

An Exploratory Study of Effects of Lecture's Acoustic Features on L1 Chinese
Listeners' Listening Comprehension of Online L2 English Lectures

Jingjing Huang

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in
partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Curriculum and Instruction

Kenneth Potter, Chair
Katherine Cennamo
Barbara Lockee
Alicia Johnson

April 22, 2021

Blacksburg, Virginia

Keywords: Speech rate, duration, listening comprehension, L1 Chinese learner, L2 English
online academic listening

An Exploratory Study of Effects of Lecture's Acoustic Features on L1 Chinese Listeners'
Listening Comprehension of Online L2 English Lectures

Jingjing Huang

ABSTRACT

This study explored the effects of lecture speech rate and duration on L1 Chinese listeners' listening comprehension of online L2 English lectures on free-recall and cued-recall questions. There are many factors that may affect a Chinese listener's listening comprehension of online English lectures. The current study focused specifically on the effects of speech rate and duration, and explored the possibility of identifying tools that listeners might use to control these factors during online lectures in order to achieve better learning experiences. Using an exploratory quasi-experimental approach, this study explored the research questions in three phases: 1) An initial quasi-experiment was conducted via Qualtrics and applied as a pilot study; 2) The quasi-experiment was revised, with a new link sent out to potential participants from both the United States and China; and 3) Quasi-experiment responses were collected and analyzed. Based on 28 participant responses, the findings from the current study suggest that the interaction between speech rate and duration impacted L1 Chinese listener's comprehension of online L2 English lectures. The findings also suggest that study participants' responses on various types of questions were affected differently by the speech rate and duration of the lectures.

An Exploratory Study of Effects of Lecture's Acoustic Features on L1 Chinese Listeners'
Listening Comprehension of Online L2 English Lectures

Jingjing Huang

GENERAL AUDIENCE ABSTRACT

It is assumed that the slower and the shorter the listening material is, the better the foreign language learner's listening comprehension will be. Is this true? This study explored how a lecture's speech rate and duration may affect L1 Chinese listeners' listening comprehension of online L2 English lectures on free-recall and cued-recall questions. It discusses the possibility of identifying tools for listeners to control the speech rate and duration of online lectures in order to help listeners achieve better learning experiences. Using an exploratory quasi-experimental approach, this study explored the research questions in three phases: 1) An initial quasi-experiment was conducted via Qualtrics and applied as a pilot study; 2) The quasi-experiment was revised, with a new link sent out to potential participants from both the United States and China; and 3) Quasi-experiment responses were collected and analyzed. Based on 28 participant responses, the findings from the current study suggest that speech rate and duration worked together to impact L1 Chinese listeners' comprehension of online L2 English lectures. The findings also suggest that study participants' responses on various types of questions were affected differently by the speech rate and duration of the lectures.

ACKNOWLEDGEMENTS

First of all, I'd like to thank my family. Thank you to my husband, Dr. Zhang, who has supported my decision to start graduate school again, ten years later, and has never complained. Thank you to my daughters, Moonsie and Lusie, who have always loved me and believed that Mom could make it. Thank you to my parents, Dr. Huang and Mrs. Xiao, who have encouraged me to fulfill my dreams whenever it's possible, and thank you to my brother, Dr. Huang, who has offered me precious ideas and support during my study.

I also want to thank my committee for their support. Thank you Dr. Potter, for patiently guiding me through all the exams and writings. Thank you Dr. Johnson, for your suggestions and encouragement in and out of all the courses. Dr. Cennamo, I have enjoyed your courses, and I am encouraged by your great attitude towards the profession. Finally, Dr. Lockee, though I didn't have the chance to take your courses, I have really appreciated your fierce support during my journey through the VT IDT program.

Special thanks to Dr. Kibong Song, who created and revised the Qualtrics quasi-experiment for the current study. Without Dr. Song's help, I would not have been able to carry out the experiments for my study.

There are many others, such as my classmates, the professors and GAs from the EDRE program, the librarians, and others who have participated in my study and helped me out. Thank you to my friends in SOE who have encouraged me, supported me, and worked with me. It has been my pleasure to join the VT SOE family and spend time with the wonderful people there.

TABLE OF CONTENTS

ABSTRACT	II
GENERAL AUDIENCE ABSTRACT	III
ACKNOWLEDGEMENTS	IV
TABLE OF CONTENTS	V
TABLES	VIII
FIGURES	VIII
CHAPTER ONE.....	1
INTRODUCTION	1
STATEMENT OF THE PROBLEM.....	3
PURPOSE OF THE STUDY	5
ORGANIZATION OF THE STUDY	6
CHAPTER TWO.....	7
REVIEW OF THE LITERATURE	7
NNS LISTENING PROCESS	7
<i>Listener's Memory in Listening Process</i>	11
<i>Cognitive Load in NNS Listening Process</i>	13
<i>Automatic Processing and Controlled Process</i>	15
<i>Summary</i>	16
SPEECH RATE, SPEECH DURATION, AND NNS LISTENING COMPREHENSION.....	16
<i>Speech Rate in Listening Comprehension</i>	17
<i>Duration of Speech in Listening Comprehension</i>	23
<i>Summary</i>	24
NNS LEARNERS' LISTENING COMPREHENSION IN ONLINE L2 ENGLISH ACADEMIC LISTENING.....	25
<i>Listening Difficulty for L2 English Learners</i>	26
<i>L1 Chinese Listeners' Difficulties in L2 English Academic Listening</i>	27
<i>Difficulties in Online Academic Listening</i>	29
<i>L2 English Learners' Listening Strategies</i>	29
<i>Summary</i>	32
ASSESSING LISTENING COMPREHENSION.....	33
<i>Summary</i>	37
SPEECH RATE AND DURATION IN LISTENING COMPREHENSION IN INSTRUCTIONAL DESIGN AND TECHNOLOGY	38
<i>Summary</i>	39
SUMMARY OF LITERATURE REVIEW	39
RESEARCH QUESTIONS.....	39
CHAPTER THREE.....	41
METHODOLOGY	41
STUDY DESIGN	42

DATA COLLECTION METHOD	42
PILOT STUDY	45
PARTICIPANTS	46
DATA COLLECTION PROCEDURE	48
DATA CODING	49
CHAPTER FOUR	51
RESULTS AND FINDINGS	51
FINDINGS FROM THE RESPONSES	51
<i>Findings - Types of Questions</i>	52
Free-recall Questions.....	52
SPSS Results	53
Long Duration	57
Medium Duration	57
Short Duration	57
Cued-recall Questions.....	58
SPSS Results	58
Slow Speech Rate.....	62
Normal Speech Rate.....	63
Fast Speech Rate	63
SUMMARY	64
CHAPTER FIVE	65
DISCUSSION AND CONCLUSION	65
DISCUSSION OF FINDINGS	66
<i>Effect of Speech Rate on Listening Comprehension Performance</i>	67
<i>Effect of Duration on Listening Comprehension Performance</i>	68
<i>Interaction of Speech Rate and Duration</i>	69
SUMMARY	71
PRACTICAL IMPLICATIONS.....	72
<i>Online Instructors</i>	72
<i>Instructional Designers for Online Lectures</i>	73
STUDY LIMITATIONS	73
FUTURE RESEARCH	75
APPENDICE	92
APPENDIX A: QUASI-EXPERIMENT QUESTIONS	92
APPENDIX B: VT IRB-20-514 APPROVAL LETTER	94
APPENDIX C: RESPONSES TO BACKGROUND QUESTIONS.....	95
RATING RESULTS FROM SAMPLE QUESTIONS BY TWO RATERS AND THE NEGOTIATED RATE	96
<i>Rater A</i>	97
<i>Rater B</i>	101
<i>Negotiated</i>	103
<i>Calculation of Scores for Respondents</i>	104

TABLES

Table 1: Parameters of Audios in the Study	43
Table 2: Participants' Location and Occupation	47
Table 3: Participants' English Learning Experience, Academic and Geographic Information...	47
Table 4: Participants' Self-evaluated English Proficiency	48
Table 5: Descriptive Statistics for Scores from Free-recall Questions	53
Table 6: Within-subjects Effects for Scores from Free-recall Questions	54
Table 7: Pairwise Comparison for Speech Rate - Free-recall Questions	55
Table 8: Pairwise Comparisons for Duration - Free-recall Questions	55
Table 9: Descriptive Statistics for Scores from Cued-recall Questions	59
Table 10: Within-Subjects Effects for Scores from Cued-recall Questions	59
Table 11: Pairwise Comparisons for Speech Rate - Cued-recall Questions	60
Table 12: Pairwise Comparisons for Duration - Cued-recall Questions	61

FIGURES

Figure 1: Plots – Free-recall Questions	56
Figure 2: Plots - Cued-recall Questions	61

CHAPTER ONE

Introduction

As humans with normal listening ability, we listen every day, and we listen to various types of sounds. While we may not normally pay attention to our listening, there are times when we notice that listening is important. A good example of this is when we listen to things that we are not familiar with, such as a new language.

The process of listening is to pay attention to, and get meaning from what is heard (Underwood, 1989). Listening is, “an active process in which listeners select and interpret information which comes from auditory and visual clues in order to define what is going on and what the speakers are trying to express” (Rubin, 1995, p. 7). Listening is, “both a critical communication competency and a critical life competency” (Worthington & Fitch-Hauser, 2018, p. 25). “The most widely agreed view is” (Gruba & Suvorov, 2020, p. 2) that, “listening is the process of receiving, constructing meaning from, and responding to spoken and/or non-verbal messages” (International Listening Association [ILA] 1995, p. 4).

For foreign language (FL) or second language (SL) learners, listening is a skill that both learners and teachers need to develop because, “as essential part of communicative competence, listening is the skill that deserves” attention (Flowerdew & Miller, 2005, p. xi). In this study, listening is defined as, “uni-directional listening where the person seeking comprehension does not interact with the speaker(s), as when listening to a lecture, radio broadcast or to an announcement” (Graham & Santos, 2015, p. 1). In uni-directional listening, a listener acts as an addressee “who is being spoken to directly and who has limited rights to response” (Rost, 1990, p. 5). A typical example of an addressee listener is a student who is sitting in classrooms and listening to lectures.

Listening comprehension is, “a very complex process” (Buck, 2001, p. 1) which includes various types of knowledge, such as the input to the listener, knowledge of the language used, world knowledge, context knowledge for communication, and mental representations of meaning (Buck, 2001). With the development of cognitive theories, researchers have argued that listeners are not passive receivers of information, but active interpreters working on, “inference (contextualizing those signals by a hearer)” (Rost, 1990, p. 3). There are many factors that affect listening comprehension (Lynch, 1998), such as the individual listener characteristics of language proficiency and strategies, the individual speaker characteristics of speech rate and accents, the text characteristics of text length, type, and complexity, and task characteristics (Gruba and Suvorov, 2020, p. 2).

Academic listening, in contrast to conversational listening, “has its own distinctive features” (Flowerdew, 1994, p. 7). Academic listening, “involves attending to and comprehension of spoken texts in academic settings, such as lectures, tutorials, small group discussions, and seminars” (Goh, 2018, p. 1). Most commonly, academic listening is used in English for Speakers of Other Languages (ESOL) “context to refer to a special set of skills required for processing academic spoken discourse by second language users” (Goh, 2018, p. 1). Because “within the field of academic study, ... the lecture remains the central instructional activity” (Flowerdew, 1994, p. 1), in the current study, when lecture listening is applied it means academic listening (AL), and vice versa.

Estimates of the total number of speakers and learners worldwide vary. For example, one estimate claims there are 1.2 billion English speakers, including both native speakers (NS) (L1 English speakers) and non-native (NNS) speakers (L2 English speakers) (Yaghi, 2019). According to British Council member John Knagg, “there are 1.5 billion English learners

worldwide” (Beare, 2019). Used as a lingua franca, English has been the one of the main means of communication to teach, learn, and use among speakers of various native languages (Nordquist, 2019). The British Council forecasts that by 2020, “two billion people will be using or learning to use English” (British Council, 2013, p. 2).

English language has been listed as one of the subjects included in the Chinese College Entrance Examination since 1978. From 1987, standard English tests have been adopted into university education in China. Nowadays, English is also one of the main subjects taught to K-12 Chinese students. As many as 93.8% of foreign language learners in China have studied English, which makes the number of English learners in China more than 390 million. “English has been the most popular foreign language in China” (Wei & Su, 2012, p. 11).

Statement of the Problem

The listening comprehension of English as second language (L2 English) learners in regards to academic content has been explored in many fields. For example, researchers such as Flowerdew (1994a and b), Flowerdew and Miller (2005), and Goh (2018) have contributed their thoughts about L2 English academic listening from the perspectives of Linguistics and second language acquisition (SLA).

In the field of instructional design and technology (IDT), Thompson and Rubin (1996) focused on listening strategies and studied, “the effect of both cognitive and metacognitive strategy instruction on listening comprehension performance in Russian” (p. 333). Jones and Plass (2002) studied the effectiveness of multimedia learning environments on L2 French listening comprehension, and concluded that the L1 English participants performed better with both written and pictorial annotation while listening to 2’20” historical account in L2 French. In his dissertation about the management of flow of control in computer-assisted listening

comprehension tasks for second and foreign language learners, Hoven (1997) proposed an “instructional design model for a humanistic integrated multimedia Computer-Enhanced Language Learning (CELL) package for self-access language learning.”

Although research about listening and listening comprehension has occurred with increasing frequency, it is not enough to cover the “multiple factors that affect listening comprehension” (Gruba & Suvorov, 2020, p. 2). Recognizing this situation, researchers from Linguistics and SLA, such as Foulke and Sticht (1969), Sticht (1971), Carver (1973), Derwing (1990), Griffiths (1990a, 1992), Zhao (1997), and Le (2006) have explored the relationship between speech rate and duration of lectures, and the performance of L2 learners on tests of listening comprehension. Scholars in IDT have focused on various areas like listening strategy instruction (Thompson & Rubin, 1996), listening materials construction (Ramírez Verdugo & Alonso Belmonte, 2007), listeners’ anxiety and experience in various environments (Ginther, 2002; MELANLIOĞLU, 2013), Computer-Assisted Language Learning (CALL) or CELL environments (Hoven, 1997; Guichon & McLornan, 2008), IDT tools in EFL settings (Mirzaei, 2017), and others.

With the blooming of distance learning and online learning, online L2 English academic activities have become increasingly involved in English for academic purpose (EAP) research (Bensoussan, Avinor, Ben-Israel, & Bogdanov, 2006; Agnesia, 2010; Asoodar, Atai, Vaezi, & Marandi, 2014). However, many new questions are arising. For example, how might listeners’ performance be affected if they could adjust lecture speed and duration when joining online lectures? How should we, as instructional designers, based on research about the relationship between L2 English listening comprehension and the factors of lecture speed and duration, design online lectures and tools to better promote learning?

This exploratory study has investigated the effects of various speech rates and durations of L2 English online lectures on L1 Chinese listeners' performance. The study aims to provide grounding for online lecture instructors and/or designers to determine whether and how to utilize instructional tools that include options for lecture speed and duration for L2 English online lecture listeners.

Purpose of the Study

The purpose of the current study was to explore the effects of speed rate and duration of online L2 English lectures on the listening comprehension of L1 Chinese listeners on free-recall and cued-recall questions. With the findings from this exploratory study, instructional designers for online lectures will be able to identify and employ appropriate lecture tools to promote lecture listening by L2 English learners. In addition to this, online lecturers who have L2 English learners may benefit by choosing the appropriate lecture speed and duration for their learners.

With this purpose in mind, the following research questions were addressed in the study:

1. What is the effect of variations in speech rate on listening comprehension on a test of free-recall?
2. What is the effect of variations in speech rate on listening comprehension on a test of cued-recall?
3. What is the effect of variations in duration on listening comprehension on a test of free-recall?
4. What is the effect of variations in duration on listening comprehension on a test of cued-recall?

Organization of the Study

Chapter 2 begins with a literature review of related topics such as the NNS listening process, effects of lecture speech rate and duration on L2 English listening comprehension, NNS learners' listening comprehension of L2 English academic listening, and assessment in listening comprehension. Chapter 3 details the proposed methodology to employee the answers to the research questions including study design, data collection methods, pilot study, participants, data collection procedures and data coding, and limitations. In Chapter 4, findings from the experiments are presented and the data is analyzed through various methods. Chapter 5 focuses on a discussion of the findings presented in Chapter 4, answers to the research questions, practical implications and limitations of the study, and potential directions for future research.

CHAPTER TWO

Review of the Literature

NNS Listening Process

Listening is a complex process, more complex than just using auditory organs to catch sounds and interpret the meanings. “Nowadays, listening is conceived of as an active process in which listeners select and interpret information which comes from auditory and visual clues in order to define what is going on and what the speakers are try to express” (Mendelsohn & Rubin, 1995, p. 7). For NNS speakers, comprehending information that is presented in a second language is more complex because more variables may affect listening comprehension (Rost, 1990; Rubin, 1995; Buck, 2001; Graham & Santos, 2015).

Nagle and Santer (1986) are among the scholars who have described the listening comprehension process as the way in which listeners take in acoustic signals, and then produce an understanding of the text. This can be thought of as a type of flow chart with acoustic signals entering from one end and comprehension produced at the other (Buck, 2001). Generally, listeners need to deal with several “most important knowledge sources involved in listening comprehension” (Buck, 1995, p. 117) which include: 1) linguistic knowledge, including knowledge of phonology, lexicon, syntax, morphology and discourse features of cohesive markers, various registers, and others; 2) co-text knowledge about what has already been said; 3) knowledge about the context of the situation where the communication happens, such as private talking with family or a formal job interview; and 4) general world knowledge (Buck, 1995).

With so much information in mind and limited time to process the information, how or what do listeners need to do to keep the processing smooth and functional? There are “two most important views” (Buck, 2001, p. 2) about how the knowledge of the listeners is applied to the

incoming information: the bottom-up view, and the top-down view. However, some researchers argue that there are actually three “most widely known” models to explain the listening process: the bottom-up model, the top-down model, and the interactive model (Flowerdew & Miller, 2005, p. 24).

The bottom-up view, or text-based processing, of listening states that all the knowledge “operate in a fixed serial order” (Buck, 1995, p. 118). Under this view, listeners first deal with the phonological information, then attend to the lexical meaning, then work with the syntactic knowledge. Finally, listeners combine the context and general world knowledge to understand the semantic meanings of the information (Buck, 1995). Listeners understand the input by starting with the smallest units, then progressing to bigger and bigger ones until they are able to combine individual sentences to create “ideas and concepts and relationships between them” (Flowerdew & Miller, 2005, p. 24). Listeners, “need to decode the linguistic input rapidly and accurately, and to map the input against these expectation to confirm consistencies and to refute implausible interpretation” (Tsui & Fullilove, 1998, p. 433). In listening comprehension, the bottom-up process occurs, “when the listener focuses on sounds, phonemes and parts of the words heard” (Graham & Santos, 2015, p. xii).

The top-down process, or knowledge based processing, requires listeners to, “use their pre-existing knowledge to interpret the text and to create plausible expectations of what they are about to hear” (Tsui & Fullilove, 1998, p. 433), and emphasizes, “the use of previous knowledge in processing a text rather than relying upon the individual sounds and words” (Flowerdew & Miller, 2005, p. 24). The top-down model argues that listeners do not rely solely on the acoustic signals to interpret a verbal text, but that pre-established knowledge works in the process as well. Such previous knowledge can be referred to as, “*schema, frame, script, and scenario*, although

schema is often used as a cover term” (italic originally, Flowerdew & Miller, 2005, p. 25). Buck (1995) argues that, “there is continual top-down influence on language processing” (p. 118). Expectations based on previous knowledge influence listeners all the time and throughout the process, with all kinds of knowledge interacting at all levels. Listeners synthesize every piece of knowledge possible to generate their understandings of the information (Buck, 1995).

If listeners synthesize both bottom-up and top-down processes, they are using the interactive model, as Rumelhart (1975) argues. Under the interactive model, “language is processed simultaneously at different levels... phonological, syntactic, semantic, and pragmatic information interact, although it is not clear exactly how” (Flowerdew & Miller, 2005, p. 26).

Some researchers believe that non-native speaker (NNS) listeners at different proficiencies prefer different processing methods in listening comprehension. Shohamy and Inbar (1991) did a study to investigate the effects of both text and question types on subjects’ scores on listening comprehension tests administered to 150 EFL senior middle school learners. They found out that, “while high level listeners seemed to process the text in knowledge-based manner, the low level test takers seemed to process the text in a data-driven manner” (p. 35). However, there are researchers who believe that the processing methods NNS listeners apply in order to deal with the difficulties in second language listening may be influenced by something else. In disagreement with the view that less effective/beginning level L2 listeners rely heavily upon bottom-up information processing, Field (2004) suggests that, “listening to a foreign language may be assisted by an interactive-compensatory mechanism already available in L1, which compensates for gaps in understanding” (p. 363).

Non-native speakers (NNS) are speakers, under certain circumstances, who are not speaking their native languages as native speakers (NS) do. For example, most of the

international students in the United States are NNS speakers of English except for those who come from English-speaking countries. For NNSs of English, English is the second language (L2 English) or foreign language (FL).

NNS listening comprehension involves skills on its own (Flowerdew, 1994a), which can be grouped into two sections: “real-time processing and phonological and lexico-grammatical features” (Flowerdew, 1994a, p. 10).

The feature of real-time processing in academic listening works when old information that is sent out by the speaker(s) disappears as new information emerges. Listeners have little control over, “the speed with which they process the text, nor can they backtrack to make sure that they have understood correctly.” (Flowerdew & Miller, 2005, p. 51). There are many factors that can influence the effect of such real-time processing characteristics on NNS listeners, such as the length of the lecture, the speaker(s)’s speed in lecturing, the subject/topic of the lecture, the listeners’ proficiency level with the language, and others.

Another challenge for NNS listeners in the listening process is dealing with the phonological and lexico-grammatical features in the language. NNS listeners must deal with the phonological features such as the unit boundaries, speaker pauses or hesitations, stresses, intonations, and accents. Meanwhile, NNS listeners have to recognize words, phrases, idioms, and grammar to get the information (Flowerdew, 1994a). All these things must be processed within a very short period of time, because new information is coming in for processing.

L2 academic listening comprehension also has its own unique features compared to listening in general. First of all, L2 academic listening requires a certain degree of background knowledge. Listeners in academic listening need to have some knowledge of the target subject. Secondly, listeners in academic listening need to be able to pick out what is relevant from the

information stream. For example, within a 45-minute lecture, the lecturer might use the first ten minutes to warm up and do attendance, and leave the last five minutes for questions and answers. He or she may tell jokes or refer to popular topics during the remaining thirty minutes to keep the listener focused, motivated, or relaxed. Listeners need to distinguish between this less relevant information and what is relevant to the main topic of the class. For NNS listeners, this ability is difficult but important to acquire. Thirdly, in academic listening, listeners will have limited opportunities to have turn-taking conventions unless allowed by the lecturer. Fourthly, academic listening focuses more on the information to be passed on, but not the interpersonal or illocutionary meaning (Flowerdew, 1994a).

Listener's Memory in Listening Process

One of the features of listening is that it involves real-time processing. Listeners have limited time to process one piece of information before another piece comes in. There is only one chance unless playback is possible. In lecture listening, this unique characteristic is one of the difficulties that NNS listeners encounter. The discussions about the difficulties that NNS listeners face are presented in a later section.

