The effect of delivery mode on L2 students' listening comprehension skills: Context versus animated videos

Leyla Karatay
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Bilingual, Multilingual, and Multicultural Education Commons, and the English Language and Literature Commons

Recommended Citation
https://lib.dr.iastate.edu/etd/16606

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
The effect of delivery mode on L2 students’ listening comprehension skills: Context vs animated videos

by

Leyla Karatay

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Major: Teaching English as a Second Language/Applied Linguistics (Computer Assisted Language Learning)

Program of Study Committee:
James M. Ranalli, Major Professor
Gary Ockey
Geoffrey Sauer

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this thesis. The Graduate College will ensure this thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University
Ames, Iowa
2018

Copyright © Leyla Karatay, 2018. All rights reserved.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ iv

LIST OF TABLES ............................................................................................................ v

ACKNOWLEDGEMENTS .............................................................................................. vi

ABSTRACT .................................................................................................................... vii

CHAPTER 1. INTRODUCTION ..................................................................................... 1
  Statement of the Problem ......................................................................................... 1
  Aims and Scope of the Study .................................................................................... 3
  Structure of the Study .............................................................................................. 4

CHAPTER 2. LITERATURE REVIEW .......................................................................... 5
  Defining Listening Comprehension ......................................................................... 5
  Listening Delivery Modes ...................................................................................... 7
  Audiovisual Processing ......................................................................................... 8
  Research on the Use of Visuals in Listening .......................................................... 10
    Students’ Perceptions of Using Different Listening Delivery Modes .................. 13
    Integrated Listening-Speaking Tests .................................................................... 13

CHAPTER 3. METHODOLOGY .................................................................................. 15
  Participants ................................................................................................................ 15
  Setting ....................................................................................................................... 16
  Materials .................................................................................................................. 16
  Task ........................................................................................................................... 20
    Rubric ..................................................................................................................... 20
    Rating Process ...................................................................................................... 21
  Procedures ............................................................................................................... 22
  Data Analysis .......................................................................................................... 24
CHAPTER 4. RESULTS AND DISCUSSION ................................................................. 26

Research Question 1 ......................................................................................... 26
Research Question 2 ......................................................................................... 28

CHAPTER 5. CONCLUSION .............................................................................. 36

Implications ....................................................................................................... 36
Limitations ......................................................................................................... 38
Suggestions for Future Research ...................................................................... 39

REFERENCES .................................................................................................. 41

APPENDIX A. PARTICIPANTS’ BIODATA ............................................................. 44

APPENDIX B. IRB APPROVAL ........................................................................... 46

APPENDIX C. VIDEO SCRIPTS ................................................................. 47

Video 1: Where do superstition come from? ..................................................... 47
Video 2: Should we get rid of standardized testing? ......................................... 48

APPENDIX D. SURVEY ....................................................................................... 50

APPENDIX E. KEY DETAILS IN THE SUPERSTITIONS VIDEO ...................... 51

APPENDIX F. KEY DETAILS IN THE STANDARDIZED TESTING VIDEO ........... 52

APPENDIX G. CONSENT FORM ....................................................................... 53

APPENDIX H. SCREENSHOTS OF THE TASK IN CANVAS ............................. 57
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Screenshot of the first animated video.</td>
<td>17</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Screenshot of the second animated video</td>
<td>18</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Screenshot of one of the context videos</td>
<td>19</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1  Rating scale................................................................................................................................. 21
Table 2  Overview of data collection and analysis procedures................................................................. 24
Table 3  Descriptive statistics for animated and context video groups.................................................... 26
Table 4  Survey responses regarding difficulty level of task................................................................. 29
Table 5  Survey responses regarding the frequency of looking at the screen........................................ 29
ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Jim Ranalli, and my committee members, Dr. Gary Ockey and Dr. Geoffrey Sauer for their guidance and support throughout the course of this research. Thank you, Dr. Ockey, for answering my never-ending-questions.

I also thank Dr. Yusuf Sen for his endless support and encouragement for my graduate studies.

In addition, I would also like to thank my dear friends Erin and Alif for their help in creating videos in this study and Nazlı for all the support and ideas she gave since day one. I want to also offer my appreciation to those who were willing to participate in my study, without whom, this thesis would not have been possible.

I extend my special thanks to my family and family-in-law for making me feel their love and support all the time. Thank you, mom, for teaching me that the world is limitless!

Last but not least, I would like to thank my colleague, best friend, classmate and husband, Yasin. Thank you for tolerating your wife whining about her thesis. Thank you for believing in me and cheering me up whenever I was down. Thank you for all the gummy bears. I wouldn’t be writing this acknowledgement to ‘this’ thesis if it was not for you. Keep up being such a great husband, dissertation is loading.
ABSTRACT

This study investigated the effect of different delivery modes on L2 students’ listening comprehension. A mixed-methods explanatory research design was used to identify whether animated videos or context videos facilitated better listening comprehension. Fifty-seven ESL undergraduate/graduate students were recruited from an academic speaking and pronunciation course offered at Iowa State University. The effectiveness of the delivery modes was assessed by the participant’s ability to orally summarize the videos. Two trained raters were recruited to rate participants’ audio-recorded summaries. A one-way ANOVA was used to test for a significant difference between mean scores of the two groups. Also, students’ perceptions of the delivery modes were elicited through a survey and then analyzed thematically to gain a deeper insight into how ESL students viewed each delivery mode. Results showed that students in the animated video group outperformed those in the context video group. In addition, all students in the animated group preferred animated videos. However, one third of the students in the context group reported preferring audio-only listening since they did not find the context videos helpful.
CHAPTER 1. INTRODUCTION

This study aims to investigate the effect of visuals on L2 learners’ listening comprehension skills and their perceptions of different types of visuals in language learning. More specifically, this study explores the use of context videos (i.e., live-action videos) and animation videos to see which one facilitates better listening comprehension. Although the effect of visuals on students’ L2 listening comprehension has been explored from different perspectives (Coniam 2001; Ockey, 2007; Suvorov, 2008; Wagner 2010), the question of the quality and effectiveness of the visuals deserves a fresh look, specially considering the advances in technology that offers greater quality and richer visuals. Also, the effect of visuals has been mostly assessed through multiple-choice tests while this study is concerned with integrated listening tasks (i.e., listening and then speaking to produce an oral summary) because it helps to produce real-life language use (Lewkowicz, 1997). This kind of production in integrated tasks is closer to what the students would produce in academic contexts (Lee, 2015).

Statement of the Problem

As technology provides more multimedia language learning materials for foreign and second language classrooms, the issue of designing those materials effectively has gained more importance in Computer-Assisted Language Learning (CALL). As Chapelle (2003) argues, professionals in the field of applied linguistics face a new challenge that requires them to search for empirical evidence to design CALL applications more effectively for the use of language learning and teaching tasks. To address this challenge, this study seeks to understand what kind of visual support better facilitates second language (L2) learners’ listening comprehension.
Listening practice in L2 was mostly carried out through audio-only mode before technology provided more affordances. Thanks to advancements in technology, we can now witness more inclusion of visuals in listening activities and tests. When it comes to the types of visuals, there are different views. For example, Ockey (2007) suggests that the inclusion of visuals, video or still images, might lead to more authentic practice. Likewise, Bejar et al. (2000) and Ginther (2002) classify these visuals mainly as content and context visuals. While content visuals refer to the content of spoken stimulus, context visuals are related to setting that spoken stimulus takes place (Bejar et al., 2000). In addition to these descriptions, McCuistion (1991) regards them as static (e.g., pictures) and dynamic (live-action videos and animation) visuals. As can be inferred from these different types of visuals, the conceptualization of listening skills has shifted from solely ‘listening’ to also incorporating ‘viewing’ (Baltova, 1994). This new component of listening skill offers several perspectives for researchers interested in listening.

The existing body of literature on the use of visuals in L2 listening comprehension has so far yielded contradictory findings. Ockey (2007) suggests that L2 learners do not look at still images as much as they look at videos in a language test, and similarly, Wagner (2010) proposes that watching videos yields higher scores compared to audio-only mode. In contrast, Coniam (2001) and Rashasoor et al. (2016) argue that test-takers are distracted by the video, which they suggest causes lower scores. To add to these inconclusive findings, a different kind of video, animated videos, was also compared to other modes. Essentially, these are videos in which dynamic visuals are used to illustrate the content at the same time the speaker is discussing it. While it was found to lead more effective comprehension compared to audio-only mode (Aldera, 2015), adding captions to it led to a lower score on a vocabulary test compared to animation mode without captions (Aldera & Mohsen, 2013). Thus, referring to Chapelle’s (2003)
projection about the challenge that awaits professionals in the field of CALL, several studies call for research to gain more insights into the effectiveness of visuals in different settings (Aldera, 2015; Buck, 2001; Li, 2016; Suvorov, 2013; Wagner, 2010). However, considering that there is little research done on the extent to which animated videos are effective for listening comprehension, it is important to investigate animated videos as a unique mode of presenting and conveying information.

**Aims and Scope of the Study**

This thesis is concerned with the effect of visuals on students’ L2 listening comprehension. This effect is explored through context videos in which a lecturer talks about a topic that is supported with a single picture related to the topic, and animated videos in which the topic is given with dynamic images that are manipulated to appear as moving images by visualization artists. In addition, students’ perceptions of the two video forms are explored to gain a deeper understanding of what factors influence effectiveness.

