Video or audio listening tests for English language teaching context: Which is more effective for classroom use?

¹Clara Herlina Karjo iD, **²Menik Winiharti** iD, **³Safnil Arsyad** iD

1,2English Department, Bina Nusantara University, INDONESIA
1,2*Jalan Kebon Jeruk Raya No. 27, Kebon Jeruk, Jakarta Barat 11530, Indonesia*

3English Education Postgraduate Program, Bengkulu University, INDONESIA
3*Jalan WR Supartaman Kandang Limun Kota Bengkulu 38371, Indonesia*

ABSTRACT

Multimodal inputs (both auditory and visual) in the forms of films and videos have long been used in teaching EFL listening comprehension. Previous studies have shown that listening while watching videos can significantly aid students' comprehension. However, videos were rarely used as testing materials since they contained more than aural input so they did not 'really' test listening. This study explored the extent to which multimodal testing materials can be used in testing listening comprehension for EFL students and how the results would differ from that of mono modality testing materials. The participants were 100 students of the English Department, Bina Nusantara University (henceforth Binus) University Jakarta. The researchers gave them two kinds of tests: the video listening test (VLT) and audio listening test (ALT). The materials were two short videos from YouTube. The first test, ALT was given after the participants listened to the videos twice. On the contrary, VLT was administered after they watched the videos twice. To examine the differences in the effects of VLT or ALT on EFL students' performance in listening comprehension, the data were analyzed quantitatively. The results indicate that students got better scores for VLT compared to ALT. The findings imply that students' performance in listening comprehension is significantly improved with multimodal testing materials.

Since elementary schools, English has been studied by Indonesian students as a foreign language. Nevertheless, their length of learning English

does not guarantee that they are proficient in four language skills. Ur (1984) mentioned that even though their grammar skills are good enough, they still have problems in doing listening exercises. Listening comprehension is considered the most important among the four language skills because it provides the aural input which is the foundation of language acquisition (Gao, 2012). Listening is considered difficult by most EFL learners. In literature, there are a lot of complex factors such as rate of speech (Blau, 1990; Griffiths, 1992), prosody, accent, phonological features (Matter, 1989), hesitations, complex grammar (Gao, 2012), rhetorical signaling cues (Cross, 2011), lack of background knowledge (Chiang and Dunkel, 1992) and low language proficiency (Brown, 1995) can affect listening comprehension. Despite being the most difficult skill, listening is the most frequently used language skill in the classroom, compared to the other language skills (Ferris, 1998). In learning grammar, reading, vocabulary and speaking, listening skill is applied. For example, students must have sufficient listening ability to be able to understand the materials that the teacher is teaching. Unfortunately, many teachers still have the perception that listening is a passive skill, in which the students only have to listen passively without doing anything. This perception is not accurate, since listening is an active skill, not as passive as it seems.

When a person listens to something, there are a lot of cognitive processes going on. Purdy (1997) stated that in listening there are active and dynamic processes of attending, perceiving, interpreting, remembering, and responding to verbal and nonverbal needs, concerns, and information offered by other people. Listening is also a complicated process that involves linguistics, cognitive, cultural and social knowledge (Wang & Miao, 2003). Moreover, listening comprehension is also an inferential process that involves various background knowledge (Gilakjani & Ahmadi, 2011). Similarly, Rost (2002) and Hamouda (2013) described listening comprehension as an interactive process of meaning construction that involved listeners. Listeners, according to Gilaksjani & Sabouri, 2016) should comprehend the oral input through sound discrimination, previous knowledge, grammatical structures, stress and intonation, and other linguistic or non-linguistic clues. Therefore, successful listening, according to Anderson and Lynch (1988), does not only depend on what a speaker says, but also on how the listener plays part in the process, by activating various background knowledge, and by applying his previous knowledge to what he hears and tries to understand what the speaker means. To be able to listen well, listeners must have the ability for decoding the message, apply a variety of strategies and processes to make meaning, and the ability for responding to what is being said in various ways, depending on the purpose of communication. From the above discussion, it is clear that listening comprehension requires active participation on the part

of the listeners (in this case EFL the students) as well as the speaker because listening is a highly complex problem-solving activity where the listeners interact with a speaker to construct meaning in the context of their experiences and knowledge.

