DEVELOPING AN AI-POWERED, PERSONALIZED LEARNING SYSTEM FOR EFL/ESL LISTENING COMPREHENSION

CONG DANH VU

A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in English Language Studies
Suranaree University of Technology
Academic Year 2021

การพัฒนาระบบการเรียนรู้เชิงลักษณะส่วนบุคคลที่ขับเคลื่อนด้วยปัญญาประดิษฐ์เพื่อความเข้าใจการฟังภาษาอังกฤษในฐานะภาษาต่างประเทศและหรือภาษาอังกฤษในฐานะภาษาที่สอง

นายคอง แยง หวู

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาภาษาอังกฤษศึกษา
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2564

DEVELOPING AN AI-POWERED, PERSONALIZED LEARNING SYSTEM FOR EFL/ESL LISTENING COMPREHENSION

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Thesis Examining Committee

(Assoc. Prof. Dr. Wareesiri Singhasiri)
Chairperson

(Prof. Dr. Andrew Lian)
Member (Thesis Advisor)

(Dr. Butsakorn Yodkamlue)
Member

(Dr. Nattaya Puakpong)
Member

(Dr. Suksan Suppasetseree)
Member

(Assoc. Prof. Dr. Chatchai Jothityangkoon)
Vice Rector for Academic Affairs and
Quality Assurance

(Assoc. Prof. Dr. Thara Angskun)
Dean of Institute of Social Technology

คอง แยง หวู : การพัฒนาระบบการเรียนรู้เชิงลักษณะส่วนบุคคลที่ขับเคลื่อนด้วยปัญญาประดิษฐ์เพื่อความเข้าใจการฟังภาษาอังกฤษในฐานะภาษาต่างประเทศและหรือภาษาอังกฤษในฐานะภาษาที่สอง (DEVELOPING AN AI-POWERED, PERSONALIZED LEARNING SYSTEM FOR EFL/ESL LISTENING COMPREHENSION)
อาจารย์ที่ปรึกษา : Professor Dr. Andrew Lian, 237 หน้า.

คำสำคัญ: การฟังเพื่อความเข้าใจ/ ระบบปฏิสัมพันธ์เพื่อช่วยการเรียนรู้ภาษาโดยคอมพิวเตอร์ (ICALL)/ แนวคิดการเรียนรู้แบบเฉพาะบุคคล/ แนวคิดการเรียนรู้ด้วยสื่อประสม

งานวิจัยนี้มุ่งนำเสนอแนวทางส่งเสริมความสามารถด้านการฟังเพื่อความเข้าใจของผู้เรียนภาษาอังกฤษเป็นภาษาต่างประเทศโดยพัฒนาระบบการเรียนรู้ซึ่งใช้เทคโนโลยีสมองกล (AI-powered learning system) โดยให้ความสำคัญต่อผู้เรียนเป็นหลัก งานวิจัยมีวัตถุประสงค์ 4 ประการ คือ (1) เพื่อพัฒนาระบบการเรียนการสอนรูปแบบใหม่ (2) เพื่อศึกษาประสิทธิภาพของระบบการเรียนการสอนที่พัฒนาขึ้นในการเพิ่มทักษะด้านการฟังเพื่อความเข้าใจของผู้เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ (3) เพื่อศึกษาพฤติกรรมการเรียนของผู้เรียนเมื่อใช้ระบบการเรียนการสอนดังกล่าว และ (4) เพื่อสำรวจความคิดเห็นของผู้เรียนต่อการใช้ระบบการเรียนการสอนดังกล่าว

งานวิจัยนี้เป็นการวิจัยแบบผสม (Mixed-method approach) ซึ่งเก็บข้อมูลทั้งเชิงปริมาณและเชิงคุณภาพโดยใช้วิธีที่หลากหลาย และได้ทำการวิจัยกับนักศึกษาวิชาเอกภาษาอังกฤษชั้นปีที่ 1 จำนวน 53 คนจากมหาวิทยาลัยแห่งหนึ่งในสาธารณรัฐสังคมนิยมเวียดนาม นักศึกษาถูกแบ่งออกเป็น 2 กลุ่มด้วยวิธีการสุ่มคือกลุ่มทดลอง (Experimental Group) และกลุ่มควบคุม (Control Group) นักศึกษากลุ่มทดลองเรียนทักษะการฟังภาษาอังกฤษเป็นภาษาต่างประเทศโดยใช้ระบบการเรียนการสอนที่ผู้วิจัยพัฒนาขึ้น ระบบดังกล่าวมีชื่อว่า “Listening Hacked” ในขณะที่นักศึกษากลุ่มควบคุมเรียนทักษะดังกล่าวโดยใช้วิธีดั้งเดิมคือทำแบบฝึกหัดการฟังผ่านระบบ “Google Classroom” ทั้งนี้การเรียนโดยใช้ “Listening Hacked” นั้นจะมีการชมภาพยนตร์เสียงพากย์ภาษาอังกฤษและถอดความแบบมีพักหยุดเป็นระยะ (Paused transcription tasks) ระบบดังกล่าวเปิดโอกาสให้นักศึกษาควบคุมการเรียนของตนอย่างเต็มที่ กล่าวคือนักศึกษาสามารถฝึกทักษะการฟังฯ โดยเลือกภาพยนตร์ที่ตนสนใจและใช้ตัวช่วยที่เหมาะสมเพื่อให้บรรลุวัตถุประสงค์ของการเรียน ผู้วิจัยพบว่าเมื่อนักศึกษาเรียนไป 12 สัปดาห์ มีเพียงกลุ่มทดลองเท่านั้นที่ทักษะการฟังภาษาอังกฤษเป็นภาษาต่างประเทศเพิ่มขึ้นอย่างมีนัยสำคัญและทำคะแนนแบบทดสอบหลังเรียนด้านการฟังฯ ได้สูงกว่ากลุ่มควบคุมอย่างมีนัยสำคัญ นอกจากนี้ ผู้วิจัยยังพบว่ากระบวนการเรียนของระบบการเรียนการสอนที่พัฒนาขึ้นช่วยให้นักศึกษากลุ่มทดลองมีทักษะการถอดความที่ดีขึ้นอย่างมีนัยสำคัญ ตลอดจนมีรูปแบบพฤติกรรม

การใช้สื่อและตัวช่วยในการเรียนในทิศทางเดียวกัน ผลที่ได้จากการใช้แบบสอบถามและการสัมภาษณ์บ่งชี้ว่านักศึกษาส่วนใหญ่มีความคิดเห็นเชิงบวกต่อการเรียนด้วยระบบการเรียนการสอนที่ผู้วิจัยพัฒนาขึ้นและเชื่อเป็นอย่างยิ่งว่าระบบดังกล่าวมีประสิทธิภาพในการเพิ่มทักษะการฟังฯ ของตน

กล่าวโดยสรุป งานวิจัยนี้แสดงให้เห็นว่าระบบการเรียนการสอนที่ผู้วิจัยพัฒนาขึ้นสามารถเพิ่มทักษะการฟังภาษาอังกฤษเป็นภาษาต่างประเทศของผู้เรียนได้อย่างมีประสิทธิภาพกว่าระบบการเรียนการสอนแบบดั้งเดิม ระบบการเรียนการสอนดังกล่าวสร้างโดยใช้แนวคิดการเรียนรู้แบบสร้างสรรค์นิยม (Constructivist learning) และแนวคิดการเรียนรู้แบบเฉพาะบุคคล (Personalized learning) เป็นหลัก และได้แสดงให้เห็นถึงความสำเร็จในการนำเทคโนโลยีมาแก้ปัญหาในการสอนทักษะการฟังภาษาอังกฤษเป็นภาษาต่างประเทศ นับได้ว่าเป็นการริเริ่มวางรากฐานแนวทางการสอนที่ถูกต้องซึ่งจะนำไปใช้ในการเรียนการสอนเพื่อพัฒนาทักษะการฟังภาษาอังกฤษเป็นภาษาต่างประเทศของผู้เรียนต่อไป

สาขาวิชาภาษาต่างประเทศ ลายมือชื่อนักศึกษา _____

ปีการศึกษา 2564 ลายมือชื่ออาจารย์ที่ปรึกษา _____

CONG DANH VU : DEVELOPING AN AI-POWERED, PERSONALIZED LEARNING SYSTEM FOR EFL/ESL LISTENING COMPREHENSION. THESIS ADVISOR : PROF. ANDREW LIAN, Ph.D., 237 PP.

Keyword: Listening Comprehension/I CALL/ Personalized Learning/ Multimedia Learning

The present study set out to put forward a comprehensive solution to EFL listening comprehension by developing an AI-powered learning system with a focus on personalization. With that, four key objectives were established: (1) to develop the proposed learning system, (2) to investigate the effectiveness of the learning system in developing EFL listening comprehension, (3) to explore the learning behaviors on the learning system, and (4) to study the students' opinions of the proposed learning system.

The study adopted a mixed-methods approach in which both quantitative and qualitative data were collected through different procedures. Fifty-three first-year students studying English major at a university in Vietnam participated in the experiment. They were randomly assigned into the Experimental Group and the Control Group. The students in the Experimental Group learned EFL listening on the proposed learning system, namely Listening Hacked, while those in the Control Group learned EFL listening by doing traditional listening exercises on Google Classroom. The learning on Listening Hacked generally involved viewing English-speaking movies and doing paused transcription tasks. The learning system allowed the students to fully control their learning by choosing movies of their interests and using appropriate help options as to accomplish the tasks. After 12 weeks of learning, the results show that only the students in the Experimental Group made significant improvement in their EFL listening ability and that they outperformed their counterparts in the Control Group significantly in the EFL listening posttest. Additionally, the results of the learning progress indicate that the students in the Experimental Group have improved their perceptual decoding ability significantly. The results also reveal that the students exhibited a shared behavioral pattern of using the resources and help options on the learning system. Finally, the results from the questionnaire and the interviews indicate

that most students held positive opinions towards learning on the proposed learning system and strongly believed that the learning system was effective in developing their EFL listening comprehension.

All in all, the present study shows that the proposed learning system was more effective in developing the students' EFL listening ability than the traditional approach. Constructed on a principled framework emphasizing constructivist learning and personalized learning, the learning system demonstrates a successful implementation of technology-based solutions into EFL listening pedagogy. This is an initial step in establishing the right direction which will move forward the field of EFL listening teaching and learning.

School of Foreign Languages Student's Signature _____

Academic Year 2021 Advisor's Signature _____

ACKNOWLEDGEMENTS

This research project would not be finished without the help of so many people.

I would like to thank Professor Dr. Andrew Lian, my dear supervisor. I feel privileged to be supervised by him. He has always been a great source of inspiration and encouragement for me. His great knowledge, wisdom, and care guided me through the most difficult times in my Ph.D. study. I would like to recognize the significant influence of his intellectual works on this research project.

Special thanks to Dr. Peerasak Siriyothin, who were with me in this project from the beginning and gave me invaluable advice on improving its quality. My deepest gratitude also goes to the committee, Associate Professor Dr. Wareesiri Singhasiri, Dr. Butsakorn Yodkamlue, Dr. Nattaya Puakpong, Dr. Suksan Suppasetseree, whose expertise and insightful feedback have greatly contributed to the writing of this dissertation.

I would like to thank Associate Professor Dr. Punyathon Sangarun, Dr. Sirinthorn Seepho, Assistant Professor Dr. Adcharawan Buripakdi who have instilled in me a love for research. I will never forget the beautiful memories I had with them at SUT.

I am grateful to all the participants who spent their valuable time on this research project. Without your collaboration, this research would have been impossible.