To achieve this purpose, two videos about the topics of superstition and standardized testing were retrieved from [https://ed.ted.com/](https://ed.ted.com/), the website which hosts the videos, and remade into content videos using the same scripts. TED-Ed videos served as animated videos that provided rich content. For the context videos, two Ph.D. students who were native speakers of English acted as lecturers, presenting the content of the animated videos using the same script, but a single still image related to the topic was projected on a whiteboard. Data were collected through students’ audio recordings of oral summaries which were submitted to an online course on a learning management platform (i.e., Canvas). Students were also asked to complete an online survey probing their perceptions of the two video delivery modes. The data gathered in
the study were analyzed both quantitatively and qualitatively, and conclusions were drawn on the basis of the findings.

Structure of the Study

This study has five chapters. Following this introductory chapter, Chapter 2 presents a literature review and research questions. Chapter 3 explains the methods that were used in this study to design the task, collect the data and analyze it. It also provides detailed information about the participants, setting, and materials. Chapter 4 presents the results in the form of a statistical analysis of quantitative data and descriptive analyses of qualitative data. Chapter 5 discusses the results, limitations, implications, and suggestions for future studies.
CHAPTER 2. LITERATURE REVIEW

The purpose of Chapter 2 is to discuss theoretical perspectives that form the basis of the current study. This chapter begins with defining listening comprehension and presenting challenges that contribute to comprehension problems. Then, listening delivery modes are reviewed in the existing body of scholarship. Following delivery modes, theoretical frameworks underpinning this study are presented in the section on audiovisual processing. In the next section, studies about different listening delivery modes and students’ perceptions of them are reviewed. Chapter 2 concludes with a presentation of work about integrated listening and speaking tests.

Defining Listening Comprehension

Listening is a complex mental process that is of great importance for language learning. As Pichert (2002) underlines, listening comprehension is enabled through an interaction between the material and the learner. In other words, a listener should receive input (i.e., auditory signals) and parse it into small linguistic units and connect them to the ones in their long-term memory (Becker, 2016). Although this process might seem very straightforward, this is only one aspect of listening comprehension. In another dimension, these meaning interpretations can also be matched against a listener’s prior knowledge structures in existing schema (Buck, 2001). These two ways of interaction contribute to the complexity of listening comprehension.

In addition to this complicated process, listening comprehension is also regarded as difficult for second language (L2) learners for several reasons (Huang, 2005). As an example, L2 learners might struggle to understand the input received via listening because of delivery mode, low language proficiency, unknown vocabulary, or unfamiliar pronunciation (Qiu & Huang, 2012). Similarly, other challenges that affect L2 learners’ comprehension might be the speed of
presentation, topic unfamiliarity, lack of listening strategies, and different learning styles (Schmidt & Hegelheimer, 2004). These challenges might impact an L2 learner’s performance in an academic setting in several ways, such as poor communication with classmates or professors, failure in listening tests, or lack of active participation in discussions (Ghassemi, 2013). Therefore, the search for effective ways to improve L2 learners’ listening comprehension could generate valuable implications for language instruction.

With the aim of investigating ways to improve L2 writers’ listening comprehension, several scholars point out how advancements in technology potentially affect the nature of listening. For example, Wagner (2008) underlines that language teachers around the world started using movies, television shows, or any other online multimedia in the teaching of L2 listening. The affordances of videos allow L2 learners to process both the aural and visual input similar to real-life speaking situations. As can be inferred from this view, listeners are also expected to interpret non-verbal visual cues such as body movement, gestures, facial expressions, or any other cues that are available in the context (Buck, 2001; Ockey, 2007; Rubin, 1995). This change in the nature of the listening skill was also foreseen more than two decades ago by Baltova (1994). She argued that we do not only ‘listen’ in real-life situations, but we also ‘view’ what is being conveyed to us and make interpretations based on these two kinds of information. However, as Buck (2001) argues, the effect of visual information on L2 listeners’ comprehension might change from one situation to another, influencing the listener’s interpretation in a crucial way (p. 48). Therefore, it can be implied that if the listener’s background knowledge or expectations are not met by visual cues, communication might be hindered significantly. Another possibility way that visuals might hinder a listener’s comprehension could stem from the lack of connection between the message itself and the visual cues (Rubin, 1995) such as a single image of a lecturer talking about space. To summarize,
visual information in L2 listening could potentially help listeners establish a connection between what is being said and what is being seen and contribute to better interpretation.

**Listening Delivery Modes**

One way to understand how visual cues help L2 listeners is to explore different listening delivery modes. Videos that contain both visual and audio content could be effective in listening comprehension (Ockey, 2007; Wagner, 2010). Although L2 listening practice is viewed as something done through the audio mode in most settings, Ockey (2007) suggests that the inclusion of visuals, video or still images, might lead to a more authentic practice. Similar to Ockey (2007), Celce-Murcia (2002) suggests that using videos in classrooms facilitates authenticity and exposes learners to a variety of input sources by serving as an important motivator. Also, another advantage of using videos for language learning and teaching could be that it enhances learners’ comprehension since visual cues help them make connection between the visuals and how language is used (Harmer, 2001).

Although using visuals for listening comprehension is recommended by several researchers, it should be emphasized that there are two main types of visuals: context and content (Bejar et al., 2000; Ginther, 2002; Ockey, 2007). Context visuals are related to the setting in which the spoken stimulus takes place (Ginther, 2002). This can be exemplified as a series of still photos of the speaker and setting (Ginther, 2002, p. 134). To be more explicit, a professor giving a lecture in a classroom can be categorized as a context visual. Context visuals are also categorized into three different sub-types depending on the information they convey (Bejar et al., 2000). The first type is visuals that convey information about the setting (e.g., a visual of a classroom). The second type are those that provide information about the participants who
produced the oral input (e.g., a picture of the lecturer). The last type are visuals with information about the text type (e.g., a visual of a student giving a presentation).

As for the content visuals, Ginther (2002) identifies this type of visual as “a photo, graph or drawing that is related to the content of the verbal stimulus” (p.134). Similar to context visuals, Bejar et al. (2000) also categorize content videos into different types according to their connection to oral input. The first type are those that replicate the oral stimulus (e.g., a key word or phrase projected on the whiteboard that is explained in the oral stimulus). In this type, the listener sees exactly what is heard. The second type are visuals that illustrate the oral stimulus (e.g., a picture of a historical site in a history class). The listener views a visual of what is described orally in the stimulus. Another type are those that organize information in the stimulus (e.g., an outline of the main points of a lecture). The last type are those that supplement the oral stimulus (e.g., an image of a sample essay while the oral stimulus is about L2 learners’ writing process).

In addition to this distinction, visuals can also be categorized as single still images, a series of still images, and videos (Ockey, 2007). Similar to this type of classification, McCuistion (1991) refers to static and dynamic visuals; the former includes pictures and photographs, and the latter include live-action video and animation. Finally, Ginther (2002) adds drawings and diagrams to the list of different types of visuals.

Audiovisual Processing

According to cognitive load theory, when two types of input support the same information (e.g., video captions), the listener potentially suffer from excessive cognitive load
(Mayer & Moreno, 2003). To illustrate, Mayer et al. (2014) describe an L2 student who listens to an audio podcast which includes several unknown words. This student needs to spend cognitive resources to infer the meaning of the words from context. The researchers suggest that one way to support this learner in accessing word meaning is to add a redundant video that connects the visuals and narration so that s/he spends his/her sources in understanding the content of the video rather than the specific words. However, they also hypothesize that adding captions can create extra cognitive processing for this learner. To be more explicit, learners split their visual attention between the video content and the caption, which adds to their cognitive processing (Mayer et al., 2014, p.655).

Similar to cognitive load theory, Paivio (1971) proposes dual coding theory that suggests how these separate verbal and nonverbal inputs can be collectively processed. According to this theory, the verbal system consists of written, auditory, and articulatory codes while the nonverbal system encompasses images for almost everything. Both types of input (i.e., verbal and nonverbal) are distinct and triggered by mental representations associated with them. However, irrespective of how distinct they are, they can be joined by different connections as in an image evoking a word representation. While these connections can be made consciously or subconsciously, contextual factors influence them to a large extent. For example, the nonverbal system could be triggered by pictures shown to a listener and the production of more mental images can thus be facilitated. Paivio (1971) also underscores that the effect of contextual factors depends on individual’s past experiences, which means that showing pictures, for example, might not trigger the same representations for everybody. Thanks to its giving importance to contextual factors and individual differences, theoretical perspective underpinning the present study is dual coding theory.
Research on the Use of Visuals in Listening

Investigating the use of visuals in listening tasks has mostly taken place in the testing literature (Coniam 2001; Ockey, 2007; Suvorov, 2013; Wagner 2010). For example, Ockey (2007) explored the engagement of the test-takers with still images and videos in an academic computer-based test. He concluded that while still images were on the screen, test-takers did not seem to pay attention to them whereas they did pay attention to the videos. This finding implies that while developing a test task, deciding whether to use videos or still images is of great importance. Similar to Ockey (2007), Wagner (2010) compared the performance of L2 test-takers on an audio-only listening test and on another video-based listening test. He found that the video group scored significantly higher than the audio-only group. This finding and those in Ockey (2007) imply that videos in language tests draw more of test-takers’ attention and that non-verbal cues in video facilitate their performance more. However, use of videos in language testing have also been found to be problematic. For example, Coniam (2001) conducted a case study to find the perception of the teachers of different delivery modes and concluded that the videos are distracting for the test-takers since they have to go back and forth between the screen and the questions on the paper. He suggested that only audio might be used in listening tests. Finally, Suvorov (2013) created a test to reveal the effect of content versus context visuals on the comprehension scores. His analysis showed that there was no significant effect of visuals. He also reported that the participants were mostly distracted by the context visuals because of the body movements of the speaker. By contrast, the participants felt that the content videos were helpful to answer the test questions. These contradictory findings regarding using visuals in language tests can be explained in terms of the connections between verbal and nonverbal cues.
Since all these studies were implemented in a testing situation, test delivery modes, test anxiety, or time pressure might have affected the extent to which test-takers utilized the visuals.

