size is the strongest predictor of listening comprehension, followed by general language proficiency, while metacognitive awareness is less important. Therefore, this study aims to assess whether or not the developed VALRM can improve learners' listening comprehension by giving them more right of

decision making on reviewing the passages with unfamiliar or unidentified vocabularies.

2.2 Computer-Assisted Listening Instruction

With the rapid development of information and communication technology (ICT), teaching and learning foreign languages with the assistance of computers have gained popularity (Gündüz, 2005). Many highly developed countries have adopted computer-assisted language learning (CALL) in language instruction and applied it to enhancing different language skills, including listening (Gündüz, 2005). Several studies (O'brien & Hegelheimer, 2007; Vahdat& Eidipour, 2016) confirmed the effectiveness of CALL in listening and paid attention to the development of it. In addition, learners' attitude toward CALL is relatively positive as well as learners have more time to practice listening strategies they used in class or after class if they adopt CALL in language learning (Talebinezhad & Abarghoui, 2013). They also thereby have more chances to explore to different types of spoken English by using CALL. Liu, Chen, and Hwang (2017) designed a ubiquitous fitness English listening comprehension system (UFELCS) incorporating collaborative listening activities into a fitness center to help learners improve their listening comprehension with a combination of context-aware tools and video-based materials. Their study indicated that the participants' learning performance was significantly improved. Also, O'brien and Hegelheimer (2007) conducted a structured attempt to integrate CALL activities in the form of podcasts into an academic EFL course on listening strategies, indicating that both the teacher and the students find the podcasts to be a positive component of the course.

Mohsen (2016) examined that the potential help options (HOs) in multimedia listening environment that can facilitate listening comprehension include captions, scripts, annotations and dictionary notes that learners can access to overcome the comprehension breakdown and discover the meaning of unknown words. For example, Hsu's study (2015) presented an adaptive video caption filtering mechanism to enhance EFL learners' listening training. The provision of captions or subtitles enhances not only listening competence and reading ability, but also the load of the learners during the learning process can be reduced (Hsu, Hwang & Chang, 2014). Listeners are less anxious and stressed with the help of captions or subtitles. However, some researchers argued that relying on captions or subtitles is detrimental to listening acquisition (Latifi, Mobalegh, & Mohammadi, 2011). Diao, Chandler and Sweller (2007) indicated that learners learned listening with the presence of subtitles had poor performance on the subsequent listening test than those listening with the auditory materials only. In short, the effects of captions on listening comprehension (Yang & Chang, 2014; Başaran & Köse, 2013; Leveridge & Yang, 2013; Markham, 2001; Huang & Eskey, 1999) show the effectiveness of captions on listening comprehension was influenced by many factors and was mostly positive. Therefore, the usage of captions or not was decided by learners in both the experimental and control groups in this study. Namely, learners in both groups could choose whether displaying captions or not when proceeding listening practice by using video player in Youtube or the developed VALRM.

3. Research Methodology

3.1 Research Design

This study adopted a mixed research methodology combining quantitative and qualitative approaches to examine whether the listening comprehension performance between the experimental and control groups who respectively used the developed VALRM and the SDLRM to review English listening video has significant difference. There were four listening videos chosen from Youtube website and a self-made listening video with about four to five minutes long for both groups used in the experiment. The difficulty level of each listening video is slightly more advanced than the participants' current English listening levels in order to assess their listening

comprehension performance discriminatively. In the experimental stage, both groups listened to the same listening videos and completed the same listening comprehension tests and both groups were given the same length of learning time, which conducted once a week and lasted for a month.

3.2 Research Participants

A total of 39 Grade 8 students aged 13-14 years old were recruited from two classes of a junior high school in Taipei City, Taiwan to participate in the instruction experiment. English is their second language and they had learned English for six years. The two classes were randomly assigned to two learning groups. The experimental group consisted of 18 students, whereas the control group consisted of 21 students. During the instructional experiment, both groups listened to the same listening text but reviewed it by using different listening review mechanisms. The study used the independent-samples *t*-test to confirm whether there were no significant differences between prior English proficiency of both groups according to their semester English scores in school.

3.3 Research Procedures

This study aimed to confirm the effectiveness of the VALRM on promoting listening comprehension performance. Figure 1 shows the experimental procedures of the study. First, one of the two classes from a junior high school in Taipei City, Taiwan was randomly assigned to the experimental group using the VALRM and the remaining class was assigned to the control group using the SDLRM for English listening learning. Before performing the experiment, the procedures of the entire experiment were explained in seven minutes to all the participants, and then how to operate the VALRM and the SDLRM was demonstrated for the learners in the both groups. The learners in the both groups then had three minutes to get familiar with the usage of the assigned listening review mechanisms. In sum, a total of ten minutes were performed for the introduction and orientation of the instrument.

In the formal experiment, each group listened to a four to five minutes long video and had two more minutes to review the content. That is, each learner in the both groups had about a total of seven minutes to listen to a video and review it with different review mechanisms support. The main reason that review time was set as half of listening learning time is to examine whether the VALRM could support to review the content more efficiently and effectively than did the SDLRM in a short time. During the listening practice, the experimental group could use all the functions supported by the VALRM to assist their review while the control group could only review the video by pausing or rewinding the video with the mouse. Both groups could rewind and review the content at any time during the given learning listening content. The review time for the learners in the both groups was the same except for the review mechanisms they used. After reviewing, the learners in the both groups were asked to perform the listening comprehension test with ten minutes immediately to assess their listening comprehension performance. The learning activities and the listening comprehension tests were performed once a week and were lasted for a month. That is, the entire learning activities contain a total of four listening practices with the corresponding listening comprehension tests.