I also had great pleasure of working with Cuong, who helped me with the coding process of the software used in this project. I would like to acknowledge his expertise and hard work that cannot be underestimated.

I gratefully thank my family and friends who were always with me during my study at SUT, Thailand. Without your relentless support, I could not have gone this far.

Finally, I thank the Lord Almighty for giving me the chance to meet great people in my life. You have given me great strength in my pursuit of knowledge. Thank You for making all of these happen.

Cong Danh Vu

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

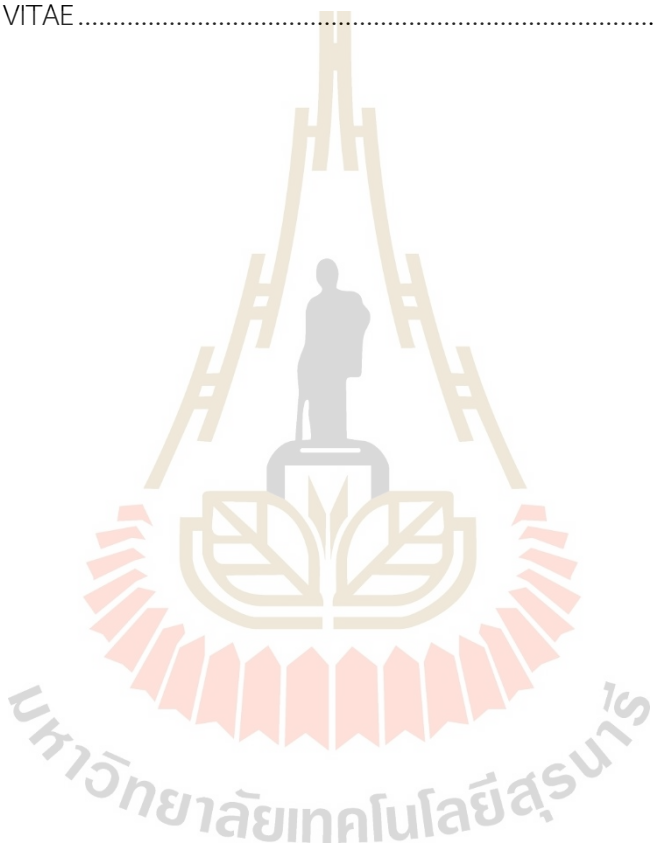
Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page



LIST OF TABLES

Table **Page**

LIST OF TABLES (Continued)

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF ABBREVIATIONS

AI	Artificial intelligence
AWE	Automatic writing evaluation
CALL	Computer-assisted language learning
CF	Corrective feedback
EFL	English as a foreign language
ESL	English as a second language
ICALL	Intelligent computer-assisted language learning
FCE	Cambridge First Certificate in English
ILTS	Intelligent language tutoring system
L2	Second language
NLP	Natural language processing
TPS	Task performance score

CHAPTER 1
INTRODUCTION

1.1 Background of the study

1.1.1 The case of teaching and learning EFL listening in Vietnam

Listening is an essential part of human communication and contributes to the development of other communication skills. From a social perspective, one needs to listen, along with other sensory observations, to experience things in the world, thus survive in it. Research in communication has unveiled that listening can take up 55% of our daily communication (Worthington & Fitch-Hauser, 2018). Listening in a language other than the native language is no less important. Learning to listen in a foreign language, especially in English, has become a necessity in most schools around the world. It is essentially important in the contexts of today's globalization and international mobility when individuals might be required to prove their English proficiency for admission to a foreign study program, or employment opportunities by foreign-invested companies, or simply for promotion in their career. The case of Vietnam well exemplifies this trend. In Vietnam, English is a compulsory school subject from Grade Three in the 12-year national education system (Dudzik & Nguyen, 2015; Hoang, 2010). Students are given instructions on the grammar, vocabulary, and four skills of English from the outset, yet assessments and examinations in the national education system have overlooked speaking and listening skills (Tran & Duong, 2020). In recent years, however, some high schools in major cities have implemented an intensive English program which includes the teaching and assessment of English listening comprehension. Furthermore, most universities in Vietnam now require their undergraduates to reach a certain level of English proficiency for graduation, certified by international qualifications like IELTS, TOEIC, or other equivalents (Dudzik & Nguyen, 2015). One harsh reality has been that despite 12 years of learning English at high school, many Vietnamese students find listening to English a big problem during their university study. English-major students are no exception (Duong & Chau, 2019; N. Q.

Nguyen & Thai, 2018; Trinh, n.d.). Many first-year English-major students at Ton Duc Thang University, the current workplace of the researcher, have been reporting their difficulties in learning English listening and many of them could not manage to get a pass for the very first listening course in their study program. Additionally, teachers complain that the instruction time in class is too limited and there is nothing much they can do about it to help their students overcome this issue.

1.1.2 The challenges of EFL listening

The challenges of learning and teaching EFL listening comprehension could be attributed to the complex nature of the skill itself (Wolvin, 2009), the unproductive tradition of instruction, and the lack of interdisciplinary applications (Field, 2008b).

The difficulty of listening to English as a foreign language (EFL) may first originate from the ephemeral nature of listening (Vandergrift & Cross, 2018; Vandergrift, 2002). When listening to a speech, the listener has to attend to the continuous streams of sounds that need to be segmented into perceptible units for various cognitive processing. For this real-time and continuous aspect inherent in the listening process, the listener does not have a chance to rewind the spoken speech; as a result, once the speech signal is gone, it is lost to perception. Another challenge of EFL listening is the issue of variability in the speech signal as opposed to the expected perceptual constancy (Barriuso & Hayes-harb, 2018). Due to the nature of the human articulatory apparatus and socio-geographical influences, pronunciations of the same word or even the same phoneme vary across speakers and contexts, often referred to as phonetic and phonological variations (Roach, 2009). These variations are caused by co-articulation processes in connected speech such as assimilation, vowel reduction, and elision (Rost, 2011). Perhaps, the most common examples among many others are the use of weak forms instead of citation forms when pronouncing pronouns and function words in sentences. For example, the prepositions *of* and *at* are often pronounced in their weak forms as /**ə**v/ and /**ə**t/ in sentences respectively. Variations in pronunciation may also come from differences in the speakers' regional accents (Roach, 2009). For instance, Scottish speakers may pronounce words like *sports, more, there* with the /r/ sound after a vowel, while this pronunciation is considered rare among the British (Robinson, 2019). The clash between the varying speech signal and

the expected perceptual constancy, unfortunately, hinders non-native listeners from recognizing words or sound patterns in various contexts while their native counterparts can perform this most of the time expeditiously and effortlessly. In most previous research on L2 listening difficulties, this issue of variability has been repeatedly reported as a primary source of listening difficulties which is associated with other types of difficulties such as word recognition skills, speaker's characteristics (Flowerdew & Miller, 1992; Graham, 2006; Hasan, 2000; Richards, 2008) and perceptual processing skill (Goh, 2000).

Viewing from a pedagogic angle, the teaching of EFL listening has also encountered numerous challenges. Traditionally, the comprehension-based approach was frequently employed in the L2 listening class (Field, 2008b, 2009). As commentators have pointed out, this comprehension-focused approach to L2 listening pedagogy "tests without teaching" (Field, 2008b, p. 6) because in this approach the teacher's main concern is whether students can get correct answers to pre-set comprehension questions. Accordingly, little is known about the process of how a student gets a right or wrong answer. In addressing the weakness of the conventional comprehension approach, two process-based approaches corresponding to the two underlying processes of listening, namely, bottom-up and top-down, were introduced (Field, 1999; Goh, 2016; Johnson, 2008; David Nunan, 2002; Rost, 2001; Vandergrift, 2011). The bottom-up approach rests on the notion that when listening to a speech, the listener utilizes his or her linguistic knowledge to recognize words and sound patterns such as phonemes, syllables, intonation patterns in the streams of speech (Richards, 2008; Rost, 2006). In this approach, teachers may have students work on word recognition and phonemic contrasts. Another alternative approach holds the view that listening can also be a top-down process that involves drawing on the listener's background knowledge to predict the incoming speech (Field, 1999). Top-down activities usually require the teacher to prepare questions and topic-related vocabulary for a given listening text so that students can discuss or brainstorm before they do the listening. The purpose of these preparations is to activate students' prior knowledge of the topic. To date, researchers have agreed that listening involves both top-down and bottom-up processes that operate simultaneously in harmony (Field,

1999; Flowerdew & Miller, 2005). It follows that an instructional approach overemphasizing only one process cannot adequately address the problem. On the one hand, while focusing on bottom-up processing may help students with their word recognition skills, this approach has been criticized for neglecting a global understanding of a speech. Too often, students are likely to focus on segmenting word boundaries in the speech streams and forget to pay attention to the gist of the speech. On the other hand, while background knowledge intervenes in building an understanding of a speech, without adequate bottom-up skills, acceptable understanding of the speech is hard to achieve, and top-down information is most of the time hence unavailable. In recent decades, there is an interest in strategy instruction approach to L2 listening; however, commentators have expressed concerns about its effectiveness and whether the time used for strategy training should be spent instead on practicing the language skills (Littlejohn, 2008; McDonough, 2006; Renandya & Farrell, 2011). Still, observations from classroom practices indicate a preference for using the comprehension approach, a product-based approach (H. Nguyen & Abbott, 2016), and thus, the much-needed focus on process has been overlooked.

As Field (2008b) and others suggest, a barrier to L2 listening research may come from the fact that the area of L2 listening has little interface with other scientific disciplines. Perhaps, there is a more ample concern, i.e., the lack of transfer from research to practices. For example, in speech perception research, high variability perceptual training has produced substantial evidence regarding the effectiveness of this training method for perceptual learning for nearly three decades (Logan et al., 1991); however, to date, many L2 listening instructors have still been unaware of it. Likewise, L2 research frequently draws on technological advances and encourages positive changes accordingly. However, teaching practices are still lagging behind necessary changes. The world has become more connected and visualized as multimedia have become dominant on the internet, yet many EFL listening classrooms today are still relying on the use of audio CDs and sound speakers manually controlled by the teacher.

1.1.3 Technology in teaching and learning EFL listening

Since the advent of the internet in the 1990s, EFL listening materials have changed dramatically, moving from analog to digital recordings, and from classroom-based tapes/CDs to online audio/video files (Clancy & Murray, 2016). To date, technologies for EFL listening have evolved and flourished, within the area of computer-assisted language learning (CALL) software.

CALL applications have contributed to two significant changes in EFL listening. First, they have given learners access to substantial quantities of authentic materials obtained from online sources for learning EFL listening, for example, English speaking podcasts on many news media websites, online videos, or free listening materials offered by EFL learning websites like ELLO (http://www.elllo.org). Second, CALL tools can facilitate EFL listening learning through providing help options, "embedded application resources that assist learners in performing computing operations and/or support language learning" (Monica S. Cárdenas-Claros & Gruba, 2009, p. 69). Typical help options in CALL contexts include dictionaries, transcripts, captions, media replay, speed control, and explanatory feedback (Pujolà, 2002). Research into help options in CALL software for L2 listening has revealed that students express preferences for learning software with help options and that, with use, their comprehension is aided (Cárdenas-Claros & Gruba, 2014; Grgurović & Hegelheimer, 2007; Rivens Mompean & Guichon, 2009). Today, as mobile devices are becoming more affordable, CALL may take the shape of mobile-assisted language learning (MALL), an environment that enables the idea of 'anytime, anywhere' learning (Reinders & Hubbard, 2013; Stockwell, 2016).