In addition to testing situations, the use of different listening delivery modes has also been widely investigated in pedagogy related studies. However, it should be noted that similar to the aforementioned assessment studies, the pedagogical studies also generated contradictory findings regarding the effectiveness of different modes. For example, Li (2016) investigated whether a visual silent film might serve as an advance organizer when provided prior to full-film presentation. Specifically, he compared using visual-only silent film to activate learners’ prior knowledge before full audiovisual presentation with full audiovisual presentation only and a visual-only silent film presentation followed by audio-only narration. After each mode, the author used a multiple-choice listening test followed by a survey and a focus group discussion to measure the students’ comprehension. He found that the first mode facilitated listening comprehension most compared to other two modes. In contrast to Li (2016), Rashasoor et al. (2016) reported that the students in their study were more successful in audio-only mode than video mode.

Listening delivery modes have also been investigated from other perspectives. In a study involving mobile-assisted language learning (MALL), Chen and Chang (2011) compared single mode (i.e., auditory materials only) and a dual mode (i.e., audio and textual input simultaneously) to see which mode maximizes learning outcomes while minimizing cognitive load. They concluded that the presence of text makes the listening passage less difficult and more comprehensible.

Similar to live-action videos, using animation has also been of interest to some researchers. Aldera and Mohsen (2013) investigated the effectiveness of animated videos across
three modes: with captions and keyword annotation; captions only; without captions. Interestingly, they reported that the students who watched animations without caption outscored the other two groups. They explained this unexpected result from the perspective of cognitive load theory. In other words, since the animation-only group was not distracted by extra text presenting the same information in the audio channel, they comprehended the video better compared to other groups. Animated videos were also reported as a facilitator in foreign language teaching (Danan, 2004). For example, Heffernan (2005) reports that the participants in her study enjoyed the animation videos a lot, which led to a significant improvement listening comprehension skills. Finally, Similar to Heffernan (2005), the students in Abuzahra et al. (2016) also believe that animated cartoon films help them more improve what has been said in the videos. Since an appreciation for both live-action videos and animation by students have been reported in the literature, it would be valuable to determine which mode is actually effective in eliciting students’ listening comprehension skills.

However, given the little research done regarding the extent to which animation videos are effective in supporting listening comprehension, drawing conclusions prematurely about animation videos should be avoided. Also, as the studies reviewed above imply, no consensus has yet emerged with regards to which mode is most effective in facilitating listening comprehension. One explanation for the conflicting findings could be differences in research settings. As suggested by Paivio (1971), the quality of associative connections between verbal and nonverbal cues, contextual factors (e.g., mobile vs. computer, visual with captions versus without captions) and the proficiency level of the learners might have led to this inconclusiveness regarding different delivery modes.
Students’ Perceptions of Using Different Listening Delivery Modes

In addition to actual performance, there is also a need to approach this issue from students’ perspectives as well. With this in mind, Sulaiman, Muhammad, Ganapathy, Khairuddin, and Othman (2017) investigated how using video in listening assessment is perceived by EFL students. They looked at 150 EFL students’ pre-test and post-test results of the two tests which included the same set of questions but delivered in two different modes. They concluded that the students highly appreciated the videos for providing authentic, meaningful and real-life situations. While Sulaiman et al. (2017) approached this issue from an assessment perspective, using videos has also been reported as effective for pedagogical reasons (Sarani et al., 2014; Woottipong, 2014).

Integrated Listening-Speaking Tests

In order to accomplish a listening task, identifying the purpose is of utmost importance. Bejar et al. (2002) propose a list of purposes: a) listening for specific information; b) listening for basic comprehension; c) listening to learn; and d) listening to integrate information. Meanwhile, Becker (2016) classifies listening comprehension according to local and global targets. Integrated tasks are intended to measure more than one subskill. Ockey and Li (2015) suggest that developers of these tasks regard oral communication as two-way speech. They also describe these tasks as requiring a listener to provide an extended response after listening to an oral stimulus (e.g., videos) or reading a long text. In this type of task, a listener orally summarizes what s/he just listened to, watched, or read for a hypothetical audience (Ockey & Li, 2015, p.13). One example of this kind of test task is the TOEFL iBT speaking test, in which test-takers are presented with aural or textual input. Although this kind of summary-type task has recently attracted the interest of task developers, it is questionable whether these tasks assess interactional
competence (Ockey & Li, 2015) and whether working memory capacity affects a listener’s oral performance. However, these concerns might be overlooked depending on the purpose of the test.

As Lee (2015) suggests, the goal of these tasks is to measure a listener’s ability to appropriately use key ideas and details of the presented stimuli in their oral summary. For this reason, the present study used integrated tasks (i.e., listen and produce an oral summary) in order to elicit students’ listening comprehension. Although listening and speaking are two integrated skills, there is little research on assessing L2 listening comprehension through speaking. One example of such study is that of Lee (2015), who investigated how authentic integrated listening-speaking tasks are by examining test-takers’ views. Questionnaire responses revealed that the participants perceived the integrated listening-speaking tasks as more authentic. Apart from the perception of students, Frost et al. (2012) designed a study in which the participants listened to a stimulus and then presented an oral summary of what they heard. The researchers found that the accuracy of participants’ summaries distinguished the different levels of students.

Based on the contradictory findings of the studies reviewed in this chapter regarding the effectiveness of different listening delivery modes, this study was grounded in the following research questions:

1) To what extent do different listening delivery modes affect students’ L2 listening comprehension?

2) What are students’ perceptions of the effectiveness of the delivery mode?
CHAPTER 3. METHODOLOGY

The aim of this chapter is to present information about the methodology employed to investigate the effects of visuals on L2 learners’ listening comprehension skills. Specifically, this chapter begins with an introduction of the participants of the study and the setting in which the study took place. Then, a detailed description of materials used and the procedure that was followed for data collection is provided. Finally, this chapter explains the data analysis procedure to answer the research questions of the study.

A mixed method explanatory research design (Mackey & Gass, 2016) was used in this study to investigate the effects of different delivery modes on students’ L2 listening comprehension and the students’ perceptions of the effectiveness of the delivery mode. Adopting mixed method increases the validity of findings by helping to triangulate data by comparing quantitative and qualitative findings (Dörnyei, 2007).

Participants

The study was conducted with 57 non-native speakers of English who were enrolled in four ENGL 99S Academic Speaking and Pronunciation course at Iowa State University (ISU). Although 64 students signed the consent forms to participate in the study, seven of the participants had some technical problems during the data collection (such as turning off the recording button accidentally, not saving the recording properly, or not being able to upload the recording). Because of the aforementioned reasons, 57 students’ participation was recorded for this study. The majority of the students were aged between 19 to 21 years old ($M=22.1$, $SD=3.68$) and they were native speakers of Chinese (n=30). Out of 57 participants, 43 were males and 14 were females (see Appendix A). The study was classified as exempt by the
Institutional Review Board (see Appendix B) since the research was conducted in commonly accepted educational setting and data was obtained in a way that human subjects could not be identified.

**Setting**

Each of the ENGL 99S classes from which the participants of this study were recruited had 16-17 students enrolled. The students are placed in this class on the basis of their performance in the English Placement Test (EPT) Speaking test. This test is given at the beginning of the semester to the students whose native language is not English. The purpose of ENGL 99S is to develop students’ oral communication skills in English to be efficient speakers while communicating with their classmates, teachers, and university staff. The main objectives of the course are to develop pragmatic skills, listening strategies and skills, academic presentations skills, and oral communication skills (both in academic and non-academic situations) that are required to successfully complete the classes at the university.

Students who are taking ENGL 99S are evaluated based on their performance on the assignments (25%), discussions (20%), two discussion tests (30%), one final group presentation (15%) and attendance and preparation (10%). The class meets three times a week for 14 weeks, with each section lasting 50 minutes. All the course materials and assignment submissions are uploaded to Canvas, an online learning management system used at ISU. Canvas was also used in this study for data collection because of students’ familiarity with the platform.

**Materials**

With the aim of exploring the effects of visuals on students’ listening comprehension skills, four different videos were used in this study. Two of the videos were taken from the
website https://ed.ted.com/ which were classified as animated videos. The first animated video, *Should we get rid of standardized testing?* by Arlo Kempf (https://ed.ted.com/lessons/should-we-get-rid-of-standardized-testing-arlo-kempf), discussed whether standardized tests measure what test designers intend to measure (see screenshot in Figure 1). The video was 4.51 minutes long. The primary purpose of selecting this video was because of students’ familiarity with standardized tests in general. Although the participants had a variety of cultural backgrounds, it was assumed that they experienced standardized tests, such as TOEFL or IELTS, to some extent prior to their admission to ISU.

![Screenshot of the first animated video](image)

*Figure 1 Screenshot of the first animated video.*

The second video used in this study was about superstitions: *Where do superstitions come from?* by Stuart Vyse (https://ed.ted.com/lessons/where-do-superstitions-come-from-stuart-vyse) (see Figure 2). The video was 4.56 minutes long. This short lesson mainly provided information about specific origins of some of well-known superstitions, such as number 13 and knocking on
the wood. The rationale behind selecting this video was to bring variety to topics discussed to capture the interests of as many students as possible.

Figure 2 Screenshot of the second animated video.