If we look at the real-time feature of listening from another perspective, what is important becomes the ways that listeners deal with the vanishing information before processing it. From previous discussion, it is known that once listeners receive incoming information, they need to process the information at various levels regardless of the approach they may follow.

Ehrlich (1982) performed reading comprehension tests with four groups to investigate the relationship between comprehension and memorization for a test. She found that there was, “a strong relationship exists between comprehension and memorization of a text” (Griffiths, 1990a,

p. 314). Though Ehrlich's (1982) study is about reading comprehension, her results are still useful for us to look at when discussing memory in listening comprehension.

Atkinson and Shiffrin (1968) divide memory into, "three structural components: the sensory register, the short-term store, and the long-term store" (p. 90). In the listening process, sensory register or sensory memory is also called echoic memory (Neisser, 1967) or pre-categorical acoustic storage (Crowder & Morton, 1969), "though a boundary between purely sensory retention (which involves no processing) and short-term retention (in which items are subject to processing) has been difficult to establish" (Nagle & Sanders, 1986, p. 15).

Incoming sensory information first enters the sensory register, where it resides for a very brief period of time, then decays and is lost. The short term store is the subject's working memory; it receives selected inputs from the sensory register and also from long-term store. Information in the short-term store decays completely and is lost within a period of about 30 seconds, but a control process called rehearsal can maintain a limited amount of information in this store as long as the subject desires. The long-term store is a fairly permanent repository for information, information which is transferred from the short-term store (Atkinson & Shiffrin, 1968, p. 90-91).

Though Cowan (2017) points out that the terms of short and long-term stores by Atkinson and Shiffrin, "were theoretical mechanisms not to be equated with short and long-term memory" (p. 1167), and argues that "working memory ... a system of temporary memory that is multifaceted" (Cowan, 2014, p. 201), the Atkinson/Shiffrin model of human memory as a flow chart still works as one approach to display how human memory works.

Underwood (1989) also agrees that, "there are three distinct stages in the aural reception of an utterance" (p. 2). These three stages are the echoic memory, the short-term memory, and

the long-term memory. Sounds enter the echoic memory, but remain for a very short time (lasting about a second). NNS listeners often make errors when they attempt to create meaningful units based on these sounds before new information goes into the echoic memory. The short-term memory is where the information is processed within a few seconds. “At this point, words or groups of words are checked and compared with information already held in the long-term memory and the meaning is extracted from them. Once the meaning has been grasped, the actual words are generally forgotten.” (Underwood, 1989, p. 2). At this stage, the time for processing is still limited, and the speed of processing is important because new information is coming. If listeners fail to process the first chunk of information successfully, confusion will appear, and the processing system will be overloaded. Low-skilled NNS listeners who are at a low proficiency level for the language often cannot process information fast enough at this stage because they need a relatively longer time to search and match what is heard with what is stored. Most of the difficulties in L2 listening happen here as well. Once the listeners have constructed meanings from the information, they may store the information into the long-term memory for later.

Cognitive Load in NNS Listening Process

In previous sections of the literature review, there is one factor that is mentioned often and implicitly in reference to both general listening and listeners’ memory. This factor is time. In real-time listening, the listeners have very limited time to process the current incoming information before new information is added for processing. This requires effective processing abilities in listeners, and brings difficulties for NNS listeners especially.

The fact that listeners have to use their background knowledge to process a piece of information within a few seconds leads us to consider listeners’ cognitive load during the

listening process. Cognitive load, or cognitive capacity or workload, refers to a person's cognitive abilities in regards to problem-solving skills, expert-novice distinctions, working memory, categorizing problems, and others (Sweller, 1988).

“Cognitive load theory is concerned with working memory limitations and strategies to overcome those limitations” (Cottam & Savenye, 2014, p. 84, from Sweller 1999, 2005). Sweller's (1988) cognitive load theory states that, “people have a working memory (WM) capacity that is accessible at any given point in time ... consists of 7 ± 2 spaces” (Wagner-Loera, 2018, p. 170) and, “is very limited in capacity (about seven items can be held in working memory at a time) and duration (items remain for only a few seconds)” (Cottam & Savenye, 2014, p. 85). Humans use these spaces to initially process information, then decide to either move the information to long-term memory for storage and future retrieval, or reject and lose it. Sweller (2005) believes that a person's long-term memories are stored as schemas that are, “cognitive constructs that allow multiple elements of information to be categorized as a single element” (p. 21). There are three types of load within cognitive load theory: intrinsic cognitive load, extrinsic cognitive load, and germane cognitive load (Sweller, 1999, 2005). Germane cognitive load is the load that is created by effort, and that is used to create and make schema automatic (Sweller 2005). Automatic schema refer to information that is transferred to long-term memory, which requires less space and time. For listening instruction, the goal is to, “increase germane cognitive load” (Cottam & Savenye, 2014, p. 85).

According to cognitive load theory, listeners will have 5-9 spaces to deal with incoming information. However, listeners cannot listen to lectures with empty minds. One or more of the spaces might already be occupied when listeners are processing incoming information. For NNS listeners, “it is proved that ESL/EFL students experience higher cognitive load than their

counterparts due to cultural and linguistic adjustments and subsequent discomfort” (Wagner-Loera, 2018, p. 171, from Miller & Endo, 2004). Limited time for processing during real-time listening, and the large amounts of information required for academic listening create heavy cognitive load for NNS listeners, especially when they are less skilled.

Automatic Processing and Controlled Process

Of the three types of cognitive load, germane cognitive load is used to make schema automatic. In the field of human information processing, Schneider and Shiffrin (1977) and Shiffrin and Schneider (1977) have proposed, “a two-process theory of human information processing” (p. 1). The two processes are controlled processing and automatic processing.

Automatic processing is activation of a learned sequence of elements in long-term memory that is initiated by appropriated inputs and then proceeds automatically – without subject control, without stressing the capacity limitations of the system, and without necessarily demanding attention. Controlled processing is a temporary activation of a sequence of elements that can be set up quickly and easily but requires attention, is capacity-limited (usually serial in nature), and is controlled by the subject (p. 1).

It is quite normal that speakers speak at a speed of, “three words a second” (Buck, 2001, p. 6), which leaves very little time for the listeners. It is almost impossible for listeners, especially NNS listeners, to catch the precise meaning of every word, to recognize the precise structure of all the clauses, to match the pronouns to each corresponding character, to analyze various intonations and accents, and to do other actions within a few seconds. Remember that listeners have to synthesize all of their knowledge to acquire their own interpretation of one piece of information before another piece flashes in. For less skilled NNS listeners, most of the listening processing is controlled processing, which requires attention and occupies cognitive

load capacity. However, if listeners can learn to automatize the listening process or parts of it, the burden on cognitive load will lessen since automatic processing is, “activated without the necessity of active control or attention by the subject” (Shiffrin & Schneider, 1977, p. 155). This will leave more cognitive capacity for listeners. Alternatively, if listeners must expend cognitive load on controlled processing, but they are allotted a longer time for the process, the effect of cognitive load needs to be considered carefully.

Summary

In this section, I have discussed literature focused on NNS listening, academic listening, and features of listening comprehension. Literature about listeners’ memory, cognitive load, and human information processing methods were discussed to emphasize that time is a critical factor in the listening process, leading to most of the listening difficulties faced by NNS listeners.

Speech Rate, Speech Duration, and NNS Listening Comprehension

In this section, I will discuss literature about the Speech Rate and Duration of listening materials, and the effects of Speech Rate and Duration on NNS’s listening comprehension. This review aims to present researchers’ work about listening comprehension from another angle.

The reason for discussing literature related to speech rate in this section is that, in L2 teaching, “it would be useful if teachers knew the rate at which listening comprehension is maximized to make the best use of that scarce resource, time” (Griffiths, 1990a, p. 312). This is because, “speech rate variability affects the mental workload of speed processing” (Baldwin, 2012, p. 714). Previous discussions about listeners’ memory, cognitive load, and processing methods in academic listening have shown that time in listening comprehension is crucial for listeners, and especially NNS listeners. It is assumed that a slower speech rate may be beneficial for listening, and especially for NNS listeners. However, this is just an assumption.

The duration of the lectures prepared for NNS listeners is another factor that is related to the time for listening comprehension. Though it is a given that longer lectures will tire listeners more easily because of heavier cognitive load and more information, attention is paid to related literature to clarify the relationship between NNS listeners' listening comprehension and the duration of listening materials.

Speech Rate in Listening Comprehension

In the literature related to phonetics, terms like “speaking rate”, “speech rate”, and “articulation rate” are, “tended to be used interchangeably to indicate speech tempo, i.e. the pace at which a stretch of connected discourse is delivered by the speaker” (Jacewicz, Fox, O'Neill, & Salmons, 2009, p. 234). Nowadays, both speaking/speech rate and articulation rate are defined as, “the number of output units per unit of time (e.g., syllables or words per minute)” (Tsao, Weismer & Iqbal, 2006, p. 1156). Grosjean and Lane (1981), and Miller, Grosjean, and Lomanto (1984) both concluded that there were two parameters to describe variations in speaking rate: “the speed of articulatory gestures throughout an utterance and pause frequency (the number of pauses) and pause intervals that typically separate uninterrupted articulatory sequences” (Tsao, Weismer & Iqbal, 2006, p. 1156).

In linguistic fields such as sociolinguistics, there are two different views about what is included in speech rate ranges. The first view regards speech rate as the articulation rate plus pauses (Miller, Grosjean, & Lomanto, 1984; Derwing, 1990). Miller, Grosjean, and Lomanto (1984) argue that, “speaking rate is a complex variable” and is, “composed both of the rate at which the speech itself is produced – articulation rate - and the number and duration of pauses in the utterance – pause rate” (p. 216). Goldman-Eisler (1956) and Grosjean (1980) believe that changes in speaking rate, “are largely due to the pause rate” (Miller, Grosjean, & Lomanto, 1984,

p. 216), while the articulation rate remains relatively stable. The other view takes the opinion that speech rate is calculated via the number of words articulated divided by the duration of the speech (literally how fast the speakers talk). Baldwin (2012) states that, “speech rate includes pauses and is typically measured in words per minute (wpm)” (p. 174).

In L1 pausology and L2 input studies about listening comprehension, some scholars consider speech rate to be one of the temporal variables inclusive of, “speech rate (SR) and pause phenomena (PP): pause duration, distribution, and frequency” (Dechert & Raupach, 1980; Griffiths, 1990a, p. 312; Griffiths, 1990b, p. 106). Research focused on pausology is related to hesitation in articulation, anthropology, generation of conversation, and others (O’Connell & Kowal, 1983; Watanabe & Rose, 2010).

In the current study, the discussion of the literature will be focused on Speech Rate as Baldwin (2012) has described it. The study of pause phenomena will not be included. The determination of pauses including the frequency, location, and length, is beyond the scope of this study, and will not be reflected in the experiments.

Research about speech rate reveals that, “there is tremendous variability both within and between speakers in both the number and duration of pauses and the articulation rate” (Baldwin, 2012, p. 175, from Miller et al., 1984). For the current study, the speech rate is considered to be the articulation rate plus the pause rate, and the measurement method of speech rate as wpm has been adopted. The manipulation of speech rate for the quasi-experiment audio files has been applied to entire segments in order to adjust the rates of listening materials for the participants.

There were at least four units used to calculate speech rate in L1 literature: words per second, syllables per second, phonemes per second, and syllables per minute (Griffiths, 1990a). However, in L2 research about speech compression, Griffiths (1990a) recommends that words

per minute (wpm) should be applied alongside other measures. Buck (2001) also agrees that wpm is, “the most common measure of speech rate” (p. 39). Therefore, in the current study, wpm is applied to measure the speech rate of lecture materials for participants in experiments.

Carpenter and Just (1977) have stated that, “normal speech (and hence auditory comprehension) proceeds at a rate of about 160 words per minute” (p. 110). Tauroza and Allison (1990) studied the average speech rate of British speakers within various types of texts. The average speech rate was about 170 wpm, the rate of conversation was, at the highest end, about 210 wpm, and the rate of lectures to NNS speakers was at the slowest end of about 140 wpm. Buck (2001) states that such a giant difference, “suggests that the lecturers are probably slowing their speech to aid comprehension” (p. 39). Buck (2001) reports that the average speech rate of a native English speaker is around 165 to 180 wpm. Rubin (1994) has also concluded that for NS of English, a normal speech rate is around 165 wpm to 180 wpm (p. 200).

What interests the researchers in second language acquisition (SLA) are: 1) whether a threshold of speech rate exists at which comprehension changes at certain levels, and 2) whether it is true that slower speech rates are better for listening comprehension?

Foulke and Sticht (1969) and Sticht (1971) argue that such a speech rate threshold occurs at about 250 wpm. Carver (1973) provides the following explanation regarding the phenomenon of speech rate threshold:

the data suggest that when an individual is presented with speech at increasingly high rates, there will be little decrease in his understanding of the thoughts until the minimum amount of time required to process each word, each group of words, or each thought is exceeded. Then, the decrease in thoughts understood will drop abruptly because little or nothing is understood or comprehended below the threshold. (Carver, 1973, p.124)

Carver (1973) launched an investigation to find out whether, “there is a threshold in speech rate beyond which comprehension drops abruptly” (p. 118). Carver first re-investigated Foulke’s (1968) study of 360 NS English subjects listening to a 2925-word passage at varying rates between 125 and 400 wpm, finding that, as measurement approaches vary, the relationship between speech rate and listening comprehension varies. Carver then reanalyzed data from previous investigations. 54 subjects were asked to listen to 10 passages at varying rates between 75 and 400 wpm. Results show that there is a speech rate threshold at 150 wpm beyond which the, “percentage of understanding dropped abruptly” (p. 123).

Derwing (1990) set up a study to determine whether, when Native English speakers slow down, NNS listeners will understand better. 16 NS watched a short film and served as narrators to 16 NNS and NS partners one by one. The NSs were adults who spoke Standard Canadian English, and NNS participants were, “12 Vietnamese, 1 Roumanian, and 3 Cantonese speakers” (p. 306). Narrators are asked to describe the film to their partners. The experimenter then asked the partners to respond to comprehension questions orally. Based on the partners’ performances, narrators were divided into a successful group and an unsuccessful group. The results showed that, “the subjects who successfully communicated the story to NNSs did not adjust their speech rate, while those who had difficulty communicating with NNSs increased pause time significantly.” In Derwing’s study, the speech rate could be considered similar to articulation rate. This was because she found that, “those narrator who slowed their speech did not articulate significantly more slowly, but rather made greater use of pause time when talking to NNSs” (p. 308). Therefore, it was not the speech rate of the narrators, but the pauses added that affected the NNS listeners’ listening comprehension. Thus, it was not proven that a slower speech rate would result in better listening comprehension in NNS listeners.

Griffiths (1990a) performed an investigation into the relationship between various speech rates and listening comprehension with 15 lower-intermediate NNS Japanese listeners. The rates were: 200 wpm (marked as the fast speech rate), 150 wpm (average speech rate), and 100 wpm (slow speech rate). Results showed that the fast speech rate (200 wpm), “resulted in a significant reduction in comprehension” (p. 311). However, no significant differences were reported between the scores on texts delivered at an average speech rate (150 wpm) and at a slow speech rate (100 wpm). In Griffiths’ study, the threshold of speech rate in NNS listening comprehension was around 200 wpm.

Griffiths (1992) launched a further study to find out whether NNS listeners could comprehend at a greater range of rates. In the study, the slow speech rate was about 127 wpm, the average speech rate was about 188 wpm, and the fast speech rate was approximately 250 wpm. 24 Japanese subjects with estimated proficiency levels between upper elementary and intermediate participated in the study. Results showed that mean test scores for text at the slow speech rate were significantly higher than that of the fast speech rate, and mean test scores of text at the slow speech rate were significantly higher than that of the average speech rate. There were no significant differences between the mean scores for texts at the average speech rate and at the fast speech rate. In this study, the threshold of speech rate for NNS listeners was around 250 wpm, which is similar to Buck’s (2001) belief that the threshold of speech rate at which comprehension decreases is about 250-275 wpm.

Perhaps because of conflicting results within the literature (Rubin, 1994), Zhao (1997) looked at, “the effects of listener’s control of speech rate on comprehension” (p. 51) to gain new insights into the effects of speech rate on listening comprehension. Zhao believed that such a setting would exclude the effects of individual differences (e.g. language proficiency, previous

knowledge, and prior experiments), leaving the differences in subjects' listening comprehension to external factors such as speech rate and delivery style. 15 adult subjects were invited to participate in four different listening conditions, all NNS speakers of English from six different countries including China, Korea, Taiwan, Turkey, and Venezuela. "What was manipulated in the four conditions was the amount of control the listeners had over speech rate and repetition" (p. 53). Subjects in condition 1 listened to 20 sentences presented once at a fixed rate of approximately 185 wpm, and answered multiple choice questions following each sentence. Subjects in conditions 2, 3, and 4 had the opportunity to choose the speed, dynamics, and amount of repetitions for the texts they listened to. They also listened to coherent passages instead of separate sentences. In condition 2, subjects were able to identify their ideal speech rates after trying a calibration passage. They then took the test passage using that speed. The speech rates that subjects could choose from varied from 75 to 200 wpm. In condition 3, subjects were allowed to, "change the speech rate while listening to a passage by clicking on the 'Faster' or 'Slower' button" (p. 55). They were able to listen to the whole passage, or part of the passage at their desired speed. Results affirmed that, "when given control, students' listening comprehension improved, and improved listening comprehension was achieved by slowing down the speech rate" (p. 49).

Johnson (1998) studied the effects of speech compression on recall in a multimedia environment. A major research purpose was to find out, "how the rate of speech, word per minute (wpm), may influence learning in a multimedia setting" (p. 1). 192 college students were assigned to one of three treatment groups with audio at 1) normal speech (125 wpm), 2) compressed speech (175 wpm, 40% compression), and 3) compressed speech (200 wpm, 60% compression). The participants were asked to solve puzzle problems after listening to solutions

presented in one of the three treatment groups. The number of attempts and duration of time to finish the puzzles were recorded and scored. The results showed that, “the recall of information at various levels of compressed speech decreases when audio delivery rate (words-per-minute) increases” (p. 23). Johnson (1998) concluded that participants, “performed better, which means that they had a lower number of attempts in treatment group 1” (p. 35) in which the audio was presented at a normal speed of 125 wpm.

Le (2006) has conducted experiments to determine whether various speech rates (slower, normal, and faster) may have different effects on high-intermediate ESL learners’ academic listening comprehension. Eleven high-intermediate level international students (6 Korean, 4 Chinese, and 1 Lebanese) participated in the study by listening to, “three long academic lectures on unfamiliar topics and answer multiple choice comprehension questions” (p. vi). The normal SRs for the lectures were 157wpm/3.40 sps (syllabus per second), 168 wpm/3.34 sps, and 173 wpm/3.33 sps. The faster SRs were 181 wpm/3.91sps, 193 wpm/3.84 sps, and 199 wpm/3.83 sps. The slower SRs were 134 wpm/2.89 sps, 143 wp/2.83 sps, and 147 wpm/2.83 sps. Le concluded that 1) no significant differences were found between all three sets of SRs although participants scored highest at the slower SRs, 2) participants appeared to prefer a slower SR, and 3) speech rate was only one of the major causes of incomprehensibility in academic listening (p. 29).

Duration of Speech in Listening Comprehension

Few studies have mentioned the effect of speech duration on NNS listeners’ listening comprehension for lectures. In standard L2 English listening tests, such as the listening section for the TOEFL, listeners will listen to 3-4 short lectures and answer 6 questions from each to test their “basic comprehension” and “pragmatic understanding (speaker’s attitude and degree of

certainty) and connecting and synthesizing information” (ETS, 2021, para. 1). The lectures are 3-5 minutes long in duration.

Rahmatian and Armiun (2011) compared the effectiveness of audio and video documents in L2 French listening comprehension by 44 adult learners at universities in Iran. The proficiency levels for the learners were not provided. However, the authors stated that the participants were “M. A. students of French language” and, “having a fairly well knowledge of French language” (p. 118). The authors argued that “duration of the document,” meaning the duration of the materials that were presented to the listeners, was an “important parameter” because, “a long document would tire the participants and a short one would not differentiate well” (p. 119). The durations for the videos in the study were 3 minutes and 7 seconds (p. 119), which fits within the typical duration range for the previously mentioned TOEFL lecture listening materials.

Summary

In Carver (1973)’s study, the subjects were NS English listeners, and there was a threshold for speech rate at 150 wpm. In Derwing (1990)’s studies, no threshold of speech rate for NNS listeners was found. However, results showed that it was not the speech rate, but the greater use of pause time that affected NNS listeners’ comprehension. None of the NNS listeners in Derwing’s study were L1 Chinese speakers. In Griffiths (1990a, 1992)’ studies, all the NNS listeners were L1 Japanese speakers. The threshold was found to be 200-250 wpm. However, the proficiencies of the subjects in Griffith’s studies were at the elementary and intermediate levels. The lack of advanced subjects in both studies, and small number of subjects may have affected the validity of the results. Zhao (1997)’s study minimized the effect of individual differences on listening comprehension, but the small sample size of 15 subjects was not large enough to give

the study validity. Le (2006) only had 11 subjects in his study, and all of the subjects were at the same proficiency level, making the validity of Le's study questionable.

Griffiths (1990b) once criticized that, "the basic relationship between speech rate and comprehension by L2 learners" had not been clarified. Recent research has been focused on children's language development, and adults with diseases like hearing loss and Parkinson's among others (Hustad, Jones, & Dailey, 2003; Love, Walenski, & Swinney, 2009; Chiu & Neel, 2020). However, little has been focused on the relationship between speech rate and listening comprehension in online L2 English academic listening by L1 Chinese learners. There is a lack of agreement on how to study the effects of speech rate on listening comprehension, and what speech rates can be considered slow, average or fast. All of these unsolved questions have led me to explore the results from previous work to clarify the factors affecting listening comprehension for L1 Chinese listeners of online academic L2 English.

Similarly, very limited research has been focused on the relationship between NNS listeners' listening comprehension of lectures and lecture duration. However, this investigation into lecture duration is necessary, as it is related to the critical factor of time in listening comprehension, and falls within the scope of cognitive load.

NNS Learners' Listening Comprehension in Online L2 English Academic Listening

Research about L2 English learners' listening comprehension is mainly focused on NNS learners' listening difficulties, listening strategies, and instructional strategies that are meant to improve listening comprehension. In this section, literature focused on L2 English in each of the three areas is presented.

Listening Difficulty for L2 English Learners

DeCarrico and Nattinger (1988) have stated that L2 learners' difficulties in comprehending academic lectures are caused by, "lexical phrases, chunks of language of varying length" because even advanced L2 learners still have difficulties in understanding "the points" (p. 91) of the lectures.

Vandergrift (1999) states that beginning learners of a foreign language have difficulty listening for, "accurate meaning and learning to produce correct sounds at the same time" because their short-term memory, "is not capable of retaining all of this information" (p. 169). Listeners' lack of capacity to process all the information drives them to focus mainly on words and a bottom-up approach, thus causing them to miss general information. This makes it difficult to understand the whole picture.

In Osada's (2004) literature review of listening comprehension research, it was concluded that, based on Underwood's (1989) work, there are seven types of difficulties in listening to English. First, listeners have no control over the speed of delivery. Second, listeners are not able to have what they've heard repeated. Third, listeners have vocabulary problems. Fourth, listeners may lack listening strategies. Fifth, listeners need contextual or background knowledge. Sixth, listeners find that it is difficult to concentrate while listening to a foreign language. Seventh, listeners' pre-existing listening habits could be a problem (Osada, 2004, pp. 62-63).