There is an outstanding technological advance in the digital era, which is worth mentioning, i.e. Artificial Intelligence (AI). AI has been defined in various ways and it is still an evolving term. Kurzweil (1990, cited in Moral Robots, 2019) defines AI as "the act of creating machines that perform functions that require intelligence when performed by people". This definition looks at the equivalence of machine behavior to that of humans. Similarly, Bellman's (1978, cited in Moral Robots, 2019) AI definition also focuses on thinking ability: "The automation of activities that we associate with human thinking, activities such as decision-making, problem-solving, learning."

According to McCarthy (2007), AI is “the science and engineering of making intelligent machines”. There are dozens of other AI definitions in different fields as well, however, most definitions supposedly express the core feature of an AI system, which is also included in its name, intelligence or capability of cognizing. In language education, AI-based technologies are usually referred to as intelligent computer-assisted language learning (ICALL) (Azizinezhad & Hashemi, 2013; Heift, 2012). The integration of AI in CALL often involves using natural language processing (NLP) which works with natural language understanding and natural language generation (Tschichold and Schulze, 2016). ICALL tools supported by NLP are generally used to analyze students’ input, spoken or written, and then generate individualized corrective feedback and/or instructional guidance as appropriate (Tschichold & Schulze, 2016). For example, speech recognition powered by NLP can be integrated into a CALL application for evaluating the student’s translation appropriateness (Bouillon et al., 2016; C. Wang & Seneff, 2007). Alternatively, when combining NLP and corpora, ICALL tools can generate automated feedback and evaluation for student’s essay writing, for example, *Criterion* developed by Educational Testing Service for detecting and giving feedback on student’s grammatical errors at the sentence level, and *Intelligent Academic Discourse Evaluator* developed by Iowa State University for providing feedback on student’s research article writing (Chapelle et al., 2015). There is a concern, though, that ICALL systems are only mentioned in research and still rarely found in the language classrooms (Tschichold & Schulze, 2016), a commonly observable gap between research and practice. Furthermore, ICALL has been mostly applied to writing, grammar, vocabulary and speaking instructions; research on and incorporation of ICALL systems in L2 listening; however, is still in its infancy (Cardenas-Claros & Gruba, 2013).

1.2 Rationale of the study

A commonly observable issue in foreign language pedagogy has been the undesirably rare interaction between research and practice (Block, 2000). Despite decades of research resulting in large quantities of research findings and instructional approaches, it is ironic that when it comes to answering the simple, yet practical, question of what should be done to improve language learning and teaching, many

practitioners and teachers are unable to provide a direct answer. Consider a situation when a student approaches their English teacher, asking about how to improve their EFL listening comprehension. What the student is likely to be advised are such general answers as increasing their English listening time, doing more listening practice, expanding vocabulary, trying to practice useful strategies, and the like. In the end, the student is left with vague instructions, which in turn gives the student no idea of how to handle a myriad of problems occurring in EFL listening. In fact, research findings over the last few decades have shown us various methods for addressing specific problems of language learning, and perhaps, among those methods, there may be some worth a try in the language classroom. However, those findings have remained theories and have not effectively made their way into practice (Ellis, 2010; Erlam, 2008; Nassaji, 2012). Had it been otherwise, EFL teachers would have found no problem in giving a quick answer to the above question. Indeed, this indicates that there is currently no practical solution to learning EFL listening.

With the belief that scientific knowledge is built on past research achievements but, as Einstein suggests, a problem cannot be solved by the same level of thinking that created it, therefore the present study was an effort to provide a workable and practical solution to learning and teaching EFL listening comprehension in the Vietnamese context by developing a technology-based, personalized learning system constructed on a sound theoretical framework of language learning and listening comprehension. Radical constructivism was used as the main theoretical foundation for this study because the study took a cognitive approach to language learning, and because listening comprehension is essentially a cognitive process which remedial activities should tap into to address problems arisen from it. It is also noted that the approach proposed here is applicable for EFL contexts such as Vietnam or Thailand, and ESL contexts such as Singapore or English-speaking countries, since there may be some differences in teaching practices, but the nature of language learning is not different in those contexts.

1.3 Statement of the problem

Since the 1980s, L2 specialists and practitioners have recognized the critical role of listening in L2 learning (Rost, 2011). However, learning to listen to a foreign language is never an easy task. Graham and Macaro (2008) maintain that listening is still a problematic skill for many L2 learners these days. As discussed earlier, the unfruitful practice of L2 listening instruction is due to the difficulties inherent in the skill itself, the lack of interdisciplinary applications, and the unfriendly relationship between research and practice.

There is, in fact, another issue that further complicates EFL listening instruction, i.e., individual differences. There are various factors that influence how individuals listen including, but not limited to, gender (Phillips et al., 2001), age (Wolvin, 2018), linguistic knowledge (Long, 1990), and working memory (Sakai, 2018). These individual variables present considerable challenges to the success of any instructional approach to EFL listening, and what is more, they fail the one-size-fits-all instruction. This has been seen in the case of Vietnam where the prevalence of large-sized classes still exists. It is not clear whether ineffectual instruction in large-sized classes is still better than no instruction, though often assumed by many EFL listening instructors in Vietnam. What is clear is that L2 learning needs to be personalized as ways to address individual variability, and this is where technology comes in to offer workable solutions.

In fact, the much problematic issue of individual differences in large-sized language classrooms has existed for ages. It was traditionally impossible for a single teacher to address the diverse needs, suppose, of 40 students in the classroom; conversely, today with ICALL software, it is technically achievable because ICALL tools are steadily more capable of generating detailed feedback and proper guidance meeting specific needs of learners in a timely and increasingly more precise manner. However, the existence of ICALL software is still scarce, both in research and reality, and mostly deals with writing and grammar instructions, if any. Apparently, ICALL systems developed for EFL listening have not been identified in the literature, perhaps because they do not exist, or their availability is too limited.

In a nutshell, three main problems have been discussed: the challenging nature of listening skill, the complications caused by individual differences, and the rare

existence of ICALL systems for EFL listening despite their promising capabilities. The arguments presented thus far suggest that any effective solution should directly address the difficulties of listening by tapping into its nature, promote personalization in learning, and apply technological advances in creating practical learning solutions. Unfortunately, there has been no such satisfactory solution for EFL listening available at the moment.

1.4 Purpose of the study

The main purpose of the study was to develop a technology-based learning system for EFL listening, namely *Listening Hacked*. The learning system, based on a solid theoretical foundation, utilized AI technology and other computing technologies to create a personalized learning environment and assist learners in dealing with their listening difficulties. Next, the study sought to investigate the effectiveness of the proposed system. Specifically, the study investigated whether Vietnamese undergraduates' EFL listening comprehension was improved as a result of the training program using Listening Hacked. The study was also interested in exploring what behaviors the students exhibited when learning on a technology-based autonomous learning environment as such. More importantly, the study attempted to get an understanding of the students' reasons for having those learning behaviors. As well, the study was concerned about the students' opinions of the learning system and its features, the knowledge of which will pave the way for future developments.

In short, the purposes of the study were fourfold as follows.

1. To develop a technology-based, personalized learning system for EFL listening comprehension.
2. To investigate the effectiveness of the training using Listening Hacked on EFL listening comprehension performance of Vietnamese undergraduates.
3. To explore the students' learning behaviors, i.e., how the students interacted with Listening Hacked and what influenced their decisions in those situations.
4. To examine the students' opinions of Listening Hacked and its features.

1.5 Research questions

To those ends, the study needs to develop the learning system, the theoretical foundation and the technical descriptions of which will be discussed in detail in Chapter 2 and section 3.4. The study also attempted to answer the following research questions.

1) What was the effectiveness of Listening Hacked on Vietnamese English-major undergraduates' EFL listening comprehension?
2) What happened when students encountered a listening comprehension problem on Listening Hacked? What were the students' reasons for exhibiting such behaviors?
3) What were the students' opinions of the usefulness of Listening Hacked and its features in the development of their EFL listening comprehension?

1.6 Significance of the study

The study contributed to the research and development of technology-based learning approaches to L2 listening, which may benefit interested researchers and practitioners who are also seeking ways to implement technology in EFL/ESL listening instruction.

First, the study presents a framework for designing a technology-based learning system for L2 listening. This framework specifies the appropriate theories of learning and language learning, models the listening comprehension process, and introduces remedial methods for individual problems in learning L2 listening. Regarding the methods in this framework, it is noted that the study devised novel listening tasks as they have not been used elsewhere. The listening task used in this study was Paused Transcription which was inspired by the task proposed by John Field (2008); however, it was largely modified to properly assist with learning EFL listening rather than testing it. Additionally, the remedial tasks for dealing with the issue of variability in speech while rooted in high-variability phonetic training research (Barriuso & Hayes-harb, 2018) were developed to take a new shape, shifting away from decontextualized presentations of separate words to contextualized input. The study, therefore,

introduced original methods for addressing EFL listening problems, rather than just exploitation of existing methods.

Second, the study can expand our knowledge about the technical process of developing such an ICALL system with the aim of giving rise to more ICALL applications in the future. The development process of the learning system in this study involved 7 stages, from constructing theoretical principles to translating those principles to various functionalities of the software, preparing the data for its operation, programming the software, implementing the software, testing and modifying its functionalities, and maintaining its operation (see detailed descriptions at Section 3.3). Furthermore, the study documented the development and operation of the software which future CALL researchers and CALL software developers who are interested in building similar CALL applications can use for their own projects.

Third, the study provides valuable insights into the effectiveness of this ICALL system for L2 listening as well as the ways that students engaged with various features of such a system in their independent learning. This type of information will be useful for interested researchers and practitioners in designing robust ICALL systems in the future. Moreover, the students' opinions about various features of the learning system, both positive and negative opinions, provided practical and valuable implications for improving the design and effectiveness of future CALL and ICALL applications.

1.7 Definitions of key terms

1. AI-powered software: in this study, refers to the type of Computer-Assisted Language Learning (CALL) software which utilizes Natural Language Processing (NLP) for natural language understanding and natural language generation. The software developed in this study used NLP for language pattern identification.
2. Radical constructivism: a theory of learning which holds the view that knowledge is a person's cognitive structures, an organization of one's pre-existing experiences. Knowledge is subjectively constructed, and even sensory perception is also constructed (von Glasersfeld, 1991).
3. Personalized learning: in this study, refers to the effort to respond as accurately and timely as possible to individual needs arising in learning.

4. EFL listening comprehension: in this study, refers to the student's ability to understand spoken English language in a communicative event that provides both purposes and contexts for understanding, by drawing on their existing knowledge and skills. EFL listening comprehension is indicated by their performances in the pretest and posttest, and their learning progress.
5. Multimedia learning: a learning environment that presents multimedia listening materials, for example, audios and visuals or audios, visuals, and texts.
6. Corrective feedback: in this study, refers to the computer-generated information on student's errors based on the student's input in a listening task.
7. High variability training: a training method that utilizes speaker variability for perceptual learning. In this study, it refers to the presentations of contextualized, audio-visual input with speaker variability for dealing with specific word recognition problems.
8. Listening Hacked: the name of the learning system developed in this study. The system is an online, autonomous learning environment for practicing English listening comprehension through movies. It personalizes learning by providing students with immediate corrective feedback and various help options, using AI for analyzing students' errors, and automatically suggesting follow-up practice for individual students (using a feature named Listening Boosters).
9. Paused transcription tasks: refers to the tasks students had to do as they were viewing a movie on Listening Hacked. In this task, students were required to view a movie and occasionally transcribe some parts of it when encountering a forced pause.
10. Technology-based learning system: in this study refers to computer-assisted language learning (CALL) software for EFL listening which utilizes various technological tools or digital resources for its operations, such as natural language processing, multimedia, websites,...