Regarding the teaching of listening comprehension in the EFL classroom, Gilakjani and Ahmadi (2011) propose three activities consisting of pre-listening, while listening and post-listening activities. Pre-listening activities include the outline of the listening text and teaching the key concepts. Here, a teacher can give a general description about the topic or theme of the text that the students will listen to, discuss some difficult words or concepts, and probably discuss the accent of the speakers, etc. These activities are used to activate students' prior knowledge and expectation and to provide the necessary context for specific listening tasks. While listening activities are aimed to construct clear and accurate meaning as they interpret the speaker's verbal message and nonverbal cues. The purpose is to focus the comprehension of speaker ideas, to focus on organizational patterns, to encourage students' reaction to the speaker's ideas and the use of language. The activities include open-ended activities like question and answer, filling in the missing words, or pair work activity. Post-listening activities have the purpose of connecting what the students heard with their experience, encouraging critical listening and reflective thinking. Post listening activities include checking students' comprehension and clarifying understanding, answering questions that have not been understood.

Traditionally, listening comprehension materials are given in audio-only input. However, if learners can see how the language is used in an actual situation, listening activity can be more meaningful. To put it in another way, when learners see how the speakers speak the language, they can learn the language from both audio and visual inputs. Harmer (2007) cited several materials that can give access to both audio and visual inputs such as DVD, film clips on video, or online watch while listening. By watching videos, some paralinguistic behaviors such as how intonation matches the facial expression and gestures can be seen. The availability of the various materials for listening comprehension will enable teachers to give students continuous and theoretical directions that will make students change their previous notion that listening course is only teacher's playing audio and students repeat to improve their linguistic skills through listening (Ruan, 2015). Providing multimodal inputs (visual and auditory) in the form of materials for listening comprehension has been done for quite a long time. Wagner (2008) claimed that language teachers have already utilized movies, TV shows, and other sources of audiovisual media in their teaching, especially in teaching L2 listening. Wagner (2008) further asserted that the use of multimodal media

such as videos allows the listener to process both the aural and visual information communicated by the speakers as in real-life situations.

Compared to audio-only listening materials, multimodal presentation (i.e., movies or videos with subtitles or captions) is found to improve EFL learners' listening comprehension (Zareian, Adel &Noghani, 2015). This is because, multimodal presentation (particularly in the case of captioned video) can aid comprehension for language learners who are "hard of listening" and find the speech of foreign language TV, films, and videos difficult to follow and understand (Vanderplank, 2016). Captions or subtitles which accompany the video have been found highly effective in promoting listening comprehension (Yang & Chang, 2014). Thus, it will be highly likely that videos will be used in the teaching of L2 listening in the future because of considering the encompassing influence of video technology and the internet in daily life. In the field of instructional design, the focus of future studies will rely on multimedia learning regarding the effective use of video to enhance learning and the impact of video on learning (Guo, Kim & Rubin, 2014; Kay, 2012).

However, irrespective of the popular use of videos as teaching materials for L2 listening, they are not equally popular as testing materials. This notion was affirmed by Wagner (2010) who said that video is commonly employed in L2 classrooms, but test developers have been reluctant to use videotexts on the tests of L2 listening ability. The test developers' reluctance for using videos in a listening test might be related to technology and practicality. Previously, Wagner (2008) stated that more resources were needed to create a video listening test than the more traditional audio-only test. Some researchers are concerned that visual channels may affect the test-taker performance during the video listening test. In other words, whether the students who take the video listening test will be more influenced by the visual factors rather than concentrating on the aural content of the video.

According to Taylor and Garenpayeh (2011), test takers' performance may be affected by external contextual factors and individual characteristics. Visual images may distract test takers. Moreover, even though the students' language proficiency is the same, not all students may understand he content. Test-takers' performance may also be affected by their internal cognitive factors through a loading effect while they are processing information.

Despite the concerns regarding the effect of visual images on test-takers' performance, previous studies of video listening tests (VLT) and audio-only listening tests (ALT) showed various results.In the study of Basal, Gulozer, and Demir (2015), even with the visual elements of the videos, the ALT group performed significantly higher than the VLT group. On the contrary, Shin's (1998) study discovered that participants of video tests performed significantly (around 25%) better compared to an audio test

group.However, Gruba (1993) found out that the scores of video-mediated and audio-mediated groups did not show a statistically significant difference.

Nevertheless, further research still needs to be carried out, either to corroborate or to refute the previous studies. The present study is aimed to compare the EFL students' performance in listening comprehension tests using mono-modality presentation (audio-only test) and multimodality presentation (audio, visual and textual) using video clips from YouTube. The research is also aimed at finding the possibility of using multimodal materials for the teaching of listening comprehension for EFL students. As a guideline, the following hypotheses are addressed in this study:

1) The use of multi-modality presentation (audio, visual and textual) can significantly improve student's listening ability, and
2) The use of multi-modality presentation cannot significantly improve students' listening ability.