Smit (2006) studied ESL students' assignments and found that students' difficulties in academic listening are due to a couple of reasons. One difficulty is that the students devote much of their attention to understanding the words and grammar for the language used, "rather than focusing on the message conveyed by the speaker" (p. 12). Such students fill their working memory capacity with word-level information, thus failing to build "words into higher-level

meaning” (Field, 2004, p. 365). The other difficulty is that the students are exposed to lecturers “with different English accents and pronunciation features” (p. 12). Lecturers who are not native English speakers bring additional difficulties in the form of cultural and/or ethnic differences to L2 English listeners.

Le (2006) presented the problems listed by the 11 subjects in his study about the effects of speech rate on ESL learners’ academic listening comprehension and reasons for lecture incomprehensibility. Though fast speech rate was listed as one of the major causes for lecture incomprehensibility, it was not the only one. In Le’s study, the leading cause of difficulties in participants’ listening was unfamiliar topics (Le, 2006, p. 28).

L1 Chinese Listeners’ Difficulties in L2 English Academic Listening

As mentioned in previous discussion, NNS listeners face various difficulties in academic listening. One of these difficulties is based on the listener’s background. L1 Chinese listeners who are from, “a different educational system and cultural environment, experience particular challenges in English academic listening” (Huang, 2005, p. 553).

Wong (1988) published a critical review of research on Chinese speakers learning English, focusing on the difficulties and the variables determining their ability to understand successfully. Difficulties were found in the areas of phonology, morphology, and syntax. Phonologically, the following problems were noted: “final consonant deletion, consonant cluster simplification, vowel simplification, unreduced vowels in unstressed positions, and epenthesis” (p. 3). The problems in morphology and syntax, though mostly within writing, were caused by the “uninflected nature of Chinese,” which impacts learning English, “a moderately inflected language” (p. 4). Wong’s study focused on the difficulties faced by Chinese speakers learning English, caused by the differences between the Chinese language and English language.

Contrastive studies about linguistic differences between Chinese and English provide possible reasons for the difficulties that Chinese listeners have with English. However, as discussed in previous sections, listening is a complex process which involves knowledge from many different areas. Linguistic differences or linguistic knowledge of English makes up just one type of knowledge that L2 English listeners must acquire.

Goh (2000) reported on real-time listening difficulties experienced by 40 ESL Chinese adult learners, examining data from learners' self-reports, group interviews and immediate retrospective verbalizations (p. 55). The five most common listening comprehension problems listed in Goh's (2000) study were: 1) quickly forgetting what was heard, 2) not recognizing known words, 3) understanding words, but not the intended message, 4) neglecting the next part when thinking about meaning, and 5) being unable to form a mental representation based on the words heard (Goh, 2000, p. 60). All these problems occurred during the listening process, and were related to the time and processing capacity that listeners had when they were exposed to real-time listening. Goh claimed that there were similarities in the difficulties reported by the effective listeners and the less-effective listeners, but less-effective listeners experienced more difficulties with perception and parsing.

In Huang's (2004, 2005) study, 78 L1 Chinese students were asked to complete questionnaires in which they responded using a five-point scale to report the factors that affect English academic listening. The results showed that, "both linguistic and non-linguistic factors were reported to affect Chinese students' comprehension of English academic lectures. Non-linguistic factors included instructional, psychological, and individual factors" (2005, p. 553). Among the suggestions the subjects made, two suggestions were related to the speech rate. One was that, "the teacher should vary the pace of the lesson and break up content into accessible

units.” The other was that, “the teacher should slow down a little bit when teaching to make it easier for ESL students” (Huang, 2005, p. 553). Both items emphasized the importance of appropriately slowing down speech rate and adding necessary pauses into lectures.

Difficulties in Online Academic Listening

During face-to-face listening, listeners can benefit from extralinguistic clues, such as interactions, body language, facial expression, and prepared aid, to enhance listening comprehension. However, in online courses, especially those that are completely asynchronous, listeners may need to listen to pre-recorded videos or audio files (Cottam & Savenye, 2014).

Liu’s (2002) study was conducted to, “identify the online processing problems encountered by Hong Kong students of differing L2 proficiency” (p. i) and to investigate students’ compensatory schema use in problem solving. The results from the interviews stated that students experienced problems caused by unfamiliar topics, vocabulary, accents, and “acoustic blur” in which they could hear the English sounds, but could not identify the words. Liu also studied the performances of students at various proficiency levels, finding that high proficiency students transcribed more than less proficient students and that, “listeners’ online processing problems are spread across a string of two or more words” (p. i). Liu confirms that “students tend to blame fast speed or lack of concentration for their online processing problems” (p. ii).

L2 English Learners’ Listening Strategies

Listening comprehension is a cognitive process that requires certain strategies (Rost, 2007). Therefore, much attention has been devoted to the listening strategies used by L1 Chinese learners of L2 English when this group of learners is involved. Among the limited research about L2 English listening by L1 Chinese learners, applied listening strategies are the main focus (Liu,

2011; Liu & Thondhlana, 2015). Listening comprehension strategies are defined as, “conscious plans to manage incoming speech, particularly when the listener knows that he/she must compensate for incomplete input or partial understanding” (Rost, 2013, p. 236).

According to Oxford (1990), there are six groups of language learning strategies: memory, cognitive, compensation, metacognitive, affective, and social. Three of the six strategies are adopted for listening comprehension specifically, and they are cognitive, metacognitive, and socio-affective strategies (O’Malley & Chamot, 1990; Vandergrift, 1990). Cognitive strategies are separate learning activities that a learner implements in order (Serri, Boroujeni, & Hesabi, 2012) to manage or complete a particular learning task. This may involve inferencing, elaboration, summarization, translation, transfer, repetition, resourcing, grouping, note-taking, substitution, and deduction/induction (Vandergrift, 1997). Metacognitive strategies are, “the ability of learners to control their thoughts and to regulate their own learning” (Vandergrift and Goh, 2012, p. 5). These may include planning, monitoring, evaluation, and problem identification (Vandergrift, 1997, p. 396). Xiao & Sidhu (2018) found that metacognitive strategies were more frequently applied by the adult Chinese learners of English, followed by cognitive and socio-affective strategies (Xiao & Sidhu, 2018, p. 1270). Socio-affective strategies are activities initiated by the learners to strengthen positive emotional attitudes toward language learning (Chamot & O’Malley, 1987). They can be divided into five categories: questioning for clarification, cooperation, lowering anxiety, self-encouragement, and taking emotional temperature (Vandergrift, 1997, p. 395).

Researchers want to identify the listening strategies that are applied by listeners for instructional purposes, and which may help L2 listeners to improve their listening (Oxford, 1990).

Liu (2011) has attempted to determine whether Chinese and Korean students achieve academic listening comprehension through specific listening strategies. Liu groups the first three strategies of the Oxford (1990) model (memory, cognitive, and compensation strategies) as “direct strategies,” and the latter three (metacognitive, affective, and social strategies) as “indirect strategies (Liu, 2011, p. 23). Liu argues that the order of the strategies indicates the increasing complexity level. She has recategorized strategies from Oxford (1990) and Vandergrift (1997) into four groups: memory, cognitive, metacognitive, and socio-affective strategies. One hundred and sixty-six Chinese and Korean college students participated in the study. Their TOEFL listening scores were used to differentiate skilled and less-skilled listeners with a threshold score of 570. Liu’s results showed that her participants tended to, “employ memory strategies as a means of achieving listening comprehension” (p. iv).

Liu and Thondhlana (2015) examined the relationship between students’ foreign language listening anxiety and foreign language listening strategy use among 1702 first and second-year undergraduate EFL learners at Chinese universities. The 26-item Foreign Language Listening Strategy Use Scale (FLLSUS) was adopted for the study in order to test seven dimensions: 1) negotiation for meaning while listening, 2) fluency-maintaining strategies, 3) scanning strategies, 4) getting the gist strategies, 5) nonverbal strategies while listening, 6) less active listener strategies, and 7) word-oriented strategies (p. 10). The results showed that, “more than half of the correspondents moderately use different types of listening strategies,” such as FLLSUS1, FLLSUS2, FLLSUS3, and FLLSUS4 (p. 16). It also proved that the students’ FLLSUS was, “significantly correlated with the students’ listening comprehension performance” and that, “FLLSUS6,..., and FLLSUS1 were good predictors of English listening comprehension performance” (p. 18).

Since it has been proven that listening strategy use is correlated with listeners' performance, researchers have been working hard to identify appropriate instructions for listening strategies to improve language learners' listening comprehension.

As early as the academic year 1987-1988, Rubin (1990) had applied experiments to study the effects of various listening strategies on high school L2 Spanish learners' performance in the comprehension of video. Results showed that, "the use of some listening strategies can help students work with more difficult material" (Vandergrift, 1999, p. 171). Rubin agreed that, "the combination of well-selected video and training in effective listening strategies can improve student affect and motivation" (p. 315).

Thompson and Rubin (1996) reported the results from a classroom-based, longitudinal study on the effects of instruction centered on learner strategies. The impacts were examined on the listening comprehension of subjects from a third-year Russian course at a university. They investigated the effects of instruction for metacognitive and cognitive strategies on the listening comprehension of the subjects. The listening materials were video segments, Russian television, and movies. It was confirmed that, "systematic instruction in the use of strategies will result in the improvement of listening comprehension" (p. 331).

Unfortunately, limited research has been done about the relationship among L1 Chinese listeners' L2 English listening comprehension, the speech rate of the listening materials, and the instructional strategies applied. This gap has inspired the start of the current study.

Summary

In this section of the literature review, much attention was given to the listening difficulties and listening strategies of L1 Chinese learners. Much work has been done within each of the areas of listening processing, temporal variables, listening difficulties, listening

strategies, and instructional strategies for L2 English listening. However, limited studies have reported on interactions among these areas.

Assessing Listening Comprehension

One of the steps in the current study is to assess the participants' listening comprehension in order to know about their understanding of the listening materials. Therefore, a literature review focused on the assessment of listening comprehension is needed.

According to Ehrlich (1982), the term "comprehension" refers to both the activity in which the subject is involved while processing information, as well as the product of this activity. The product may be analyzed as a mental structure or observing the subject's behavior while engaged in different tasks (Griffiths, 1990b, p. 104). However, the ways in which listeners comprehend auditory information through active cognitive processing are invisible to others. To know whether listeners comprehend the information that they listen to, researchers must depend on the reconstructions of meaning created by listeners during listening comprehension assessments. The purpose of listening assessments in the current study is to determine listeners' comprehension of specific academic listening materials under certain circumstances. Therefore, this section of the literature review will focus on how to design and construct listening comprehension assessments to understand listeners' performance in academic listening.

In his book *Assessing Listening*, Buck (2001) outlined how to design listening tests for academic listening. First of all, test-designers need to know the specificity of the test. What is the situation within which the target language will be used? For example, we might want to assess whether L1 Chinese learners of L2 English have sufficient ability to follow the instructor in an undergraduate math course. Secondly, test-designers generate language use tasks based on observations, recordings of listening activities, and information provided by the subject

specialists. Thirdly, test-designers operationalize the construct to, “develop a set of test tasks that will reflect the target-language use tasks as closely as possible” (p. 206). For example, it is possible that there will be several types of speech events in the math course, such as an introduction and warm up, formal lecture, group discussions, and presentations. Various tests must be developed for each of the speech events. For the introduction and warm up, test-takers can be offered cultural scenarios to practice greetings and social interactions. For formal lectures, test-designers might develop tests of multiple choice questions, diagrams, open-ended summaries, or multiple cloze to assess test-takers’ understanding of short talks (p. 207).

Buck (2001) provided three different approaches to assessing listening: the discrete-point approach, the integrative approach, and the communicative approach (p. 62). At the time when the audio-lingual method, structuralism, and behaviorism dominated the field, the “discrete-point approach was the most common approach to language testing” (Buck, 2001, p. 62). Common discrete-point tests are true/false and multiple-choice tests. In listening assessments, the most common discrete-point tests are “phonemic discrimination tasks, paraphrase recognition and response evaluation” (Buck, 2001, p. 63).

Arguing that the whole is greater than the sum of its parts in language use, the integrative testing approach puts emphasis on “assessing the processing of language” rather than, “assessing knowledge about the elements of the language” as the discrete-point approach does (Buck, 2001, p. 67). Common listening tests that adopt the integrative approach are listening cloze, dictation, and translation.

The communicative testing approach argues that it is the use of language for its communicative function, rather than the usage of the language that should be tested (Buck, 2001). Carroll (1980) stated that, “the *use* of language is the *objective*, and the mastery of the

formal patterns, or *usage*, of that language is a *means* to achieve that objective” (p. 7, italics in original). The communicative approach argues for authentic texts, communicative purposes, and authentic tasks.

Cheng (2004) has compared the effects of assessment methods, multiple-choice (MC), multiple-choice cloze (MCC), and open-ended (OE) questions, on L1 Chinese subjects’ listening performance in English. One hundred fifty-nine L1 Chinese college students participated in the study to take part in all three formats through three balanced treatments. The results showed that there were significant differences between the MC and the OE test scores, and between MMC and OE test scores. No significant differences were found between the scores of the MC and MCC tests. Subjects got the highest scores on the MCC test, followed by the MC test, and with the lowest scores on the OE test. Post-test surveys revealed that 97% of the subjects admitted that guessing was an important strategy. They preferred the MC tests because, through the given possible answers, they could recall the spoken stimuli more effectively and achieve better accuracy in guessing. Subjects in Cheng’s (2004) study were an average age of 18.5 years with at least 7 years of experience learning English as a foreign language in Taiwan. None of them had passed the Intermediate Level of the General English Proficiency Test (p. 546). Though Cheng does not provide the proficiency levels of the subjects, it can be concluded that all the subjects were at the beginning level of proficiency in L2 English. Therefore, the results were only relevant to the beginning listeners’ performance with various types of tests. For less effective listeners, use of memory and cognitive strategies might lead them to rely on guessing more frequently. Therefore, MC tests that offer most possible answers and stimulus can be listeners’ choices. At this point, subjects at various proficiency levels are needed to find out what types of assessments are appropriate to test target listeners’ performance. However, the results showing

that Cheng's subjects performed best on, and most preferred MC tests does not mean that the MC test is the best choice to assess subjects' listening performance in the study.

According to Dick, Carey, and Carey (2005), there are basically four types of criterion-referenced tests: entry behaviors tests, pretests, practice tests, and posttests (p. 146). Each type of test has its own purpose to serve within the assessment design process. The entry test is applied to learners before they start the instruction. The pretest is to determine whether the learners have mastered the skills before the instruction starts. The practice test provides practice for learners during the instruction, and the posttests are given after the instruction to measure instructional objectives that were covered. (p. 147).

With this in mind, what types of items will best assess the listening comprehension of L1 Chinese listeners of L2 English academic listening? Dick, Carey, and Carey (2005) provide a summary of types of behavior and related item types, with "the behavior specified in the objective provides clues to the type of item or task that can be used to test the behavior" (p. 155),

Normally, academic lectures at the university level have various assessment requirements for listeners. Particularly for the current study about L1 Chinese listeners' comprehension of L2 English academic lectures within the conditions of manipulated acoustic features, the focus was centered on the participants' general understanding of the provided manipulated listening materials, and comprehension of specific statements made by the lecturer. Therefore, the types of behaviors stated in the objectives included evaluate/judge, solve, and discuss among others. The test items were mainly essay based. Specifically, free-recall questions and cued-recall questions were adopted as assessment methods in the current study for participants' listening comprehension.

Prince (2012) analyzed the assessment and teaching of listening comprehension and comprehension-restitution (C-R) procedures based on the performance of 8 adult L1 French learners of L2 English. Prince argues that discrete tests (multiple-choice tests and closing tests) are different from tests that require continuous responses, such as C-R procedures. This is because they require different types of processing. C-R procedures “call for the production of meaningful sentences” (p. 66). Three C-R tasks were involved: paragraph restitution, sentence completion, and summarizing. These tasks were used to test the three criteria of word recognition, intelligibility of answers, and task accomplishment (p. 71). The participants’ levels of proficiency were determined based on their previous semester’s exam results, ranging from the lowest proficiency to the highest (p. 71). The participants were allowed to take the tests at home, and to listen to the materials as many times as they wanted. Results showed that, “a high percentage of words needs to be recognized before the task can be accomplished, but that top-down processes can help in reaching an intelligible answer” (p. 65). In Prince’s study, word recognition ability was a priority for L2 English listeners, perhaps due to the study setting.

Summary

In this section of the literature review, listening comprehension assessments were the focus. Literature comparing various types of test tasks were discussed to show why the design of listening assessments is essential for researchers to understand listening comprehension. Based on the work of Buck (2001) and Dick, Carey, and Carey (2005), the possible testing items of free-recall and cued-recall questions for the current study were determined.

Speech Rate and Duration in Listening Comprehension in Instructional Design and Technology

Galbraith and Spencer (2001) evaluated the utility of variable speed playback (VSP) and 2x AV Plug-Ins to undergraduate students in an accounting course watching recorded lectures. The 2x AV Plug-In allows players to modify “1/3 to 2 ½ times the recorded speed” (p. 3). Participants reported “improved attentiveness and interest,” and that “the ability to individually and dynamically control playback rate” was valuable (p. 8). “Most users preferred and accelerated speed of at least 1.5 times normal” (p. 2). The researchers argued that, “the ability to continually (or intermittently) decelerate or accelerate lectures according to individual comprehension level and interest fundamentally changes the nature of video” (p. 9). The results match the arguments about “learner-centered language learning” (Hoven, 1997), and that “the instructional design had a significant impact on students’ listening comprehension” (Jones, 2008, p. 404) However, Galbraith and Spencer (2001) did not provide information about the normal speech rate of the lectures and participants’ background in their study.

In order to study the problem of length caused by long pauses in recorded lectures, Shenoy (2016) and Shenoy, Amresh, and Femiani (2017) removed long pauses and modified the rate of lectures by playing “uneventful sections faster and significant sections slower” (p. 267). They then conducted a user study among 202 computer science students to measure their engagement. The results showed no significant differences in the participants’ engagement, compared to the original lecture. The results about the playback rate indicate, “a dissatisfaction of playback rate range selected compared to the original video” (p. 274) where the rate of the edited lecture varies, “between 254 wpm and 354 wpm” (p. 274). The researchers admitted that “... giving the user the ability to vary the playback rate ranges would have helped” (p. 274).

Summary

It is a current trend that researchers from instructional design and technology and listening comprehension are working together to look at the listening comprehension for English lectures. Still, more needs to be done to explore how listeners from various backgrounds perform when listening to lectures bearing different designs for lecture speed and duration.

Summary of Literature Review

In this literature review, research from various topic areas related to the listening comprehension of L2 English online academic lectures by L1 Chinese listeners was discussed. It was shown that the overall time that this group of listeners have plays an critical role in their listening comprehension. Specifically, more attention should be paid to how the lengths of the lectures may impact the listening performance of this population. Therefore, the current study is focused on exploring how the targeted group of listeners' performance may be affected by the acoustic features (speech rate and duration) of listening materials. With the answers to the following research questions, designers and researchers in the IDT community may be able to better design lecture speed and duration tools for learners of online L2 English lectures.

Research questions

Research questions are proposed as follows:

1. What is the effect of variations in speech rate on listening comprehension on a test of free-recall?
2. What is the effect of variations in speech rate on listening comprehension on a test of cued-recall?
3. What is the effect of variations in duration on listening comprehension on a test of free-recall?

4. What is the effect of variations in duration on listening comprehension on a test of cued-recall?

CHAPTER THREE

Methodology

The purpose of the current study was to explore the effects of lecture speed rate and duration on the listening comprehension of L1 Chinese listeners of online L2 English lectures. The findings provided by the study may contribute to the literature base for L2 English listening comprehension research, provide a basis for instructional designers to identify feasible tools for online L2 English learners, and may be useful to ESL teachers and lecturers who have learners with various linguistic backgrounds. These findings may enable them to better help L2 English listeners through designing courses with appropriate parameters for lecture materials.

The following research questions were addressed in the study:

1. What is the effect of variations in speech rate on listening comprehension on a test of free-recall?
2. What is the effect of variations in speech rate on listening comprehension on a test of cued-recall?
3. What is the effect of variations in duration on listening comprehension on a test of free-recall?
4. What is the effect of variations in duration on listening comprehension on a test of cued-recall?

Chapter 3 focuses on the methodology used for this study. It includes discussions of study design, participants, data collection methods, the pilot study, data collection procedures, data analysis technique, and limitations to investigating the relationship between speech rate and duration of lectures and listening comprehension for online lectures.

Study Design

The study applied an exploratory quasi-experimental approach to investigate the relationships between L1 Chinese listeners' listening comprehension of L2 English lectures and two acoustic features of the lectures (speech rate and lecture duration). The study asked participants to respond to listening comprehension questions after listening to all audio included in the study.

Data Collection Method

In this study, a quasi-experiment was conducted to test the L1 Chinese participants' listening comprehension of provided lectures in L2 English. The quasi-experiment consisted of two parts. The first section consisted of English learning questions for the participants. Information collected in this section included the number of years that participants had spent learning English, experiences with academic programs, location, and self-evaluated L2 English proficiency levels.

The second section of the quasi-experiment consisted of 9 audio segments and a total of 26 questions. Each audio segment presented a different combination of speed rates and durations. After listening to each audio clip, all participants answered two types of essay questions: a single free-recall question, and multiple cued-recall questions on the contents from the clip. Participants were asked to type in their answers. Table 1 presents the parameters for the audio segments used in the study. The audio tags were used to indicate the levels for the factors, S1 indicating the slow Speech Rate, S2 indicating the normal Speech Rate, S3 indicating the fast Speech Rate, D1 indicating the short Duration, D2 indicating the medium Duration, and D3 indicating the long Duration.

Table 1*Parameters of Audios in the Study*

Audio Tag	S1D3	S2D3	S3D3	S3D2	S2D2	S1D2	S2D1	S3D1	S1D1
Speech Rate (wpm)	105	163	300	300	160	100	160	297	100
Duration	5'45"	4'12"	4'16"	2'07"	2'42"	2'41"	1'04"	1'	1'30"
Num. of words	607	687	1278	635	432	269	158	297	150
Number of listening questions	4	5	4	3	2	2	2	2	2

The original audio material dealt with myths regarding how people learn. The subject to the material was chosen because it did not appear to represent common knowledge to the general public. In this way, prior knowledge was less likely to be a factor that influenced the listening comprehension results for this study. Additionally, the lecturer selected for the study adopted a relatively colloquial style in presenting the topic, without applying many terms. Her presentation method excluded possible difficulties caused by vocabulary for the listeners. Please see Appendix A for the listening comprehension questions created for each audio segment in the order of presentation for the study.

The speed rates for samples of standard English tests for learners, such as the TOEFL (Exam English Ltd, 2019), are between 160 wpm and 186 wpm. However, the rate of lectures to NNS speakers can be as low as 140 wpm (Tauroza & Allison, 1990). The lecturer used an original speech rate of around 160 wpm. This is within the range for speed rates recommended by other researchers.

The parameters for each audio segment are presented above in Table 4. The software Audacity was used to modify the speed rate of segments from the original audio by using the command 'Change Tempo...'. Changes were made in the tempo without changing pitch by

choosing the ‘Change Tempo...’ drop-down selection from the Effect menu instead of the ‘Change Speed...’ option. This is because the ‘Change Speed...’ command modifies the pitch of the audio, resulting in a chipmunk-like high pitch sound which is not natural, and difficult for the listener to understand. By doing this, the speeds of the audio segments were adjusted, “without affecting the pitch of the speaker” (Jones, 2008, p. 401) to provide the listeners with a comfortable listening experience.