1.8 Chapter summary

The chapter begins with the discussions about the challenges of learning and teaching EFL listening comprehension with reference to the educational context in Vietnam which provides background for the study. The chapter then calls attention to the unavailability of practical, ready-made solutions for learning EFL listening, especially those with technological applications. The purposes of the study, the research questions, the significance of the study are then discussed. Finally, definitions of key terms are presented. The next chapter will review the related theories and relevant research studies to formulate the theoretical framework for this study.

CHAPTER 2
LITERATURE REVIEW

This chapter presents the conceptual framework used in the present study. It reviews and discusses the relevant theories in language learning, some important issues in language pedagogy, and potential methods for solving these issues. At the end of the chapter, a coherent theoretical framework on which the development of the learning system is based is proposed.

2.1 Constructivist theories of learning

Learning is normally associated with knowledge delivery. Yet we understand that knowledge cannot be transmitted intact to the learner by the teacher in any circumstances. This is a fundamental principle of constructivism, a learning approach developed in the 20th century which posits that knowledge is constructed, and thus, cannot be transferred (Hendry, 1996; Mascolo & Fischer, 2005). While this view is accepted among most constructivists, the matter of how knowledge is constructed has been hotly debated over many decades. The chapter starts by answering this epistemological question from a constructivist standpoint as it is particularly relevant to the quest for finding a solution to learning a foreign language, or in fact, learning anything.

Constructivist theories are diverse, with two major dimensions, one focusing on cognitive aspects, e.g. Jean Piaget's theory of cognitive development or von Glasersfeld's radical constructivism, and the other emphasizing social impacts on learning process, e.g. Vygotsky's (1978) social constructivism (Aljohani, 2017; Brau, n.d.). The current review will present relevant arguments and discussions in selected constructivist theories for the conceptualization of knowledge and learning in this study.

Piaget pioneered cognitive constructivism with his theory about children's cognitive development. His proposition is that knowledge is cognitive structures

constructed to fit an experience a person has had with the world (von Glasersfeld, 1982). Inspired by Piaget's theory, Von Glasersfeld pushed the idea of cognitive constructivism further with his theory of radical constructivism. Von Glasersfeld defines knowledge as conceptual structures that organize a cognizing subject's pre-existing experiences to fit new experiences (von Glasersfeld, 2007). It means that we perceive anything through our mental structures if and only when it fits our mental structures intended for it, otherwise, it does not exist in our perception. Von Glasersfeld's theory is radical in ways that he completely rejects the notion of objectivity in knowledge construction. He contends that knowledge constructions are not reflection of 'objective' realities, and they are not either true or near copies of external entities in the world because, in fact, we have no means to validate the resemblance of our constructed mental structures with reality (von Glasersfeld & Cobb, 1983). For instance, a few years ago, a picture of a dress sparked much controversy on the internet about the actual colors of the dress, whether being black and blue, or white and gold. The interesting thing was that even family members seeing the same picture on the same device failed to see the same colors. Certainly, nobody won the debate over the actual colors of the dress, but the story makes a good point that what we perceive is not necessarily the real reflection of reality but simply what our conceptual structures allow us to see.

Nevertheless, our knowledge constructions must be viable as we are living in a world of constraints. An EFL learner may know that the sequence of letters, e.g., '*ch*' in English is usually pronounced as /tʃ/, as in *child, church, teacher*. Thus, when encountering a word like '*archive'*, he or she may think that the word '*archive*' should be pronounced with the sound /tʃ/ instead of /k/. He or she may keep that thinking and apply that pronunciation for the word until he/she encounters some people pronouncing it with the sound /k/. This may change his/her way of thinking about how to pronounce the word, especially when the former way of pronouncing often leads to misunderstanding or low intelligibility in communication. The concept of viability of knowledge essentially relates to the concept of fit. Note that the concept of *fit* is often compared with *match*. The distinction between *match* and *fit* is that in a match we seek the absolute correspondence between what we construct and reality, whereas a

fit means what works. Von Glasersfeld (1983, p. 8) used an example of the key which a burglar may use to enter a house to illustrate the concept of fit.

> *The 'key' with which he successfully opens the door might be a paper clip, a bobby pin, a credit card, or a skillfully crafted skeleton key. All that matters is that it fits within the constraints of the particular lock and allows the burglar to get in.*

Bodner (1986, p. 5) restates that idea as follows.

> *Each of us builds our own view of reality by trying to find order in the chaos of signals that impinge on our senses. The only thing that matters is whether the knowledge we construct from this information functions satisfactorily in the context in which it arises.*

When a person confronts an experience that does not fit within the pre-existing cognitive structures, there are two possibilities. One possibility is, as mentioned above, that the experience will come out of his or her perception, being non-existent to that person. The other is that there will be a re-construction or re-organization of the pre-existing mental structures so that the new experience can be perceived.

The two possibilities can be explained by two biological processes as suggested by Piaget: assimilation and accommodation. To understand these two processes, it is vital to get the idea of a cognitive structure that Piaget and Von Glasersfeld termed *scheme* (Bodner, 1986; von Glasersfeld, 1982). An action scheme comprises of a trigger, a reaction (or an action), and a result of the action. In the example about the pronunciation of the word '*archive*', the trigger is the learner's recognition of the sequence '*ch*' in the word, while the application of an existing scheme for other words containing this sequence is the reaction. The result is that the learner pronounces '*ch*' in the word '*archive*' as /tʃ/, which, in turn, may cause confusion or decrease intelligibility of the word in oral communication. As Bodner (1986, p. 2) explains, "*assimilation* involves applying a pre-existing scheme to interpret sensory data". However, if the application of a pre-existing scheme does not bring out an expected result, such as the pronunciation causing confusion to the listener(s), the scheme will be modified to fit the constraints of the new experience, in this case, that of English

pronunciation and communication. This latter process is called *accommodation* (Bodner, 1986; von Glasersfeld, 1982).

Certainly, in education we do not expect the first possibility to happen; however, admittedly it happens quite frequently even without any pedagogical influences. The task of distinguishing phonemes in minimal pairs in EFL learning, say /θ/ and /ð/, exemplifies this matter. For native speakers of English, distinguishing the two phonemes causes no problem, that is not to say it is effortless. Yet for many non-native EFL learners, like Vietnamese EFL learners, it might not be an easy task. Many of them may not recognize the distinctive feature of the two phonemes. The reason is that their cognitive structures possess few or almost no prior experiences with these phonemes, consequently, listening to these phonemes is a novel experience and thus it does not fit well within the learners' current cognitive structures. It also applies to situations when we listen to conversations in a foreign language that we have never had experience with. We will probably feel that everything we hear seems meaningless and that it sounds like strings of noise. The major difference between a native speaker and an EFL learner in the example about the minimal pair task is that the conceptual structures of the native speaker include many past experiences with the usage of the two phonemes, whereas that of the non-native learner may contain very few or none.

Returning to the argument of the two possibilities when the learner encounters a new experience, for the second possibility to happen, it is important that there is a re-organization of one's mental structures which accommodates the perception of new experience. In learning, self-evaluation of the viability of knowledge may initiate this process. If learners are aware that their mental structures clash with the contexts and that the clash creates unexpected results, they may feel a need for re-organizing their mental structures so that its viability will be improved. It is a matter of biological adaptation, so as to survive in a world of constraints. Obviously, it is not easy to get it done, however, it can be done with awareness-raising activities according to Lian (2014), or reflection as Von Glasersfeld's (2007) termed it. In short, learning is an adaptive process in which clashes with the constraints in the experiential world are noticed and pre-existing conceptual structures are re-organized to accommodate new experiences.

Regarding the importance of awareness in learning, it is crucial that educators provide conditions and situations for fostering self-evaluation of the viability of knowledge, in other words, the recognition of possible clashes between the learner's application of pre-existing schemes and contexts. One common technique in L2 pedagogy has been the use of corrective feedback. Corrective feedback is "any indications of learners' non-targetlike use of the target language" (Kim, 2003, p. 1). Corrective feedback can be in written form or spoken form or can be pictures or videos with annotations. Von Glasersfeld (2007) insists that visual feedback is effective for learning, at least, for learning new moves in athletics. When watching slow-motion videos of their own performance, athletes were visually aided in self-evaluating their performance of a specific move, then by comparing it with the way that they suppose the move should be performed, they may have proper adjustments in subsequent practice.

Another example in language learning is the pronunciation training mobile app Elsa. In this app, the learner has to pronounce some given words containing the target phoneme and the app will use speech recognition technology to assess the accuracy of his or her pronunciation of the words based on American English or British English standard. Every time the learner pronounces a word as prompted; the app will immediately display feedback in terms of percentage of how well their pronunciation matches the standard so that they may have to adjust in subsequent attempts. The role and effectiveness of corrective feedback have been investigated quite extensively in L2 research, and it is agreed that corrective feedback is necessary in the learning process. Section 2.3 provides fuller discussions on this matter.

Returning to the debate on the origin of knowledge, radical constructivists view knowledge as cognitive structures, independent of the objective world and even of our senses. Knowledge is all about a subject's organization of pre-existing experiences, given that such an organization does not produce any clashes with the experiences. It is worth noting that radical constructivism does not reject the existence of objective realities in the world; however, it postulates that we cannot have an absolute or even near accurate perception of reality, even with our senses. Critics of radical constructivism suggest that we cannot truly be radical, rejecting the contribution of

our sensory system. Neisser proposed a theory about perception in which a person's sensory faculties can offer direct information for structuring or changing pre-existing conceptual structures which he called schemata. It is these schema structures that then guide further exploration of sensory information from the object. This idea was first introduced in his book *Cognitive Psychology* (2014) and an article about the perceptual cycle (Neisser, 1978), see Figure 2.1. According to Neisser, perception is a source of self-knowledge. He took issue with the view that perception is an end product of a constructive process carried out inside one's head. While Neisser agreed that perception is actually a constructive process, he did not suggest that it happens only inside our heads. He maintained that perception, supported by our sensory system, is a continuous cycle taking place both inside and outside our minds. From the outside, we receive external information of the object through our senses. From the inside, we use schemata established from our existing experiences with the world. The schemata provide anticipations, directing selective exploration of sensory information. More importantly, this theory posits that new sensory data can be obtained directly and can change the currently active schema, and thus, changing subsequent anticipations and exploration of information. This cycle continues until we feel satisfied that our perception is adequate, and we stop the exploratory collection of sensory information. Therefore, perception does not just happen inside one's mind, but both the inside and outside.

For illustration, Neisser (1978) gave an example of tracking capability of human eyes, which even a few month old babies are capable of doing. When observing a running animal, for example, our vision will be directed according to the movement of the animal. Doing this certainly involves the influence of relevant schemata. It is the schemata that anticipate the direction of the movement of that animal so that the tracking and collecting visual information become successful, otherwise, we cannot see anything we need. Another illustration is the ability to follow a conversation in a crowded party. It is quite easy for an adult to ignore all other conversations from all around him and to focus only on the one talking to him.