METHOD

Research design

The present study utilized the post-test only control group design (Creswell, 2009). Two types of test modality (audio or video) in the listening comprehension test were used to measure participants' performance. These two different modalities were the Audio listening test (ALT) and the Video listening tests (VLT) were administered to the same group of participants. The data were analyzed quantitatively using SPSS software. The statistical calculation used was paired samples T-test to find out the answer to the hypotheses of whether the use of multi-modality presentation can significantly improve the students' listening ability which is reflected by their listening test scores.

Participants

The participants of the research were 100 students of the English Department of Bina Nusantara University, Jakarta. They were purposely chosen from semesters 3, 5, and 7, with the assumption that they have sufficient English knowledge to participate in this research. However, the research did not differentiate the students' results based on their years of study (the length of semesters they have taken) and their English proficiency levels, but only measured the results based on the modes of presentation (audio and visual) used. Listeners did not demonstrate hearing and vision problems, nor did they stay in English-speaking countries before. The following table 1 shows the summary of participants' data concerning their age, semester, and gender.

Table 1: Summary of Participant Data

Gender	Semester			Average Age	Total
	3	5	7		
Male	5 (18.5 %)	15 (55.5 %)	7 (25.9 %)	20	27
Female	21 (28.8 %)	39 (53.4 %)	13 (17.8 %)	19.5	73

Instruments

The topics for ALT and VLT listening comprehension tests were taken from the video clips of Ted Talks from YouTube. TED (Technology, Entertainment, Design) is a global set of conferences run by the private nonprofit Sapling Foundation, with the slogan "Ideas Worth Spreading". TED emphasizes on the educational aspects. There were two topics taken as the materials for the tests, i.e. *The language of lying* and *The effect of sleep deprivation.* These topics were chosen randomly and were not related whatsoever to the participants' study course. The duration of each video is around 5 minutes. The first video was used to devise the listening test in the Audio (ALT) modality, while the second video was used to devise the listening test in Audio Visual (VLT) modality. The test for each modality consists of 20 question items, which were grouped into three types of test items. Table 2 below shows the representation of the topics and the type of test items.

Table 2: Topics and test items.

Testing Modality	Topic	Type of test	Number of items
ALT	The language of lying	Word listing	7
		Cloze summary	8
		Multiple choice	5
VLT	The effect of sleep deprivation	Cloze summary	7
		Number listing	4
		Multiple choice	9
Total			40

Data collecting procedure and analysis

This study used a post-test-only design as proposed by Creswell (2009) to measure the learners' performances in two types of modality in listening tests. Participants were given two types of tests:

(a) Audio listening test (ALT)

This test was administered after the participants *listened* to the first video *The language of lying* for 5 minutes approximately. In this kind of test, the screen was turned off so the participants only got the audio input. After a three-minute break, they listened to the audio content of the video once again. Finally, they were given a written test regarding the content of the video.

(b) Video listening test (VLT)

The second test was administered after the participants finished doing their ALT. The procedure was similar. This time, they watched the second video *The effect of sleep deprivation* with the screen turned on. Thus they got audio as well as visual input to help them understand the materials in the video. After watching the video two times, they were given a written test regarding the content of the video.

FINDINGS

The descriptive statistic calculation gives the following results for both samples. The average means score of VLT is 11.80. This means that VLT participants answered 11.80 questions correctly out of 20 items. The minimum score gotten is 4 and the maximum score gained is 20. On the other hand, the ALT participants only answered 10 out of 20 questions correctly, with the minimum score and the maximum score of 19. In general, the results indicate that VLT generated a better score than ALT. These scores were further analyzed for finding the correlation between the VLT and ALT. The analysis yielded the Pearson correlation coefficient of 0.755 with a probability value of 0.000, far below the significant level $\alpha = 0.05$. This shows a significant correlation between VLT and the ALT scores.

Table 3 below provides the results of the paired samples test. Paired samples t-test was chosen because the same participants were tested using two modes of presentation (VLT and ALT). Thus, the purpose of this test is to find out the significance value of the differences between VLT and ALT. The mean difference between VLT and ALT is 1.8. T-test for the hypothesis of *H0: VLT = ALT* gave the score of *t* = 11.808. The p-value for the two sides was 0.000, which is less than $\alpha = 0.05$. This result is convincing evidence to reject the hypothesis that VLT gives the same result as ALT. The conclusion that can be drawn is that VLT gives a better score than ALT.

Table 3: Paired Sample T-Test

Testing Modality	Paired differences		t	Sig. (2-tailed)
	Mean Differences	Std. Error Mean		
VLT vs ALT	1.80	.294	11.808	.000

Besides the modality of testing, the types of questions might also affect the scores attained by the students. Since there are three types of questions in both VLT and ALT, table 4 below shows the interaction between the question types with the modality. The questions are divided into three types: summary, listing and multiple choices. The dependent variable is the score gained by the

participants. The statistical calculation is done using Two-way ANOVA. There are two kinds of outputs given, the test between-subject effects and post hoc comparison.