When the listening materials were ready, listening comprehension questions were created based on the content of the lectures. As discussed in the section on listening assessment in Chapter 2, assessment methods such as multiple choice, true or false, and yes or no questions, were not adopted in this study. The main type of test item adopted for the current study was the essay, and participants were asked to write down their answers for each question in the quasi-experiment. The study website was created with the Qualtrics software via Virginia Tech. The audio segments were presented in the same order for all the participants.

For each audio, two types of essay questions were created. The first one was a free-recall question asking the participants to write down anything that they recalled from the audio. The second group of questions were about details within the clip. Participants were asked to write down answers to the questions without help from response options. The number of questions for each audio segment were not always the same because the longer audios presented more content. The participants had unlimited time to finish questions for each audio segment before they clicked the ‘next’ button on the screen to continue the quasi-experiment, however they could only listen to the audio once.

Pilot Study

A pilot study was applied after the study website was initially created to test the procedures and obtain feedback from a targeted group of participants. Six current Chinese graduate students from the School of Education at Virginia Tech participated and finished the pilot study. Willis and Artino (2013) stated that, “as few as 5 or 6 subjects may provide useful information to improve survey items” (p. 355) in their discussion about the practical considerations for small-scale projects. Therefore, for the current study, six participants were chosen for the pilot study.

The final question of the quasi-experiment asked participants to share general thoughts and feelings towards the quasi-experiment itself. The comments were used for the modification of the study design and procedure. Two major modifications were adopted.

First, clearer instructions were included for the participants. Also, it was clearly stated that the quasi-experiment could take around one hour to complete, and that it was recommended that participants find themselves a quiet, distraction-free place to complete it. Participants were asked to avoid taking notes during listening as they normally might. Carrell, Dunkel, and Mollaun (2004) reported that listeners performed differently on short and longer lectures when they were allowed to take notes. In the current study, notetaking was not allowed because it was out of the researcher’s control to define the degree to which participants would behave while taking notes. This could have resulted in potential differences among participant performance when answering the listening comprehension questions. It was also possible that some participants might have taken notes while others may not have, and that participants who took notes may not have taken notes for all audio segments. The other reason was that the length setting in the current study was different from that in the study of Carrell, Dunkel, and Mollaun

(2004). In the current study, there were three audio lengths: short (1' – 1'30'), medium (2'-3'), and long (4'-6'). In the study of Carrell, Dunkel, and Mollaun (2004), the short lecture was 2.5 minutes long, and the long lecture was 5 minutes long. Different settings for audio lengths might require different study settings.

Second, the order in which the audio segments were presented was changed to present long audio segments first, with the short ones presented toward the end. Such an arrangement was meant to take advantage of the fact that the participants were full of energy at the beginning of the quasi-experiment, and to help the participants feel comfortable when they were tired of listening towards the end of the study.

Participants

The study was conducted completely online to match the settings for online lectures. During December 2020, the quasi-experiment link was sent to Chinese communities within various contexts to recruit native Chinese speakers living in a Northeast university town, studying in academic programs in the United States, or studying English at Chinese universities in China. By the end of January 2021, 308 responses were recorded. Although 308 people initially attempted to take part in the quasi-experiment, after figuring out how long it would take, most of them declined to continue. As a result, only 28 of them completed the whole quasi-experiment.

As mentioned previously, the first part of the study asked participants to answer some background questions about their English learning experience, the number of years of participation in academic programs, self-evaluation of English proficiency level, and/or years of living in the United States. The complete responses to the background questions can be found in Appendix C.

Since the quasi-experiment link was sent to universities and Chinese communities in both China and the United States, the information about the respondents is complex and needs to be presented in detail. In the following tables, Table 2 presents the information regarding participants' location and occupation, Table 3 shows the distributions of respondents' English learning experiences and academic situations (in years), and Table 4 tells us how the respondents self-evaluated their English proficiency.

Based on the location information provided by Qualtrics, it can be concluded that half of the respondents were in China when they finished the experiment, and the second half were in the United States. About half of the respondents were enrolled in academic programs (46%), and all others were not students at the time that they finished the experiment.

Table 2

Participants' Location and Occupation

	In China	In the U.S.	Enrolled in academic program	Not enrolled in academic program
Numbers	14	14	13	15
percentage	50%	50%	46.4%	53.6%

Most of the respondents had been learning English for more than 10 years (57%). For those who were in academic programs, most of them had been students for less than 5 years. Most of the respondents who were in the United States had been living in an English environment for more than 10 years (70%).

Table 3

Participants' English Learning Experience, Academic and Geographic Information

		Number	percentage
Years of learning English	<=10	12	43%
	11-20	5	18%
	21-30	4	14%
	>=30	7	25%
years in academic program	<5	6	67%

	6-10	1	11%
	>=10	2	22%
years of living in the States	<=10	3	30%
	11-20	2	20%
	>=20	5	50%

The majority of the respondents evaluated themselves to be at the intermediate level of English proficiency. All of the 4 respondents who marked themselves as beginners were in China, and one of them was in an academic program. All of these 4 respondents reported that they had less than 15 years of experience with learning English. All the 6 respondents who self-evaluated themselves as being at an advanced level of English proficiency were in the United States, with one of them reporting that they had been in an academic program for 10 years. All of these 6 respondents reported that they had more than 25 years of experience with learning English. Ten of the 18 respondents who evaluated themselves as intermediate for English proficiency were in China with around 10 years of experience in learning English, and 8 of the 10 were in academic programs. Seven of 8 respondents who were in the United States and evaluated themselves at an intermediate level of English proficiency reported more than 25 years of experience with learning English. The one exception reported that s/he had been learning English for 2 years, and had been in an academic program for 1 year.

Table 4

Participants' Self-evaluated English Proficiency

	Beginning	Intermediate	Advanced	total
Numbers of participants	4	18	6	28
percentage	14%	65%	21%	100%

Data Collection Procedure

Before distributing a quasi-experiment link, a proposal was sent to the Virginia Tech Institutional Review Board for human project permission. Upon receiving the determination letter (VT IRB – 20 -514) (Appendix B) indicating that the current study was not research

involving human subjects, a follow up data collection procedure was adopted. After the quasi-experiment was modified to reflect results from the pilot study, the link was emailed to the potential participants.

Once a participant finished all of the 26 questions and clicked the submit button to finish the experiment, a reference number was generated, and stored with the answers the participants provided as a PDF file. Two coders worked on the data coding for the current study. The researcher served as Coder A, and Coder B was another coder who was trained and coded independently. After both the independent coder and the researcher finished coding, each participant's final score for audio was generated and used for data analysis based on discussions and agreements between the coders.

Data Coding

Coding in this study involved three sets of data. The first one (Appendix C) contained information about participants' English, including self-evaluated L2 English proficiency level, years learning English, information about current academic situation, and geographic information. The second data set contained the respondents' answers to the questions from the listening comprehension process in respondent ID order. The last data set presented each respondent's answers in question ID order.

Both the independent coder and the researcher first worked on the third data set (responses in question ID order) independently. Categories for responses were generated, and scores corresponded to the extent to which a respondent discussed a category within his/her answers. Scores were marked numerically from 0 to 3, with 0 standing for "not mentioned at all" and 3 for "very detailed mention." Such an arrangement for the coding method made it possible

for more objective analysis of the respondents' answers to identify information important for answering the research questions.

After both coders finished coding the categories and determining the scores for each respondent, they discussed their work with each other, and generated the final sets of categories and scores for data analysis (Appendix D).

Utilizing the results from the inter-rater negotiations, further data processing steps were applied. First, for each question, scores from all categories were added together for each of the 28 respondents. Second, for each of the 9 audio segments, the 28 respondents' scores were calculated by summing up individual scores from questions related to the audio. Third, for each of the 28 respondents, scores from the second step were divided by the number of questions to achieve an average score for each audio segment. The results achieved from the third step were adopted for further analysis.

In the following chapters, Chapter 4 focuses on presenting the findings from the study, and Chapter 5 presents a discussion of the findings, implications, and areas for future study.

CHAPTER FOUR

Results and Findings

This chapter presents the results from the study, and an in-depth analysis of the findings. As presented in Chapter 3, free-recall questions and cued-recall questions were designed to explore possible differences in L1 Chinese participants' comprehension of the provided audio segments to explore the effects of speech rate and duration. Respondents' answers were rated by one independent rater and the researcher. After the independent rater and the researcher finished categorizing and scoring, they discussed and finalized the categories and scores (0-3) that were applied for each answer.

The following research questions were addressed:

1. What is the effect of variations in speech rate on listening comprehension on a test of free-recall?
2. What is the effect of variations in speech rate on listening comprehension on a test of cued-recall?
3. What is the effect of variations in duration on listening comprehension on a test of free-recall?
4. What is the effect of variations in duration on listening comprehension on a test of cued-recall?

Findings From the Responses

The findings from this study were examined according to different types of questions: the free-recall questions that asked the participants to write down anything they could remember from the audio, and cued-recall questions that required correct comprehension of the audio content. Results associated with these two perspectives were presented in the order of Duration

group, Speech Rate group, and SPSS results. (Appendix E contains detailed presentations of the respondents' performances on various types of questions.)

Findings - Types of Questions

Two types of questions were utilized for this study. The free-recall questions asked participants to write down anything they could remember from the audio. There were no right or wrong answers for these type of questions. The participants could put down as much as they could recall. The more that the answers fit into the categories, the higher the number of points that were given. It was possible that a respondent could achieve relatively high scores on these types of questions. The cued-recall question asked the participants to answer questions about specific points that the lecturer stated in the lecture, which meant that there were right or wrong answers. Wrong answers meant no points. Only correct answers would earn points. Differing numbers of cued-recall questions were presented depending on the various audios. More questions were designed for longer audios since they included more content. Therefore, it was possible that respondents may have had a tendency to achieve higher scores on audio segments that carried more questions (though the final score that each respondent achieved on each audio was adjusted through division by the number of questions associated with the audio). It was also possible that the respondents might have performed differently on different types of questions. Therefore, in this section, findings are presented according to the types of questions. Charts and tables were also created to visually present the results.

Free-recall Questions. The first question presented to the respondents each time that they finished listening an audio was a free-recall question that asked them to write down anything they could recall. There were no right or wrong answers for this type of question. The

independent rater and the researcher worked together to determine the categories and scales. The following section will present the respondents' performance on free-recall questions only.

SPSS Results. The following tables and figure present results from the SPSS analysis (version 26), such as the mean scores achieved by the respondents on the free-recall questions, the within-subjects effects of Speech Rate, Duration, and Speech Rate*Duration, pairwise comparisons of Speech Rate and Duration, and the plot results.

Table 5 presents the mean scores, standard deviations, and numbers in each group for the analysis for scores from free-recall questions. The highest mean score for free-recall questions was achieved for S2D3 (4.54), and the lowest one was achieved for S1D1 (1.25).

Table 5

Descriptive Statistics for Scores from Free-recall Questions

Descriptive Statistics			
	Mean	Std. Deviation	N
S1D3	2.71	2.733	28
S2D3	4.54	4.702	28
S3D3	2.18	2.212	28
S3D2	1.39	1.370	28
S2D2	2.25	2.238	28
S1D2	2.00	1.563	28
S2D1	1.57	.920	28
S3D1	1.96	1.598	28
S1D1	1.25	.967	28

Table 6 presents the within-subjects effects of Speech Rate, Duration, and Speech Rate*Duration from the repeated measurement SPSS analysis.

Table 6*Within-subjects Effects for Scores from Free-recall Questions*

Tests of Within-Subjects Effects							
Measure: MEASURE_1							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
S	Sphericity Assumed	113.937	2	56.968	7.498	.001	.217
	Greenhouse-Geisser	113.937	1.192	95.595	7.498	.007	.217
	Huynh-Feldt	113.937	1.216	93.729	7.498	.007	.217
	Lower-bound	113.937	1.000	113.937	7.498	.011	.217
Error(S)	Sphericity Assumed	410.286	54	7.598			
	Greenhouse-Geisser	410.286	32.180	12.750			
	Huynh-Feldt	410.286	32.821	12.501			
	Lower-bound	410.286	27.000	15.196			
D	Sphericity Assumed	63.865	2	31.933	13.094	.000	.327
	Greenhouse-Geisser	63.865	1.756	36.365	13.094	.000	.327
	Huynh-Feldt	63.865	1.869	34.175	13.094	.000	.327
	Lower-bound	63.865	1.000	63.865	13.094	.001	.327
Error(D)	Sphericity Assumed	131.690	54	2.439			
	Greenhouse-Geisser	131.690	47.418	2.777			
	Huynh-Feldt	131.690	50.456	2.610			
	Lower-bound	131.690	27.000	4.877			
S * D	Sphericity Assumed	39.683	4	9.921	3.798	.006	.123
	Greenhouse-Geisser	39.683	2.449	16.205	3.798	.020	.123
	Huynh-Feldt	39.683	2.711	14.636	3.798	.017	.123
	Lower-bound	39.683	1.000	39.683	3.798	.062	.123
Error(S*D)	Sphericity Assumed	282.095	108	2.612			
	Greenhouse-Geisser	282.095	66.115	4.267			
	Huynh-Feldt	282.095	73.203	3.854			
	Lower-bound	282.095	27.000	10.448			

The results from the two-way repeated measures ANOVA revealed that there was a significant main effect of Speech Rate on respondents' scores for the free-recall questions assessing listening comprehension ($F(2,54) = 7.498, p < .05, \eta_p^2 = .217$), there was a significant main effect of Duration on respondents' scores for the free-recall questions assessing listening comprehension ($F(2,54) = 13.094, p < .05, \eta_p^2 = .327$), and there was a significant interaction between Speech Rate and Duration ($F(4,108) = 3.798, p < .05, \eta_p^2 = .123$) (here the Greenhouse-Geisser approach was adopted).

Table 7 presents the pairwise comparison results for Speech Rate. It suggests that there were significant differences between the S1 and S2 groups, and between the S1 and S3 groups ($p < .05$), but no significant differences between the S2 group and S3 group ($p > .05$). It suggests

that the respondents' performance on normal speed and fast speed audio were not significantly different.

Table 7

Pairwise Comparison for Speech Rate - Free-recall Questions

Pairwise Comparisons						
Measure: MEASURE_1						
(I) S	(J) S	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
					Lower Bound	Upper Bound
1	2	1.262 [*]	.493	.049	.002	2.521
	3	1.548 [*]	.516	.017	.230	2.866
2	1	-1.262 [*]	.493	.049	-2.521	-.002
	3	.286	.180	.375	-.175	.746
3	1	-1.548 [*]	.516	.017	-2.866	-.230
	2	-.286	.180	.375	-.746	.175

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Table 8 presents the pairwise comparison results for Duration. It suggests that there were significant differences between the D1 and D2 groups, and between the D2 and D3 groups ($p < .05$), but no significant differences between the D1 and D3 groups ($p > .05$). It suggests that the respondents' performance on short length and long length audios were not significantly different.

Table 8

Pairwise Comparisons for Duration - Free-recall Questions

Pairwise Comparisons						
Measure: MEASURE_1						
(I) D	(J) D	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
					Lower Bound	Upper Bound
1	2	-1.024 [*]	.254	.001	-1.673	-.375
	3	.083	.192	1.000	-.407	.574
2	1	1.024 [*]	.254	.001	.375	1.673
	3	1.107 [*]	.269	.001	.419	1.795
3	1	-.083	.192	1.000	-.574	.407
	2	-1.107 [*]	.269	.001	-1.795	-.419

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Figure 1 shows the plots for the free-recall questions, with Duration as the X axis and various colored lines representing respondents' scores for audio segments with various Speech Rates.

Figure 1

Plots – Free-recall Questions

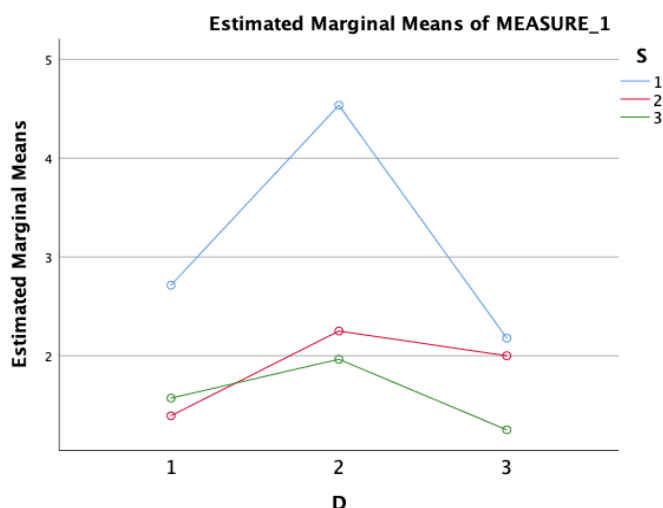


Figure 1 shows that all three lines have a positive slope from D1 to D2, and a negative slope from D2 to D3. Such a tendency suggests that respondents' performance at all SRs, were best at the D2. The slow Speech Rate line stands above the other two lines without any crossover, indicating higher scores in the S1 group than in the other two groups. The S3 line stays on the bottom of the chart, suggesting that the respondents' performance on the fast Speech Rate audio segments was the worst, with the exception of the D1 group. Within the D1 group, respondents' best performance was on S1, followed by S3, and finally S2. This suggests that within the short Duration group, respondents' performance on free-recall questions was better at the slow SR. Within the D2 group, respondents' best performance was on S1, followed by S2, and then S3. This suggests that within the medium Duration group, respondents' performance on

free-recall questions was worst at the fast Speech Rate. Similar results were found within the D3 group.

Long duration. Table 5 shows that the highest scores the respondents achieved on free-recall questions within the D3 group were on S2D3 (4.54). The lowest mean score that the respondents achieved was on S3D3 (2.18). This suggests that at long durations of 4-6 minutes long, L1 Chinese respondents performed best on free-recall questions for L2 English online lectures at the Speech Rate of 160 wpm, and the worst on the questions from the fast Speech Rate audios.

Medium duration. Within the D2 group shown in Table 5, the mean scores from S2D2 (2.25) were the highest among the three. The lowest mean scores that respondents achieved were on S3D2 (1.39). It suggests that at medium durations of 2-3 minutes long, L1 Chinese respondents performed best on free-recall questions for L2 English online lectures at the Speech Rate of 160 wpm and worst at the Speech Rate of 300 wpm.

It can be seen that there are similarities among the respondents' performance on questions for long and medium audio segments. In both groups, the best mean scores occurred with the audios presented at a normal Speech Rate of 160 wpm, and the worst mean scores were observed for the audios presented at a fast Speech Rate of 300 wpm. Performance scores for the slow Speech Rate audios, S1D3 and S1D2, were in the middle.

Short duration. Within the D1 group shown in Table 5, the mean score from S3D1 (1.96) was the highest among the three. The lowest mean score was for S1D1 (1.25). This suggests that at short durations of 1-1.5 minutes long, L1 Chinese respondents performed best on free-recall questions for L2 English online lectures at the Speech Rate of 300 wpm, and worst at the Speech Rate of 100 wpm.

It is noticeable that the respondents' performance on D1 group audios were different from that of D2 and D3 group audios in the current study. Respondents in the short Duration group in the current study performed relatively better on audio segments with a fast Speech Rate, and worse on audio segments with a slow Speech Rate. However, all respondents performed better on audios with a normal Speech Rate, and worse on audios with a fast Speech Rate in both long Duration and medium Duration groups.

Cued-recall Questions. Except for every first question, all the other questions presented to the respondents when they finished listening to an audio segment were cued-recall questions. For these, participants had to provide specific answers, and only right answers led to points. The independent rater and the researcher worked together to determine the categories and scores for data analysis (see Appendix D). The following section presents the respondents' performance on cued-recall questions only. The scores shown in the following charts were adjusted by dividing by the number of questions for each audio segment in order to limit the effect of the different number of questions for each audio.

SPSS Results. The following tables and figure show different results from the SPSS (version 26) analysis, such as the mean scores achieved by the respondents on the cued-recall questions, the within-subjects effects of Speech Rate, Duration, and Speech Rate*Duration, pairwise comparisons of Speech Rate and Duration, and the plot results.

Table 9 presents the mean scores, standard deviations, and numbers in each group for the analysis for scores from cued-recall questions. The highest mean score for cued-recall questions was achieved on S3D1 (2.2857), and the lowest was on S1D1 (.8214).

Table 9*Descriptive Statistics for Scores from Cued-recall Questions*

Descriptive Statistics			
	Mean	Std. Deviation	N
S1D3	1.4168	.76273	28
S2D3	2.2500	1.67636	28
S3D3	1.0004	.92936	28
S1D2	1.4286	1.66508	28
S2D2	1.8214	1.80644	28
S3D2	1.4821	1.44326	28
S1D1	.8214	.90487	28
S2D1	1.8571	1.69344	28
S3D1	2.2857	1.53616	28

Table 10 presents the within-subjects effects of SR, D, and SR*D from the repeated measurement SPSS analysis.

Table 10*Within-subjects Effects for Scores from Cued-recall Questions*

Tests of Within-Subjects Effects							
Measure: MEASURE_1							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
S	Sphericity Assumed	.455	2	.228	.165	.848	.006
	Greenhouse-Geisser	.455	1.674	.272	.165	.811	.006
	Huynh-Feldt	.455	1.772	.257	.165	.823	.006
	Lower-bound	.455	1.000	.455	.165	.688	.006
	Error(S)	74.528	54	1.380			
D	Sphericity Assumed	23.879	2	11.939	7.935	.001	.227
	Greenhouse-Geisser	23.879	1.371	17.419	7.935	.004	.227
	Huynh-Feldt	23.879	1.420	16.820	7.935	.003	.227
	Lower-bound	23.879	1.000	23.879	7.935	.009	.227
	Error(D)	81.252	54	1.505			
S * D	Sphericity Assumed	33.074	4	8.269	7.304	.000	.213
	Greenhouse-Geisser	33.074	2.442	13.541	7.304	.001	.213
	Huynh-Feldt	33.074	2.703	12.234	7.304	.000	.213
	Lower-bound	33.074	1.000	33.074	7.304	.012	.213
	Error(S*D)	122.264	108	1.132			
	Sphericity Assumed	122.264	65.947	1.854			
	Greenhouse-Geisser	122.264	72.993	1.675			
	Huynh-Feldt	122.264	72.993	1.675			
	Lower-bound	122.264	27.000	4.528			

The results from the two-way repeated measures ANOVA suggested that there were no significant main effects of Speech Rate on respondents' scores for the cued-recall questions assessing listening comprehension ($F(2,54) = .165, p > .05, \eta_p^2 = .006$), there was a significant main effect of Duration on respondents' scores for the cued-recall questions assessing listening comprehension ($F(2,54) = 7.935, p < .05, \eta_p^2 = .227$), and there was a significant interaction between Speech Rate and Duration ($F(4,108) = 7.304, p < .05, \eta_p^2 = .213$) (here the Greenhouse-Geisser approach was adopted).

Table 11 presents the pairwise comparison results for Speech Rate. It suggests that the respondents' performance on questions from all of the audio segments were not significantly different ($p > .05$).