After four weeks, the learners in the both groups listened to a five minutes long self-made video that includes the contents of the previous ones without reviewing this time and took the final listening comprehension test with ten minutes. That is, a total of five listening comprehension tests were designed to assess the listening comprehension performance of research participants in the entire experimental procedures. This experiment aims to access whether the learners in the both groups have significant difference in the listening comprehension performance after using different review mechanisms. Finally, a semi-structured interview was conducted to collect qualitative data from eight of the research participants in the both groups. The interviewees consisted of four high

achievers and four low achievers from the both groups. Before the interview, all of the interviewees from the control group were taught how to use the VALRM and they were given sufficient time to experience how to use it to support promoting listening comprehension. All of the data collected from the experiment process were analyzed and discussed in the later sections.

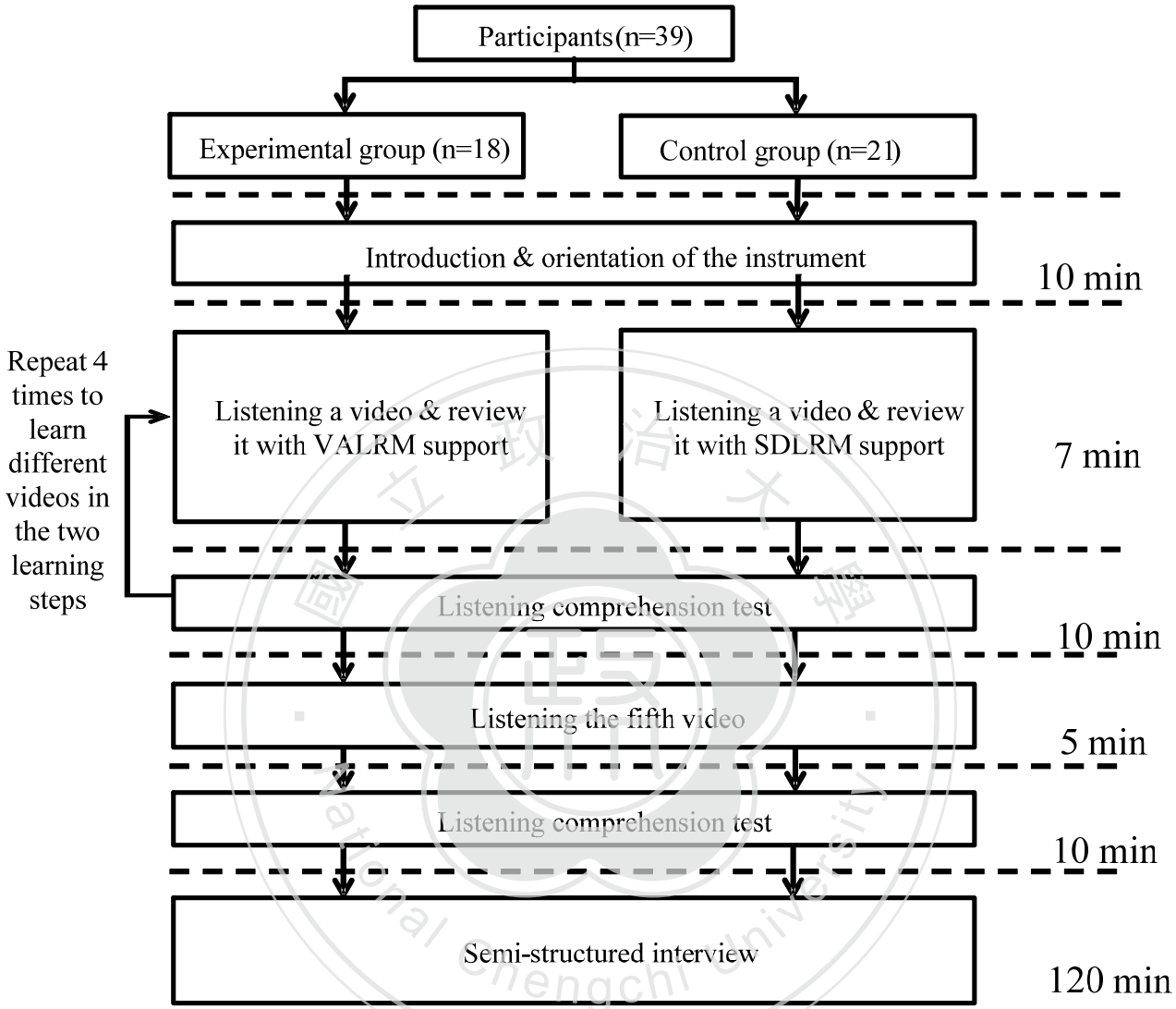

Figure 1. Experimental procedures of the study

3.4 Research Tools

3.4.1 Video-Annotated Listening Review Mechanism (VALRM) for the Experimental Group

The VALRM is an online listening comprehension learning system which needs to be executed with the Google Chrome browser. It can play a video with English and Chinese captions from Youtube on its own platform. The VALRM can support customized learning which provides the listening learning mechanisms where learners have the choice to determine when and how they learn and review the videos for promoting English listening comprehension performance. When a learner watches and listens to a video through the VALRM, she/he can mark the video sections not understanding well at any time according to their listening comprehension statuses and review these marked sections with bilingual captions immediately or later. The main functions of the proposed VALRM are detailed in the following subsections.