Table 4 below shows the comparison and the correlation between the modality and the type of test items. The results indicate that there are no significant differences between the types of test items and the modality of presentation. This is not a surprising result because Brindley and Slatyer (2002) reported that learners' performance in the competency-based listening assessment is affected by the test item format and item difficulty, not by the modes of presentation. Thus, the findings indicate that modality and the types of test items do not have interaction even though the mean scores for each modality (VLT and ALT) differ significantly. Or, in other words, modality of presentation does not affect the attainment of scores for different types of test questions.

Table 4: Test between Subject Effects

Source	Dependent Variable: Score		
	Mean Square	*F*-stat	Sig.
Corrected Model	1.764	3.488	0.035
Intercept	202.608	400.698	0.000
Modality	6.044	11.952	0.005
Type	1.387	2.743	0.104
(Modality x type)	0.000	0.000	1.000

a. R-squared = .592 (Adjusted R Squared = .423)

The above table of ANOVA shows the statistical values for the main effect as follows:
a. For the modality factor: the F value is 11.952 with a degree of freedom *(df)* = 1 and *p*= 0.000, which is less than *a* =0.05, then H0 : ALT = VLT is rejected. The conclusion for this factor is that ALT scores differ from VLT scores.
b. For the question type factor: the F value is 2.743 with a degree of freedom (df) = 2 and *p*= 0.005, which is less than α =0.05, then H_0 : summary = listing = multiple choices is rejected. Then it can be concluded that each question type gives diverse results.
c. For the interaction factor: the F value is 0.000 with a degree of freedom (df) = 2 and *p* = 1.000, which is bigger than *a* =0.05, then H_0 : (μ summary – modality) = (μ listing – modality) = (μ choice – modality) cannot be rejected. This means that the modality (VLT or ALT) does not affect the attainment of scores for each type of question.

To see which test type shows different means, a *Post Hoc Multiple Comparison* was done. The summary can be seen in table 5. The results clearly show that test types do not influence the score gained. Again, these results confirm that the score difference results from the modality used in testing listening, not the types of test items given in the testing. Thus, it can be concluded that test results depend on the testing modality, not on the test items. The results are in line with Jafari and Hashim (2012) who no interaction effect between the test format and the students' listening proficiency level.

Table 5: The Multiple comparisons

(I) types	(J) type	Mean Difference (I-J)	Std. Error	Sig.
cloze	listing	-.0250	.41054	1.000
	choice	-.8450	.41054	.186
listing	cloze	.0250	.41054	1.000
	choice	-.8200	.41054	.207
choice	cloze	.8450	.41054	.186
	listing	.8200	.41054	.207

The multiple comparisons in table 5 show that the mean of cloze type differs from the mean of listing type and multiple-choice items. The mean differences between each type are very small, for instance between cloze and listing, the difference is only 0.025. The highest mean difference occurs between multiple-choice and summary, which is 0.8450, followed by multiple-choice and listing, which is 0.8200. The results indicate that multiple-choice items get the higher correct answers compared to listing and cloze. Further analysis of the results demonstrates that these differences can also be attributed to the modality used in the tests. Table 6 below displays a descriptive distribution of each type of question shown in percentages.

Table 6: Percentage Of Each Type Of Questions

	Video Listening Test		Audio Listening Test	
Cloze Summary	3.64 / 7	52.00 %	2.50 / 8	31.25%
Listing	-	-	2.49 / 7	35.57%
Multiple Choice	4.49 / 9	49.89%	3.33 / 5	66.60%
Cloze Number	3.67 / 4	91.75 %	-	-
Mean score	11.80/ 20	59.00%	10 / 20	50.00%

When the results of the test are broken down into different test types, several things can be noted. In VLT, the total mean score for all items is 11.80 or around 59% of the test items can be answered correctly. However, for each part, students show diverse results. The highest score is achieved for the cloze number test that reaches 91.75%. This means that the students can answer

accurately 3.67 from 4 questions regarding numbers. It is understandable since in the video the numbers are visually shown. Numbers are relatively harder to memorize if they are just spoken and not seen. The second type, which is also benefited from the video, is the close summary test. In this type of test, students have to fill in the summary with one or two words they heard from the video. The students got 3.64 out of 7 items. Again, by watching the video, they can see the visualization of the keywords, which appear on the screen. However, out of 9 items, students were only able to correctly answer 4.49 items in multiple-choice question type, thus making it the worst result.