Table 11

Pairwise Comparisons for Speech Rate - Cued-recall Questions

Pairwise Comparisons						
Measure: MEASURE_1						
(I) S	(J) S	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
1	2	-.022	.202	1.000	-.536	.493
	3	-.099	.136	1.000	-.445	.247
2	1	.022	.202	1.000	-.493	.536
	3	-.077	.199	1.000	-.585	.430
3	1	.099	.136	1.000	-.247	.445
	2	.077	.199	1.000	-.430	.585

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Table 12 presents the pairwise comparison results for Duration. It suggests that there were significant differences between the D1 and D2 groups, and between the D1 and D3 groups ($p < .05$), but that there was no significant difference between the D2 and D3 groups ($p > .05$). It suggests that the respondents' performance on cued-recall questions for medium length audios and long length audios was not significantly different.

Table 12*Pairwise Comparisons for Duration - Cued-recall Questions*

Pairwise Comparisons						
Measure: MEASURE_1						
(I) D	(J) D	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
					Lower Bound	Upper Bound
1	2	-.754 [*]	.216	.005	-1.306	-.202
	3	-.367 [*]	.108	.006	-.642	-.092
2	1	.754 [*]	.216	.005	.202	1.306
	3	.387	.222	.277	-.179	.953
3	1	.367 [*]	.108	.006	.092	.642
	2	-.387	.222	.277	-.953	.179

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Figure 2 shows the plots for the cued-recall questions, with Duration as the X axis and various colored lines representing respondents' scores for audios with various Speech Rates.

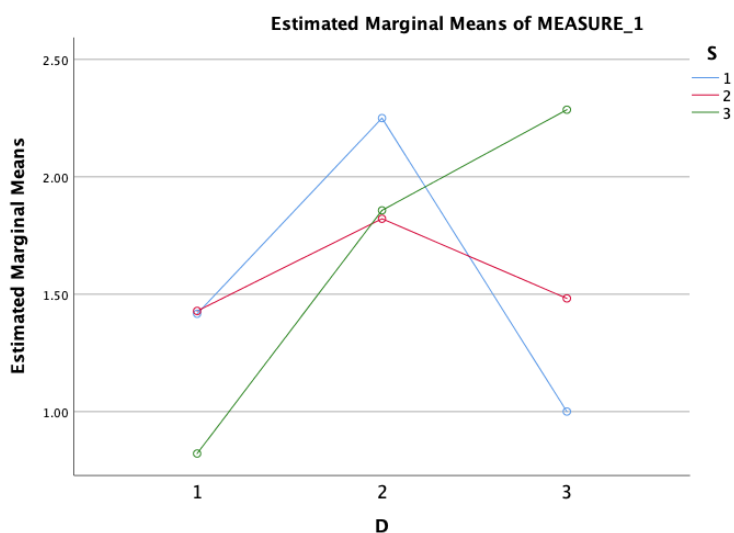
Figure 2*Plots - Cued-recall Questions***Profile Plots**

Figure 2 shows that the S1 and S2 lines have a positive slope from D1 to D2, and a negative slope from D2 to D3. The line representing the respondents' performance on cued-recall questions for the S3 group shows a positive slope all the way from D1 to D3. Such a tendency suggests that respondents' performance on cued-recall questions was best at the D2, with slow Speech Rate and normal Speech Rate audios, and at D3 with fast Speech Rate audios. Within the D1 group, respondents performed best on S2, followed by S1 and S3, suggesting that within the short Duration group, respondents' performance on cued-recall questions was better at a normal Speech Rate. Within the D2 group, respondents performed best on S1, followed by S3 and S2, suggesting that within the medium Duration group, respondents' performance on cued-recall questions was best at the slow Speech Rate. Within the D3 group, respondents performed best on S3, followed by S2 and S1, suggesting that within the long Duration group respondents' performance on cued-recall questions was best at the fast Speech Rate.

The results from the two-way repeated measures ANOVA suggested that there were significant main effects of Duration on respondents' performance when free-recall questions and cued-recall questions were considered separately.

Slow Speech Rate. Table 5 suggests that the highest scores the respondents achieved on free-recall questions within the S1 group was on S1D3 (2.7154). The mean scores that the respondents achieved on S1D1 were the lowest (1.25). This suggests that at a slow SR, respondents performed best on free-recall questions with a length of 4-6 minutes, and worst on the questions from the short duration. This conclusion aligns with the previous conclusion about performance from all questions.

Table 9 suggests that the highest scores the respondents achieved on cued-recall questions within the S1 group was on S1D2 (1.4286). The mean scores that the respondents

achieved on S1D1 were the lowest (.8214). This suggests that at a slow SR, the respondents performed best on cued-recall questions for L2 English online lectures with a length of 2-3 minutes, and worst on the questions from the short duration.

It can be concluded that regardless of the question type, respondents performed the worst on slow audios with short lengths (1-1.5 minute), and performed better on long audios.

Normal Speech Rate. Table 5 suggests that the highest scores the respondents achieved on free-recall questions within the S2 group were on S2D3 (4.54). The mean scores that the respondents achieved on S2D1 were the lowest (1.57). This suggests that at a normal SR, the respondents performed best on free-recall questions from audios at the length of 4-6 minutes long, and the worst on the questions from short audios. This conclusion aligns with the previous conclusion about performance from all questions.

Table 9 suggests that the highest scores the respondents achieved on cued-recall questions within the S2 group were on S2D3 (2.25). The mean scores that the respondents achieved on S2D2 were the lowest (1.8214). This suggests that at a normal SR, the respondents performed best on cued-recall questions for audios at the length of 4-6 minutes long, and the worst on the questions from medium length audios.

It can be concluded that regardless of the question type, respondents performed best with normal speed audios with a long length (4-6 minute), and worse with short audios.

Fast Speech Rate. Table 5 suggests that the highest scores the respondents achieved on free-recall questions within the S3 group was on S3D3 (2.18). The mean scores that the respondents achieved on S3D2 were the lowest (1.39). This suggests that at a fast SR, the respondents performed best on free-recall questions from lectures with a length of 4-6 minutes

long, and the worst on the questions from the medium duration. This conclusion is different from the previous conclusion about performance for all questions.

Table 9 suggests that the highest scores the respondents achieved on cued-recall questions within the S3 group were on S3D1 (2.2857). The mean scores that the respondents achieved on S3D3 were the lowest (1.0004). This suggests that at a fast Speech Rate, the respondents performed best on cued-recall questions from lectures with a length of 1-1.5 minutes long, and the worst on the questions from the long duration.

Summary

Findings from the study results are summarized as follows:

Respondents' performance on free-recall questions suggested that both Speech Rate and Duration had a significant main effect on L1 Chinese listening comprehension for online L2 English lectures, and that there was a significant interaction between Speech Rate and Duration impacting participants' performance on free-recall questions (Table 6).

Respondents' performance on cued-recall questions suggested that the factor of Duration had a significant main effect on L1 Chinese listening comprehension for online L2 English lectures, but that the factor of Speech Rate did not. There was a significant interaction between Speech Rate and Duration on the participants' performance on free-recall questions in this study (Table 10).

A discussion of the findings, as well as the implications, limitations, and directions for possible future studies are presented in Chapter 5.

CHAPTER FIVE

Discussion and Conclusion

An exploratory quasi-experimental design was adopted to explore the research questions, “What information might suggest that the speech rate of L2 English academic lectures affects the listening comprehension of L1 Chinese learners when they take online lectures?” and, “What information might suggest that the duration of L2 English academic lectures affects the listening comprehension of L1 Chinese learners when they take online lectures?” This study consisted of three phases. In Phase I, a quasi-experiment instrument was developed in Qualtrics. Six Chinese graduate students from the School of Education at Virginia Tech participated in the pilot study and provided feedback. The quasi-experiment then was modified based on comments and recommendations to create a final version for Phase II. In Phase II of the current study, the link to the online quasi-experiment was sent to Chinese communities both in the United States and in China in order to recruit participants. A total of 28 L1 Chinese learners of L2 English with various academic and geographic backgrounds completed the study. In Phase III, the responses were collected and analyzed. Findings and results previously reported in Chapter 4 will be discussed from different perspectives, followed by discussion about the practical implications for instructional design, limitations of the study, and suggestions for future research.

The following research questions were addressed:

1. What is the effect of variations in speech rate on listening comprehension on a test of free-recall?
2. What is the effect of variations in speech rate on listening comprehension on a test of cued-recall?

3. What is the effect of variations in duration on listening comprehension on a test of free-recall?

4. What is the effect of variations in duration on listening comprehension on a test of cued-recall?

Discussion of Findings

The focus of the research questions was on the effects of speech rate and lecture duration on L1 Chinese listeners' listening comprehension of online L2 English lectures. Speech rate and lecture duration are just two of the factors that affect NNS listeners' comprehension. As discussed in Chapter 2, there are many factors that might affect NNS listeners' comprehension including the listeners' English abilities, listeners' listening skills and strategies, listeners' background knowledge, lecture topics, acoustic features of the lectures, lecturers' English, and others. Many of the factors are related to the time that listeners have for listening. The two factors that this study focused on were the Speech Rate and Duration of online L2 English lectures.

Respondents' performance on free-recall questions suggested that both Speech Rate and Duration had significant main effects on L1 Chinese listening comprehension of online L2 English lectures, and that there was a significant interaction between Speech Rate and Duration on the participants' performance with free-recall questions in the study.

Respondents' performance on cued-recall questions suggested that the factor of Duration had a significant main effect on L1 Chinese listening comprehension of online L2 English lectures, but the factor of Speech Rate did not. There was a significant interaction between Speech Rate and Duration on the participants' performance with free-recall questions in the study.

In the following sections, the effects of Speech Rate, Duration, and the interactions between Speech Rate and Duration on the respondents' performance in the study will be discussed from several perspectives.

Effect of Speech Rate on Listening Comprehension Performance

Carpenter and Just (1977) state that, "normal speech proceeds at a rate of about 160 words per minute" (p. 110). Many researchers, such as Foulke and Sticht (1969), Sticht (1971), Carver (1973), Griffiths (1990a,1992), have studied the relationship between Speech Rate and NNS listening comprehension. The fast Speech Rates applied in their studies ranged from 200 wpm to 400 wpm. Buck (2001) has argued that the threshold for Speech Rate at which listeners' comprehension decreases is about 275 wpm. In Carver's (1973) study about the "threshold in speech rate beyond which comprehension drops" (p. 118), the rate varies between 75 wpm and 400 wpm. Griffiths (1990a) set the slow Speech Rate in his study to 100 wpm. In his 1992 study, the slow Speech Rate was about 127 wpm. Therefore, the factor of Speech Rate for the current study contained three levels: 1-slow (100 wpm), 2-normal (160 wpm), and 3-fast (300 wpm).

Griffiths (1992) found that Japanese participants achieved significantly higher mean test scores with a slow Speech Rate (about 127 wpm) than with a fast Speech Rate (about 250 wpm). In the current study, this is not the case. Participants' mean scores from S1D3 and S1D2 (slow Speech Rate at 100 wpm) were not higher than those from S2D3 and S2D2 (normal Speech Rate at 160 wpm). Findings from the current study did not support Zhao's (1997) argument that, "students' listening comprehension improved ... by slowing down the speech rate" (p. 49). Zhao's study had a different setting from the current study, and 79% of the participants reported that slower speeds helped listening. Le's (2006) study about the effects of various Speech Rates

on ESL learners' listening comprehension found no significant differences between groups. Such results were not compatible with results from the current study.

In the current study, the factor of Speech Rate did show a significant main effect on participants' listening comprehension on free-recall questions, but not on cued-recall questions. Explanations for the differences between findings from the current study and that of the previous literature could be several. First, the current study limited the respondents' L1 to Chinese, while previous studies did not select L1 Chinese listeners as the only target group. L1 influence could be the cause. Second, in the current study, participants listened to audios with various Speech Rates and Durations. The inclusion of duration as one of the independent variables for this study may have resulted in different findings. As seen in the findings from Chapter 4, the interactions between Speech Rate and Duration were significant all the time. Third, in the current study, respondents had a variety of academic and geographic backgrounds. They were not all current students in the United States. Some of the respondents lived in China, and some were not college students. A variety of learning and living experiences may have also affected the respondents' English abilities. Last but not least, in the current study two different types of listening comprehension questions were applied. However, in previous literature, no such study was applied.

Effect of Duration on Listening Comprehension Performance

The results from the two-way repeated measures ANOVA suggested that there were significant main effects of Duration on respondents' performance on both free-recall questions and cued-recall questions. It can be concluded that regardless of the question type, respondents performed better on fast audios with short lengths, and worse on long audios.

The differences between question types led to the idea that when we study L1 Chinese learners' listening comprehension of online L2 English lectures, we need to pay attention to the settings for the listening comprehension questions. It was not within the scope of the current study to discuss the assessment methods for listening comprehension, but it may shed some light on the topic.

Literature about the effects of duration on listening comprehension are very limited. It is not known where the significance really came from. It is therefore worthwhile to have future research that may investigate the effects of lecture duration on L2 English learners' performance for listening comprehension, with more detailed design and settings for the listening comprehension questions. Cognitive load research should also be included in future research to investigate the effects of lecture duration as it is related to the time that the listeners have during listening.

Interaction of Speech Rate and Duration

The results from the two-way repeated measures ANOVA suggested that there was a significant interaction between Speech Rate and Duration on respondents' scores in the study. It is suggested that respondents' performance in answering the listening comprehension questions was affected by the interaction of Speech Rate and Duration from the audios that were presented to the participants in the study, and that both Speech Rate and Duration impacted respondents' performance together.

Generally, the lowest scores occurred for S1D1, meaning that the respondents performed the worst in the conditions where the lectures were presented at a slow speed and short duration. The highest scores occurred for S2D3, indicating that respondents performed better under the conditions where the lectures were presented at a long duration and normal speech rate. The

crossing of the lines in the Profile Plots (Appendix E) also showed the interactions between Speech Rate and Duration. However, the interaction between Speech Rate and Duration worked differently on different combinations of speed and duration in the study, and on different types of questions.

From the statistical results about respondents' performance on the free-recall listening questions, pairwise comparisons showed that significant differences only happened between the S1 and S2 groups, and between the S1 and S3 groups. No significant differences were found between the S2 and S3 groups (Table 7). This suggested that respondents' performances on free-recall questions were not significantly different among various Speech Rates groups. Normal Speech Rate and fast Speech Rate did not affect the performance on free-recall questions. Similarly, though a significant main effect of Duration was found, pairwise comparisons showed that significant differences only happened between the D1 and D2 groups, and between the D2 and D3 groups. No significant differences were found between the D1 and D3 groups (Table 8). This suggested that respondents' performances on free-recall questions were not significantly different among various Durations groups. Short Duration and long Duration did not affect the performances on free-recall questions .

From the statistical results about respondents' performance on the cued-recall questions, a significant main effect of Duration was found. Pairwise comparisons showed that significant differences only occurred between the D1 and D2 groups, and between the D1 and D3 groups. No significant differences were found between the D2 and D3 groups (Table 12). This suggested that respondents' performances on cued-recall questions were not significantly different among various Duration groups. Medium Duration and long Duration did not affect the performances on cued-recall questions .

Summary

When L1 Chinese learners took online L2 English lectures on a test of free-recall questions, it was suggested that both the Speech Rate and the Duration of L2 English academic lectures affected the listeners' listening comprehension, and the interaction of Speech Rate and Duration affected listening comprehension, too. However, there were no significant differences between participants' performances with normal Speech Rate group and fast Speech Rate group on free-recall questions. And there were no significant differences between participants' performances with short Duration group and long Duration group on free-recall questions.

When L1 Chinese learners took online L2 English lectures on a test of cued-recall questions, it was suggested that the Duration of L2 English academic lectures affected the listeners' listening comprehension, and the interaction of Speech Rate and Duration affected listening comprehension, too. It was suggested that the Speech Rate of L2 English academic lectures did not affect the listeners' listening comprehension. Also, there were no significant differences between participants' performances with medium Duration group and long Duration group on free-recall questions.

Results did not show that a slower speech rate led to better listening comprehension, nor the fast speech rate to worse listening comprehension. Due to the interactions between speech rate and duration, the participants in the current study performed differently on lectures with different combinations of Speech Rate and Duration. Generally speaking, the participants performed better with lectures of normal speed (160 wpm) and durations longer than 2 minutes, and performed worse with lectures of extreme combinations such as short duration and slow speed, and long duration and fast speed. The effects of speech rate and duration for the lectures caused respondents to perform differently on various types of questions.

Practical Implications

The ultimate purpose of the current exploratory study was to investigate the feasibility and desirability of creating learner-controlled instructional tools to adjust lecture speed and duration for online lecture learners to fit their learning needs. This study's findings demonstrate the importance of the speech rate for online lectures in relation to L1 Chinese listeners' comprehension of the lectures. There are practical implications for online instructors and instructional designers.

Because of limited number of participants in the current study, it is not possible to discuss further research related to other fields beyond this study. If similar studies with more participants and random assignments are to be applied, it is worthy of further considerations to collaborate works from various fields.

Online Instructors

Findings from the current study revealed possible effects of speech rate and combined effects of speech rate and duration. These specific findings may provide online instructors with incentives to pay attention to the speech rate and duration of the lectures. When necessary, appropriate combinations of speech rate and duration may help learners to better comprehend the lectures. If the durations for lectures are not to be changed, online instructors may notice that the speech rate of the lectures affects the comprehension of the listeners, and it is not necessarily the case that a slower speech rate will lead to better listening comprehension in general. Normal or fast speech rates would be better for listeners. Instructors need to pay attention to the length of the lectures. If possible, instructors should control the length of their lectures, aiming for around two to six minutes for each topic.

Instructional Designers for Online Lectures

Findings from the current study might offer instructional designers some insight into whether and how they would employ learner-controlled tools for online lecture learners.

The research about the effects of learner control, learner's self-regulation, and learner's motivation to learn is out of the scope of the current study. However, the provision of learner-controlled tools that target speech rate and duration for learners of online L2 English lectures should be a consideration for instructional designers. Kinzie (1990) argues that, "provision of learner control allows students to tailor their instructional experience to suit personal needs and interests, thus increasing instructional relevance and encourage continuing motivation" (p. 8). The existence of individual differences in English proficiency level, prior knowledge, and personal preference also argues for the importance of self-controlled tools for learners in listening to online L2 English lectures.

Instructional designers for online L2 English lectures can adopt results from the current study to develop speed/length-control tools that are embedded into online lectures, allowing listeners to adjust the speech rate/length of lectures based on their own learning needs and listening abilities. The findings show that all three levels in the current study affect respondents' listening comprehension performance. Instructional designers might design tools with various speech rate options for learners, as all 15 participants in Zhao's (1997) study, "reacted positively to the use of computers to control speech rate" (p. 60). Therefore, online lecture learners might welcome such opportunities.

Study Limitations

This study was designed to explore the effects of speech rate and duration for online L2 English lectures on L1 Chinese learners' listening comprehension. This study was limited to

asynchronous online lectures. Therefore, it might limit the impact that findings from the current study have for synchronous online learning because all data collection procedures assumed an asynchronous environment where the respondents participated at different locations and times.

A total of twenty eight participants completed all of the study questions, and the respondents carried various English learning experiences, and academic and geographic backgrounds. This may limit the impact that findings from the current study have for L1 Chinese learners of L2 English online lectures, as different prior experience and living conditions may add variables that affect learners' English abilities. The limited number of participants that completed the study may also impact the generalizability of the findings.

Another limitation was the length of the online experiment. The current study was designed to contain nine audio segments that represented nine combinations of speech rate and duration. Such a design made it time-consuming for the participants to finish the study. The most common comment left by the 28 participants was that the study was too long. Having people do all the nine sections in the current study built up the time burden for the participants. If random assignment to smaller number of sections was adopted for the quasi-experiment, it would cut back on the time demands and might attract more participants for the current study. In addition, if the study audio segments were recorded by real lecturers instead of being created through manipulated software, the quality of the audio might have been better. Therefore, the participants would have had better experiences during the study.

The last limitation comes from the settings for the question types in the study. In the current study, the free-recall questions asked the respondents to write down anything they could recall from the audios. There were no right or wrong answers for this type of question. Any answers the participants provided might lead to a point increase. However, the other type of

question asked about details in the audio, and required the participants to write down specific answers to earn points. Technically, any answers to the free-recall questions could lead to scores recorded for the respondents' performance. However, the cued-recall questions asked for correct answers to earn points. This might have led to a tendency that respondents earned higher scores on free-recall questions, therefore performing better overall on this specific type of question.

Future Research

Findings from the current study have helped to lay a foundation for online lecturers and instructional designers of online lectures that takes into consideration the effects of speech rate and duration. This study was designed to explore the effects caused by speech rate and duration, and to provide support for online lecturers and designers when they work on online lectures. Therefore, findings from this study provide starting points for future research in several areas.

First, although findings from the current study showed significant effects of speech rate and interactions between speech rate and duration on respondents' performance in the study, it is not known how duration and speech rate work together to affect listeners' understanding of online lectures. As discussed in previous sections, participant performance on audios S3D3 and S1D1 was among the worst, but their performance on S2D3 was better. However, the duration of S2D3 was only 4 minutes and 12 seconds long. It was still relatively short compared to real-life lectures that might take up to 20 or 30 minutes. It is still not clear whether a particular, extended lecture duration would cause serious problems for listeners. To better understand the effects of speech rate and duration on listening comprehension, it is recommended that future research would include more levels of variation for the variables, and recruit more participants to explore the detailed effects.

Second, the online lecture tool literature shows very limited studies designed to investigate the effects of lecture speed and duration. It is recommended that future research provide more answers regarding the development of speed/length-adjustment tools provided for L2 English listeners.

Third, currently there are media apps such as YouTube and Bilibili that provide viewers with the option to adjust play speed to meet their needs. For example, YouTube viewers can choose from eight playback speed options. The speed can be set to as slow as 0.25, and as fast as 2 times the original. Bilibili also allows its viewers to choose from six speed options, ranging from 0.5x to 2x. The technology is available, and future research is needed to investigate how learners use the technology, as well as how to inspire learners to use the technology effectively for online lectures.

References

- Anderson, J. R. (2005). *Cognitive psychology and its implications* (6th ed.). Worth.
- Anderson-Hsieh, J., & Venkatagiri, H. (1994). Syllable duration and pausing in the speech of Chinese ESL speakers. *TESOL Quarterly*, 28(4), 807-812.
- Agnesia, R. H. (2010). Features affecting task-motivation in English for academic purposes online learning. *University of Hawai'i Second Language Studies Paper 29 (1)*.
- Asoodar, M., Atai, M. R., Vaezi, S., & Marandi, S. S. (2014). Examining effectiveness of communities of practice in online English for academic purposes (EAP) assessment in virtual classes. *Computers & Education*, 70, 291-300.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory : a proposed system and its control processes. *Psychology of Learning and Motivation, Vol. 2, Pp. 89-165*.
- Baldwin, C. L. (2012). *Auditory cognition and human performance: Research and applications*. CRC Press.
- Beare, K. (November 18, 2019). How many people learn English? Retrieved May 12, 2020, from <https://www.thoughtco.com/how-many-people-learn-english-globally-1210367>.
- Bensoussan, M., Avinor, E., Ben-Israel, B., & Bogdanov, O. (2006). CMC among multilingual students of English for academic purposes: linguistic and sociolinguistic communicative factors in online written responses. *language@ internet*, 3(2).
- British council. (2013). THE ENGLISH EFFECT. Retrieved May 12, 2020, from <https://www.britishcouncil.org/sites/default/files/english-effect-report-v2.pdf>.
- Brown, G. 1995. Dimensions of difficulty in listening comprehension.
- In Mendelsohn, D. and Rubin, J. (eds.) *A guide for the teaching of second language listening*. San Diego: Dominic Press. p. 59–73.