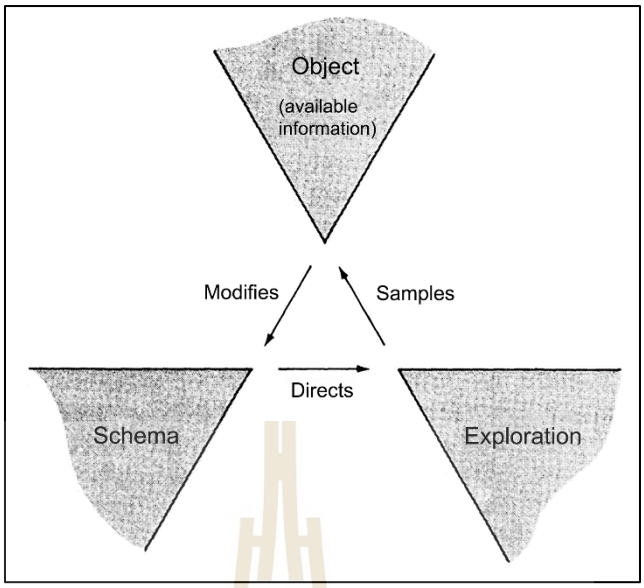


Figure 2.1 The Perceptual Cycle

To give just one more interesting illustration of the perceptual cycle, consider a unique experiment by McGurk and MacDonald (1976). They recorded a voice articulating a given consonant but presented it with a video of a person articulating a different consonant. What really strikes us is that the participants of the experiment reported hearing an unrecorded consonant. This shows that the incongruent visual information changed their original anticipation of what they would hear, therefore, they could not perceive the recorded auditory information but something else. However, suppose the recorded consonant appears in a word and in complete utterances, there may be no problem perceiving the auditory information even when there is incompatibility with visual information. This may happen because contextual information once again changes the misleading schema, and then redirects the exploration of information more accurately. This experiment also indicates that perception can take information from various sources; perception is multisensory and multimodal rather than being restricted to a single channel (Tiippana, 2014). Except for incongruence in the different sources as in the abovementioned experiment, the availability of multimodal sources of input is beneficial for accurate perception.

Neisser (1978) further contends that perception is not the result at any point in time. It is, however, the whole cycle of perception which is subject to change. He gave

an example of haptic perception. Perception of this kind cannot be achieved if we only rely on a point in time that we first touch an object. Instead, it should be taken place over a period of time regardless of short or long duration. Likewise, when we look at an object, a clock for instance, it means we have been receiving a ray of light sending to our eyes which has a shape called a clock. This act by no means can be done in a single fixation of the eyes. The idea that perception is a cycle, and new sensory data can modify pre-existing schemata and anticipations is very meaningful in education. It is because if we educators can help learners notice important sensory data that they miss previously for having misleading anticipations, there are possibilities that they may modify their currently active schemata for exploring and searching other available information that aids more accurate perception. Once again, awareness-raising activities seem necessary for this process. The issue is that as individuals develop diverse plans for their perceptions, raising awareness for different individuals is a real challenge. In the next section, I will discuss the individualization of learning as a solution to individual differences in building accurate perceptions.

The major difference between the two constructivist schools is the matter of how much sensory perception assists with accurate reflection of realities. As mentioned, radical constructivism insists that there is no direct perception of reality, everything we perceive is in our experiential world because our perception must fit the organization of our experiential world. On the other hand, perceptual cycle, a constructivist theory with a dualistic position, is rather positive about this matter, claiming that direct perception exists, and it helps us reflect reality more accurately. It is clear that, knowledge in radical constructivism exists only in the mind while knowledge in constructivist theories with a dualistic position resides both inside and outside the mind (T. M. Duffy & Jonassen, 1992), as perception of objective realities (Neisser, 1978, 2014) or as socially constructed schemata (Lyddon & McLaughlin, 1992).

From the above arguments in constructivism, the theoretical standpoint of the present study is summarized as follows.

1) Knowledge is subjectively constructed, and not transferred (von Glasersfeld, 1991). This study embraces the idea of radical constructivism proposed by Von

Glasersfeld that knowledge is a person's cognitive structures, an organization of one's pre-existing experiences.

2) Since individuals may have different histories of experiences, personal knowledge constructions may vary considerately, or in other words, our perceptions of the world are unique (Cole, 1992; Hendry, 1996; Karagiorgi & Symeou, 2005).
3) Perception of sensory information is a source of knowledge, and it is also constructed, not passively received from external objects (von Glasersfeld, 1991).
4) Knowledge is not mental copies of reality in the world, neither should it match the reality, however, it should fit the contexts where it arises (Bodner, 1986). Therefore, knowledge constructions must have viability conforming to the constraints of the world (von Glasersfeld, 2007).
5) Viability of knowledge is a decisive factor in re-constructing or re-organizing one's pre-existing cognitive structures, and the awareness of viability is a principle source for triggering learning process (von Glasersfeld & Cobb, 1983).
6) Learning is an adaptive process of re-structuring/ re-organizing a person's pre-existing cognitive structures and re-directing the exploration of new experiences (Bodner, 1986; von Glasersfeld, 1991, 2007).

The above premises from radical constructivism lend support for the first and most important principle for developing the learning system in this study, that is to respect the learner's meaning-making process. Learning activities and instructional support on the learning system, therefore, should not impose knowledge on the learner. They also indicate that the learning system should make use of awareness-raising in learning.

2.2 Personalized learning

The conceptualization of learning in constructivism presents a considerable challenge to both learners and educators: individual differences (ID). As Dörnyei (2005, p. 4) defines it,

> *ID constructs refer to dimensions of enduring personal characteristics that are assumed to apply to everybody and on which people differ by degree. Or, in other words, they concern stable and systematic deviations from a normative blueprint.*

The 'enduring personal characteristics' may imply many individual variables. Cooper (2002) classified individual variables into four groups: personality, ability, mood, and motivation. De Raad (2000, p. 41) included more categories such as "attitudes, values, ideologies, interests, emotions, capacities, skills, socioeconomic status, gender, height, and so forth". Dörnyei (2005) also discussed other individual variables specifically relevant to language learning; these are language aptitude, cognitive styles, self-regulatory capacity, willingness to communicate, creativity, anxiety, and learner beliefs. The current discussion will not elaborate on the descriptions of those individual differences; however, it is adequate to suggest that learners are distinct in many aspects, which in turn affects the efficiency and the outcomes of their language learning (Sasayama, 2018). Take age as an example. Learners at younger ages are generally in a better position than adults in terms of learning a foreign language (DeKeyser, 2012). Although research has shown that young learners may have a slower rate of language learning in the beginning, their level of second language attainment is higher than adult learners (Long, 1990). Another example is language aptitude. The language aptitude or the capacity of learning a language, alas, varies across individuals. That is to say, some people are more likely to be successful in learning a language whereas others are not. Likewise, people with high working memory capacity, a component of cognitive abilities, have an advantage in language learning (Vatz et al., 2013). Given the variability in personal knowledge constructions and the various individual differences, effective learning inevitably requires personalization.

2.2.1 Precision language education

Personalization in education is actually not a new thing, and yet, it has become a trend in the 21st century. The original goal of personalized education concerns tailoring instructions and providing learning conditions as to meet the learner needs. This idea leads to the concept of precision education and more recently precision language education proposed by A-P Lian & Sangarun (2017). The concept of

precision language education, inspired by precision education and formerly precision medicine, is a new conceptual move in which language instructions or pedagogical interventions should focus on addressing specific problems experienced by learners and moving away from group characteristics, just as precision medicine is intended for specific needs of individual patients. As Lian and Sangarun (2017, p. 3) put it,

> *The essential characteristic of precision (language) education is the desire to access information that is as detailed and accurate as possible about learner characteristics and performances in order to initiate the most effective intervention in support of the students' learning efforts.*

The move toward precision language education turns individual differences, usually considered as an unavoidable nuisance in education, to that of navigating effectual ways to language learning by responding to individual variability instead of normalizing it. Lian and Sangarun (2017) suggest that the distinction between precision education and personalized education is in the goal of getting information "as detailed and accurate as possible" (p.3) about learner variables and needs, with the former being the full personalization and "the ultimate objective of the research effort" (p.6). In precision language education principles, pedagogical interventions should aim at "providing accurate, detailed, timely, adaptive and contextualised personalised data" (A.-P. Lian & Sangarun, 2017, p. 4) so as to arrive at the most effective solutions to individual learner's problems. The premise is that we can increase chances of success in learning if 'precise' data of performances can be obtained and made available to the learners since awareness of the viability of knowledge can trigger effective learning (von Glasersfeld & Cobb, 1983). However, in normal teaching situations, the achievement of 'precision' is very rare unless there is extra assistance from available technologies, the potential of which will be discussed in later sections. Nevertheless, the precision agenda is often accompanied by awareness-raising activities. In this study, awareness-raising activities are integrated in the listening practice by means of computer-generated corrective feedback.

2.2.2 Research on personalized language learning

According to a recent review by Chen et al. (2021), scientific publications on personalized language learning have increased significantly in quantity in the last twenty years, indicating a growing interest in the L2 research community. Previous research studies have also implied the encouraging results of implementing personalization in various areas of L2 teaching and learning.

For personalizing recommendations and instructions, Xie et al. (2016), for instance, proposed a personalized system for learning English vocabulary which utilized learner-based profiles for generating personalized tasks and words for individual learners. The learner profiles receive the data from the students' learning logs and social network data. The findings show that the group studying with the full personalization (that is personalization of word coverage, task diversity, and context familiarity) achieved significantly higher results in word retention, learning completion rate, and level of enjoyment than the other groups studying with lesser degree of personalization. Similarly, Chen and Li (2010) developed a context-awareness vocabulary learning system by detecting the learner's location and suggesting corresponding learning materials for certain locations. The findings of the quasi-experiment indicate that after 2 weeks using the context-awareness learning system, the experimental group outperformed the control group on the vocabulary posttest. The researchers attributed the positive outcomes to the benefits of situational learning.

Another area of personalized learning also attracts much research attention is personalizing contents in digital language learning games. Various methods have been proposed for this purpose such as the data-driven procedural content generation method for learning reading skills proposed by Hooshyar et al. (2018) or the use of AI systems to continuously assessing the students' learning performance and generating tasks and words accordingly developed by Pereira et al. (2012).

Despite the positive findings presented above, there are three things worth mentioning. First, the effectiveness of implementing personalization in learning is inconclusive; and in fact, a few studies indicate mixed results (Enyedy, 2014; Murray & Pérez, 2015; Yarnall et al., 2016). Second, while previous investigations into the effectiveness of personalized learning are numerous, that of personalized language

learning are extremely limited (Yarnall et al., 2016). Third, the research focus of personalized language learning is predominantly on grammar learning (Fang et al., 2018; Haristiani et al., 2019), reading comprehension (Hsu et al., 2013; Lo et al., 2013; Wu et al., 2014), writing performance (Botarleanu et al., 2019; Price et al., 2017), or vocabulary learning (C. M. Chen & Chung, 2008; Y. M. Huang et al., 2012). Scientific reports on the effectiveness of personalization in L2 listening learning are relatively rare. One of them is the study by Pérez-Segura et al. (2020).

Pérez-Segura et al. (2020) conducted a quasi-experiment with 136 sixth-graders to explore the effectiveness of personalized feedback through an audience response system (ARS). The personalized feedback in their study was given in the form of listening and reading exercises based on students' problems identified in previous sessions. The findings show that the students' EFL reading and listening performances improved significantly before the training and their performances were better than that of the control group studying without personalized feedback. Nevertheless, as stated above, little has been known about the effectiveness of personalization in L2 listening learning and thus more research on this matter needs to be conducted. The present study hence contributed to the our understandings of this research area.

2.3 Corrective feedback in L2 pedagogy

The effort of collecting detailed data from the learner's performances in the end primarily serves the provision of increasingly 'precise' feedback for individual learners with diverse needs and difficulties. Gass (1997) claims that language learners can be exposed to two types of input: positive evidence and negative evidence. Positive evidence entails "the set of well-formed sentences to which learners are exposed" (Gass, 1997, p. 36). By contrast, negative evidence, also known as corrective feedback (CF), is "any indications of learners' non-targetlike use of the target language" (Kim, 2003, p. 1)

CF is broadly presented either in oral form or written form. Ranta and Lyster (2007) classified oral feedback into six types, i.e. recasts, explicit correction, elicitation, metalinguistic clues, clarification requests, and teacher repetition; these types can be further classified as reformulations or prompts. The distinction between a

reformulation and a prompt is that the former provides positive evidence whereas the latter withholds the targetlike forms and only prompts learners to retrieving them from their knowledge (Lyster, 2018). In a meta-analysis comprising 15 studies on oral feedback, Lyster and Saito (2010) concluded that learners may benefit more from negative evidence as found in prompts, though they may also benefit from positive evidence such as that in recasts.