Meanwhile, ALT results showed a different image. In total, out of 20 items, the students managed to correctly answer 10. Thus, the total mean percentage for all items was only 50%. In two types of test items, word listing and cloze summary, students only get 2.49 out of 7 and 2.5 out of 8 items. In these types of questions, students have to rely on their short-term memory to recall the words that they have to fill. Some of the words in this test are unfamiliar scientific terms that are quite difficult to catch such as *electroencephalograph* and *convoluted.* Visual showing of new vocabulary on the screen will help students to comprehend their meaning. This finding was similar to Winke et al. (2010) who stated that captions contributed better to learning than no caption concerning novel vocabulary recognition. Surprisingly, in ALT, students were able to achieve 66.6% correct answers in multiple choices items. This may be because, in multiple-choice questions, students can make informed guessing.

DISCUSSION
The statistical results above corroborate the indication that by listening to the audio and watching the video (which also includes subtitles or captions) at the same time, the students can perform and comprehend better. Visual modality enhances the students' understanding of the materials. Yang and Chang (2014) confirmed that the application of visual materials was very effective in promoting learning, particularly the use of caption which may enhance listening comprehension. Huang and Eskey (1999-2000) argued that captions and subtitles make audiovisual input more accessible and comprehensible to L2 learners, which is in line with Krashen's (1985) input hypothesis. The findings of these studies also revealed the students who were exposed to more than one modality, i.e., audio and visual can improve their performance compared to when they were only exposed to one modality such as to audio-only. Behroozizad & Majidi (2015) claimed that the affordance of three channels or modalities (aural, visual, and textual) might reduce the students' listening anxiety. In turn, these three channels will lead to better listening performance and more confidence in students' listening ability. The audio mode only provides students with aural input, so they have

to fully concentrate on the sounds provided. Consequently, they had to remember a lot of sound-based information. On the contrary, through audiovisual mode, students were equipped with more various information in addition to sound, such as pictures and also captions which occasionally occur in the video. These additional elements may have assisted the students' comprehension of the contents. Students' comprehension might be more thorough with the assistance of these three modalities compared to that supported by the audio mode only. The effects of these three modalities on students' comprehension were corroborated by the studies of Basal, Gulozer and Demir (2015), and Shin (1998).

Another reason for the higher scores gained by the VLT group is because according to Kruger and Doherty (2016), the capacity of working memory can be significantly enhanced by multimodal presentation as learners can process information in both channels (auditory narration and visual text). This statement is affirmed by Goh (2000) who argued that in general, EFL students had a limited capacity in their short-term memory, and they had a tendency to forget immediately what they heard previously because they were hasty to understand the new input. This means that students will easily fail to recall the materials if they are only presented in audio mode. On the other hand, if the information they receive is given in various modes, students can remember better. Consequently, they can perform better in video listening tests rather than in audio listening tests. The same result was confirmed by Chang, Lei and Tseng (2011) who found that multi modes (text plus sound), were more effective than single-mode listening instruction.

However, even though the modality of presentation did affect the attainment of the overall scores, the results did not show any significant interaction between the modes of presentation and the types of test items. Thus, in this case the students' listening proficiency level is not determined by the test format. The results are in line with Jafari and Hashim (2012) who found that there was no interaction between the test format and the students; listening proficiency level. Meanwhile, Brindley and Slatyer (2002) found different results. They reported that learners' performance in competency based listening assessment are affected by the test item format and item difficulty, not by the modes of presentation.

Nevertheless, the findings in Table 6 indicate that students performed differently for each type of question in VLT and ALT. In general, the students performed better in listening comprehension using the VLT mode. Students' understanding may be supported by the visual elements of the video such as the images and the subtitles/captions even though they might not understand the contents of the video. In other words, the video gives them multimodal inputs, which are beneficial for the understanding of the learning materials.

The visualization of difficult words on the screen also helps students to connect the pronunciation, and the spelling with the meaning, which in turn can make them retain the words better. For example, the words *electroencephalograph* was spoken, shown in the picture and written on the screen. Jing (2010) also confirmed that subtitles or captions in the video could help students with the spelling of difficult words and writing a summary after listening. That is why 80% of the students were able to recall the difficult words in the test of VLT. On the other hand, the difficult word *convoluted* found in ALT can only generate 5% correct answers. Markam, Peter and McCarthy (2001) claimed that L2 learners generally have higher reading comprehension skills than listening comprehension skills; thus, subtitles can be beneficial when they listen to L2 reading materials. The lower rate of word recognition in ALT suggests that students need visual input besides aural input to recall L2 listening materials better. These results confirm several researchers' findings that bimodal input (audio and visual) can speed up the recognition of words, and the comprehension of content (Chung, 1999; Guillory, 1998; Koolstra & Beentjes, 1999). Another interesting finding relates to the results for multiple choice questions. Unexpectedly, the VLT group only achieved 50% accuracy for multiple-choice questions compared to the 66.6% of the ALT group. This might be caused by the possibility for the students to make informed guessing for this type of question. Hence, although the multiple-choice question type is the most widely used format to measure listening ability it does not necessarily offer the best result in the listening comprehension test (Hemmati & Ghaderi, 2014).