- Buck, G. (1995). How to become a good listening teacher. *A guide for the teaching of second language listening*, 113-131.
- Buck, G. (2001). *Assessing listening*. Cambridge University Press.
- Carpenter, P. A., & Just, M. A. (1977). Reading comprehension as eyes see it. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 109-140). Hillsdale, New Jersey: Lawrence Erlbaum Associate.
- Carrell, P. L., Dunkel, P. A., & Mollaun, P. (2004). The effects of notetaking, lecture length, and topic on a computer-based test of ESL listening comprehension. *Applied Language Learning*, 14, 83–105.
- Carroll, B. J. (1980). *Testing communicative performance : an interim study* (1st ed., Ser. Language teaching methodology series). Pergamon Press.
- Carver, R. (1973). Effect of increasing the rate of speech presentation upon comprehension. *Journal of Educational Psychology*, 65, 118- 126.
- Chaudron, C. 1995. Academic Listening. In Mendelsohn, D. and Rubin, J. (eds.) *A guide for the teaching of second language listening*. San Diego: Dominie Press. p. 74–96
- Chamot, A. U., & O'malley, J. M. (1987). The cognitive academic language learning approach: A bridge to the mainstream. *TESOL quarterly*, 21(2), 227-249.
- Chang, C.-C., Tseng, K.-H., & Tseng, J.-S. (2011). Is single or dual channel with different english proficiencies better for english listening comprehension, cognitive load and attitude in ubiquitous learning environment? *Computers & Education*, 57(4), 2313–2321.
<https://doi.org/10.1016/j.compedu.2011.06.006>

- Cheng, H.-fang. (2004). A comparison of multiple-choice and open-ended response formats for the assessment of listening proficiency in english. *Foreign Language Annals*, 37(4), 544–553.
<https://doi.org/10.1111/j.1944-9720.2004.tb02421.x>
- Chi, Y. (2011). *Validation of an academic listening test: Effects of "breakdown" tests and test takers' cognitive awareness of listening processes* (Doctoral dissertation, University of Illinois at Urbana-Champaign).
- Chiu, Y.-F., & Neel, A. (2020). Predicting intelligibility deficits in Parkinson's disease with perceptual speech ratings. *Journal of Speech, Language, and Hearing Research*, 63(2), 433–443.
- Cottam M.E., Savenye W.C. (2014) The Effects of Visual and Textual Annotations on Spanish Listening Comprehension, Incidental Vocabulary Acquisition, and Cognitive Load. In: Orey M., Jones S., Branch R. (eds) *Educational Media and Technology Yearbook. Educational Media and Technology Yearbook*, vol 38. Springer, Cham
- Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, 26, 197-223.
- Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin Review*, 24, 1158-1170.
- Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage (PAS). *Perception and Psychophysics*, 5, 365-373.
- De Bodt, M., Patteeuw, T., & Versele, A. (2015). Temporal variables in voice therapy. *Journal of Voice*, 29(5), 611–617. <https://doi.org/10.1016/j.jvoice.2014.12.001>
- Dechert, H. W., & Raupach, M. (1980). *Temporal Variables in Speech : Studies in Honour of Frieda Goldman-Eisler* (Ser. Janua linguarum. series maior, v. 86). De Gruyter Mouton. INSERT-MISSING-URL.

- DeCarrico, J., & Nattinger, J. R. (1988). Lexical phrases for the comprehension of academic lectures. *English for Specific Purposes*, 7(2), 91–102.
- Derwing, T. M. (1990). Speech rate is no simple matter: rate adjustment and NS-NNS communicative success. *Studies in Second Language Acquisition*, 12(3), 303–313.
- Dick, W., Carey, L., & Carey, J. O. (2005). *The systematic design of instruction* (6th ed.). Pearson/Allyn and Bacon.
- Ehrlich, M.-F. (1982). An experimental study of the relationship between comprehension and memorization of a text. In J.-F. Le Ny & W. Kintsch (Eds.), *Language and comprehension* (pp. 157-178). Amsterdam, The Netherlands: North Holland Publishing Co.
- Eltawila, M. (2009). The neglect of listening activities among Egyptian preparatory school EFL teachers.
- ETS. (2021). *TOEFL iBT Listening Section*. <https://www.ets.org/toefl/test-takers/ibt/about/content/listening/>.
- Exam English Ltd. (2019). *TOEFL Listening*. Exam English. https://www.examenglish.com/TOEFL/toefl_listening.html
- Fahmy, J. J., & Bilton, L. (1989). Using authentic texts: a study skills module. *Guidelines: A Periodical for Classroom Language Teachers*, 11(2), 74–87.
- Field, J. (2004). An insight into listeners' problems: too much bottom-up or too much top-down? *System: An International Journal of Educational Technology and Applied Linguistics*, 32(3), 363–377.
- Fischer, R. (1996). Interactive video listening comprehension in foreign language instruction: Development and evaluation. San Marcos, TX: Southwest Texas State University.

- Flowerdew, J. (1994 a). Academic listening: research perspectives. Cambridge: Cambridge University Press.
- Flowerdew, J. (1994 b). Research of relevance to second language lecture comprehension: An overview. *Academic listening: Research perspectives*, 7-29.
- Flowerdew, J., & Miller, L. (2005). *Second language listening: Theory and practice*. Cambridge University Press.
- Flowerdew, J., & Miller, L. (2014). Dimensions of academic listening. In Celce-M. Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (4th ed.) (pp. 90- 103). Boston, MA: National Geographic Learning/ Cengage Learning.
- Foulke, E., and Sticht, T. (1969). Review of the research on the intelligibility and comprehension of accelerated speech. *Psychological Bulletin*. 72 (1), 50 -62.
- Galbraith, J. D., & Spencer, S. G. (2001). Variable speed playback of digitally recorded lectures: Evaluating learner feedback. *Provo, Utah: Center for Instructional Design, Brigham Young University*. Retrieved November 18, 2020 from <https://www.enounce.com/docs/BYUPaper020319.pdf>.
- Gerald E. Schneider. *Brain Structures and Its Origins*. Spring 2014. Massachusetts Institute of Technology: MIT OpenCouseWare, <https://ocw.mit.edu/>. License: Creative Commons BY-NC-SA.
- Ginther, A. (2002). Context and content visuals and performance on listening comprehension stimuli. *Language Testing*, 19(2), 133-167.
- Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28(1), 55–75.

- Goh, C.C.M. (2018). Academic Listening. In *The TESOL Encyclopedia of English Language Teaching* (eds J.I. Lontas, M. DelliCarpini, W. A. Renandya, and G. Hu).
doi:[10.1002/9781118784235.eelt0595](https://doi.org/10.1002/9781118784235.eelt0595)
- Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech, Language, and Hearing Research*, 36(6), 1276-1285.
- Graham, S., & Santos, D. (2015). *Strategies for second language listening: Current scenarios and improved pedagogy*. Springer.
- Griffiths, R. (1990a). Speech rate and NNS comprehension: A preliminary study in time-benefit analysis. *Language Learning*, 40(3), 311-336.
- Griffiths, R. (1990b). Pausology and Listening Comprehension: theory, research and practice. *JALT Journal*, Vo. 12, No. 1 (May 1990), 99-120.
- Griffiths, R. (1992). Speech rate and listening comprehension: Further evidence of the relationship. *TESOL quarterly*, 26(2), 385-390.
- Grosjean, F., & Lane, H. (1981). Temporal variables in the perception and production of spoken and sign languages. *Perspectives on the study of speech*, 207-237.
- Gruba, P., & Suvorov, R. (2020). Technology and Second Language Listening. Retrieved November 10, 2020, from https://minerva-access.unimelb.edu.au/bitstream/handle/11343/240689/Gruba-Suvorov2020_ReferenceWorkEntry_TechnologyAndSecondLanguageLis.pdf
- Guichon, N., & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. *System*, 36(1), 85-93.

- Hildyard, A., & Olson, D. R. (1982). On the comprehension and memory of oral vs. written discourse. In Deborah Tannen (ed.) *Spoken and Written Language: Exploring Orality and Literacy* (pp. 19–33). Norwood.
- Hoven, D. 1997. Improving the management of flow of control in computer-assisted listening comprehension tasks for second and foreign language learners. Brisbane: Unpublished doctoral dissertation, University of Queensland. <https://auspace.athabascau.ca/handle/2149/2294>
- Huang, J. (2004). Voices from Chinese students: professors' use of English affects academic listening. *College Student Journal*, 38(2), 212–212.
- Huang, J. (2005). Challenges of academic listening in English: reports by Chinese students. *College Student Journal*, 39(3), 553–553.
- Hudson, T., & Llosa, L. (2015). Design issues and inference in experimental L2 research. *Language Learning*, 65(S1), 76–96. <https://doi.org/10.1111/lang.12113>
- Hustad, K. C., Jones, T., & Dailey, S. (2003). Implementing speech supplementation strategies: effects on intelligibility and speech rate of individuals with chronic severe dysarthria. *Journal of Speech, Language, and Hearing Research*, 46(2), 462–74.
- International Listening Association (ILA). (1995). A ILA definition of listening. *The Listening Post*, 53(1), 4–5.
- Jacewicz, E., Fox, R. A., O'Neill, C., & Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language variation and change*, 21(2), 233–256.
- Jensen, C., & Hansen, C. (1995). The effect of prior knowledge on EAP listening-test performance. *Language Testing*, 12(1), 99–119.
- Johnson, M. E. D. (1998). The Effects of Speech Compression on Recall in a Multimedia Environment (Doctoral dissertation, Virginia Tech).

- Jones, L. (2008). Listening comprehension technology: building the bridge from analog to digital. *Calico Journal*, 25(3), 400–419.
- Jones, L. C., & Plass, J. L. (2002). Supporting listening comprehension and vocabulary acquisition in french with multimedia annotations. *The Modern Language Journal*, 86(4), 546–561.
- Kim, Y. S., & Phillips, B. (2014). Cognitive correlates of listening comprehension. *Reading Research Quarterly*, 49(3), 269-281.
- Kinzie, M. B. (1990). Requirements and benefits of effective interactive instruction: Learner control, self-regulation, and continuing motivation. *Educational Technology Research and Development*, 38(1), 5-21.
- Le, F. (2006). Faster, normal or slower? - the effects of speech rates on high-intermediate ESL learners' listening comprehension of academic lectures (Master Thesis, Iowa State University).
- Liu, N. F. (2002). *Processing problems in L2 listening comprehension of university students in Hong Kong* (Doctoral dissertation, The Hong Kong Polytechnic University).
- Liu, M., & Thondhlana, J. (2015). A study of Chinese university EFL learners' foreign language listening anxiety, listening strategy use and academic listening performance. *Indonesian Journal of English Language Teaching*, 10(1), 30-47.
- Liu, Y. C. (2011). *The utilization of listening strategies in the development of listening comprehension among skilled and less-skilled non-native English speakers at the college level* (Doctoral dissertation, Texas A & M University).
- Love, T., Walenski, M., & Swinney, D. (2009). Slowed speech input has a differential impact on on-line and off-line processing in children's comprehension of pronouns. *Journal of Psycholinguistic Research*, 38(3), 285–304.
- Lynch, T. (1997). Life in the slow lane: observations of a limited l2 listener. *System*, 25(3), 385–98.

- Lynch, T. (1998). Theoretical Perspectives on Listening. *Annual Review of Applied Linguistics*, 18, 3-19. doi:10.1017/S0267190500003457
- McGregor, G. (1986). Listening outside the participation framework. In McGregor, G. and White, R. S. (eds) *The art of listening*, 55-72.
- MELANLIOĞLU, D. (2013). Impacts of authentic listening tasks upon listening anxiety and listening comprehension. *Educational research and reviews*, 8(14), 1177-1185.
- Mendelsohn, D. J., & Rubin, J. (1995). *A guide for the teaching of second language listening*. San Diego, CA: Dominie Press.
- Miller, P., & Endo, H. (2004). Understanding and meeting the needs of ESL students. *Phi Delta Kappan*, 85(10), 786-791.
- Miller, J. L., Grosjean, F., & Lomanto, C. (1984). Articulation rate and its variability in spontaneous speech: A reanalysis and some implications. *Phonetica*, 41(4), 215-225.
- Mirzaei, M. S., Meshgi, K., Akita, Y., & Kawahara, T. (2017). Partial and synchronized captioning: A new tool to assist learners in developing second language listening skill. *ReCALL*, 29(2), 178-199.
- Nagle, S. J., and Sanders, S. L. (1986). Comprehension theory and second language pedagogy. *TESOL Quarterly*, 20(1), 9-26.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Nordquist, R. (April 01, 2019). Definition and example of a lingua franca. Retrieved May 12, 2020, from <https://www.thoughtco.com/what-is-a-lingua-franca-1691237>.
- O'Connell, D. C., & Kowal, S. (1983). Pausology. *Computers in language research*, 2(19), 221-301.
- O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition* (Ser. The cambridge applied linguistics series). Cambridge University Press.

- Osada, N. (2004). Listening comprehension research: A brief review of the past thirty years. *Dialogue*, 3(1), 53-66.
- Oxford, R. L. (1990). *Language Learning Strategies: What Every Teacher Should Know*. New York: Heinle & Heinle.
- Pan, Z. (2017). *Assessing listening for Chinese English learners : developing a communicative listening comprehension test suite for CET* (Ser. China perspectives series). Routledge. INSERT-MISSING-URL.
- Park, S. (2016). Measuring fluency: Temporal variables and pausing patterns in L2 English speech.
- Prince, P. (2012). Writing it down: issues relating to the use of restitution tasks in listening comprehension. *Tesol Journal*, 3(1), 65–86.
- Rahmatian, R., & Armiun, N. (2011). The effectiveness of audio and video documents in developing listening comprehension skill in a foreign language. *International Journal of English Linguistics*, 1(1), 115.
- Ramírez Verdugo, D., & Alonso Belmonte, I. (2007). Using digital stories to improve listening comprehension with Spanish young learners of English. *Language Learning & Technology*, 11(1), 87-101.
- Raupach, M. (1980). Temporal variables in first and second language speech production. *Temporal variables in speech*, 263-270.
- Rost, M. (1990). *Listening in Language Learning*. Harlow: Longman
- Rost, M. (2013). *Teaching and researching: Listening*. Routledge.
- Rubin, J. (1990). Improving foreign language listening comprehension. *Georgetown University round table on languages and linguistics*, 309-316.

- Rubin, J. (1994). A Review of Second Language Listening Comprehension Research. *The Modern Language Journal*, 78(2), 199-221. doi:10.2307/329010
- Rumelhart, D. E. (1975). Notes on a schema for stories. In *Representation and understanding: Studies in Cognitive Science*. By D. G. Bobrow & A. Collins. (pp. 211–236). Elsevier Science & Technology. <https://doi.org/10.1016/B978-0-12-108550-6.50013-6>
- Schneider, W., and Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and identification. *Psychological Review*, 84 (1), 1-66.
- Serri, F., Boroujeni, A. J., & Hesabi, A. (2012). Cognitive, metacognitive, and social/affective strategies in listening comprehension and their relationships with individual differences. *Theory & Practice in Language Studies*, 2(4).
- Shenoy, S. P. (2016). An adaptive time reduction technique for video lectures (Order No. 10110873). Available from ProQuest Dissertations & Theses Global. (1797593365). Retrieved November 18, 2020, from <http://login.ezproxy.lib.vt.edu/login?url=https://www-proquest-com.ezproxy.lib.vt.edu/docview/1797593365?accountid=14826>
- Shenoy, S. P., Amresh, A., Femiani, J., & 12th International Conference on e-Learning, ICEL 2017 12 2017 06 01 - 2017 06 02. (2017). An adaptive time reduction technique for video lectures. *Proceedings of the International Conference on E-Learning, ICEL*, 267-276.
- Shiffrin, R. M., and Schneider, W. (1977). Controlled and automatic human information processing: II. Detection, search and identification. *Psychological Review*, 84 (2), 127-190.
- Shannon, C. & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.
- Shang, H. F. (2008). Listening strategy use and linguistic patterns in listening comprehension by EFL learners. *The International Journal of Listening*, 22, 29-45.

- Shohamy, E., & Inbar, O. (1991). Validation of listening comprehension tests: the effect of text and question type. *Language Testing*, 8(1), 23–40. <https://doi.org/10.1177/026553229100800103>
- Singh, S., Bucks, R., & Cuerden, J. (2001). Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech. *Aphasiology*, 15(6), 571–583.
- Smit, T. C. (2006). *Listening comprehension in academic lectures: A focus on the role of discourse markers* (Doctoral dissertation, University of South Africa).
- Sticht, T. (1971). *Learning by listening relation to aptitude, reading, and rate-controlled speech: additional studies*. Human Resources Research Organisation. Technical Report
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257-285.
- Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, VIC: ACER Press.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19-30). New York: Cambridge University Press.
- Sweller, J. (2010). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257-258 (1988). doi:10.1207/s15516709cog1202_4.
- Tauroza, S., and Allison, D. (1990). Speech rates in British English. *Applied Linguistics*, 11, 90-105.
- Temple, L. (2000). Second language learner speech production. *Studia Linguistica*, 54(2), 288–297. <https://doi.org/10.1111/1467-9582.00068>
- Thompson, I., & Rubin, J. (1996). Can strategy instruction improve listening comprehension? *Foreign Language Annals*, 29(3), 331–42.
- Tsao, Y. C., Weismer, G., & Iqbal, K. (2006). Interspeaker variation in habitual speaking rate: additional evidence. *Journal of Speech, Language, and Hearing Research : Jslhr*, 49(5), 1156–64.

Tsui, A. B. M., and Fullilove, J. (1998). Bottom-up or top-down processing as a discriminator of L2 listening performance. *Applied Linguistics* 19, 4: 432-451.

Tyler, L. K., & Warren, P. (1987). Local and global structure in spoken language comprehension. *Journal of Memory and Language*, 26(6), 638–657.
[https://doi.org/10.1016/0749-596X\(87\)90107-0](https://doi.org/10.1016/0749-596X(87)90107-0)

Two-Way Repeated Measures ANOVA.

<https://www.open.ac.uk/socialsciences/spsstutorial/files/tutorials/two-repeated-measures-ANOVA.pdf>

Underwood, M. *Teaching Listening*. New York: Longman, 1989.

Vandergrift, L. (1997). The comprehension strategies of second language (French) listeners: A descriptive study. *Foreign Language Annals*, 30(3): 387-409.

Vandergrift, L. (1999). Facilitating second language listening comprehension: acquiring successful strategies. *ELT Journal*, 53(3), 168–76.

Vandergrift, L. (2006). Second language listening: listening ability or language proficiency. *Modern Language Journal*, 90(1), 6–18.

Vandergrift, L. and Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge, New York.

Voss, B. (1984). *Slips of the ear : investigations into the speech perception behaviour of German speakers of English* (Ser. Tübinger beiträge zur linguistik, 254). G. Narr.

Wagner-Loera, D. (2018). Flipping the ESL/EFL Classroom to Reduce Cognitive Load: A New Way of Organizing Your Classroom. In *Innovations in Flipping the Language Classroom* (pp. 169-184). Springer, Singapore.

- Wang, H. (2018). *Testing Lecture Comprehension Through Listening-to-summarize Cloze Tasks : The Trio of Task Demands, Cognitive Processes and Language Competence*. Springer.
- Watanabe, M., & Rose, R.L. (2010). Pausology and Hesitation Phenomena in Second Language Acquisition. Retrieved May 28, 2020, from <https://www.semanticscholar.org/paper/Pausology-and-Hesitation-Phenomena-in-Second-Watanabe-Rose/528aa5a14e6ecfa08d17b559c0d2e3fb2c9f51bd>
- Wei, R., & Su, J. (2012). The statistics of English in China: An analysis of the best available data from government sources. *English Today*, 28(3), 10-14. Retrieved from https://www.researchgate.net/publication/236986651_The_statistics_of_English_in_China
- Willis, G. B., & Artino, A. R. J. (2013). What do our respondents think we're asking? using cognitive interviewing to improve medical education surveys. *Journal of Graduate Medical Education*, 5(3), 353–6. Retrieved March 22, 2021, from <https://doi.org/10.4300/JGME-D-13-00154.1>
- Wong, S.-L. C. (1988). What We Do and Don't Know about Chinese Learners of English: A Critical Review of Selected Research. *RELJ Journal: A Journal of Language Teaching and Research in Southeast Asia*, 19(1), 1–20. Retrieved June 8, 2020, from <https://doi-org.ezproxy.lib.vt.edu/10.1177/003368828801900101>
- Worthington, D. L., & Fitch-Hauser, M. E. (2018). *Listening : Processes, Functions, and Competency: Vol. Second edition*. Routledge.
- Xiao, C., & Sidhu, G. K. (2018). LISTENING COMPREHENSION STRATEGIES EMPLOYED BY EFL TERTIARY STUDENTS IN CHINA: A CASE STUDY. Retried from <https://pdfs.semanticscholar.org/9f9c/1bb0177065fb7241e3ea2ea7dad4ee683bf7.pdf>

Yaghi, J. (November 17, 2019). Most spoken languages in 2020 and why you need to know them.

Retrieved May 12, 2020, from <https://cudoo.com/blog/most-spoken-languages-in-2020-and-why-you-need-to-know-them/>.

Zellner, B. (1994). Pauses and the temporal structure of speech. In E. Keller (ed.) *Fundamentals of speech synthesis and speech recognition*. (pp. 41-62). Chichester: John Wiley. (pp. 41-62). John Wiley.

Zhao, Y. (1997). The effects of listeners' control of speech rate on second language comprehension. *Applied Linguistics*, 18(1), 49–68.

Appendix A: Quasi-Experiment Questions

The link to the quasi-experiment is

https://virginiatech.qualtrics.com/jfe/form/SV_cTrjBlDszo8oCk5.

Tag: S1D3 Duration: 5'45" SR: 105

Questions:

Q01: Please write down anything you can remember from the lecture you just listened to.

Q02: Why didn't the clerk at the grocery store recognize you when you changed your glasses?

Q03: The lecturer states that we are not completely good at judging our own learning. Please list at least one of the reasons that she holds for the statement.

Q04: What kinds of moments are necessary before we achieve a new level of understanding?

Tag: S2D3 Duration: 4'12" SR: 163

Questions:

Q05: Please write down anything you can remember from the lecture you just listened to.

Q06: What role does emotion take during learning?

Q07: What is the surprise that she mentions when she talks about interest?

Q08: The lecturer argues that some careful thoughts are needed when we say that people learn from rewards and punishments, please list at least one of her arguments.

Q09: According to the lecturer, what does a higher IQ actually mean?

Tag: S3D3 Duration: 4'16" SR: 300

Questions:

Q10: Please write down anything you can remember from the lecture you just listened to.

Q11: The lecturer provides an example of golfers and non-golfers to support what statement she makes about learning?

Q12: Why it is not true that people at all ages learn basically the same way?

Q13: What can be concluded from the examples of her 50 year old Ph. D student and 70 year old undergraduate student?

Tag: S3D2 Duration: 2'07" SR: 300

Questions:

Q14: Please write down anything you can remember from the lecture you just listened to.

Q15: What is the difference between learning and development according to the lecturer?

Q16: What is the difference between learning and memory according to the lecturer?

Tag: S2D2 Duration: 2'42" SR: 160

Questions:

Q17: Please write down anything you can remember from the lecture you just listened to.