3.4.1.1 Function setting interface

Figure 1 shows the function setting interface of the VALRM. As shown in Fig. 2, the proposed VALRM allows learners to set whether the English captions or Chinese captions in a video will be displayed or not during performing English listening practice. After performing the function setting, learners can type an URL address of Youtube into the address bar, and then click on the "send" button to listen the video for proceeding English listening practice.

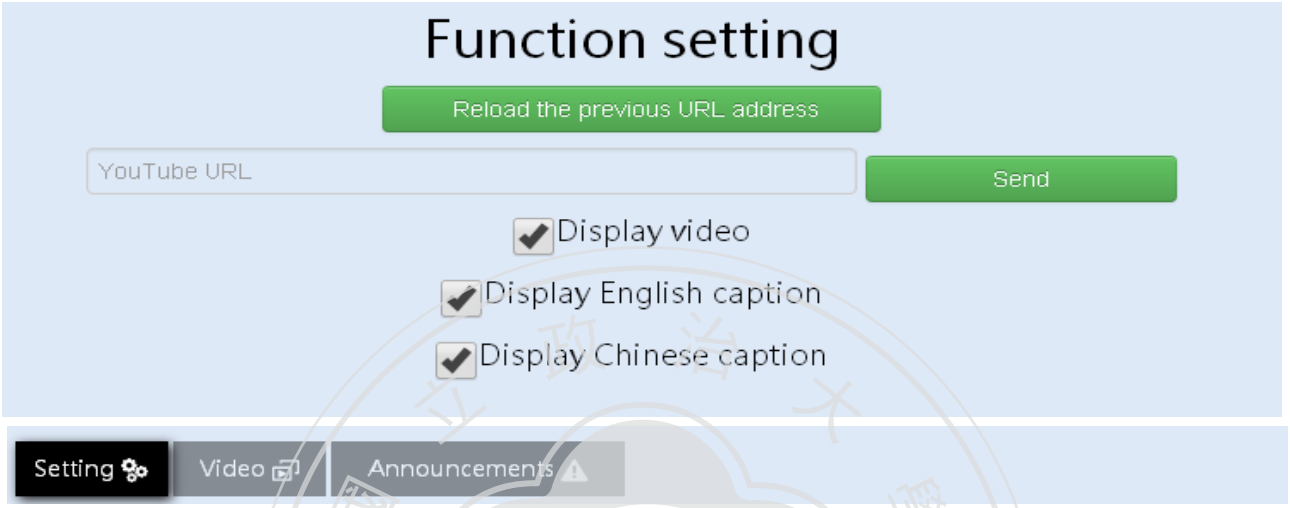

Figure 2. The user interface of function setting

3.4.1.2 The customized review mechanism

The VALRM developed in this study can play any Youtube videos with captions, such as English and Chinese captions, and provides two annotation functions— "add the mark" and "mark and replay" for supporting two review modes— delayed and immediate review to aid learners to promote their listening comprehension abilities. These two functions allow a learner to replay the marked section of a video starting at the previous five seconds from that the time of video was marked, and the learner can go back to review it again readily without wasting time to recall or look for their confusing parts. Figure 3 shows the user interface of customized review mechanism of the VALRM. When a learner watches and listens a Youtube video for promoting English listening comprehension ability by using the function of "add the mark" in the VALRM, she/he can mark the video sections that do not understand well at any time, and these marked video sections will be recorded into the customized review list with the corresponding time stamps. The learner can review them later by clicking on the customized review list according to the time stamps. The time stamps can directly link to the corresponding video sections so that the learner can review these video sections efficiently. After the learner has already comprehended a certain reviewed video section, she/he can click on "delete the mark" to remove the time stamp of the reviewed video section from the customized review list. The review mode was called as the delayed review mode in this study. The delayed review mode aims to support a learner to record the video sections that cannot comprehend well by self-judgment while listening a video to train listening skills as well as assist the learner to review these video sections efficiently without taking extra efforts to recall or look for after finishing the listening practice of the whole video. Moreover, if a learner prefers immediately reviewing the unfamiliar video section that she/he marked, she/he can click on "mark and replay", and then the marked video section will be immediately repeated until the function is disabled. This review mode was called as the immediate review mode in this study. The immediate review mode aims to support a learner to immediately review a video section that cannot

comprehend well by self-judgment during listening a video to train listening skills.