The findings of this study bring about several implications either for the teaching or testing of listening comprehension. The first implication is the use of videos in teaching listening. Vandergrift (2011) has predicted that the language learning environment will be transferred into a new era of teaching listening based on the use of authentic audiovisual materials, because of the emergence of technology. Videos offer multimodal inputs that will enhance both situational and interactional authenticity and aid learners' comprehension (Wagner, 2007). Visual elements in videos can also activate the background knowledge of the listeners (Ockey, 2007). Thus, it is highly recommended to use videos in teaching listening for EFL students. Hosogoshi (2016); Danan (2004), and Vanderplank (2013) have explored some positive effects of using audiovisual materials for L2 learners, among others it can improve listening comprehension, foster vocabulary learning, develop oral production skills, and lower the learners' anxiety.

However, care must be taken when choosing the materials to be used in the classroom. Teachers should previously select and watch the videos themselves before using them as the materials for listening subjects. Videos should be correlated with the student's proficiency in English. The topics

should be within the students' understanding and interest. Students may have been involved in choosing the materials for their learning.

Moreover, teachers should not forget that the purpose of video viewing is for teaching listening comprehension. Thus, teachers should create some activities, which can provoke engagement and expectation (Harmer, 2007) on the part of the listeners or students. Some of the activities, among others, are picture less listening (similar to the audio listening procedure in which the teacher covers or turns off the screen and the students only listen to the dialogue or the talks) and using subtitles (the sound is turned off and the students try to construct the dialogue based on the subtitles). There are many more activities using audiovisual materials that can be done to improve students listening comprehension.

The second implication is the use of videos in testing listening. Although there has been no conclusive opinion, videos begin to be used in testing listening comprehension especially for EFL students. Sovorov (2015) believes that videos have the possibility of providing multimodal inputs which will result in a greater level of authenticity of test tasks. They will also create testing conditions that may be closely similar to the situation of the target language domain. Videos are readily available materials that can be taken (mostly freely) from the Internet or language learning websites. Again, care must be taken in choosing the materials for testing listening comprehension. The same principles apply that the materials should be appropriated with the student's proficiency level and the test should be carefully prepared so that it fulfills the purpose of testing students' listening ability.

CONCLUSIONS

This study confirms previous studies that multimodal presentation can improve EFL students' listening comprehension. The findings show that the video listening test produced a mean score of 11.80 compared to the mean score of the audio listening test, which is only 10.00. This result indicates that the video listening test increased the students' comprehension of the lesson materials. Videos enhance learners' comprehension because they give multimodal inputs in the forms of visuals (context, subtitles, pictures, etc.) as well as auditory input. Multimodal inputs are considered to enhance working memory and comprehension. Therefore, multimodal presentation is highly recommended to be used in the teaching of EFL listening comprehension. Even, it is also suitable for teaching and learning other language skills (reading, writing and speaking). With the advance of technology, language teaching and learning materials can be easily obtained. However, care should be taken in selecting and designing the instruction and testing materials to achieve the desired goal. Thus, the fact that the materials for this study were

chosen randomly from TED Talks recording became the limitation of the present research. For future research, therefore, the researcher can carefully choose the materials which can fulfill the testing objectives. Moreover, larger sample size can be used to examine the impact of multimodalities on the comprehension of various text types, such as lectures, dialogues, and authentic listening materials.

This study only used a posttest only control group design in which the students' listening improvement may be assisted by other individual learning activities done by the students outside the research treatments. Therefore, future studies should use a pretest and posttest controlled and experimental group design to get more valid and reliable data on the students' listening ability improvement. This will give more convincing evidence on the effective use of multi-modal presentation in listening tests to use in English as a foreign language class.

ACKNOWLEDGMENTS

REFERENCES

Anderson, A., & Lynch, T. (1988). *Listening*. Oxford: Oxford University Press.

Başal, A., Gülözer, K. & Demir, İ. (2015). Use of Video and Audio Texts in EFL Listening Test. *Journal of Education and Training Studies,* 3(6), 83-89.