The other two videos, which were classified as context videos, were created by two native speakers of English who were graduate students in the English department at ISU. With the aim of a better comparison of the two types of videos (i.e., animated vs. context), the scripts of those videos were exactly the same with the scripts of animated videos. The scripts were taken from www.YouTube.com (See Appendix C for scripts). For this purpose, the graduate students acted as a lecturer in a classroom and a picture related to the script was projected to the board (see Figure 3). The lecturers stood in front of the board and the scripts were shown to the lecturer by means of two laptops used as teleprompters. The lecturers read the scripts acting as giving a lecture in a classroom. The purpose of including two laptops were to help the lecturers to look at different directions and create a more authentic atmosphere that lecturers acted as if they were
looking at the students in a classroom. The videos were recorded using a professional camera and were edited using Windows Movie Maker. The video about standardized test was 5.06 minutes long and the video about superstitions was 5.04 minutes long. These videos can be reached through these links: [https://www.youtube.com/watch?v=Xu35_RaIoPM](https://www.youtube.com/watch?v=Xu35_RaIoPM) and [https://www.youtube.com/watch?v=qwaevX4vzP0](https://www.youtube.com/watch?v=qwaevX4vzP0).

![Figure 3 Screenshot of one of the context videos.](image)

In addition to the videos, a survey, which was adapted from Suvorov (2008), was created to reveal the students’ perception of the videos (see Appendix D). This survey was created on Google Forms and students were provided with the link to it on Canvas at the end of the listening activities. In addition to items about student’s background information (e.g., name, native language, gender, age, etc.), there were 13 questions which aimed to reveal students’ opinions about the usefulness of the visuals, problems, topic familiarity, and their preferences regarding each video. For yes/no questions, follow up questions were added to obtain detailed explanations.
Task

In this study, each student watched two videos (either context or animated), recorded their oral summary, submitted their audio files to each assignment, and completed the survey. First, the students started watching the first video which was about superstitions. After the video ended, the students opened the Voice Recorder application on the computers and started summarizing the video for three minutes. The students used the voice recording timer on the screen and the researcher walked around the classroom to make sure they did not exceed the allocated time. When they finished their summaries, they uploaded their audio recording files to Canvas. The same procedure was repeated for the second video which was about standardized testing. The average time students in animated video group spent on summarizing the videos is 2.00 minutes while the context video group summarized the video in 1.74 minutes in average. After the students finished watching both videos, they were instructed to click the survey link and submit their responses. Once a student competed these three steps of the study, they were allowed to leave the room.

Rubric

In order to assign scores to students’ performances, a task-specific four-point rating scale was designed to find how successfully they could summarize the videos. As Brookhart (2013) suggests, since the purpose of creating these rubrics is to reveal if the visuals helped students’ oral summaries, linguistic features (i.e., grammar, vocabulary, pronunciation and fluency) were not included in any of the rubrics.

Before creating the rubrics, the main ideas, detailed information, and conclusions were listed for each video. The video about superstitions had one main idea, six key details supported by two examples for each detailed information, and a conclusion (see Appendix E). In total, this
video had fourteen important points that students were expected to mention in their oral summaries. The second video, which was about standardized tests, had two main ideas, three key details supported by one example each, and four other key details without examples (see Appendix F). In total, this video had 12 key details which the students were expected to mention in their oral summaries. However, during the pre-analysis of the oral summaries, it was realized that only three students could mention all six key details with examples in the first video and only three students could mention all four key details without examples in the second video. For this reason, the points which were not mentioned by significant numbers of students were not taken into account (G. Ockey, personal communication, February 12, 2018). In the end, each video had 10 key details which were expected from students to mention in their oral summaries. Based on these adjustments, a four-point rating scale which was adapted from Frost et. el. (2012) was used to assess students’ summaries (see Table 1).

<table>
<thead>
<tr>
<th>Rating scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>4</strong></td>
</tr>
<tr>
<td><strong>3</strong></td>
</tr>
<tr>
<td><strong>2</strong></td>
</tr>
<tr>
<td><strong>1</strong></td>
</tr>
<tr>
<td>● Talks about the main idea skillfully</td>
</tr>
<tr>
<td>● Gives almost all of the details and examples</td>
</tr>
<tr>
<td>● Talks about conclusion skillfully</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rating Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>The researcher and a second rater were involved in the rating process. The second rater was a PhD student who was majoring in Applied Linguistics at ISU. He was familiarized with</td>
</tr>
</tbody>
</table>
the videos that the students watched and the rubric was explained in detail. The raters scored one of the samples obtained in the piloting session and discussed the rating process. Then, they started rating the summaries. The point a student received from rater one for the first video was added to the point the same student received from rater two. The same procedure was repeated for the second video and the average of these two scores was assigned to the student. The consistency among raters was determined by performing interrater reliability analysis using Cronbach’s Alpha ($\alpha = .852$).

**Procedures**

With the aim of explaining the study to the students and recruiting participants, two of the classes were visited prior to data collection and the other two classes were met only on the data collection day. The students who were willing to share their data were asked to read and sign the consent form (see Appendix G). Visiting the classes prior to study or meeting the students on the data collection day was organized depending on the lecturers’ choice.

Also, before data collection, the students in each section were randomized using SPSS to assign each to one of two groups. Those in Group 1 were assigned to watch the two animated videos and those in Group 2 were assigned the two context videos. Since each class was divided into two different groups, it would be impossible to show the videos to the whole class through a projector by the researcher. For this purpose, an online class, called ‘Listen and Speak Activity’ was created in the Canvas learning management system. Four different assignments were created on Canvas where the participants were supposed to watch videos and submit their recordings in audio format. Thanks to the feature of Canvas, the researcher could assign each student to different assignments based on their groups.
The data was collected at a computer lab in the department of English at ISU. The lab had 16 desktop computers (Windows) and 16 headsets were connected to the computers for the study. Prior to the data collection day, both the computers and the headsets were checked to see if they were functioning well. In order to make sure that the headsets were functioning well, a power point was created, with screenshots showing how to set up the headsets. Also, the researcher and another PhD student were available to help students with the technical issues during the data collection.

On data collection day, the students were asked to sit at any computer and log in. While they were doing so, the students were assigned to one of the groups and the course was made available for them. The purpose of doing this on the data collection day was because a notification would be sent to the students by Canvas when they were added to the course, which might lead to confusion. When all the students were ready and logged in to the computers, they were instructed about how to set up the headsets thanks to the PowerPoint presentation. They were also asked to find the built-in application ‘Voice Recording’ of Windows by using the ‘Search’ function on the computers. The students were told they would need to use this application while recording their summary after watching the video. After making sure each student set up the headsets, they were instructed to log in to their Canvas account and click on the course called ‘Listen and Speak Activity.’ It was made sure that they could see three items under the assignments on the course: Listen and Speak Test 1, Listen and Speak Test 2, and a survey (see Appendix H). They were also instructed that they should upload the recordings to Canvas by opening the file location of the recording and dragging and dropping it to the ‘Upload Media’ window which would be opened when they clicked on the ‘submit assignment’ after they watched the video. This procedure was repeated for all sections. After each section completed
the test, the recordings were downloaded to the researchers’ computer to store on CyBox, which is a secure file storage system available at ISU, and the students were removed from the Canvas course. This procedure was repeated for all sections.

**Data Analysis**

In this study, audio recordings of students’ orally summarizing the videos constituted the quantitative data providing evidence of the extent to which delivery mode facilitated more listening comprehension. Survey with open-ended questions constituted both quantitative and qualitative data by helping to gain insight into students’ perceptions of the delivery modes. Table 2 below illustrates an overview of data collection and analysis procedure.

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Data collection</th>
<th>Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ1.</td>
<td>Recordings of students’ oral summaries</td>
<td>One-way ANOVA</td>
</tr>
<tr>
<td>RQ2.</td>
<td>Survey</td>
<td>Descriptive statistics and thematic analysis</td>
</tr>
</tbody>
</table>

To address the effects of different delivery modes on students’ L2 comprehension (RQ1), a one-way ANOVA was conducted to determine if there was any significant difference between means of the scores of the students’ oral summaries in animated video group and the context video group (p<.05 level). Also, eta squared was calculated to find the effect size.

The second research question which addressed the students’ perceptions of the delivery modes was answered through descriptive and thematic analyses of survey responses. For the first two questions, which were the only five-point Likert scale items in the survey, and for two yes/no questions (Q3 and Q4), only percentages were reported. This procedure was repeated
separately for each group in the study. The rest of the questions were also yes/no questions but required an explanation. Those questions were categorized in terms of visuals, difficulties met during the task and the preference of the modes. In addition to percentages, common comments were reported.
CHAPTER 4. RESULTS AND DISCUSSION

Research Question 1

The first research question addressed the effect of listening delivery modes on students’ L2 listening comprehension. Table 3 provides descriptive statistics of the scores of test-takers in the animated and context video groups.

Table 3  Descriptive statistics for animated and context video groups

<table>
<thead>
<tr>
<th>Groups</th>
<th>Video 1*</th>
<th>Video 2*</th>
<th>Both videos</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Animated Video Group</td>
<td>27</td>
<td>5.59</td>
<td>1.69</td>
</tr>
<tr>
<td>Context Video Group</td>
<td>30</td>
<td>4.27</td>
<td>1.59</td>
</tr>
</tbody>
</table>

* Video 1: Superstitions, Video 2: Standardized Tests

The results in Table 3 shows that the mean for animated video group ($M=5.29$) is higher than the mean for context video group ($M=3.83$). Similarly, both for superstition video and standardized test video, means for animated video group ($M=5.59$ and $M=5.00$ respectively) are higher than the means for context video group ($M=4.27$ and $M=3.4$).