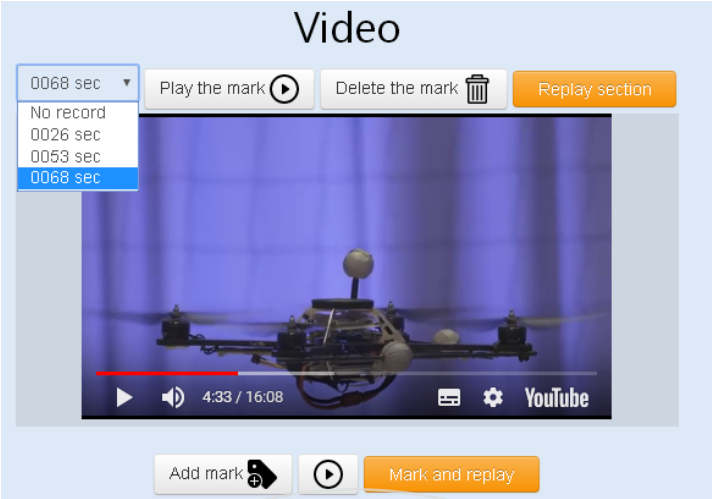

Figure 3. The user interface of customized review mechanism

3.4.1.3 The bilingual caption display

In addition, the VALRM also provides the function that can automatically retrieve and display the bilingual captions marked as the video sections that learners need to review on the user interface of the VALRM. This function aims at assisting the learner to diagnose the unfamiliar vocabularies that appear in the marked video section and to confirm whether the listening comprehension to the marked video section is correct or not while performing English listening comprehension practice. Şevik (2017) indicated that bimodal subtitling has positive effects on listening comprehension and that caption reliance did not have any negative effects for the participants. Moreover, several studies (Hsu, 2015; Hsu, Hwang & Chang, 2014) indicated that appropriately providing captions can enhance listening competence. Figure 4 shows the display interface of bilingual caption for the video section marked by a learner due to poor listening comprehension. The bilingual captions were extracted from the back end of Youtube and were shown on the display area of bilingual caption. The function of bilingual caption display can be turned on or off according to the need of individual learner while proceeding listening practice. Learners can choose either to hide or to display captions of the chosen section before playing the video.

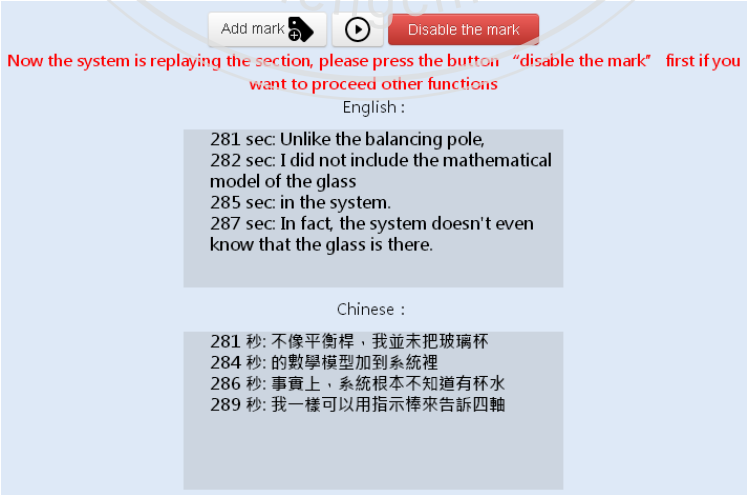

Figure 4. The display interface of bilingual caption

3.4.2 Self-Determined Listening Review Mechanism (SDLRM) for the Control

Group

The so-called self-determined listening review mechanism (SDLRM) is to use the functions provided by the Youtube player to replay a video section that a learner wants to review based on self-determined judgment by dragging a mouse to the video section. The SDLRM can support the learner to review the video section, but it cannot replay it automatically. If the learner would like to review the video section again, she/he has to drag the mouse to the video section again. In other words, the SDLRM has not the automatically immediate review mode like the VALRM provided. Furthermore, the SDLRM cannot record the time stamps for the corresponding video sections that the learner would like to review into a customized review list for the delayed review mode. Table 1 shows the comparison of review functions provided by SDLRM and VALRM.

Table 1. Comparison of review functions provided by SDLRM and VALRM

Comparison of functions	**Review mode**	
	SDLRM	**VALRM**
Review method	Self-determined judgment by dragging a mouse to the video section that a learner would like to review when playing a video with the Youtube player for listening practice	The immediate review mode- supporting a learner to immediately review a video section that cannot comprehend well by self-judgment when listening a video for listening practice The delayed review mode- supporting a learner to record the video sections that cannot comprehend well by self-judgment in a listening list while listening a video for listening practice as well as assisting the learner to review these video sections efficiently by clicking on the time stamp in the listening list without taking extra efforts to recall or look for after finishing the listening practice of the whole video
Automatically replaying function	No, but a learner can manually drag the mouse to the video section that would like to review again	Yes, and the automatically replaying function is provided for both the delayed and immediate review modes
Bilingual caption display for native language Chinese and foreign language English	No, but a learner can select a single language caption of Chinese or English provided by the Youtube player	Yes, but a learner can turn on or turn off the Chinese caption, English caption, or both by himself/herself

3.4.3 Listening Videos

There were five listening videos and the length of each video is about four to five minutes long used in the study for both groups. Among the five listening videos, there were four listening videos chosen from Youtube website by an English teacher who was the English teacher of the research participants and has already had over ten years experiences in English instruction in a Taiwan's junior high school for the listening learning and review process. Moreover, there was a self-made listening video that contained the contents of the previous four listening videos in terms of vocabulary words and text contents designed by the same English teacher for the listening learning process without reviewing after learning. The difficulty level of each listening video is slightly more advanced than the participants' English listening levels in order to assess their listening comprehension performance discriminatively. Mohamadi (2013) indicated that the difficulty level of a language listening task refers to the intrinsic ―cognitive load of a listening, its linguistic and informational complexity. Moreover, Buck (2001) identified numerous difficulties which can be confronted in listening tasks including unknown vocabularies, unfamiliar topics, fast speech rate, and unfamiliar accents. This study invited the English teacher who choice and designed the five listening videos to judge whether these five listening videos are appropriate to be served as the listening materials of the study with slightly more advanced difficulty than the research participants' English listening levels based on considering the difficulty factors that affect a listening task－the

possible unknown vocabularies, topics' familiarity, speakers' speech rates and accents.