Behroozizad, S.; Majidi, S. (2015). The effect of different modes of English captioning on EFL learners' general listening comprehension: Full text vs. keyword captions. *Advances in Language and Literary Studies*, 6(4), 1670-1677.

Blau, E. K. (1990). The effect of syntax, speed and pauses on listening comprehension. *TESOL Quarterly,* 24, 746-753.

Brindley, G. & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing,* 19(4), 369-394

Brown, G. (1995). dimensions of difficulty in listening comprehension. In D. Mendelshohn & J. Rubin (Eds). *A Guide for the Teaching of Second Language Listening*, 59-73. San Diego: Domine Press.

Chang, C. C., Lei, H. and Tseng, J. S. (2011) Media presentation mode, English listening comprehension and cognitive load in ubiquitous learning environments: Modality effect or redundancy effect? *Australasian Journal of Educational Technology,* 27(4), 633–654

Chiang, C. S. & Dunkel, P. (1992). The effect of speech modification, prior knowledge and listening proficiency on EFL lecture learning. *TESOL Quarterly,* 26, 345-374.

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), 296–308

Creswell, J. (2009). *Research design: Qualitative, Quantitative and Mixed Methods Approaches*, 3rd Edition. London, United Kingdom: SAGE.

Cross, J. (2011). Comprehending news videotexts: the influence of visual contents. *Language Learning and Technology 15(2)*, 42-68.

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta: Translators' Journal*, 49(1), 67–77.

Ferris, D. (1998). Students' view on academic aural/oral skills: A comparative needs analysis. *TESOL Quarterly*, 289-318.

Gao, Y. (2012). Effects of speaker variability on learning spoken English for EFL learners. *Faculty of Arts and Social Sciences,* 59-67.

Gilakjani, A., & Ahmadi, M. (2011). A study of factors affecting EFL learners' English listening comprehension and the strategies for improvement. *Journal of Language Teaching and Research,* 2(5), 977-988.

Gilakjani, A. P., & Sabouri, N. B. (2016). The Significance of listening comprehension in English language teaching. *Theory and Practice in Language Studies*, 6(8), 1670–1677.

Griffith, R. (1992). Speech rate and listening comprehension: Further evidence of the relationship. *TESOL Quarterly*, 26, 385-391.

Gruba, P. (1993). A comparison study of video and audio in language testing. *JALT Journal 15*, 85-88.

Gruba, P. (1997). The role of video media in listening assessment. *System*, 25(3), 335–345.

Guillory, H. G. (1998). The effects of keyword captions to authentic French video on learner comprehension. *Calico Journal*, 15(1-3), 89–108.

Guo, P.J., Kim, J. & Rubin, R. (2014). How video production affects students engagement: An empirical study of MOOC videos. *Proceedings of the First ACM Conference on Learning @ Scare Conference.* Atlanta, Georgia.

Harmer, J. (2007). *The Practice of English Language Teaching.* Harlow: Pearson Longman.

Hemmati, F., & Ghaderi, E. (2014). The Effect of Four Formats of multiple-choice questions on the listening comprehension of EFL learners. *Procedia - Social and Behavioral Sciences*, 98, 637–644

Hosogoshi, K. (2016). Effects of captions and subtitles on the listening process : Insights from EFL learners ' listening strategies. *Jalt Call Journal*, 12(3), 153–178.

Huang, H. C., & Eskey, D. E. (1999-2000). 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.

Jafari, K. and Hashim, F. (2012). The effects of using advance organizers on improving EFL learners' listening comprehension: A mixed-method study. *System,* 40(2), 270–281.

Jewitt, C. (2013). Multimodal Teaching and Learning. In C. Chapelle, *The Encyclopaedia of Applied Linguistics* (pp. 1-5). Chichester: Blackwell Publishing.

Jing, Z. (2010). Testing via news videos: An exploratory study. *International Journal of Applied Linguistics*, 20(2), 178–205

Kay, R. H. (2012). Exploring the use of video podcast in education: A comprehensive review of the literature. *Computers in Human Behavior,* (28) 3, 820-831

Kelly, R. (1991). Lexical ignorance: The main obstacle to listening comprehension with advanced FL learners. *IRAL*, 29, 135-150.

Koolstra, C. M. & Beentjes, J. W. J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research & Development,* 47(1), 51–60.

Krashen, S. (1985). *The input hypothesis.* London, England: Longman

Kruger, J. & Doherty, S. (2016). Measuring cognitive load in the presence of educational video: Towards a multimodal methodology. *Australasian Journal of Educational Technology,* 32(6), 19-31.

Markham, P. L., Peter, L. A., & McCarthy, T. J. (2001). The effects of native language vs. target language captions on foreign language students' dvd video comprehension. *Foreign Language Annals,* 34(5), 439–445.