A one-way ANOVA test was conducted to determine if there was a significant difference between the means of the scores of the animated video group and context video group in this study. The results indicated that there is a significant difference between the mean scores of the animated video group and the context video group, $F (1,55) = 14.868$, $p = .000$. In addition to the one-way ANOVA test, eta squared was calculated to find out the effect size between the means of the groups. The effect size ($\eta^2 = .213$) showed that there is a large effect size (Mackey & Gass, 2016). This means that 21% of the score variance in overall performance of the students is due to students' groups. The results suggested that the students’ performance in the animated video...
group was significantly better than the students’ performance in the context video group. In addition to descriptive analysis of the groups in general, a one-way ANOVA was also conducted to determine the differences of the scores of the first video in both groups and the scores of the second video in both groups. The results indicated a significant difference between the mean scores of the animated video group and the context video group for superstitions video, $F(1,55) = 9.260$, $p = .004$ and for standardized testing video, $F(1,55) = 15.250$, $p = .000$. The effect sizes for both superstition video ($\eta^2 = .144$) and standardized tests video ($\eta^2 = .217$) were large. This means that 14% of the score variance obtained from the first video (i.e., superstition) resulted from students' being in either animated or context video group. Similarly, 21% of the differences in students' oral performances in the second video (standardized tests) is because of the students' groups.

According to the statistical analyses, the animated video group, which received rich visuals compared to context video group, did better at summarizing and being able to give more details. This means that the richness of the visuals might have an effect on students’ L2 listening comprehension (Danan, 2004; Heffernan, 2005). These findings appear to support the results of the previous research (Suvorov, 2013). Although Suvorov (2013) found no statistical difference in listening comprehension scores in both tests that used content and context visuals, the scores of content videos were slightly higher than the context videos. One potential explanation for the lack of a statistically significant difference has to do with the types of visuals used. Suvorov's content videos included an instructor and a still image in a classroom environment in the background while the context videos included only an instructor without any images. However, the higher mean scores of the content videos in Suvorov (2013) imply that richer visuals might
generate higher scores, which is an argument that is supported by the findings of the present study.

Another possible explanation may be related to the type of assessment format used to measure the learners’ ability to understand given input. In this study, a summary task was used in consideration of the integrated nature of listening and speaking skills (Frost et al., 2012; Lee, 2015; Ockey & Li, 2015). However, in previous studies, researchers often tended to isolate listening from speaking and to use multiple-choice test formats, which could lead learners to use guessing strategies (Carr, 2011). This issue might also raise questions about the validity of the results obtained in such studies. While there is little research on the difference between animated and non-animated videos, which could be considered context visuals, the present findings suggest that the degree of richness of visuals may play a role in the comprehension of listening input.

**Research Question 2**

The second research question addressed the participants’ perceptions of the two delivery modes using descriptive statistics and thematic analyses. Table 4 and 5 below present the statistical results for the five-point Likert scale items in the survey. Participants’ perceptions regarding the difficulty of the task were measured on a scale from very easy to very difficult. As seen in Table 4, a majority in both groups (58%) found the task difficult or very difficult while only 10% found the task easy. Considering each group separately, almost half of the participants (44%) in the animated video group found the task difficult, while the rest found it either normal or easy (37% and 19% respectively). As for the context video group, a large majority of the participants (70%) found the task difficult or very difficult whereas only 3% of the participants in this group found the task easy. No participants reported the task to be very easy. These results
indicate that the students who were exposed to rich visuals (i.e., the animated video group) experienced less difficulty compared to students who were not exposed to them (i.e., context video group).

Table 4 Survey responses regarding difficulty level of task

<table>
<thead>
<tr>
<th></th>
<th>Very Easy</th>
<th>Easy</th>
<th>Normal</th>
<th>Difficult</th>
<th>Very Difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both groups</td>
<td>0</td>
<td>6</td>
<td>18</td>
<td>28</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>10%</td>
<td>32%</td>
<td>49%</td>
<td>9%</td>
</tr>
<tr>
<td>Animated video</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>group</td>
<td>0.00%</td>
<td>19%</td>
<td>37%</td>
<td>44%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Context video</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>group</td>
<td>0.00%</td>
<td>3%</td>
<td>27%</td>
<td>53%</td>
<td>17%</td>
</tr>
</tbody>
</table>

The frequency with which students looked at the screen during the task was measured on a scale from never to all of the time. Results are presented in Table 5 below.

Table 5 Survey responses regarding the frequency of looking at the screen

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Most of the time</th>
<th>All of the time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both groups</td>
<td>0</td>
<td>5</td>
<td>13</td>
<td>23</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>0.00%</td>
<td>9%</td>
<td>23%</td>
<td>40%</td>
<td>28%</td>
</tr>
<tr>
<td>Animated video</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>group</td>
<td>0.00%</td>
<td>11%</td>
<td>15%</td>
<td>37%</td>
<td>37%</td>
</tr>
<tr>
<td>Context video</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>group</td>
<td>0.00%</td>
<td>7%</td>
<td>30%</td>
<td>43%</td>
<td>20%</td>
</tr>
</tbody>
</table>

All participants reported looking at the screen to some extent. A majority (68%) in both groups reported that they looked at the screen most of the time or all of the time. Looking at each
group separately, a large majority of participants (74%) in the animated video group reported looking at the screen all of the time or most of the time. Interestingly, 11% of students in this group indicated that they rarely looked at the screen. This finding is surprising considering that the visuals of the animated videos were richer than the context videos. On the other hand, a similar percentage (63%) in the context video group reported looking at the screen most of the time or all of the time. Since the purpose of this study was to study the effects of visuals, this is an important finding indicating that the students were engaged in looking at the visuals most of the time.

In addition to the Likert-scale items, two yes/no questions (Q4 and Q5) were posed to probe the participants’ familiarity with the topics addressed in the two videos. For the superstitions video, 43 students (74%) indicated that the topic was familiar to them while 14 (25%) said it was not. Similarly, for the standardized testing video, 34 students (60%) said that topic was familiar; however, more students (40%) compared to the first video reported that the topic was new to them. These findings are not surprising considering superstitions are part of every culture to some extent, and it is highly likely that students experienced standardized tests at least once in their lives. However, this does not mean that they were familiar with the specific content of the videos as well.

The survey also included open-ended questions that elicited more detailed information about the participants’ perceptions of the delivery modes. The questions were categorized under the following topics: visuals, task difficulty, and delivery-mode preference.
Visuals

a) Whether the visuals help to understand the topic

Regarding the extent to which visuals helped the participants in both groups to understand the topics, for the superstitions video, 47 students (83%) reported that the visuals helped while only 10 (17%) said they did not utilize the visuals. All 27 students (100%) in the animated video group reported that visuals helped them to understand the topic. One of the students in the animated video group who found the videos helpful mentioned:

The video presented visual images that helped understand some of the superstitions origins.

However, a student in context video group commented:

In the video, there is just the visuals of the speaker. It might be better if it has visual regarding the information the speaker provide (Context Group).

For the standardized testing video, 39 students (68%) in both groups indicated that the visuals helped them while 18 (32%) said that they did not. Similar to the first video, 25 students (97%) in the animated group reported that the visuals helped them in this video as well. These findings regarding the animated video group also align with the one-way ANOVA results showing that the animated video group outperformed the context video group. Example comments about the helpfulness of videos from both groups included:

I didn't have a clear idea about standardized test before. But the way it was presented to us, is really good. I like this video. (Animated Group)

While watching the videos, without too many slides it makes me easier to focus the place I should focus. (Context Group)

However, a student in the context video group who did not find the videos helpful mentioned:

There were no visuals. Just a speaker that did not look at the camera.
b) Whether the visuals help to remember and speak more

Regarding the extent to which the visuals help students remember the content of the videos while summarizing, 39 students (68%) indicated that visuals helped them while 18 (32%) pointed out they did not utilize visuals much. All but one student in animated group reported that visuals helped them remember and thus summarize better. A supporting comment regarding the animated videos was:

Although I did not talk much, most of what I said was based on the parts of the video that were fresh in my mind.

Another important finding regarding this issue is that more than half of the students (56%) in context video group complained about the lack of visuals that could facilitate their performance in the oral summary. One representative example said:

There is not much visual that could help to remember.

Overall, these finding also align with the students’ higher average scores in the animated group compared to context video group for both videos.

c) Whether visuals were distracting

Regarding whether any of the visuals were distracting, 39 students (68%) in both groups said the visuals were not distracting while 18 (32%) reported the visuals to be distracting. In the animated group, six students (22%) reported the visuals to be distracting. For example, a student in animated group commented:

Some picture is so funny that I was absorbed by it instead of listening what speaker said.

On the other hand, more than half of the students in the context video group (60%) reported finding the visuals distracting. For example:

The lecturers’ mimics distracted me.
This finding is interesting since, as the comment above indicates, although there is not much action in the context videos compared to animated videos, consistent movement of a single speaker could divert students’ attention, resulting in the loss of interest in the lecture content.

**Task Difficulty**

Regarding the difficulties that students in both groups experienced with the task, eight out of 57 students (14%) mentioned the ‘no note-taking’ rule, 19 (33%) mentioned the challenges of summarizing, and 14 (24%) mentioned remembering what they had watched. Example comments included:

There is not much hard thing about the task but if we were allowed to take notes, I would have spoken even more than what I did.

For me, hardest thing is to remember and summarize. There are many information in the video. So, I have to remember a lot more things at a time. But the visuals actually help us to summarize.

Remembering all the important points of the video is difficult.

Although the problems pointed out by the students are important to discuss as limitations of the study, it should also be noted that few students experienced the same difficulty during the task.

**Delivery-mode Preferences**

The last topic of the survey was about the students’ preferences regarding delivery mode. Forty-seven students (82%) preferred having an accompanying video while 10 (18%) said they would prefer audio only. Out of 47 students who preferred having video, 26 were in the animated video group and 21 were in the context video group. Out of 10 students who preferred audio-only, only one student was in the animated video group. Example comments included:

I would choose video. Because it will help me to understand what the person talk about. (Context video)
I prefer video listening. Because in video listening, I can see the visuals which actually help me to remember. (Animated video)

Only listening help me concentrate. (Context video)

In both groups, the majority of students were in favor of using videos in listening tasks over audio only. This implies that students were making use of verbal and non-verbal features of both kinds of videos.