Q18: What topics will be covered in the course? Please list at least two of them.

Tag: S1D2 Duration: 2'41" SR: 100

Questions:

Q19: Please write down anything you can remember from the lecture you just listened to.

Q20: What is this course about?

Tag: S2D1 Duration: 1'04" SR: 160

Questions:

Q21: Please write down anything you can remember from the lecture you just listened to.

Q22: According to the lecturer, what is learning?

Tag: S3D1 Duration: 1' SR: 297

Questions:

Q23: Please write down anything you can remember from the lecture you just listened to.

Q24: What does learning depend on?

Tag: S1D1 Duration: 1'30" SR: 100

Questions:

Q25: Please write down anything you can remember from the lecture you just listened to.

Q26: What the lecturer would cover in the course?

Appendix B: VT IRB-20-514 Approval Letter



Division of Scholarly Integrity and
Research Compliance
Institutional Review Board
North End Center, Suite 4120 (MC 0497)
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-3732
irb@vt.edu
<http://www.research.vt.edu/sirc/hrpp>

MEMORANDUM

DATE: June 30, 2020
TO: Ken Potter
FROM: Virginia Tech Institutional Review Board (FWA00000572, expires October 29, 2024)
PROTOCOL TITLE: A Quantitative Research About ESL Instructional Strategies for L1 Chinese Listeners of Online L2 English Lectures
IRB NUMBER: 20-514

Based on the submitted project description and items listed in the Special Instructions section found on Page 2, the Virginia Tech Human Research Protection Program (HRPP) has determined that the proposed activity is not research involving human subjects as defined by HHS and FDA regulations.

Further review and approval by the Virginia Tech Human Research Protection Program (HRPP) is not required because this is not human research. This determination applies only to the activities described in the submitted project description and does not apply should any changes be made. If changes are made you must immediately submit an Amendment to the HRPP for a new determination. Your amendment must include a description of the changes and you must upload all revised documents. At that time, the HRPP will review the submission activities to confirm the original "Not Human Subjects Research" decision or to advise if a new application must be made.

If there are additional undisclosed components that you feel merit a change in this initial determination, please contact our office for a consultation.

Please be aware that receiving a "Not Human Subjects Research" Determination is not the same as IRB review and approval of the activity. You are NOT to use IRB consent forms or templates for these activities. If you have any questions, please contact the Virginia Tech HRPP office at 540-231-3732 or irb@vt.edu.

PROTOCOL INFORMATION:

Determined As: **Not Human Subjects Research**
Protocol Determination Date: **June 30, 2020**

ASSOCIATED FUNDING:

The table on the following page indicates whether grant proposals are related to this protocol, and which of the listed proposals, if any, have been compared to this protocol, if required.

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution

Appendix C: Responses to Background Questions

Respondent ID	years of learning English	self-evaluation	In academic program	years in academic program	years of staying in States
01	25	Intermediate	No	N/A	20
02	26	Advanced	Yes	10	
03	41	Advanced	No	N/A	7
04	18	Intermediate	Yes	3.5	
05	26	Advanced	No	N/A	15
06	30	Intermediate	No	N/A	20
07	35	Advanced	No	N/A	20
08	10	Intermediate	No	N/A	In China
09	49	Advanced	No	N/A	30
10	30	Intermediate	No	N/A	6
11	2	Intermediate	Yes	1	
12	26	Intermediate	Yes	3.5	
13	10	Intermediate	Yes	0	In China
14	9	Beginner	No	N/A	In China
15	11	Beginner	Yes	1	In China
16	10	Intermediate	Yes	6	In China
17	10	Intermediate	Yes	N/A	In China
18	14	Intermediate	Yes	N/A	In China
19	10	Intermediate	Yes	4	In China
20	30	Advanced	No	N/A	20
21	10	Intermediate	Yes	1	In China
22	44	Intermediate	No	N/A	15
23	10	Intermediate	No	N/A	10
24	13	Beginner	No	N/A	In China
25	10	Beginner	No	N/A	In China
26	9	Intermediate	Yes	N/A	In China
27	8	Intermediate	Yes	14	In China
28	12	Intermediate	No		In China

Appendix D: Raters Negotiations

This appendix describes the negotiations that were conducted to reach agreement regarding the results from the quasi-experiment and presents the procedures to be used to convert them to a numeric format. As presented in Chapter 3, free-recall questions and cued-recall questions were designed to explore possible differences in participants' comprehensions about the provided audios that were presented at various speech rates and durations to explore the effects of speech rate and duration on the listening comprehensions of online English lectures by target group of L1 Chinese listeners. Participants' responses to the questions were exported into individual PDF files and shared with an independent rater by the researcher. The researcher (as rater A) and the independent rater (as rater B) independently generated categories from the answers from the respondents for each question and rated 0 to 3 in each category for each response, with 0 indicating 'no mentioned at all', 1 as 'very briefly mentioned', 2 as 'detailed', and 3 as 'very detailed'. After independent data sets containing categories and rates were generated, the independent rater and the researcher exchanged opinions upon unagreeable parts and got agreements on final data set of categories and rates for further analysis.

Rating Results From Sample Questions By Two Raters and the Negotiated Rate

In the current exploratory study, there were two factors that were paid attention to, the Speech Rate and the Duration of the online lecture. Three levels were assigned to each factor. The three levels for factor Speech Rate were fast, normal, and slow. The three levels for factor Duration were long, medium, and short. Therefore, there were nine audios each of which carried one of the nine possible combinations. For each audio, one free-recall question and several cued-recall questions were assigned. Appendix D Table 1 presents the features of the nine audios, listening comprehension questions for each audio, and features of the questions.

Appendix D Table 1

Feature of Audios and Comprehension Questions

Audio Tag	Factor – SR	Level	Factor – D	Level	Question ID	Question Type
S1D3	slow	1	long	3	Q1	Free-recall
					Q2, Q3, Q4	detailed
S2D3	normal	2	long	3	Q5	Free-recall
					Q6, Q7, Q8, Q9	detailed
S3D3	fast	3	long	3	Q10	Free-recall
					Q11, Q12, Q13	detailed
S3D2	fast	3	medium	2	Q14	Free-recall
					Q15, Q16	detailed
S2D2	normal	2	medium	2	Q17	Free-recall
					Q18	detailed
S1D2	slow	1	medium	2	Q19	Free-recall
					Q20	detailed
S2D1	normal	2	short	1	Q21	Free-recall
					Q22	detailed
S3D1	fast	3	short	1	Q23	Free-recall
					Q24	detailed
S1D1	slow	1	short	1	Q25	Free-recall
					Q26	detailed

Rater A

Responses of the following three questions, Q10, Q18, and Q26, are chosen to present how Rater A worked on the categories and scores. Among the three questions, Q10 was an free-recall question for S1D3 (Audio 1, the fast speed and long duration audio), Q18 was a question about details in S2D2 (Audio 5, normal speed and medium duration) , and Q26 was a question about details in S1D1 (Audio 9, slow speed and short duration) .

Q10. Q10 was an free-recall question that asked the participants to write down anything they could recall from an audio that was long and presented at fast speed. The audio associated with S1D3 contained 1278 words, lasted 4 minutes and 16 seconds, and was presented at a speed

of 300 wpm. Q10 asked the respondents to write down anything they could recall from the lecture after they listened to it once.

Rater A generated 7 categories based on responses by the 28 respondents. The categories were ‘myth’, ‘passive’, ‘transformation’, ‘different person/age’, ‘prior/new knowledge’, ‘learning process/method’, and ‘example of golfer/non-golfer’. Appendix D Table 2 shows the numbers and percentages at various rates for the seven categories.

Appendix D Table 2

Numbers of Respondents at Various Rates for Categories of Q10 – Rater A

	0: not mentioned	1: very brief	2: detailed	3: very detailed
myth	24 (85.7%)	1 (3.6%)	2 (7.1%)	1 (3.6%)
passive	25 (89.3%)	3 (10.7%)	0	0
transformation	24 (85.7%)	1 (3.6%)	2 (7.1%)	1 (3.6%)
different person/age	16 (57.2%)	4 (14.3%)	6 (21.4%)	2 (7.1%)
prior/new knowledge	21 (75%)	3 (10.7%)	4 (14.3%)	0
learning process/method	25 (89.3%)	2 (7.1%)	1 (3.6%)	0
example of golfer/non-golfer	25 (89.3%)	3 (10.7%)	0	0

For each of the seven categories, a majority of the respondents did not provide useful responses that aligned with categories identified by the raters. In fact, between 57% to 89% of the responses were rated as 0 (not mentioned). Of course, the numbers of relevant responses changed from category to category. The lowest percentage of 0 occurred in category ‘different person/age’ and the highest percentage of 0 happened in categories ‘passive’, ‘learning process/method’ and ‘example of golfer/non-golfer’.

Remember that Q10 was an free-recall question that asked the participants to write down anything they could recall from the audio that was long and presented at fast speed. The lecturer talked about the last three myths held about learning and applied several examples of how

various people learn. The words ‘people’ and ‘age’ occurred many times, and that probably is the reason why in category ‘different people/age’ the rates were most dispersed. Similar explanation can be applied to results for the category ‘prior/new knowledge’. Among those two categories, ‘different people/age’ and ‘prior/new knowledge’, multiple numbers of responses appeared on a continuum from 1 to 3, indicating that respondents provided their answers at various degree of details. A similar pattern but with fewer respondents was found from the responses in category ‘myth’. Though 24 out of 28 respondents did not mention ‘myth’ in their answers, 4 of them did and answered at various degrees of details.

Q18. In S2D2, the lecturer introduced five topics that would be covered in the course, followed by brief presentations of each topic. S2D2 contained 438 words, lasted 2 minutes and 42 seconds, and was presented at a speed of 160 wpm. The speed of 160 wpm was the original speech rate of the lecturer and the duration of between 2 and 3 minutes was viewed as the normal duration of the audio.

Q18 asked the respondents to write down at least two of the topics that would be covered in the lecture after they listened to it once. Rater A generated 5 categories based on responses by the 28 respondents, they were ‘history’, ‘how is not passive’, ‘what’, ‘how and why’, and ‘individual difference/who’. Table 3 shows the numbers and percentages at various rates for the five categories.

Appendix D Table 3

Numbers of Respondents at Various Rates for Categories of Q18 – Rater A

	0: not mentioned	1: very brief	2: detailed	3: very detailed
history	21 (75%)	5 (17.9 %)	2 (7.1%)	0
How is not passive	27 (96.4%)	0	0	1 (3.6%)
what	23 (82.2%)	3 (10.7%)	0	2 (7.1%)

How and why	17 (60.7%)	4 (14.3%)	6 (21.4%)	1 (3.6%)
Individual difference	19 (67.9%)	6 (21.4%)	2 (7.1%)	1 (3.6%)

The lowest percentage of 0: not mentioned occurred on the category ‘how and why’, the second lowest was category ‘individual difference’, and the highest percentage of 0 happened to the category ‘how is not passive’ in that 27 of 28 respondent provided nothing related to the category.

The responses to this question were relatively more distributed than those for Q10, more responses were categorized along a continuum from 1 to 3. Also, more responses resided in the zone of degree 1 and degree 2. All that means more respondents were able to provided something in each category for Q18.

Q26. Q26 involved an audio of slow speed (100 wpm) and short duration (1 minutes and 30 seconds). The lecturer used this audio as an opening for the course to gain the listeners’ attention by asking them to think “what is learning?” The respondents were asked to write down what the lecturer covered in the audio after they listened to the piece of lecture once.

Rater A generated 3 categories based on responses by the 28 respondents: they were ‘previous knowledge and skills’, ‘concepts relevant for learning’, and ‘overviews of main point: complex process, what, how, why and whom’. Table 4 shows the numbers and percentages at various rates for the three categories.

Appendix D Table 4

Numbers of Respondents at Various Rates for Categories of Q26 – Rater A

	0: not mentioned	1: very brief	2: detailed	3: very detailed
Previous knowledge and skills	22 (78.5%)	4 (14.3%)	1 (3.6%)	1 (3.6%)
Concepts relevant for learning	26 (92.9%)	0	2 (7.1%)	0

Overviews of main point: complex process, what, how, why and whom	19 (67.9%)	9 (32.1%)	0	0
---	------------	-----------	---	---

It is shown in Table 4 that 32.1% of the responses mentioned the category ‘overviews of main point: complex process, what, how, why and whom’, though very briefly. Numbers for category ‘previous knowledge and skills’ showed that about 21.5% of responses talked about knowledge and skills, at various degrees. Most of the responses for category ‘concepts relevant for learning’ were marked as ‘not mentioned’.

Rater B

The categories and rates made by Rater B on the same three questions, Q10, Q18, and Q26, were presented in following Table 5, Table 6, and Table 7.

Appendix D Table 5

Numbers of Respondents at Various Rates for Categories of Q10 – Rater B

	0: not mentioned	1: very brief	2: detailed	3: very detailed
Different ages	19 (67.9%)	6 (21.4%)	2 (7.1%)	1 (3.6%)
Pre-knowledge	24 (85.8%)	2 (7.1%)	2 (7.1%)	0
Different things	25 (89.3%)	3 (10.7%)	0	0
Golf/er	25 (89.3%)	2 (7.1%)	1 (3.6%)	0
transformation	23 (82.1%)	3 (10.7%)	2 (7.1%)	0
passive	25 (89.3%)	1 (3.6%)	2 (7.1%)	0
New knowledge	23 (82.1%)	5 (17.9%)	0	0
myths	23 (82.1%)	3 (10.7%)	1 (3.6%)	1 (3.6%)

Rater B generated eight categories upon responses to Q10, among that seven were the same as what were generated by Rater A. They were categories ‘myth’, ‘passive’, ‘transformation’, ‘different ages’, ‘prior knowledge’, ‘new knowledge’, and ‘golfer’. Numbers of responses in these seven categories set for scores by Rater B were similar to that of Rater A. The only difference between the category generated by Rater A and Rater B was the category ‘learning process/method’ from Rater A and category ‘different things’ from Rater B. However,

rater A and rater B provided the same numbers of responses at degree 0 and degree 3, only one difference on degree 2. This showed that though the category title might be different, both raters agreed on rating the responses.

Appendix D Table 6

Numbers of Respondents at Various Rates for Categories of Q18 – Rater B

	0: not mentioned	1: very brief	2: detailed	3: very detailed
history	21 (75%)	6 (21.4%)	0	1 (3.6%)
difference	23 (82.1%)	4 (14.3%)	1 (3.6%)	0
Questions on learning	18 (64.3%)	1 (3.6%)	6 (21.4%)	3 (10.7%)

Rater B generated three categories upon responses to Q10, among that two were the same as what were generated by Rater A. They were categories ‘history’, and ‘difference’. Numbers of responses in these two categories set for scores by Rater B were similar to that of Rater A. Rater B generated a category titled as ‘questions on learning’, while Rater A provided two similar categories of ‘what’ and ‘how and why’. The fifth category Rater A generated was ‘how is not passive’ that only one response described in great details. In this case, such a category (‘how is not passive’) was not of great representativeness.

Appendix D Table 7

Numbers of Respondents at Various Rates for Categories of Q26 – Rater B

	0: not mentioned	1: very brief	2: detailed	3: very detailed
How & why & what & who	19 (67.9%)	9 (32.1%)	0	0
Dependence of learning	13 (46.4%)	11 (39.3%)	4 (14.3%)	0

Rater B generated two categories upon responses to Q26, and one of the two – category ‘how & shy & what & who’ meets one of the three categories generated by Rater A. And the numbers of responses at each degrees set by both the independent rater and researcher were the same, 19 (67.9%) at degree 0 (not mentioned) and 9 (32.1%) at degree 1 (very brief).

The other category generated by Rater B was ‘dependence of learning’, and Rater A provided a similar category ‘previous knowledge and skills’. However, the numbers set at the degrees by the raters varied a lot.

It can be concluded from aforementioned discussions about responses to example questions that both raters generated similar categories and close numbers for various scores to determine how a respondent performed on the listening comprehension questions.

Negotiated

In order to resolve those differences of categories and rates between the work of the two independent raters, the independent rater and I worked together to discuss the differences and generated categories and scores for each question. The independent rater and the researcher agreed that 1) combine categories generated by different raters to generate a new category of larger scope if both agreed; 2) if one category from any one of the two work contained very limited responses, it should be removed; 3) the determinations of scores 0-3 were made upon agreements of both the independent rater and the researcher.

The rates achieved after discussions between Rater A and Rater B on the same three questions, Q10, Q18, and Q26, are presented in following Table 8, Table 9, and Table 10 as examples to present how the negotiated rates look.

Appendix D Table 8

Numbers of Respondents at Various Rates for Categories of Q10 – Negotiated

	0: not mentioned	1: very brief	2: detailed	3: very detailed
Different ages	16 (57.1%)	5 (17.9%)	5 (17.9%)	2 (7.1%)
Pre/new knowledge	20 (71.4%)	2 (7.1%)	5 (17.9%)	1 (3.6%)
Example of golfer transformation	24 (85.7%)	3 (10.7%)	0	1 (3.6%)
passive	24 (85.7%)	1 (3.6%)	2 (7.1%)	1 (3.6%)
	25 (89.3%)	3 (10.7%)	0	0

myths	24 (85.7%)	1 (3.6%)	2 (7.1%)	1 (3.6%)
-------	------------	----------	----------	----------

Appendix D Table 9

Numbers of Respondents at Various Rates for Categories of Q18 – Negotiated

	0: not mentioned	1: very brief	2: detailed	3: very detailed
history	21 (75%)	4 (14.3%)	3 (10.7%)	0
Individual difference	21 (75%)	4 (14.3%)	2 (7.1%)	1 (3.6%)
Learning: what, how, why, not passive	14 (50%)	1 (3.6%)	10 (35.7%)	3 (10.7%)

Appendix D Table 10

Numbers of Respondents at Various Rates for Categories of Q26 – Negotiated

	0: not mentioned	1: very brief	2: detailed	3: very detailed
Main point: complex process, what, how, why, and whom	19 (67.9%)	8 (28.6%)	1 (3.6%)	0
Dependence on previous knowledge and skills	19 (67.9%)	7 (25%)	1 (3.6%)	1 (3.6%)

By combining similar categories generated by rater A and rater B into categories of larger scope, and removing categories that contained limited numbers of responses, it is shown that the distributions of responses on each category and every degree are improved. It is proved that the negotiated categories and scores present the responses from the 28 respondents well, and it is reliable for further analysis.

Calculation of Scores for Respondents

As discussed in previous sessions, respondents' answers were rated by the researcher and one independent raters. After both rater A and rater B finished categorizing and scoring, we discussed and finalized the categories and scores (0-3) applied for each answer.

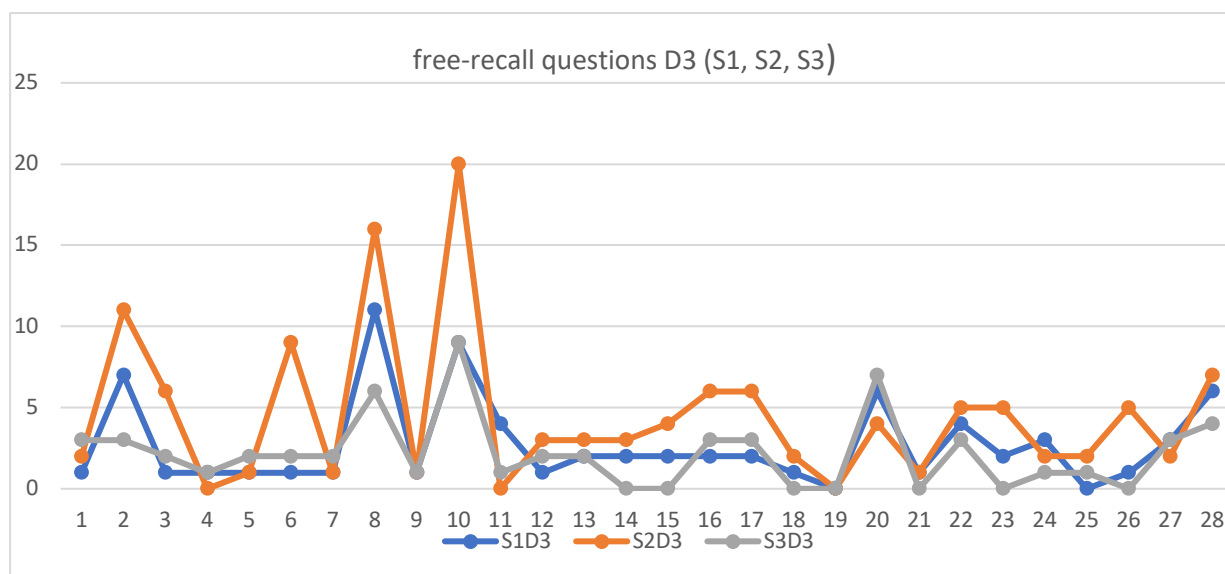
Appendix E: Respondents' Performances

Free-recalled Questions

Same Duration, Various Speech Rates. The first research question was ‘What is the effect of variations in speech rate on listening comprehension on a test of free recall?’ Appendix E Figure 1 shows the scores that the respondents achieved on free-recall questions from S1D3, S2D3, and S3D3, which were all tagged as long duration, but presented at various Speech Rates. Appendix E Figure 2 shows scores from S1D2, S2D2, and S3D2, which were all tagged as medium duration but presented at various Speech Rates. Appendix E Figure 3 shows scores from S1D1, S2D1, and S3D1, which were all tagged as short duration, but presented at various Speech Rates.

Appendix E Figure 1

Scores from S1D3, S2D3, and S3D3 – Free-recall Questions



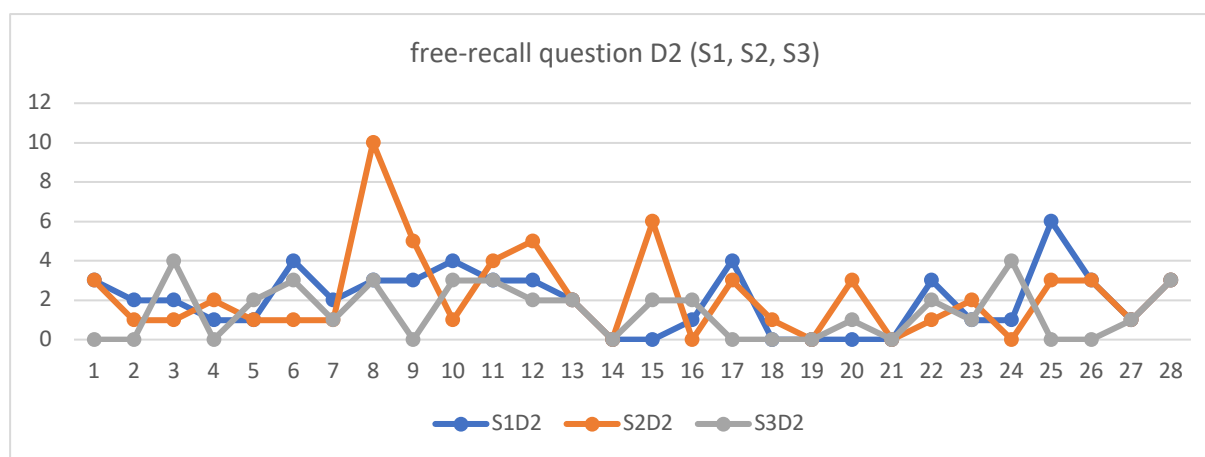
In most cases the orange line representing respondents' performance on S2D3 stays above the S1D3 and S3D3 lines, indicating that most of the respondents performed better on the free-recall questions from S2D3 than those from S1D3 and S3D3. Within the three long duration

audios, the respondents performed best with the normal speed. Regarding the S1D3 line and the S3D3 line, it appears that the S1D3 line remains higher than the S3D3 line, suggesting that respondents performed better on S1D3 (slow speed).