3.4.4 Listening Comprehension Tests

Nowadays, multiple choice questions are commonly used to measure students' learning performance, especially in the listening assessment. Therefore, this study invited the same English teacher who choice and designed the five listening videos to design the corresponding five listening comprehension tests that respectively contain ten multiple choice questions in order to assess the effects of two different review mechanisms on participants' listening comprehension. Among the five listening comprehension tests, there were four listening comprehension tests designed based on the contents of each listening video. Specially, the fifth listening comprehension test was designed based on the self-made listening video that contained the contents of the previous four listening videos in terms of vocabulary words and text contents. All of the multiple choice questions in the five listening comprehension tests were designed as 5W's and 1H type questions aiming to evaluate whether examinee can successfully identify characters, places, events, time factors, or causes for the events that appeared in the listening videos. The reliabilities of the five English listening tests were assessed by using the internal consistency of Cronbach's alpha. Analytical results show that the Cronbach's alpha values of the five English listening tests were 0.816, 0.818, 0.822, 0.832, and 0.805, respectively. The reliability of an English listening test is excellent if its Cronbach's alpha value is higher than 0.90, while the reliability of an English listening test is good if its Cronbach's alpha value is between 0.80 and 0.90 (Cronbach & Shavelson, 2004). Therefore, the five English listening tests developed in this study have all the good reliability. The five listening comprehension tests were respectively assigned for the research participants of the both groups after they finished practicing the assigned listening video.

3.4.5 Data Analysis Methods

Aiming to investigate whether or not significant differences existed in the learners' English listening comprehension performance between the experimental group using the VALRM and the control group using the SDLRM, this study used Statistical Package for Social Science (SPSS) to analyze the collected data from the instruction experiment. The data collected from the research participants' English sectional examinations in their school and the five listening comprehension tests in the instruction experiment were used to assess the differences of listening comprehension performance between both the groups based on the analysis of covariance (ANCOVA) in this study.

4. Experimental Results

4.1 Analysis of Initial English Listening Abilities between Two Groups before the Experiment

To assess whether there were significant differences in initial English listening abilities for learners in the both groups before the experiment, the independent-samples *t*-test analysis was employed to compare difference in learners' grades on school sectional examinations between both groups. Table 2 shows the results. Analytical results show that the pretest score of the experimental group (M=82.778, SD=15.36) and that of the control group (M=84.738, SD=8.17) did not differ significantly (t=.082, p=.615 >.05), implying that both groups had equivalent initial English listening abilities before performing the instructional experiment.

Table 2. The independent-samples *t*-test results of the pretest for both groups

Assessment Item	Groups (N)	M (SD)	*t*	*p* (2-tailed)
Sectional examination	VALRM (18)	82.778 (15.36)	.082	.615
	SDLRM (21)	84.738 (8.17)		

4.2 Comparison of Listening Comprehension Performance between Two Groups with Different Review Mechanisms

To conduct ANCOVA for comparison of the listening comprehension performance in four weeks' experiment between two groups with different review mechanisms, the first step was to analyze the homogeneity of regression coefficients, and learners' sectional examination scores of listening comprehension were used as covariates in the analyses. The *F* test result (F=0.717, p=.403; F=0.087, p=.769; F=0.015, p=.902; F=0.205, p=.653; F=0.016, p=0.901) did not reach the significant level, thus it means the regression slopes of two groups is equivalent. This result confirmed the assumption of homogeneity of coefficients, and so this study further preceded the ANCOVA.

Table 3 shows the results of one-way ANCOVA in the four listening comprehension performance for the both groups. Analytical results show the mean scores of the learners in the experimental group with the VALRM support for listening review during the first and third week's experiments were not significantly different from those of the control group (F=0.589, p=.447>.05; F=1.93, p=.170>.05). These results indicated that the learners in the both groups had the same listening comprehension performance in the first and third listening tasks. However, the mean scores of the learners in the experimental group were significantly higher than those of the control group (F=4.22, p=.046<.05; F=5.427, p=.025<.05) during the second and four week's experiments. These results indicated that learners' listening comprehension performance of the experimental group was significantly superior to that of the control group in the second and four listening tasks. However, the average score of the experimental group in the four listening comprehension tests during the four weeks' experiment was significantly higher than that of the control group (F=5.321, p=.027<.05). In other worlds, the average listening comprehension performance of the learners in the experimental group who performed review for the four listening videos by using VALRM is remarkably superior to learners in the control group using SDLRM to perform review process. This study inferred that the possible reason of why the listening comprehension performances of the learners in the experimental group during the first and third week's listening tasks were not significantly different from those of the control group is the insufficient review time. In other words, giving enough review time to perform listening review process with the VALRM support should be considered.