It is notable that some students in the context video group reported finding the task very difficult while no students in animated video group reported the same perception. This might be because of the effect of the visuals on understanding the topic. In terms of the source of their difficulties, the participants mentioned the “no note-taking” rule; however, since the aim of this study was to investigate the effect of visuals, note taking may have distracted the students from looking at the screen. In addition, it should be noted that Bloomfield et al. (2010) and Rubin (1994) proposed that note-taking might be one of the factors that increases the complexity of listening, thus hindering comprehension. In terms of visuals’ facilitating students’ understanding, the more familiar the students were with the topic, the less they reported finding the visuals helpful. In the present study, the average scores of the students in both groups were closer to each other in the case of the superstition video ($M = 5.59$ vs. $M = 4.27$) compared to the second video ($M = 5.00$ vs. $M = 3.4$). Since most students (74%) also reported a familiarity with the topic of superstitions, this raises the question of topic familiarity. If the students were already familiar with the topic, it would not matter much if they were in the animated video group or context video group. However, this shows that inclusion of animated videos into the curricula, particularly while teaching unfamiliar topics, would help student understand better since visuals would give them a chance to establish a connection between what is being said and what is being seen, thus contributing to more accurate interpretations.
Although it was hypothesized that visuals would help students understand better, some students reported the visuals to be distracting (Ginther, 2002). Some in the context video group mentioned that they felt uncomfortable because the speaker was not making eye contact in the video. Such comments reveal potential differences in students’ learning styles (Dunn et al., 2009). This is also supported by the survey responses regarding students’ delivery-mode preferences. Although many preferred to have visual support, some preferred audio-only since they could concentrate better on listening. Thus, the results show that the richness of visuals can enhance some students’ L2 listening skills while causing other students to be distracted. Still, many students in both groups preferred videos to audio-only, which might be considered an indicator of the need to integrate visuals into L2 listening instruction.

The differences in student perceptions could also be explained by dual coding theory, which posits that information is coded in our minds either verbally (i.e., text and sounds) or non-verbally (i.e., picture and objects) (Paivio, 1971). These two systems trigger each other when the input is received by one system and activated by the other system. Therefore, in some cases, visuals might not be effective enough to activate both systems and make the most of the cues, which results in distraction during the task or underappreciation of one mode.
CHAPTER 5. CONCLUSION

As presented in the previous chapter, it was found that the richness of visuals affects students’ understanding of the video, which was the tool to assess students’ L2 listening comprehension. While the students who received the L2 listening task through animated videos, which were rich in terms of visuals, received higher scores from the task, the students who received the L2 listening task through context videos, which included only the instructor and a still image in terms of visuals, received lower scores from the task. Also, the students in the animated video group reported the visuals to be helpful whereas the students in the context video group did not find visuals facilitative for their listening comprehension. Based on the findings, the last chapter of this thesis starts with implications, followed by the limitations of the study, and ends with the suggestions for the future research on L2 listening delivery modes.

Implications

This study investigated the effect of video delivery mode on students’ L2 listening comprehension. It also explored students’ perceptions of the delivery modes. Based on the obtained results, there are several implications. First, since the ability to receive high scores from the task was based on being able to give more details of what the students listened to, materials designers and language teachers who are teaching L2 listening courses might consider supporting their material with rich visuals especially if they intend to focus on listening for details in their courses.

Second, although multiple-choice test formats have been largely favored for measuring listening comprehension in many classroom-based assessment contexts, the summary task in this study appears to provide a viable alternative assessment method. However, practitioners should
be cautious that students’ lack of speaking skill might prevent them from displaying listening comprehension which might raise potential validity concerns. This testing format is suggested to be implemented especially when listening and speaking are taught as integrated skills.

Third, as the survey revealed, some students liked the videos while others reported them to be distracting. This raises the issue of individual differences. Task designers and instructors should take into account students’ different learning styles and learning preferences. In this study, there were only a few students who mentioned the visuals to be distracting; however, the more diverse the learning environment, the more instructors might need to implement different kinds of tasks.

Fourth, using rich visuals such as animated videos could also be considered in content-based instruction. To be more explicit, since the students in such settings are expected to fully grasp the content of the course or catch specific details, animated videos could offer invaluable affordances for both instructors and students. Similarly, instructors teaching an English for Specific course (e.g., English for computer engineering, medicine, or agriculture) could potentially make use of using animated videos to teach specific course contents.

Finally, topic familiarity should be taken into account when designing tasks. As the present study revealed, the average scores of the students received for the video ‘Where do superstitions come from?’ were closer to each other in both groups. This was also supported by the students’ comments on the survey about the familiarity of the topics. Considering the fact that this topic is about general knowledge, the more familiar it is to students, the less they will need the support of visuals. In addition to basic familiarity with the topic of superstitions, this video also use examples from Chinese culture. As mentioned in Chapter 3, most of the participants were Chinese; even if they did not understand the whole thing from the videos, they
might have summarized the video based on pre-existing cultural knowledge. These facts suggest there might be value in language instructors adding rich visuals to listening tasks addressing unfamiliar topics. In addition to topic familiarity, instructors in ESL situations should also take into account students’ varying cultural backgrounds when selecting topics and listening texts.

**Limitations**

Some limitations must be borne in mind in interpreting the findings of this study. First, as mentioned in the survey results, study design did not allow note-taking during the listening tasks. It might be argued that those students who did poorly on the task did so not because of low listening ability but because they had forgotten some of the video content, despite having understood it.

The second limitation has to do with the integrated, listen-and-speak nature of the tasks. It is possible that students might have felt uncomfortable speaking into a microphone, which could have hindered the performance of some who nevertheless understand everything perfectly. Therefore, it is difficult to attribute poor performance solely to poor listening skills. It should also be noted that this study was administered in the third week of semester when some students might not yet have adjusted to the environment and become comfortable with their production skills.

A third limitation relates to the fact that only two videos were shown to each group. Considering the fact that students were mostly familiar with one of the videos, and that that video had examples from Asian culture, it would be better to have more than two videos and/videos that were known to be unfamiliar, in order to find out the effect of visuals more
clearly. However, since the data was collected during class, there was not enough to include more than two videos.

A final limitation is that, despite being instructed to talk for a maximum of three minutes, some students provided summaries that were longer than three minutes. This may have led to an unfairly positive evaluation of their summaries by virtue of having had more time in which to provide details.

**Suggestions for Future Research**

The findings suggest several possible focuses for future research. First, based on the students’ delivery-mode preferences, audio-only mode could be included in future studies to see to what extent animated videos affect the L2 listening comprehension compared to a no-visual condition. Although audio-only mode was already compared to content or context visuals, including animated videos which are rich in terms of visuals could yield to a deeper understanding of the effectiveness or distraction level of animated videos compared to audio-only mode. Second, future studies should find ways to control time on the summary tasks to provide equal opportunities for participants to demonstrate understanding. Third, future studies might consider including a learning-preferences survey to better interpret student perception data and explore additional potential reasons for getting high or low scores. In this study, a single link to the survey was provided for both groups, which later required the researcher to separate survey responses for each group (e.g., animated video) manually. A more practical alternative would be to provide a different link to the same survey for each group. Fourth, students’ engagement duration with the visuals was revealed based on their responses in the survey. Including eye-tracking would answer this question more reliably. Lastly, future studies should take into account the topics and the vocabulary students learned at their current level when
deciding on the content of the videos, especially considering that some students had difficulties with vocabulary during the tasks in the present study.
REFERENCES


Lee, Hye Won, "Innovative assessment tasks for academic English proficiency: an integrated listening-speaking task vs. a multimedia-mediated speaking task" (2015). Graduate Theses and Dissertations. 14584.