Appendix E Figure 2 below presents the results from free-recall questions for audios S1D2, S2D2, and S3D2.

Appendix E Figure 2

Scores from S1D2, S2D2, and S3D2 – Free-recall Questions

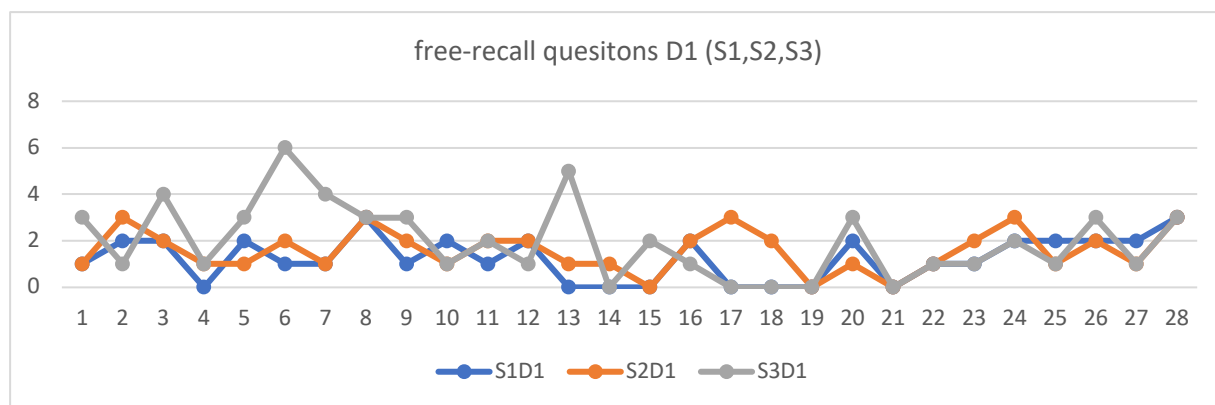


Similar to results shown in Figure 2, respondents' performance on S1D2 (medium duration and slow speed), S2D2 (medium duration and normal speed), and S3D2 (medium duration and fast speed) were similar. It is difficult to determine if any line from the D2 group remains higher than the others. Some respondents, such as participants numbered 8, 9, 11, 12, 15, and 20, achieved their highest scores on S2D2, but did not perform as well on other audios.

Appendix E Figure 3 presents the results from free-recall questions for audios S1D1, S2D1 and S3D1.

Appendix E Figure 3

Scores from S1D1, S2D1, and S3D1 – Free-recall Questions

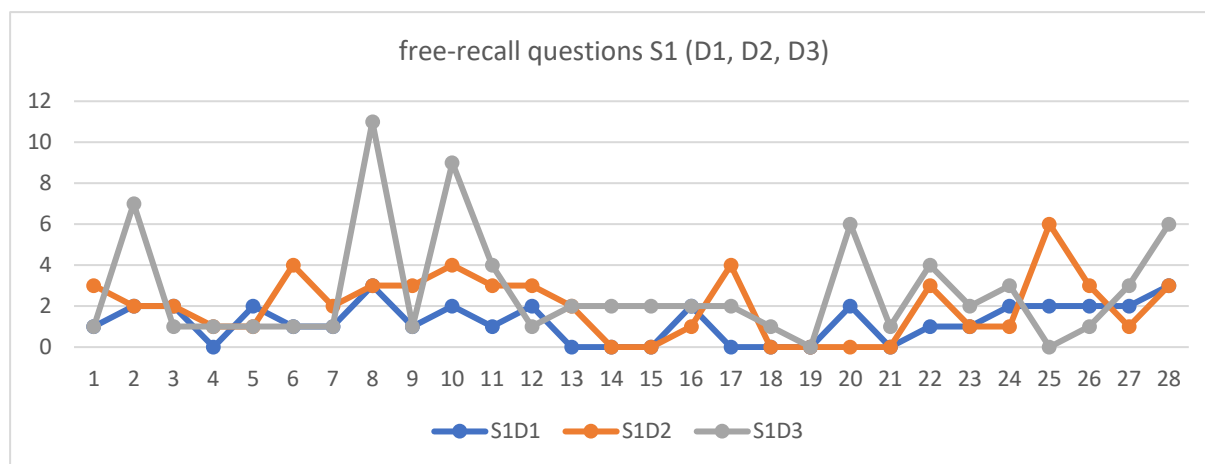


Appendix E Figure 3 suggests that the S3D1 line stood out most of time (10 out of 28). However, the three lines were close to each other in most cases, showing that respondents' performance on free-recall questions for these three short audios was similar.

Same Speech Rate, Various Durations. The third research question was 'What is the effect of variations in duration on listening comprehension on a test of free recall?' Appendix E Figure 4 shows scores from S1D1, S1D2, and S1D3, which were all presented at a slow Speech Rate, but delivered at various Durations. Appendix E Figure 5 shows scores from S2D1, S2D2, and S2D3, which were all presented at a normal Speech Rate but delivered at various Durations. Appendix E Figure 6 shows scores from S3D1, S3D2, and S3D3, which were all presented at a fast Speech Rate but delivered at various Durations.

Appendix E Figure 4

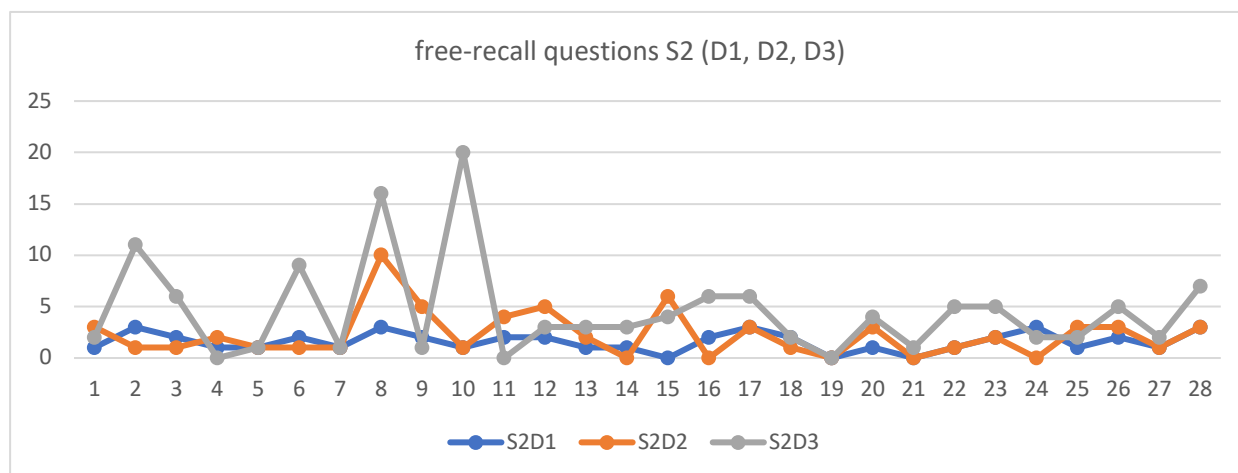
Scores from S1D1, S1D2, and S1D3 – Free-recall Questions



The S1D3 line shows better performance by the respondents, suggesting that respondents performed better on free-recall questions for the slow, long audio than on questions for the other two slow audios. The performance for S1D1 and S1D2 were close, as indicated by the proximity of their respective lines.

Appendix E Figure 5

Scores from S2D1, S2D2, and S2D3 – Free-recall Questions

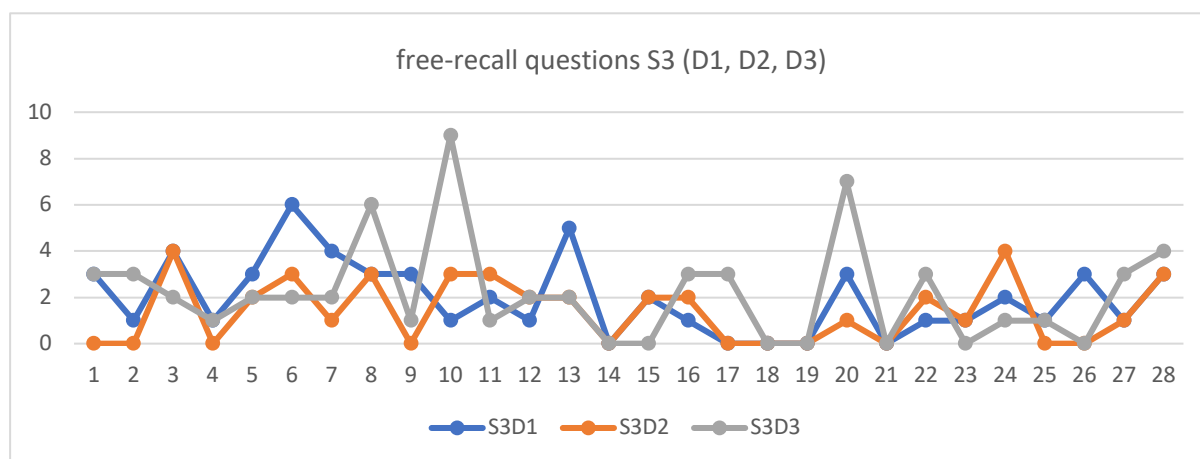


For respondents numbered 11 to 28, performance on free-recall questions from the three S2 audios were similar as indicated by the proximity of the lines in Appendix E Figure 5.

However, respondents numbered 2, 3, 6, 8, and 10 achieved higher scores on questions from S2D3, meaning that they performed much better with the normal speed, long audio.

Appendix E Figure 6

Scores from S3D1, S3D2, and S3D3 – Free-recall Questions



Appendix E Figure 6 shows differing results from Appendix E Figures 4 and 5. It suggests that the range in Appendix E Figure 6 is smaller than that of the previous two, and that there is no clear pattern for how the three lines work in the chart. It suggests that respondent performance on free-recall questions for the three audios from S3 groups was similar, and that it was difficult to determine the audio with which they performed better.

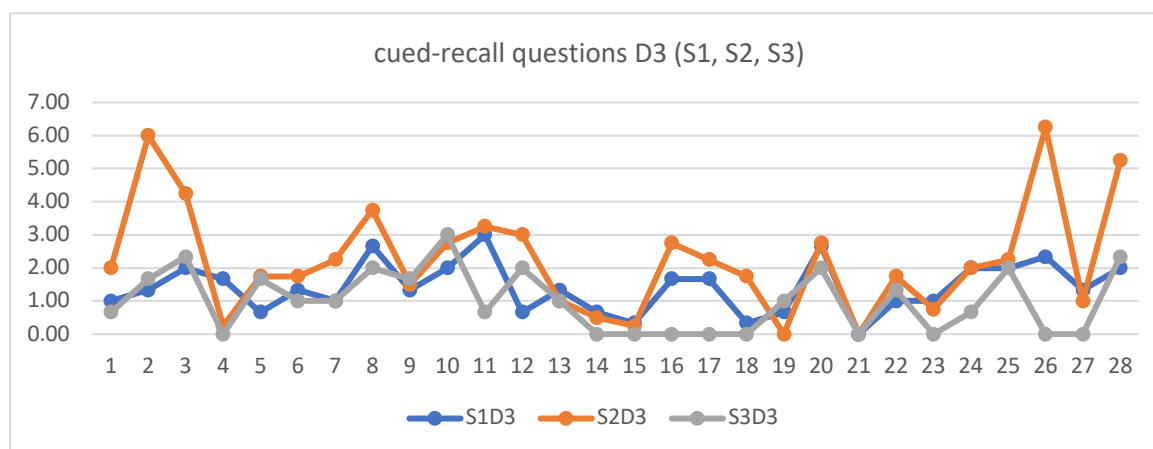
Free-recalled Questions

Same Duration, Various Speech Rate. The second research question was ‘What is the effect of variations in speech rate on listening comprehension on a test of cued-recall?’ Appendix E Figure 7 shows scores that the respondents achieved on cued-recall questions from S1D3, S2D3, and S3D3, which were all tagged as long duration, but presented at various Speech Rates. Appendix E Figure 8 shows scores from S1D2, S2D2, and S3D2, which were all tagged as medium duration, but presented at various Speech Rates. Appendix E Figure 9 shows scores

from S1D1, S2D1, and S3D1, which were all tagged as short duration, but presented at various Speech Rates.

Appendix E Figure 7

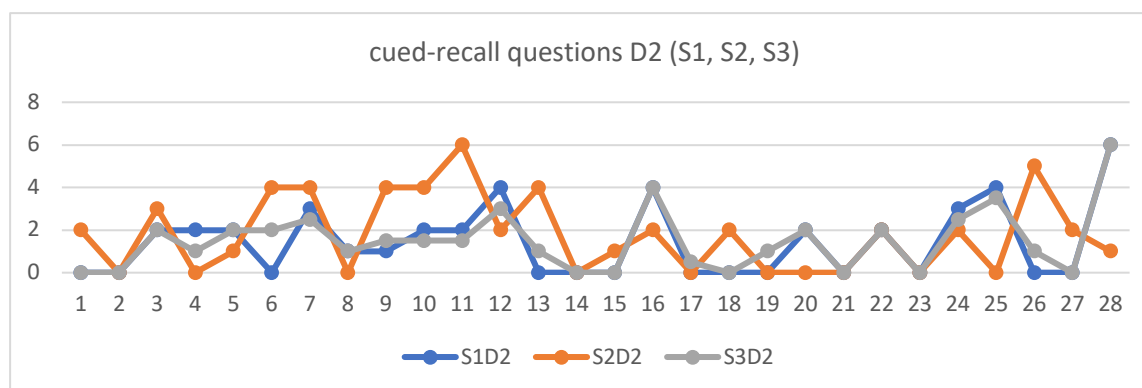
Scores from S1D3, S2D3, and S3D3 - Cued-recall Questions



The three lines remain close to one another in Figure 7, suggesting that most of the respondents achieved similar performance on questions about audio segments in the D3 group except for participants numbered 2, 3, 26, and 28. These four respondents achieved much higher scores on questions from S2D3, with the normal Speech Rate and long audio.

Appendix E Figure 8

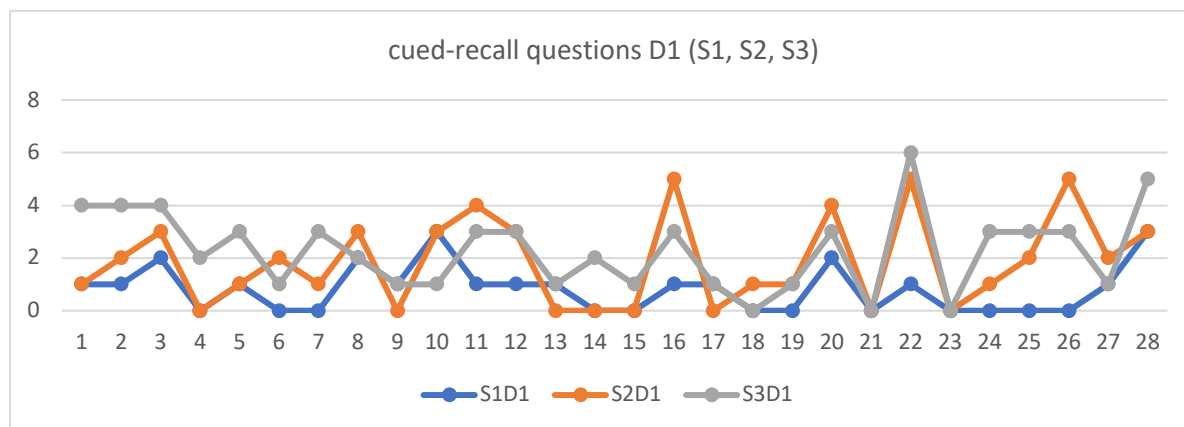
Scores from S1D2, S2D2, and S3D2 - Cued-recall Questions



Respondents' performance on cued-recall questions in D2 audios was different from that of the D3 group audios. The S2D2 line remained overall higher than the other two lines, suggesting that respondents achieved relatively higher scores on S2D2 than on S1D2 and S3D2.

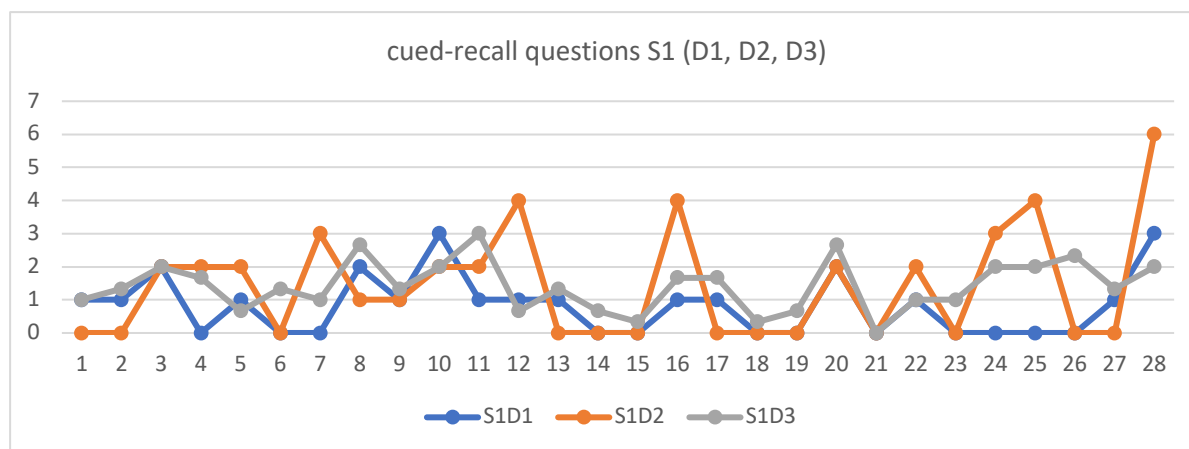
Appendix E Figure 9

Scores from S1D1, S2D1, and S3D1 - Cued-recall Questions

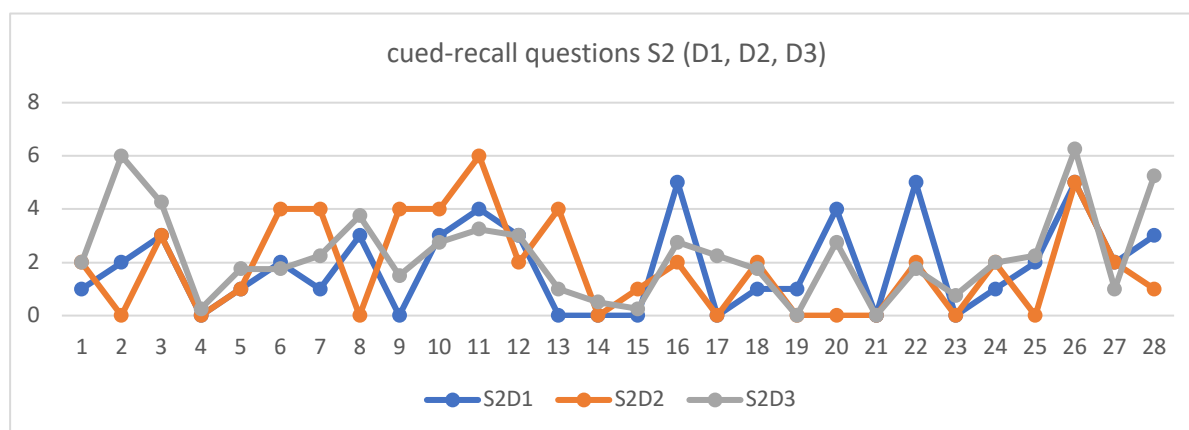


Respondents' performance on cued-recall questions for D1 audio segments were different from that of D2 and D3 group audios. It was found that the orange line, representing the scores from S3D1, stood overall higher than the other two lines, suggesting that respondents achieved relatively higher scores on S3D1 than on S1D1 and S2D1.

Same Speech Rate, Various Durations. The forth research question was 'What is the effect of variations in duration on listening comprehension on a test of cued-recall?' Appendix E Figure 10 shows scores from S1D1, S1D2, and S1D3, which were all presented at a slow Speech Rate but delivered at various Durations. Appendix E Figure 11 shows scores from S2D1, S2D2, and S2D3, which were all presented at a normal Speech Rate, but delivered at various Durations. Appendix E Figure 12 shows scores from S3D1, S3D2, and S3D3, which were all presented at a fast Speech Rate, but delivered at various Durations.

Appendix E Figure 10*Scores from S1D1, S1D2, and S1D3 - Cued-recall Questions*

The three lines remained close to each other in Appendix E Figure 10, suggesting that most of the respondents achieved similar performance on questions about audio segments in the S1 group with the exception of participants numbered 7, 12, 16, 24, 25, and 28. These six respondents achieved observably higher scores on questions from S1D2, with the slow Speech Rate and medium audio duration.

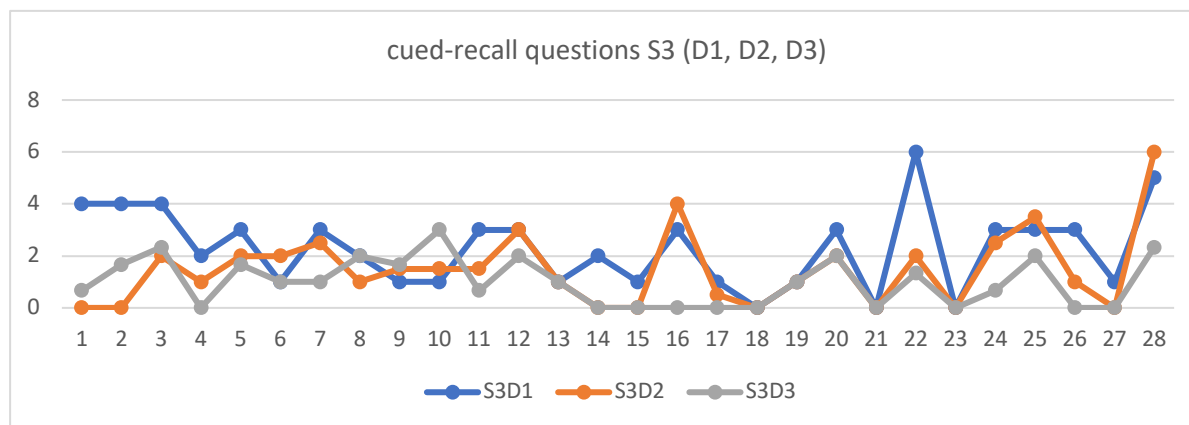
Appendix E Figure 1*Scores from S2D1, S2D2, and S2D3 - Cued-recall Questions*

Respondents' performance on cued-recall questions for S2 audios was different from that of S1 group audios. It was difficult to determine whether any line stood out, indicating better

performance. The three lines tangled, suggesting that respondents' performance on the three S2 audio segments was variable. Some of them achieved high scores on short audios, some on medium audios, and others achieved high scores on long audios.

Appendix E Figure 12

Scores from S3D1, S3D2, and S3D3 - Cued-recall Questions



The first thing to note in Appendix E Figure 12 was that the S3D1 line remained above the other two lines, suggesting that respondents performed better on S3D1 audio questions. The S3D2 and S3D3 lines stayed close to each other, suggesting that respondents' performance on S3D2 and S3D3 questions were similar.