With narrower applications than that of oral feedback, written CF is predominantly used in grammar and writing instructions for indicating learner grammatical ill-formedness in foreign language writing (Kang & Han, 2015). Investigations on the usefulness of written CF have yielded relatively congruent results revealing the positive effects of written CF on writing accuracy improvement (Bitchener, 2012). For instance, Kang and Han conducted a meta-analysis study on the effectiveness of written CF and found a moderate effect size of *.54* on which they claim that "written corrective feedback does have a substantive effect on L2 written accuracy" (Kang & Han, 2015, p. 10).

By and large, there has been a consensus that providing CF results in greater improvements in language learning than withholding it (Lyster, 2018). Notwithstanding, the effectiveness of CF necessarily involves several issues.

First of all, CF may vary on the degree of explicitness, either being implicit or being more explicit, which influences the effectiveness of CF. With several studies accumulatively showing the greater benefits of explicit feedback, learning outcomes are expected to improve provided that feedback is conveyed explicitly (Nassaji & Kartchava, 2017). Moreover, research has shown that learners benefit more in terms of noticing their own problems when being exposed to explicit feedback (Ellis, 2017; Heift & Hegelheimer, 2017), and that they enjoy feedback being explicit and direct (Li, 2017; Quinn, 2014).

The second issue concerns that of timing. Due to the mixed results from past research, there has been no conclusion as to whether immediate CF or delayed CF would benefit learners more (Quinn & Nakata, 2017). For example, Li, Zhu, and Ellis (2016), Li (2010), Lyster and Saito (2010) assert that immediate CF is more effective than delayed CF. By contrast, Quinn (2014) claims that instructions with delayed CF

would be more beneficial to L2 learners. Be that as it may, the literature seems to have preference for immediate CF since there are some theoretical grounds for using immediate CF (Ellis, 2017). One of the theoretical claims backing immediate CF might draw on Long's (1996) Interaction Hypothesis in which immediate CF creates an opportunity for learners to engage in the reflection process, or to put it another way, a process of comparing what they performed in the target language and what is accepted in the target language (Li et al., 2016). Immediate CF thus facilitates the noticing of the viability of knowledge, as discussed in 2.1, a crucial element in the learning process. Another theoretical account is that when CF is provided immediately, it is directly connected with a context, hence, it can trigger contextualized learning which is claimed to assist learners with performing better in communication (Li et al., 2016).

The third issue affecting the usefulness of CF is precision, or how precisely CF addresses learners' problems. Surprisingly enough, this issue has rarely received inquiries over many decades. In a case study, Han (2008) proposed the notion of 'fine-tuning' CF which is a process of probing primary causes of learner's errors in attempting to provide more quality CF. Han (2008) reported in one case of the study that the CF provided did not help the learner notice the problem at all and did not bring about any change in the learner's target language knowledge structures. In another case, however, the researcher employed the process of fine-tuning CF, i.e. investigating the causes of the learner's errors by looking at the learner's invariants in performances and the learner's perspectives. Han then provided remedial actions responding to the revealed causes. Han (2008, p. 590) reported that the process was "extremely successful" and that the learner in the second case could eventually overcome her once persistent errors. Han went on to suggest that "the provision of finely-tuned feedback, which successfully brought about sustained – as opposed to superficial – change in learner knowledge and behavior" (2008, p. 591). The process of fine-tuning CF seems, to a certain extent, identical to the agenda of precision language education as mentioned earlier, in which learner's performance and learner's perspectives should be obtained to inform as detailed and accurate evidence as possible the provision of feedback. This principle of learner-oriented CF assures (a) that there is a

correspondence between what problems learners experience in learning and what may cause them, and (b) that the provision of CF is intended on an individual basis, as opposed to one-size-fits-all feedback. Failure in effecting CF, hence, possibly originates from a lack of precision.

The fourth issue concerns the type of interaction in which CF happens. Obviously, the kind of face-to-face interaction featuring in teacher-learner feedback has been predominant in L2 classrooms ever since the early days of foreign language teaching and learning. Recently however, researchers have also advocated another kind of face-to-face feedback, namely peer corrective feedback (Kim, 2003; Sato, 2017). Sato (2017) posits that peer CF can assist in L2 development given that the social elements affecting the peer interactions are carefully considered. That is to say, peer CF effects only when learners can "view correction positively, trust each other's linguistic knowledge, and are comfortable with correcting and being corrected by a peer" (Nassaji & Kartchava, 2017, p. 177). Peer CF is advocated over teacher-provided CF for that it maximizes the learner exposure to the target language through peer exchanges and that it increases opportunities for modifying output production (Sato et al., 2017).

It is, however, admitted that in real teaching situations like the case of Vietnam where there can be up to 50 students attending a 90-minute English class, face-to-face CF of any kind may fail to address at least two issues mentioned above, i.e. timing and precision. The fact that there is only one teacher in charge of a whole group of 50 students means that the teacher is unlikely to provide face-to-face CF on an individual basis and that he or she has little control over the peer exchanges to ensure peer CF works in a desirable manner. The next issue, as regards the effectiveness of CF, of course, concerns the practicality of ways for individualizing CF. Responding to this issue, computer-mediated CF has emerged as a potential solution.

Heift and Hegelheimer (2017) identified two CALL environments for CF provision: Tutorial CALL and Automatic Writing Evaluation (AWE). Traditionally, Tutorial CALL applications are capable of generating immediate, error-specific feedback by comparing learners' input with the pre-set answers. This matching technique, however, may exclude some variations of learners' correct answers when the teacher cannot anticipate all the possible answers. Sometimes a variation in spelling may be treated

as incorrect. Recognizing the possible deficiencies of string-matching techniques in traditional Tutorial CALL, another type of Tutorial CALL has been developed, namely Intelligent Language Tutoring System (ILTS) or Intelligent CALL (ICALL). ILTS and ICALL are more effective than traditional Tutorial CALL programs in that they can categorize errors by incorporating Natural Language Processing (NLP) and Artificial Intelligence (AI) techniques in replacement for the traditional string-matching techniques (Heift & Hegelheimer, 2017). With the incorporation of NLP techniques, ICALL applications are capable of identifying various types of errors including those being unanticipated by the teacher, thus more accurate feedback and metalinguistic guidance. At the present, ICALL applications are normally designed as web-based tools, with some offering a plug-in for offline document processing software, such as the renowned Grammarly application (www.grammarly.com).

Tutorial CALL and ICALL applications are generally restricted to CF provision on morphosyntactic errors at sentence level. AWE applications, on the other hand, encompass a wider range of spelling, morphological, syntactic and semantic features as well as rhetorical structures and discourse features in essay writing. AWE applications such as WriteToLearn (Pearson Education) and Research Writing Tutor (Iowa State University) can generate on-demand, individualized CF on learners' essay writing by incorporating NLP and statistical techniques. Building on a corpus of genre-specific written texts, for instance, Research Writing Tutor application can point learners to any instances in which there exists a rhetorical problem or erroneous grammatical construction or inappropriate word choice. Furthermore, most AWE systems also provide learners with a general evaluation like an overall score for learners' writing products.

Although ICALL and AWE generated CF, both relied on NLP and AI, has been corroboratively proven more effective than the traditional binary CF, such as *Correct* or *Incorrect*, (Bowles, 2005; Heift & Rimrott, 2008; Petersen, 2010), the reliability of the generated CF remains inconclusive. Some studies reported a low to moderate positive correspondence between the scores generated by AWE and that by teachers (James, 2006; Li, Link, Ma, Yang, & Hegelheimer, 2014), whereas other studies yielded a high reliability (Attali et al., 2010; Burstein et al., 2004; El-Ebyary & Windeatt, 2010).

Therefore, incorporating computer-mediated CF in L2 learning may increase opportunities for individualizing CF, yet this should be done with cautions regarding the type of CF to be provided, and the reliability of the matching techniques.

One last issue worth mentioning is that there is an overwhelming research emphasis on productive skills. Previous research has frequently examined the effectiveness of CF on L2 grammar, writing, and oral production development. It is surprising, however, that little research has been probed into the effects of CF irrespective of types, on receptive skills, especially L2 listening skills. To date, most research if any attempted to do this has been related to speech perceptual training (See Hardison, 2003; Lee & Lyster, 2015; Wang & Munro, 2004). Lee and Lyster (2016), for example, investigated the comparative effects of various types of CF on speech perception training. They targeted at training EFL Korean students to recognize two English vowel contrasts, i.e. /ɪ/ versus /i/ and /æ/ versus /ɛ/ in minimal pairs by exposing them to auditory CF and visual CF. The auditory CF contained the rejection and one of following: (a) the target form, (b) the non-target form, (c) a combination of both (a) and (b). In contrast, the visual CF contained only the orthographic text on the computer screen: *Wrong*. After eight training sessions over two weeks, they discovered that all participants exposed to CF of any kind improved in perception of the target contrasts, however, those who were provided with auditory CF outperformed their counterparts presented with visual CF. One interesting thing in this study is that Hee and Lyster included speaker variability in their perceptual training and assessment. They found that speech perception training with CF provision can transfer to unfamiliar voices, but only restricted to the trained words. When it came to untrained words, there was no significant effect. They went on to suggest that this method of perception training may be word specific unless the training period is extended. Nonetheless, to date previous research seems to put little effort in investigating the effects of CF on L2 listening comprehension development.

In a nutshell, the above issues each influences to a certain degree the effectiveness of CF provision in support for L2 development. Careful considerations of those factors may enable us to seize opportunities for effectively individualizing CF. Additionally, the last issue indicates that further research needs to strive for

incorporating CF into receptive skills, particularly L2 listening development. This present study, as an attempt to fil this research gap, will integrate individualized CF into the learning environment and it is hoped that this effort of individualization will result in positive effects in EFL listening comprehension.

So far, the theoretical accounts on which this study is based, including how learning is viewed, the conditions for learning to take place, and possible ways for creating these conditions by individualizing CF provision have been discussed. The chapter will now consider a specific area of interest in this study, i.e. L2 listening comprehension.

2.4 Second language listening comprehension

Listening was once viewed as a passive process in which the job of the listener was to unpack the speaker's encoded messages. Today listening is generally considered as an interpretive process (Richards, 2008) which means that it is a meaning-making process. It follows that assisting learners with their learning involves creating a personalized learning environment in which they are confronted with their own problems during a listening event and are provided with feedback and remedial support. To come up with any pedagogically sound interventions for L2 listening improvement, it is crucial that the nature of listening comprehension be clearly understood. Hence, the next section aims to provide conceptual and operational definitions of the construct of listening comprehension.