## APPENDIX A. PARTICIPANTS’ BIODATA

<table>
<thead>
<tr>
<th>Student #</th>
<th>Age</th>
<th>Gender</th>
<th>Native Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>Male</td>
<td>Arabic</td>
</tr>
<tr>
<td>2</td>
<td>31</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>Male</td>
<td>Korean</td>
</tr>
<tr>
<td>4</td>
<td>19</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>Male</td>
<td>Malay</td>
</tr>
<tr>
<td>7</td>
<td>21</td>
<td>Male</td>
<td>Indian</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>12</td>
<td>19</td>
<td>Male</td>
<td>Arabic</td>
</tr>
<tr>
<td>13</td>
<td>25</td>
<td>Male</td>
<td>China</td>
</tr>
<tr>
<td>14</td>
<td>21</td>
<td>Female</td>
<td>Malay</td>
</tr>
<tr>
<td>15</td>
<td>24</td>
<td>Female</td>
<td>Chinese</td>
</tr>
<tr>
<td>16</td>
<td>29</td>
<td>Female</td>
<td>Farsi</td>
</tr>
<tr>
<td>17</td>
<td>18</td>
<td>Male</td>
<td>Malay</td>
</tr>
<tr>
<td>18</td>
<td>21</td>
<td>Male</td>
<td>Malay</td>
</tr>
<tr>
<td>19</td>
<td>23</td>
<td>Male</td>
<td>Indian</td>
</tr>
<tr>
<td>20</td>
<td>19</td>
<td>Female</td>
<td>Local Congo language</td>
</tr>
<tr>
<td>21</td>
<td>19</td>
<td>Female</td>
<td>Chinese</td>
</tr>
<tr>
<td>22</td>
<td>20</td>
<td>Female</td>
<td>Chinese</td>
</tr>
<tr>
<td>23</td>
<td>26</td>
<td>Male</td>
<td>Vietnamese</td>
</tr>
<tr>
<td>24</td>
<td>21</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>25</td>
<td>22</td>
<td>Male</td>
<td>Indian</td>
</tr>
<tr>
<td>26</td>
<td>22</td>
<td>Female</td>
<td>Chinese</td>
</tr>
<tr>
<td>27</td>
<td>19</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>28</td>
<td>20</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>29</td>
<td>26</td>
<td>Female</td>
<td>Korean</td>
</tr>
<tr>
<td>30</td>
<td>29</td>
<td>Male</td>
<td>Portuguese</td>
</tr>
<tr>
<td>31</td>
<td>24</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>32</td>
<td>22</td>
<td>Female</td>
<td>Chinese</td>
</tr>
<tr>
<td>33</td>
<td>18</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>34</td>
<td>19</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>35</td>
<td>18</td>
<td>Male</td>
<td>Turkish</td>
</tr>
<tr>
<td>36</td>
<td>28</td>
<td>Male</td>
<td>Spanish</td>
</tr>
<tr>
<td>37</td>
<td>23</td>
<td>Male</td>
<td>Malay</td>
</tr>
<tr>
<td>38</td>
<td>19</td>
<td>Male</td>
<td>Taiwanese</td>
</tr>
<tr>
<td>39</td>
<td>21</td>
<td>Male</td>
<td>Malay</td>
</tr>
<tr>
<td>40</td>
<td>22</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>41</td>
<td>26</td>
<td>Female</td>
<td>Chinese</td>
</tr>
<tr>
<td>42</td>
<td>20</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>No.</td>
<td>Age</td>
<td>Gender</td>
<td>Ethnicity</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>--------</td>
<td>-----------</td>
</tr>
<tr>
<td>43</td>
<td>22</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>44</td>
<td>19</td>
<td>Male</td>
<td>Arabic</td>
</tr>
<tr>
<td>45</td>
<td>23</td>
<td>Female</td>
<td>Taiwanese</td>
</tr>
<tr>
<td>46</td>
<td>20</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>47</td>
<td>28</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>48</td>
<td>24</td>
<td>Female</td>
<td>Korean</td>
</tr>
<tr>
<td>49</td>
<td>19</td>
<td>Female</td>
<td>Malay</td>
</tr>
<tr>
<td>50</td>
<td>19</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>51</td>
<td>20</td>
<td>Male</td>
<td>Malay</td>
</tr>
<tr>
<td>52</td>
<td>24</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>53</td>
<td>24</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>54</td>
<td>36</td>
<td>Male</td>
<td>Korean</td>
</tr>
<tr>
<td>55</td>
<td>18</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>56</td>
<td>22</td>
<td>Male</td>
<td>Chinese</td>
</tr>
<tr>
<td>57</td>
<td>20</td>
<td>Female</td>
<td>Chinese</td>
</tr>
</tbody>
</table>
APPENDIX B. IRB APPROVAL

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Date: 12/8/2017
To: Leyla Karatay
2348 Edenburn Dr Apt D
Ames, IA

CC: Dr. James Ranalli
319 Ross Hall
Yasin Karatay
2348 Edenburn Dr Apt D

From: Office for Responsible Research

Title: The effect of listening delivery modes on listening comprehension

IRB ID: 17-570

Study Review Date: 12/8/2017

The project referenced above has been declared exempt from the requirements of the human subject protections regulations as described in 45 CFR 46.101(b) because it meets the following federal requirements for exemption:

• (1) Research conducted in established or commonly accepted education settings involving normal education practices, such as:
  — Research on regular and special education instructional strategies; or
  — Research on the effectiveness of, or the comparison among, instructional techniques, curricula, or classroom management methods.

• (2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey or interview procedures with adults or observation of public behavior where
  — Information obtained is recorded in such a manner that human subjects cannot be identified directly or through identifiers linked to the subjects; or
  — Any disclosure of the human subjects' responses outside the research could not reasonably place the subject at risk of criminal or civil liability or be damaging to their financial standing, employability, or reputation.

The determination of exemption means that:

• You do not need to submit an application for annual continuing review.

• You must carry out the research as described in the IRB application. Review by IRB staff is required prior to implementing modifications that may change the exempt status of the research. In general, review is required for any modifications to the research procedures (e.g., method of data collection, nature or scope of information to be collected, changes in confidentiality measures, etc.), modifications that result in the inclusion of participants from vulnerable populations, and/or any change that may increase the risk or discomfort to participants. Changes to key personnel must also be approved. The purpose of review is to determine if the project still meets the federal criteria for exemption.

Non-exempt research is subject to many regulatory requirements that must be addressed prior to implementation of the study. Conducting non-exempt research without IRB review and approval may constitute non-compliance with federal regulations and/or academic misconduct according to ISU policy.

Detailed information about requirements for submission of modifications can be found on the Exempt Study Modification Form. A Personnel Change Form may be submitted when the only modification involves changes in study staff. If it is determined that exemption is no longer warranted, then an Application for Approval of Research Involving Humans Form will need to be submitted and approved before proceeding with data collection.

Please note that you must submit all research involving human participants for review. Only the IRB or designees may make the determination of exemption, even if you conduct a study in the future that is exactly like this study.
Video 1: Where do superstition come from?

Are you afraid of black cats? Would you open an umbrella indoors? And how do you feel about the number thirteen? Whether or not you believe in them, you're probably familiar with a few of these superstitions. So how did it happen that people all over the world, knock on wood, or avoid stepping on sidewalk cracks? Well, although they have no basis in science, many of these weirdly specific beliefs and practices do have equally weird and specific origins. Because they involve supernatural causes, it's no surprise that many superstitions are based in religion. For example, the number thirteen was associated with the biblical Last Supper, where Jesus Christ dined with his twelve disciples just before being arrested and crucified. The resulting idea that having thirteen people at a table was bad luck eventually expanded into thirteen being an unlucky number in general. Now, this fear of the number thirteen, called triskaidekaphobia, is so common that many buildings around the world skip the thirteenth floor, with the numbers going straight from twelve to fourteen. Of course, many people consider the story of the Last Supper to be true but other superstitions come from religious traditions that few people believe in or even remember. Knocking on wood is thought to come from the folklore of the ancient Indo-Europeans or possibly people who predated them who believed that trees were home to various spirits. Touching a tree would invoke the protection or blessing of the spirit within. And somehow, this tradition survived long after belief in these spirits had faded away. Many superstitions common today in countries from Russia to Ireland are thought to be remnants of the pagan religions that Christianity replaced. But not all superstitions are religious. Some are just based on unfortunate coincidences and associations. For example, many Italians fear the number 17 because the Roman numeral x v ii can be rearranged to form the word veecee, meaning my life had ended. Similarly, the word for the number four sounds almost identical to the word for death in Cantonese, as well as languages like Japanese and Korean that have borrowed Chinese numerals. And since the number one also sounds like the word for must, the number fourteen sounds like the phrase “must die”. That's a lot of numbers for elevators and international hotels to avoid. And believe it or not, some superstitions actually make sense, or at least they did until we forgot their original purpose. For example, theater scenery used to consist of large painted backdrops, raised and lowered by stagehands who would whistle to signal each other. Absentminded whistles from other people could cause an accident. But the taboo against whistling backstage still exists today, long after the stagehands started using radio headsets. Along the same lines, lighting three cigarettes from the same match really could cause bad luck if you were a soldier in a foxhole where keeping a match lit too long could draw attention from an enemy sniper. Most smokers no longer have to worry about snipers, but the superstition lives on. So why do people cling to these bits of forgotten religions, coincidences, and outdated advice? Aren't they being totally irrational? Well, yes, but for many people, superstitions are based more on cultural habit than conscious belief. After all, no one is born knowing to avoid walking under ladders or whistling indoors, but if you grow up being told by your family to avoid these things, chances are they'll make you uncomfortable, even after you logically understand that nothing bad will happen. And since doing something like knocking on wood doesn't require much effort, following the superstition is often easier than consciously resisting it. Besides, superstitions often do seem to work. Maybe you remember hitting a home run while
wearing your lucky socks. This is just our psychological bias at work. You're far less likely to remember all the times you struck out while wearing the same socks. But believing that they work could actually make you play better by giving you the illusion of having greater control over events. So, in situations where that confidence can make a difference, like sports, those crazy superstitions might not be so crazy after all.

Video 2: Should we get rid of standardized testing?