Table 3. Results of one-way ANCOVA in four listening comprehension performance for both groups

Assessment	Groups (N)	M (SD)	*F*	*p* (2-tailed)
Listening comprehension test 1	VALRM (18)	72.778 (24.9247)	0.589	.447
	SDLRM (21)	70.000 (18.1659)		
Listening comprehension test 2	VALRM (18)	82.778 (19.6456)	4.22	.046*
	SDLRM (21)	75.238 (21.3586)		
Listening comprehension test 3	VALRM (18)	85.000 (10.4319)	1.93	.170
	SDLRM (21)	82.381 (9.9523)		
Listening comprehension test 4	VALRM (18)	70.000 (15.7181)	5.427	.025*
	SDLRM (21)	62.381 (16.0950)		
Average Listening comprehension test	VALRM (18)	77.640(14.963)	5.321	.027*
	SDLRM (21)	72.500(12.990)		

*p < .05

4.3 Comparison of the Listening Comprehension Performance between Two Groups without Review Mechanisms

The fifth listening comprehension test was to assess each group's listening comprehension performance after the use of different listening review mechanisms for four weeks. The learners in the both groups were asked to complete the final comprehension test without reviewing the listening material. To conduct ANCOVA for comparison of the fifth listening comprehension performance between two groups, the first step was to analyze the homogeneity of regression coefficients, and learners' sectional examination scores of listening comprehension were used as covariates in the analyses. The *F* test result (*F*=1.792, *p*=.189) did not reach the significant level, thus it means the regression slope of two groups is equivalent. This result confirmed the assumption of homogeneity of coefficients, and so this study further preceded the ANCOVA. Table 4 shows the results of one-way ANCOVA in the fifth listening comprehension performance for both groups. Analytical results show the mean score of the experimental group was significantly higher than that of the control group (*F*=4.399, *p*=.043<.05). This result indicated that the listening comprehension abilities of the learners in the experimental group was significantly promoted due to adopting more effective review mechanism than that of the control group used.

Table 4. Results of one-way ANCOVA in the fifth listening comprehension performance for both groups

Assessment	Groups (N)	M (SD)	*F*	*p* (2-tailed)
Listening comprehension test 5 (Posttest)	VALRM (18)	72.778 (19.3438)	4.399	.043*
	SDLRM (21)	66.190 (18.5678)		

*p < .05

4.4 Interview Results

Interview results show six of the interviewees agreed that the VALRM was useful in English listening learning because they could find out the video sections they were not clear comprehension and replay them for review easily and efficiently. Moreover, most of the interviewees praised the

customized review mechanisms of the VALRM. By clicking on “play the mark” for the delayed review mode or “mark and replay” for the immediate review mode, learners could replay certain video sections they wanted to review many times. One of the interviewees from the control group admired the function of setting caption that provides whether or not displaying English and Chinese captions of the chosen review video sections. All of the interviewees agreed that the VALRM was easy to use. They got used to the VALRM soon after several time practices. One of the interviewees preferred “play the mark” for the delayed review mode because he could decide which video section to replay after finishing listening to the whole video. However, all the other interviewees liked “mark and replay” for the immediate review mode the most because they preferred repeating the video sections they were not clear with and figured it out immediately. All of the interviewees expressed they are willing to use the VALRM as a supported learning tool for promoting English listening comprehension ability in the future. They all commended its convenience for being able to listen to the unclear parts many times. Besides, all of the interviewees suggested the function buttons should be more easily-recognized because they sometimes got confused about these two buttons “play the mark” and “mark and replay.”

5. Discussion

This study adopted a mixed research methodology combining quantitative and qualitative approaches to examine how the VALRM influences learners’ listening comprehension performance. The analysis results show that the listening comprehension performance of the learners in the experimental group with VALRM support for listening review was significantly superior to that of the control group. The finding echoes several previous studies (Chen, Wang & Lin, 2017; Lin & Chen, 2018), revealing that performing review task with the technology support outperforms human’s autonomous review. This study inferred that the main reasons are that the proposed VALRM can support learners to proceed more effective and efficient listening comprehension training as well as the captured captions provided by the VALRM helped learners discern precisely what was being said on the video. The findings are consistent with several studies (Hsu, 2015; Hsu, Hwang & Chang, 2014), indicating that appropriately providing captions can enhance listening competence. Additionally, Wolfgramm, Suter and Göksel (2016) confirmed that the most important predictor of both listening and reading comprehension was vocabulary. Moreover, the proposed VALRM can automatically retrieve and display bilingual captions marked as the sections that learners need to review on the user interface of the VALRM. This function provides benefits to help the learners in the experimental group identify unfamiliar or unknown English vocabularies and their corresponding translations to Chinese that is the mother language of the learners while hearing a listening video for training listening competence. More importantly, the VALRM helped learners find certain vocabulary words that they cannot correctly identify by their hearing from the English captions extracted from a video so that they can more care about the pronunciation of these vocabulary words or even check the pronunciation of these vocabulary words through an online dictionary.

Encouragingly, the learners in the experimental group had better listening comprehension performance in the fifth listening learning task even though removing the VALRM support for individual review than those of the control group removing the SDLRM support. The results confirms that the listening comprehension abilities of learners in the experimental group was significantly promoted due to adopting more effective review mechanism than that of the control group used. The possible reasons are that the number of the unfamiliar or unknown vocabulary words of the learners in the experimental group was decreased and their listening abilities in linking sounds between words were promoted due to the enhancement of listening abilities caused by more effective review mechanism.