2.4.1 Conceptual definition

Speech contains continuous streams of acoustic signals which are sent to the listener's ears for perception. At this stage, a distinction between hearing and listening may be necessary. Both listening and hearing involve receiving the signals reaching the eardrums, but only listening involves a certain degree of attention. In fact, we are constantly receiving sound signals around us but not all of them are equally attended and processed (Worthington & Bodie, 2018). It is noted that even though being attended the continuous stream of speech signals is not normally processed as a whole, but rather in smaller units. These units of processing can be features, phonemes, syllables, words, chunks of words, or other sound patterns. There are

various theoretical accounts regarding how a listener stores those units of perception in their long-term memory, however, in this study the exemplar models are preferred because this kind of perceptual models can explain why sources of variability in the speech signal can be handled efficiently (Keith Johnson, 1997). Unlike prototype theory which centralizes the most typical example, in exemplar models whenever a listener encounters an exemplar of a unit of perception, this exemplar is stored as one example in a set of multiple examples for that unit of perception. For example, every time a listener hears a word, say *puppy*, uttered in different occasions and/or speakers, each encounter with the word creates an exemplar of this perceptual unit. Furthermore, the listener can generate shared values or features among the examples of a given unit (Field, 2004), which explains why variations of a phonological word or multiple meanings of a word can be recognized in listening comprehension.

Units of perceptions can be related to each other when they frequently co-occur in certain situations. For example, the words *school, student, teacher, blackboard*... can be associated to one another because they frequently co-occur in the contexts of schooling or education. These associations are stored in the listener's mental structures called *schemata* (Rost, 2006; Vandergrift et al., 2012).

According to the Parallel Distributed Processing model (Scovel, 1998), when we listen to a speech, the speech signal activates potentially corresponding perceptual units at different levels simultaneously (e.g. phoneme level or/and word level) and the schemata relevant to the activated perceptual units. Alternatively, schemata can be activated by non-verbal cues such as situational contexts or visual information (C.-D. Nguyen & Newton, 2018). In the comprehension process, this activation of schemata, either provoked by verbal or nonverbal cues, serves as a structure of meaning representation onto which activated perceptual information are mapped, according to the structure-building model (Gernsbacher, 1990). The 'goodness of fit' will suppress all activated units except for only one. The 'best fit' perceptual information will then be enriched by further evidence from later processing (Rost, 2011). The structure-building model of comprehension, while originating in reading comprehension research, claims to be cognitive general and is not tied to a specific modality (Gernsbacher, 1990), and thus offers an explanation for individual differences in

comprehension skill (McNamara & Magliano, 2009). The distinction between a competent and a less-skilled listener lies in their efficiency of suppressing 'unfit' activated units of perception to arrive at the 'best fit' quickly, rather than that of enriching contextual information (Gernsbacher & Faust, 1991).

The process of activating, enriching, and suppressing perceptual units is supported by the information obtained from both bottom-up and top-down processing. As mentioned earlier, bottom-up and top-down processing refer to the directions of processing. Bottom-up processing, also known as data-driven processing, derives information from the stimuli of the speech signal to activate relevant perceptual units and schemata while top-down processing utilizes the schematic knowledge to influence the perception of the incoming speech signal. This is known as the *interactive model* of listening (Field, 1999; J. Flowerdew & Miller, 2005; C. Goh, 2016).

In a nutshell, listening comprehension is viewed as a process of constructing coherent mental structures of an attended speech, by activating relevant perceptual units which are then enriched or suppressed in relation to their goodness of fit for the structures (Gernsbacher, 1990, 1991, 1997). Figure 2.2 illustrates the cognitive processes in listening comprehension.

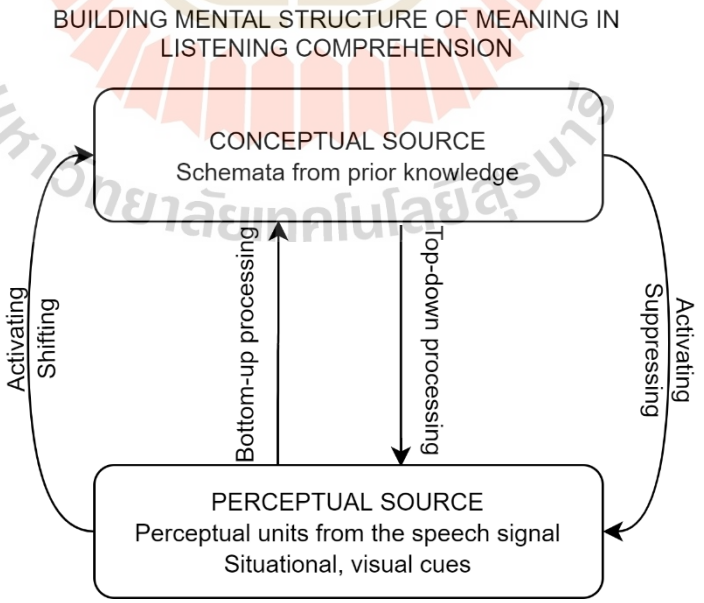


Figure 2.2 The processes in Listening Comprehension

As clearly shown in Figure 2.2, to comprehend a speech the listener must draw on two sources: perceptual and conceptual information (Field, 2011). The perceptual information is primarily derived from the speech signal in the speaker's utterances while the conceptual information is provided by the activation of relevant schemata. The interactions of the two sources of information generate evidence provided for building the structure of meaning of the attended speech. Failure to obtaining information from one source may lead to failure in comprehension. It is important to stress that while perceptual information derives from the speech signal, the resulting perception is still a product of a combined processing of both perceptual and conceptual information, that is to say, no direct perception is obtained from the speech signal.

2.4.2 Operational definition

Operationalization of a construct concerns how a construct is defined in relation to its measurement (Worthington & Bodie, 2018). Traditionally, L2 listening comprehension has been measured in three ways: the discrete-point, the integrative, and the communicative approach (Buck, 2001).

The discrete-point approach, proposed by Robert Lado (1961), purports that listening involves the ability to recognize isolated linguistic elements in the orally presented input. Tasks such as phonemic discrimination, paraphrase recognition, and response evaluation are generally recognized as discrete-point-based tasks for listening assessments. Buck (2001) pointed out two issues with this approach. Firstly, it assumes that language is comprised of divisible elements and skills and that knowledge of each element can be tested independently of others. Secondly, spoken language and written language differs only to the extent of the medium of delivery. However, it becomes clear that understanding an element of interest requires processing and understanding other elements as well. Moreover, it is undoubted that spoken language is distinct from written language by nature (J. Flowerdew & Miller, 2005).

The integrative approach to listening assessment grew out of the criticism of the discrete-point tradition. This approach views language as a system of related elements and language competence is an integrated facility (Brunfaut & Rukthong, 2018). Listening is thus assessed as an integral part of the system and comprehension

is dependent on various knowledge sources such as lexical, syntactic, paralinguistic, discourse and world knowledge. It follows that performing an integrated listening task requires using more than just listening skills but integrating knowledge of different types as well. Cloze, summary completion, dictation, and statement evaluation are generally identified as integrative tests for listening (Buck, 2001). Obviously, taking these tests necessitates not only listening skills but also reading skills, in the case of cloze or summary completion, or writing skill, given the task of dictation. This approach is concerned more with the assessment of language use and not language elements in isolation, thus an emphasis on the processing of language in comprehension (Buck, 2001).

While there was a shift from testing knowledge of isolated linguistic elements to testing processing of several related elements, the integrative approach has been said to disregard the communicative purpose (Buck, 2001). With the influence of the communicative language teaching approach, language is thought to be used within a certain situation and for certain communicative purposes. Accordingly, knowing a language implies not only the knowledge of its elements but rather the ability to use it appropriately in real-life communication, or also known as communicative competence. This has a methodological effect in listening assessment, by shifting away from language performance testing to language performance testing in a specific, real-life, communicative event. In other words, a candidate of a communicative listening test should have an authentic need for listening and a clear goal of what to listen for. Here arises an issue of authenticity, features of the listening texts that simulate real-life target-language speech (Buck, 2001). Note that even though a text can be genuine given that it is taken from a real communicative situation, it may not be authentic to the listener as it may not match the listener's communicative purposes. Admittedly, most listening tasks assumed to be communicative tend to be very contrived in the sense that it provides the listener with a hypothetical purpose of listening. Additionally, Flowerdew and Miller (2005) also took issue with this approach regarding the transferability of the assessment. While communicative tests attempt to probe the listener's communicative performance in a specific situation, there is no guarantee that the listener can transfer it to other real-life events as there are countless

communicative events in the real world. However, as an effort to make listening more communicative, thus meaningful, the positive effects of this approach are undeniable.

The above discussions on the three main approaches to the measurement of listening comprehension lay the foundation for the operationalization of listening comprehension in this study. With careful considerations of the requirements for authenticity as well as how listening operates, listening comprehension is operationalized as the listener's ability to understand the target spoken language in a communicative event, which provides both purposes and contexts for understanding, by drawing on the listener's existing knowledge and skills.

2.5 Issues in L2 listening pedagogy

Despite having an important role in language development and communication, listening is not always viewed as a skill to be taught. As Gilman and Moody (1984) points out, there are many possible reasons for the neglect of teaching L2 listening. First, learning a language is often associated with the capability of speaking that language rather than comprehension. Second, much emphasis on listening instruction and practice is thought to possibly discourage the development of speaking as it reduces the amount of speaking time. Third, because people often go to school to learn reading, writing, literature, and other skills but rarely listening skill in their native language, this somehow influences the belief of L2 teachers of what should be taught in class.

Evidently, those beliefs were reflected in the marginal roles of listening in the early days of L2 pedagogy. In the grammar translation approach prevalent from the late 19th century to the early 20th century, listening was not considered as a skill in L2 learning and was not taught in the classroom (J. Flowerdew & Miller, 2005). In the grammar translation method, learning a language involves learning a set of finite rules which was made possible through learning grammar rules deductively and translating sentences to L1. Reading comprehension and writing were the learning outcomes while L1, not L2, is the medium of instruction. In the 1920s, direct method gave listening an important role in L2 classroom. The medium of instruction was completely in the target language and listening was the first skill to be learned. However, the teaching of

listening did not actually happen in direct method as the method rested on the premise that learners can gradually understand the spoken language merely through direct exposure (J. Flowerdew & Miller, 2005). Although it may be the case when learners immerse in an L2 speaking environment in which they are exposed to L2 speech in almost any form of real-life communication, in most L2 classroom contexts exposure to L2 spoken texts alone does not guarantee improvement in comprehension, especially for less-skilled listeners (Field, 2009). Like direct method, audio-lingual method gave listening a strong position in L2 teaching. Audio-lingual method with its root in behaviorism emphasized the use of aural and oral drills in which listening had the first place in language learning, allowing the teaching and learning of pronunciation and grammar rules. Despite having a good position, listening only had an instrumental role to learning pronunciation and grammatical structures, and was not necessarily the main goal of teaching. Since 1970s, with the enlightenment of Communicative Language Teaching (CLT) approach, listening started to gain a more prominent role in communication as it is one element of 'communicative competence' which is the goal of the communicative approach (J. Flowerdew & Miller, 2005). Since 1980s listening has been corroboratively considered as a primary source for L2 acquisition as influenced by Krashen's input hypothesis (Weaver, 1972; Wolvin et al., 2000), and hence, the very first issue concerning whether L2 listening skill should be taught finally came to an end.

As the status of listening has changed, various instructional approaches to L2 listening have come into being over the last 50 years, although commentators contend that L2 listening has undergone "an unproductive tradition of instruction" (Field, 2008b, p. 2). Field (2009) suggests that major approaches to teaching L2 listening may have either a product orientation or a process orientation. The product-oriented approach places great emphasis, of course, on the product of listening which is comprehension. A typical listening lesson in the comprehension approach involves three stages: pre-listening stage, while-listening stage, and post-listening stage (See Figure 2.3). In the pre-listening stage, teachers prepare learners for the listening text by familiarizing learners with the topic, creating motivation for listening, and pre-teaching a few keywords. In the while-listening stage which is at the heart of the lesson, teachers

pose comprehension questions on contexts or speaker attitudes and have learners listen to the text to find out the answers. Teachers then check learners' comprehension of main ideas and details of the text. Comprehension checking may involve repeated listening to some parts of the text. Finally, in the post-listening stage, teachers draw learners' attention to functional language and unknown words in the text, which is often carried out by using concept check questions.