The first standardized tests that we know of were administered in China over 2,000 years ago, during the Han dynasty. Chinese officials used them to determine aptitude for various government posts. The subject matter included philosophy, farming, and even military tactics. Standardized tests continued to be used around the world for the next two millennia, and today, they're used for everything from evaluating stair climbs for firefighters in France to language examinations for diplomats in Canada to students in schools. Some standardized tests measure scores only in relation to the results of other test-takers. Others measure performances on how well test-takers meet predetermined criteria. So, the stair climb for the firefighter could be measured by comparing the time of the climb to that of all other firefighters. This might be expressed in what many call a bell curve. Or it could be evaluated with reference to set criteria, such as carrying a certain amount of weight a certain distance up a certain number of stairs. Similarly, the diplomat might be measured against other test-taking diplomats, or against a set of fixed criteria, which demonstrate different levels of language proficiency. And all of these results can be expressed using something called a percentile. If a diplomat is in the 70th percentile, 70% of test-takers scored below her. If she scored in the 30th percentile, 70% of test-takers scored above her. Although standardized tests are sometimes controversial, they're simply a tool. As a thought experiment, think of a standardized test as a ruler. A ruler's usefulness depends on two things. First, the job we ask it to do. Our ruler can't measure the temperature outside or how loud someone is singing. Second, the ruler's usefulness depends on its design. Say you need to measure the circumference of an orange. Our ruler measures length, which is the right quantity, but it hasn't been designed with the flexibility required for the task at hand. So, if standardized tests are given the wrong job, or aren't designed properly, they may end up measuring the wrong things. In the case of schools, students with test anxiety may have trouble performing their best on a standardized test, not because they don't know the answers, but because they're feeling too nervous to share what they've learned. Students with reading challenges may struggle with the wording of a math problem, so their test results may better reflect their literacy rather than numeracy skills. And students who were confused by examples on tests that contain unfamiliar cultural references may do poorly, telling us more about the test-taker's cultural familiarity than their academic learning. In these cases, the tests may need to be designed differently. Standardized tests can also have a hard time measuring abstract characteristics or skills, such as creativity, critical thinking, and collaboration. If we design a test poorly, or ask it to do the wrong job, or a job it's not very good at, the results may not be reliable or valid. Reliability and validity are two critical ideas for understanding standardized tests. To understand the difference between them, we can use the metaphor of two broken thermometers. An unreliable thermometer gives you a different reading each time you take your temperature, and the reliable but invalid thermometer is consistently ten degrees too hot. Validity also depends on accurate interpretations.
of results. If people say results of a test mean something they don't, that test may have a validity problem. Just as we wouldn't expect a ruler to tell us how much an elephant weighs, or what it had for breakfast, we can't expect standardized tests alone to reliably tell us how smart someone is, how diplomats will handle a tough situation, or how brave a firefighter might turn out to be. So, standardized tests may help us learn a little about a lot of people in a short time, but they usually can't tell us a lot about a single person. Many social scientists worry about test scores resulting in sweeping and often negative changes for test-takers, sometimes with long-term life consequences. We can't blame the tests, though. It's up to us to use the right tests for the right jobs, and to interpret results appropriately.
APPENDIX D. SURVEY

Name (first, last): Nationality: Age: Gender:
1. Overall, how would you describe the difficulty level of the task: Very easy ___ Easy ___ Normal ___ Difficult ___ Very difficult ___
2. How often did you look at the computer screen when listening? Never ___ Rarely ___ Sometimes ___ Most of the time ___ All the time ___
3. Did you have problems understanding the speakers? Yes ___ No ___
   Please explain: _____________________________________________
4. Was some information from the first lecture (Where do superstitions come from?) familiar to you? Yes ___ No___
5. Was some information from the second lecture (Should we get rid of standardized tests?) familiar to you? Yes ___ No___
6. Did the visuals in the first video (Where do superstitions come from?) help you better understand the topic? Yes ___ No ___
   Please explain: _____________________________________________
7. Did the visuals in the second video (Should we get rid of standardized tests?) help you better understand? Yes ___ No ___
   Please explain: _____________________________________________
8. Did the visual information (i.e. pictures/lecturers or animation) ever distract you from listening? Yes ___ No ___
   Please explain: _____________________________________________
9. Did the quality of the video affect your understanding of the speakers? Yes ___ No ___
   Please explain: _____________________________________________
10. Did the visuals in the video help you remember and speak more? Yes ___ No ___
    Please explain: _____________________________________________
11. Was time enough for summarizing? Yes __ No __
    Please explain: _____________________________________________
12. What was the hardest thing about the task?
13. If you were to choose between video and audio-only listening, which one would you prefer?
Any other comments: _____________________________________________
## APPENDIX E. KEY DETAILS IN THE SUPERSTITIONS VIDEO

<table>
<thead>
<tr>
<th>ITEM LIST</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main idea</strong></td>
<td>Origins of superstitions</td>
</tr>
<tr>
<td><strong>Kinds of superstitions (example and origin)</strong></td>
<td><strong>RELIGIOUS</strong></td>
</tr>
<tr>
<td><strong>Examples</strong></td>
<td><strong>Origins</strong></td>
</tr>
<tr>
<td>Number 13</td>
<td>13 people at last supper when Jesus was arrested and crucified.</td>
</tr>
<tr>
<td>Knocking on wood</td>
<td>Based on Indio-European belief that trees are home to spirits and knocking on the wood invoke the spirit.</td>
</tr>
<tr>
<td><strong>NON-RELIGIOUS</strong></td>
<td></td>
</tr>
<tr>
<td>Number 17</td>
<td>Means your life ended when Roman numeral is rearranged</td>
</tr>
<tr>
<td>Number 4</td>
<td>Sounds almost identical to the word for death in Asian cultures</td>
</tr>
<tr>
<td><strong>MAKE SENSE</strong></td>
<td></td>
</tr>
<tr>
<td>Whistling at the theatre</td>
<td>Whistling was used to communicate at backstage and caused some accidents in the past, so today no whistling at the theatre</td>
</tr>
<tr>
<td>Lighting cigarette from the same match</td>
<td>In a fox hole, lighting a match for long draw attention so lighting cigarette from the same match is bad luck</td>
</tr>
<tr>
<td><strong>Conclusion</strong></td>
<td>Superstitions are based on cultural habits—we are not born with it—It is easier to believe than to avoid</td>
</tr>
</tbody>
</table>
### APPENDIX F. KEY DETAILS IN THE STANDARDIZED TESTING VIDEO

<table>
<thead>
<tr>
<th>ITEM LIST</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main ideas</strong></td>
<td>Tests are not enough</td>
<td>Tests should be designed properly</td>
</tr>
<tr>
<td><strong>History of standardized tests</strong></td>
<td><strong>Past</strong></td>
<td><strong>Today</strong></td>
</tr>
<tr>
<td></td>
<td>In China for government</td>
<td>For students-diplomats-job interviews</td>
</tr>
<tr>
<td><strong>Types of Tests</strong></td>
<td>Bell curve-Comparison</td>
<td>Criterion based- Previously set criteria</td>
</tr>
<tr>
<td><strong>Issues of Standardized Tests</strong></td>
<td><strong>Problem</strong></td>
<td><strong>Example</strong></td>
</tr>
<tr>
<td></td>
<td>Design problem</td>
<td>Ruler can’t measure circumference of an orange</td>
</tr>
<tr>
<td></td>
<td>Test anxiety</td>
<td>Feeling too nervous to share the knowledge</td>
</tr>
<tr>
<td></td>
<td>Validity- Reliability</td>
<td>Thermometer giving different reading each time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>How brave a firefighter is</td>
</tr>
</tbody>
</table>
INFORMED CONSENT DOCUMENT

Title of Study: The effect of listening delivery modes on listening comprehension

Investigator: Leyla Karatay

This form describes a research project. It has information to help you decide whether or not you wish to participate. Research studies include only people who choose to take part—your participation is completely voluntary. Please discuss any questions you have about the study or about this form with the project staff before deciding to participate.

Introduction

The purpose of this study is to investigate the effect of different delivery modes on students’ second language learners’ listening comprehension. More specifically, this study aims to identify whether different type of videos (animated vs. content) facilitate more listening comprehension skill.

You are being invited to participate in this study because you are second language learners who are enrolled in ENGL 99S Academic Speaking and Pronunciation course. However, you should not participate if you are under age of 18.

Description of Procedures

If you agree to participate, you will be asked to watch two different videos on a familiar topic and summarize them orally. First, you will be enrolled in a Canvas course that was created for this research by the researcher. When you log in Canvas, you will be asked to open the course titled as ‘Listen and Speak Test’. As the first step, you will watch a 4-minute video while taking notes if you want. Then, you will be asked to summarize what you have just watched orally. You
will record your speech on Canvas and submit it. You will be asked to do the same procedure for the second video. At the end of the test, you will fill out an online questionnaire about the test.

Your participation will last for about a class hour (i.e., 50 minutes)

**Risks or Discomforts**

While participating in this study you may experience the following risks or discomforts:

There are no known risks/discomforts in this study.

**Benefits**

If you decide to participate in this study, there may be direct benefit to you. It is hoped that the information gained in this study will benefit society by improving the second language instruction primarily at Iowa State University and any second language instruction setting. This study will also provide you an opportunity to practice your listening and speaking skills.

**Costs and Compensation**

You will not have any costs from participating in this study. You will not be compensated for participating in this study.

**Participant Rights**

Participating in this study is completely voluntary. You may choose not to take part in the study or to stop participating at any time, for any reason, without penalty or negative consequences.

Your choice of whether or not to participate will have no impact on you as a student in any way.
If you have any questions about the rights of research subjects or research-related injury, please contact the IRB Administrator, (515) 294-4566, IRB@iastate.edu, or Director, (515) 294-3115.

**Confidentiality**

Records identifying participants will be kept confidential to the extent permitted by applicable laws and regulations and will not be made publicly available. However, federal government regulatory agencies, auditing departments of Iowa State University, and the Institutional Review Board (a committee that reviews and approves human subject research studies) may inspect and/or copy study records for quality assurance and data analysis. These records may contain private information.

To ensure confidentiality to the extent permitted by law, the following measures will be taken:

All the information about your identities will be kept confidential when the results of the study are disseminated. Your course teacher will not be notified about your performance on the test.

**Questions**

You are encouraged to ask questions at any time during this study. For further information about the study, contact: Leyla Karatay, lkaratay@iastate.edu

**Consent and Authorization Provisions**

Your signature indicates that you voluntarily agree to participate in this study, that the study has been explained to you, that you have been given the time to read the document, and that your questions have been satisfactorily answered. You will receive a copy of the written informed consent prior to your participation in the study.
Participant’s Name (printed) ____________________________________________

_________________________________________  _________________________
Participant’s Signature                      Date
APPENDIX H. SCREENSHOTS OF THE TASK IN CANVAS

APPENDIX H. SCREENSHOTS OF THE TASK IN CANVAS