All of the interviewees expressed they sometimes got confused about these two functions “play the mark” and “mark and replay.” However, it is encouraged that the results of the semi-structured

interview show that most of interviewees agreed the VALRM as a more useful and convenient listening learning tool than the SDLRM. They generated positive attitudes toward being able to arrange their own learning and review processes. The findings are consistent with previous study (Al-Yaari, 2013), indicating that most learners had positive perceptions toward CALL as a learning tool and toward playing a positive role in learning.

The limitations of the study are addressed as follows. First, there were only thirty-nine Grade 8 students, who were recruited from a junior high school, participating in the experiment. The small sample size in this study thereby could not represent the whole population of junior high school students in Taiwan. Second, English listening videos used in the study were chosen based on English competence of the Grade 8 students; therefore, the results of the study may not be transferred ready to the other age groups. Third, the experimental period only lasted for four weeks owing to the limitation of school schedule. Whether or not extending the experimental period affects research outcomes needs to be further studied. Finally, this study focused on English listening learning with the proposed VALRM support; thereby, the results may not be transferred ready to other languages or subjects.

6. Conclusions and Future Works

This study aimed to examine the effects of learners who respectively used different listening review mechanisms on learners' listening comprehension. Based on the analysis of the experimental results and semi-structured interviews, several major conclusions are summarized as follows. First, analytical results show that the learners using the VALRM were significantly superior to those using the SDLRM in English listening comprehension performance. Therefore, the study confirmed that the effects of the VALRM on promoting listening comprehension are positive. Second, the results of the semi-structured interview show that most of interviewees agreed that the VALRM is a more useful and convenient listening learning tool than the SDLRM. In addition, the two review modes provided by the proposed VALRM were highly commended because of its convenience, immediacy, and efficiency for review. Encouragingly, all of the interviewees are willing to use the VALRM to enhance their listening comprehension abilities in the future. These affirmative feedbacks proved the positive influence of the VALRM on facilitating learners' English listening learning performance.

This study recommends future studies in the following areas. Restricted by practical concerns on experimental time and cost, the experimental treatments of the present study only lasted for four weeks. However, long-term experimental treatments may generate deeper and more convincing influences on learners' listening learning performance and their perceptions toward using the VALRM. Thus, future study should evaluate long-term effects of using the VALRM on learners' listening comprehension and attitudes. Moreover, since there were only thirty-nine eighth graders recruited in the study, the effects of learners using the VALRM on listening comprehension may not be fully comprehensive. Therefore, future study with a large sample size should be considered. Moreover, the results of the study merely examine the outcome of learners' listening comprehension performance; however, how learners made use of the VALRM has still not been examined in this study. By recording and analyzing learners' learning processes, different learners' learning preferences and strategies can be explored clearly, thus providing useful knowledge to modify the system functions. Hence, further study could take recording learners' learning processes into account based on learning process recording technologies, such as experience API (xAPI) developed by Advanced Distributed Learning (ADL) Initiative.

References

Al-Yaari, S. A. S. (2013). Using audio-visual aids and computer-assisted language instruction

(CALI) to overcome learning difficulties of listening in students of special needs. *Journal for the Study of English Linguistics*, *1*(2), 142-167.

Başaran, H. F., & Köse, G. D. (2013). The effects of captioning on EFL learners' listening comprehension. *Procedia-Social and Behavioral Sciences*, *70*, 702-708.

Bozorgian, H. (2012). Listening skill requires a further look into second/foreign language learning. *ISRN Education, 2012*, Article ID 810129.

Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System*, *25*(1), 39-53.

Brown, G. (1990). *Listening to Spoken English*. Second Edition. London: Longman.

Buck, G. (2001). *Assessing listenin*g. Cambridge: Cambridge University Press.

Chung, J. M. (1999). The effects of using video texts supported with advance organizers and captions on Chinese college students' listening comprehension: An empirical study. *Foreign Language Annals*, *32*(3), 295-308.

Chao, Y., & Chien, L. (2005). College English majors' listening strategies and difficulties while taking TOEFL. *The Fifteenth International Symposium on English Teaching*, Taipei: Crane, 292-301.

Chen, A. H. (2013). EFL listeners' strategy development and listening problems: A process-based study. *The Journal of Asia TEFL*, 10(3), 81-101.

Chiu, P. S., Chen, H. C., Huang, Y. M., Liu, C. J., Liu, M. C., Shen, M. H. (2016). A video annotation learning approach to improve the effects of video learning. *Innovations in Education and Teaching International*, DOI: 10.1080/14703297.2016.1213653.

Chen, C. M., Wang, J. Y., & Lin, M. (2017). Enhancement of English learning performance by using an attention-based diagnosing and review mechanism in paper-based learning context with digital pen support. *Universal Access in the Information Society*, https://doi.org/10.1007/s10209-017-0576-2

Chen, C. M. & Chung, C. J. (2008). Personalized mobile English vocabulary learning system based on Item Response Theory and learning memory cycle.' Computers & Education, *51*(2), 624-645.

Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, *64*, 391-218.

Diao, Y., Chandler, P., & Sweller, J. (2007). The effect of written text on comprehension of spoken English as a foreign language. *The American journal of psychology, 120*, 237-261.

Feyten, C. M. (1991). The power of listening ability: An overlooked dimension in language acquisition. *The modern language journal*, *75*(2), 173-180.

Field, J. (1998). Skills and strategies: Towards a new methodology for listening. *ELT journal*, *52*(2), 110-118.