Matter, J. (1989). Some fundamental problems in understanding French as a foreign language. In H.W. Dechert & M. Raupach (Eds.). *Interlingual processes.* 105-119. Gunter Narr: Tubingen.

Ockey, G. (2007). Construct implication of including still image or video in computer-based listening tests. *Language Testing,* 24, 517–537.

Plastina, A.F. (2013). Multimodality in English for specific purposes: Reconceptualizing meaning-making practices. *LFE: Revista de Lenguas Para Finas Específicos,* 19, 385-410

Purdy, M. (1997). What is listening? In M.Purdy, & Borisoff, *Listening in everyday life: A personal and professional approach* (pp. 1-20). Lanham: University Press of America.

Ruan, X. (2015). The role of multimodal in Chinese EFL studentsautonomous listening comprehension & multiliteracies. *Theory and Practice in Language Studies,* 5(3), 549-565.

Shin, D. (1998). Using videotaped lectures for testing academic language. *International Journal of Listening* 12, 56-79.

Suvorov, R. (2009). Context visuals in L2 listening test: the effects of photograph and video vs audio-only format. In C. Chapelle, H. Jun, &I. Katz, *Developing and Evaluating Language Learning Materials* (pp. 53-68). Ames: Iowa State University.

Suvorov, R. (2014). The use of eye-tracking in research on video-based second language (L2) listening assessment: A comparison of context videos and content videos. *Language Testing*, 32 (4),463-483.

Suvorov, R. (2015). Interacting with visuals in L2 listening test: An eye-tracking study. ARAGs Research Report Online: British Council.

Taylor, R. & Geranpayeh, A. (2011). Assessing listening for academic purposes: Defining and operationalizing the academic construct. *Journal of English for Academic Purposes,* 10, 89-110.

Ur, P. (1984). *Teaching Listening Comprehension.* Cambridge: Cambridge University Press.

Vandergrift, L. (2004). Listening to learn or learning to listen? *Annual Review of Applied Linguistics*, 24, 3-25

Vanderplank, R. (2013). "Effects of" and "effects with" captions: How exactly does watch a tv program with same-language subtitles make a difference to language learners? *Language Teaching,* 1-16.

Vanderplank, R. (2016). The State of the Art I: Selected Research on Listening Comprehension and Vocabulary Acquisition. In *Captioned Media in Foreign Language Learning and Teaching,* 75-104. Palgrave: Macmillan.

Wagner, E. (2008). Video listening tests: What are they measuring? *Language Assessment Quarterly,* 5/3, 218-243.

Wagner, E. (2010). The effect of the use of video texts on ESL listening test-taker performance. *Language Testing,* 27, 493-513.

Wagner, E. (2013). An Investigation of how the channel of input and access to test questions affect L2 listening test performance. *Language Assessment Quarterly*, *10*(2),178-195

Wang, J. & Miao, Y. (2003). Theory and method for EFL listening teaching. *Computer-assisted Foreign Language Teaching,* 8(2), 1-5.

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14(1), 65–86.

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

Zareaian, G., Adel, S. M. & Noghani, F. A. (2015). The effect of multimodal presentation on EFL Learners' listening comprehension and self-efficacy. *Academic Research International*, 6(1), 263-271

THE AUTHORS

Clara Herlina Karjo is an Associate Professor at English Department, Bina Nusantara University. She teaches Sociolinguistics, History of English and Research Methods. She obtained her Doctoral Degree in English Applied Linguistics from Indonesian Catholic University of Atma Jaya. Her research interests vary from English phonology, language acquisition, language teaching, translation and discourse analysis. Her research papers have been disseminated in various journals and international conferences.

Menik Winiharti is a faculty member at English Department, Bina Nusantara University. She teaches English Grammar, Writing, Syntax, and Research Methods. Her research interests include pragmatics, translation, as well as language skills. She is now pursuing her Doctoral Degree in Linguistics at Universitas Pendidikan Indonesia, Bandung. Her research for dissertation discusses the performance of online machine translation focusing on undergraduate lecturers' academic writing.

Safnil Arsyad is a professor in English Language Education at the English Department of Education Faculty of University of Bengkulu in Bengkulu Indonesia. He has published in many international journals, such as Asia-Pacific Education Researcher, Australian Review of Applied Linguistics, Journal of Multicultural Discourses, Asian ESP Journal, Asian Englishes, Discourse and Interaction, International Journal of Instruction, Language and Linguistics Studies, Studies in English Language Education, International Journal of Language Education and Malaysian Online Journal of Education Management. His research interests are on discourse analysis of academic texts and English teaching and learning materials.