Fu, X., Schaefer, J. C., Marchionini, G., & Mu, X. (2006). Video annotation in a learning environment. *The 69th Annual Meeting of the American Society for Information Science and Technology (ASIST)*, Austin, USA.

Gündüz, N. (2005). Computer assisted language learning. *Journal of Language and Linguistic Studies*, *1*(2), 193-214.

Hsu, C. K. (2015). Learning motivation and adaptive video caption filtering for EFL learners using handheld devices. *ReCALL*, *27*(01), 84-103.

Hsu, C. K., Hwang, G. J., & Chang, C. K. (2014). An automatic caption filtering and partial hiding approach to improving the English listening comprehension of EFL students. *Educational Technology & Society*, *17*(2), 270-283.

Huang, H. C., & Eskey, D. E. (1999). The effects of closed-captioned television on the listening comprehension of intermediate English as a second language (ESL) students. *Journal of Educational Technology Systems*, 28(1), 75-96.

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education*, *69*, 121-130.

Latifi, M., Mobalegh, A., & Mohammadi, E. (2011). Movie subtitles and the improvement of listening comprehension ability: Does it help?. *The Journal of Language Teaching and Learning*, *1*(2), 18-29.

Leveridge, A. N., & Yang, J. C. (2013). Testing learner reliance on caption supports in second language listening comprehension multimedia environments. *ReCALL*, *25*(02), 199-214.

Lui, H. J. (2008). A study of interrelationship between listening strategy use, listening proficiency levels, and listening style. *ARECLS, 5*, 84-204.

Liu, G. Z., Chen, J. Y., Hwang, G. J. (2017). Mobile-based collaborative learning in the fitness center: A case study on the development of English listening comprehension with a context-aware application. *British Journal of Educational Technology*, doi:10.1111/bjet.12581

Lin, Y. T. & Chen, C. M. (2018). Improving effectiveness of learners' review of video lectures by using an attention-based video lecture review mechanism based on brainwave signals. *Interactive Learning Environments*, DOI: 10.1080/10494820.2018.1451899

Markham, P. (2001). The influence of culture-specific background knowledge and captions on second language comprehension. *Journal of Educational Technology Systems*, *29*(4), 331-343.

Mohsen, M. A. (2016). The use of help options in multimedia listening environments to aid language learning: A review. *British Journal of Educational Technology*, *47*(6), 1232–1242.

Mohamadi, Z. (2013). Determining the difficulty level of listening tasks. *Theory and Practice in Language Studies*, *3*(6), 987-994,

Nunan, D. (2002). Listening in language learning. *Methodology in language teaching: An anthology of current practice*. Cambridge: Cambridge University Press.

O'brien, A., & Hegelheimer, V. (2007). Integrating CALL into the classroom: The role of podcasting in an ESL listening strategies course. ReCALL, *19*(02), 162-180.

Osada, N. (2004). Listening comprehension research: A brief review of the past thirty years. *Dialogue*, *3*(1), 53-66.

Rost, M. (1994). *Introducing listening*. London: Penguin.

Sun, K. C. (2002). Investigation of English listening difficulties of Taiwan students. *The Eleventh International Symposium on English teaching/Fourth Pan-Asian Conference* (pp. 518-525). Taipei: Crane.

Şevik, M. (2017). *Does Bimodal subtitling effect listening comprehension in EFL contexts?* Researching ELT: Classroom Methodology and Beyond, pp. 41-48.

Schnell, J. (1995). Effective listening: More than just hearing. Retrieved November 7, 2017, from the Educational Resources Information Center (ERIC # ED379691), web available at https://eric.ed.gov/?id=ED379691

Taylor, S. E. (1964). *Listening. What Research Says to the Teacher*. ERIC Clearinghouse Underwood, M. (1989). *Teaching listening*. Addison-Wesley Longman Ltd.

Talebinezhad, M. R., & Abarghoui, M. A. (2013). The Iranian high school students' attitude toward CALL and the use of CALL for EFL receptive skills. *Theory and Practice in Language Studies*, *3*(2), 329-337.

Vandergrift, L. (1999). Facilitating second language listening comprehension: Acquiring successful strategies. *ELT Journal*, *54*(4), 168-176.

Vahdat, S. & Eidipour, M. (2016). Adopting CALL to improve listening comprehension of Iranian junior high school students. *Theory and Practice in Language Studies*, *6*(8), 1609-1617.

Wu, J. R. (2004). *The effects of pictures at different time periods on the listening comprehension of junior high school students in Taiwan* (Doctoral dissertation, Department of English, National Chengchi University).

Wolfgramm, C., Suter, N., Göksel, E. (2016). Examining the role of concentration, vocabulary and self-concept in listening and reading comprehension. *International Journal of Listening*, *30* (1-2), 25-46.

Wang, Y. & Treffers-Daller, J. (2017). Explaining listening comprehension among L2 learners of English: The contribution of general language proficiency, vocabulary knowledge and metacognitive awareness. *System*, *65*, 139-150.

Yang, J. C., & Chang, P. (2014). Captions and reduced forms instruction: The impact on EFL students' listening comprehension. *ReCALL*, *26*(1), 44-61.

Zarzour, H. & Sellami, M. (2018). Effects of a linked data-based annotation approach on students' learning achievement and cognitive load. *Interactive Learning Environments*, DOI: 10.1080/10494820.2018.1446989