A Typical Format of a Listening Lesson in the Comprehension Approach

Pre-listening

Establish context

Create motivation for listening

Pre-teach only critical vocabulary

Extensive listening

General questions on context and attitude of speakers

Intensive listening

Pre-set questions

Intensive listening

Checking answers to questions

Post listening (optional)

Functional language in listening passage

Learners infer the meaning of unknown words from the sentences in which they appear

Final play; learners look at transcript

(Adapted from Field, 2009)

Figure 2.3 A typical format of a listening lesson in the comprehension approach

Partly due to the tradition of teacher training, many L2 teachers and practitioners have been using this format for constructing a listening lesson of some kind. However, this comprehension approach has some weaknesses. The most fundamental flaw of this approach, according to Field (2009), is an overemphasis on comprehension. Teachers, because of this overemphasis, usually seek evidence of learners' comprehension of a given text by asking them to provide answers to their pre-set comprehension questions. Those answers, as the paradox emerges, give little information about the process of listening. Correct answers may be deduced from their

guesses or based on a few words in the text rather than a comprehensive understanding of the text. Likewise, incorrect answers may not necessarily mean that students do not understand the text at all. The fallacy of this overemphasis is that learners' problems in the listening process are likely unattended.

The process-oriented approach, grown out of the criticisms of the comprehension approach, pays special attention to the underlying processes of listening comprehension, i.e. top-down and bottom-up processes as discussed in the conceptualization of listening comprehension. Different process approaches may vary according to the degree of favoring top-down processing or bottom-up processing.

Top-down approaches to L2 listening instruction focus on recognition and application of top-down information to aid global understanding of the text. Learners may use their existing knowledge of the topic, for instance, to predict the incoming speech. Similarly, they may use their discourse knowledge to recognize the start or the end of a proposition in a speech. It is noted that top-down listening skills are sometimes confused with listening strategies. Wilson (2018a) contends that top-down listening skills or macro listening skills differ from strategies in that skills operate unconsciously in the listening process and strategies are deliberately used as to achieve comprehension. This assumes that applying a strategy inevitably involves the operation of related skills. There have been various taxonomies of top-down listening skills such as Richards (1983), Rost (1990), and Brown (2007). Figure 2.4 is a taxonomy of macro listening skills proposed by Wilson (2018a).

Taxonomy of top-down listening skills

1. Recognize cohesive devices in speech.
2. Recognize the communicative functions of utterances.
3. Recognize where a proposition begins and ends, signaling a topic shift.
4. Use real-world knowledge to infer ideas and purposes of speech.
5. Infer links between events, for example, cause and effect, generalization and example.
6. Distinguish between main ideas and subsidiary ideas.
7. Distinguish between literal and implied meanings.
8. Predict outcomes from events described.
9. Use nonverbal clues, for example, body language, facial expression, to aid comprehension.
10. Use paralinguistic features, for example, pitch, volume, fillers like um and er to aid comprehension.
11. Guess the meaning of unknown words based on contextual clues.
12. Process speech which occurs at different rates.

(Wilson, 2018a, p.2)

Figure 2.4 Top-down listening skills

With a focus on top-down features, a special consideration of this approach necessarily lays on context. Instructions therefore often include pre-listening activities in which contexts are established. The main purpose of pre-listening activities is to activate relevant schemata for the listening text. Common techniques for schema activation include brainstorming, using visuals or realia, using texts, introducing a situation, or discussing opinions and facts (Wilson, 2008). Teachers also hope to spark learners' interest in the listening text and pre-teach some topic-related words. Teachers then have learners predict the contents of the text before the listening event happens. There are issues concerning this teaching scenario. First, listening practice in this scenario cannot model most real-life listening experiences in ways that listeners are rarely informed about the topic of a speech through discussions about it in advance. Pre-listening activities may prepare learners for the listening text but may not effectively prepare them for daily listening encounters. Second, pre-teaching of keywords in the text may distract learners from attending to other important

information in the text, and instead misdirect them to tracking occurrences of those words. This might turn a top-down activity to a bottom-up practice accidentally, with a focus on word recognition. We actually have had empirical evidence on this. Chang & Read (2006) discovered in an experimental study that pre-teaching vocabulary indeed had a negative effect on the participants' listening performance. Third, prediction may aid comprehension provided that it is accurate. In the end, a guess is merely a guess, and thus it can go wrong. When learners' prediction about the text is wrong, it may have an unfortunate effect on their listening effort because a wrong prediction may create confusion and frustration. Some skilled listeners are better at solving conflicts between their prediction and the actual contents of the text; however, it is not the case for many less-skilled listeners. For them, confusion may cost them a loss of subsequent information in the text.

Proponents of bottom-up approaches have criticized top-down approaches for relying too much on contexts and downgrading the primacy of the speech signal (Field, 2009). As opposed to top-down oriented instructions, bottom-up methodology often prioritizes practice of recognizing bottom-up features such as phoneme discrimination in minimal pairs, word segmentation, recognition of prosodic features and co-articulation processes (Wilson, 2018b). Figure 2.5 presents Wilson's (Wilson, 2018b) taxonomy of bottom-up skills which he called micro listening skills and Field termed decoding skills.

Field (2009) suggests that there is no problem when a listening lesson aims at comprehension, and in fact, every listening lesson should do so, however, comprehension should not be the end-product of listening in the case of weak learners. In support for a process approach favoring bottom-up skills, he recommends that lack of comprehension can be diagnostic sources for addressing learners' listening problems. Where comprehension is not achieved, appropriate bottom-up skills training may become effective remedies. The issue here remains how precisely teachers are informed of the learners' listening problems and in what ways they can get that information.

Taxonomy of bottom-up listening skills

1. Discriminate among the distinctive sounds of English.
2. Recognize the stress patterns of words.
3. Recognize the rhythmic structure of English.
4. Recognize the role of stress, intonation, and pitch in a speaker's intended message.
5. Identify stressed and unstressed forms of words.
6. Recognize reduced forms of words.
7. Identify where word boundaries occur.
8. Recognize typical word order patterns in English.
9. Deduce the meaning of unknown words.
10. Recognize performance variables including pauses, errors, false starts, and self-corrections.
11. Recognize grammatical word classes (verbs, nouns, etc.).
12. Recognize syntactic patterns (e.g., verb tense, noun-verb agreement).
13. Recognize ellipsis in grammatical units and sentences.

(Wilson, 2018b, p.2)

Figure 2.5 Bottom-up listening skills

There is another line of L2 listening instructions which instead relies on teaching strategies for compensating for learners' deficiencies in listening comprehension. Oxford (1990, p. 8) defines strategies as "specific actions taken by the learner to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferrable to new situations". Oxford's (1990) also offered a substantial amount of strategies used in language learning which O'malley and Chamot (1990) further categorized them into three types, including cognitive, metacognitive, and socio-affective strategies. Since then several models of strategy training have been introduced such as Chamot and O'Malley's (1996) Cognitive Academic Language Learning Approach (CALLA), Anderson (2002), and Vandergrift and Goh's (2012) Metacognitive Pedagogical Sequence. Nonetheless, research has yielded inconclusive results regarding the effectiveness of strategy training on L2 listening development. For instance, Vandergrift and Tafaghodtari (2010) implemented a metacognitive cycle, a process-based approach, in a thirteen-week listening course. The results show that the Experimental Group with the intervention of metacognitive instruction outperformed

their peers in the Control Group. Following the same approach, Zeng (2014) found that both the Control Group with normal classroom instruction and the Experimental Group with metacognitive instruction over the course of five weeks improved listening performance, while the Experimental Group significantly outperformed the Control Group. Bozorgian's (2014) study, also based on the Metacognitive Cycle, investigated the effects of metacognitive instruction on metacognitive awareness of listening and listening comprehension of a group of male adults. The results show that the participants improved their listening skills through the application of metacognitive cycle, but there was no significant gain in their metacognitive awareness. Interestingly, another study by Teng and Lin (2015) took a similar approach with a control and Experimental Group of young EFL learners. The results were rather opposed to that of Bozorgian's (2014) study since there was no significant difference in the listening performance of the two groups, but the Experimental Group gained better metacognitive awareness than the Control Group.

By and large, a perceived weakness of strategy instruction is that it possibly misleads learners to strategy learning and consequently learners may forget that they are learning a language. Some commentators insist that strategy training costs considerable time in the classroom which is better spent on practicing the language skills (Littlejohn, 2008; McDonough, 2006; Renandya & Farrell, 2011).

Of all the issues, the one pertaining to most current approaches has been an administrative matter. For long, most classes have been set up with 50 students or more, and often run within a few restricted hours. Language classes are no exception, at least from the observations of the researcher in Vietnam and Thailand. These administrative conditions restrict any current instructional approaches to do justice for individual learners in relation to their own listening problems. In addition, an approach overemphasizing comprehension, or top-down skills, or bottom-up skills, or strategies may not be adequate for an overall development of listening skills. Thus, a balanced approach which takes into account all of these issues might become an effectual solution. The present study will employ such an approach underpinning the practice of EFL listening which considers both comprehension of a text and process of comprehending it equally important. This means that while comprehension is the

objective of listening practice, failures in understanding contexts or recognizing words are attended and addressed as breakdowns in comprehension.

As L2 listening poses a myriad of challenges on the part of the listeners, the chapter will now consider a single challenge that creates the biggest barrier to the teaching and learning of L2 listening comprehension.

2.6 Issue of variability in the speech signal

As discussed in the conceptualization of listening, perceptual information is an inextricable source that the comprehension process must draw on. Unfortunately, the speech signal presents a great amount of variability originating from speakers' voices, allophonic variations, co-articulation and other simplification processes in connected speech. Moreover, for biological reasons, human articulations of the same speech elements are far from constant. This variability poses a major challenge to non-native listener's perceptual processing and thus directly affects L2 listening comprehension.

A regular way of thinking to overcome this challenge is reducing the variability or simplifying the input, yet this method has not been proven effective. On the other hand, there has been mounting evidence in the speech perception research that high variability training can improve speech perception learning including phoneme discrimination, word recognition, and indexical feature identification (Barriuso & Hayes-harb, 2018). The idea of introducing more variability in the input as opposed to reducing it originated from an influential work by Logan, Lively, and Pisoni (1991). Logan et al. tested this idea on six native Japanese speakers by presenting a set of 68 minimal pairs of /l/ and /r/ which is considered the most problematic English contrast for Japanese natives. Each minimal pair was recorded by six native English speakers and presented to the participants for a forced-choice identification task, accompanied by corrective feedback. A training session involved listening to each minimal pair of the set spoken by one speaker twice. The training program cycled through five speakers in the course of three weeks with 15 training sessions. The pretest and posttest results revealed that the training significantly improved the participants' ability to contrast /l/ and /r/. In a follow-up study, Lively, Logan, and Pisoni (1993) further found out that this type of training can even generalize to new listening contexts, and thus concluded that

inclusion of variability considerably contributes to perceptual learning and category formation which are vital in the listening process. Several later studies investigated this approach with different dimensions and confirmed the effects on speech perception improvement such as Bradlow, Pisoni, Akahane-Yamada, and Tohkura (1997); Perrachione, Lee, Ha, & Wong (2011); Lee and Lyster (2016). A recent study by Leong, Prince, Pitchford, and van Heuven (2018) discovered that when high variability training conditions were adjusted to fit individual listeners, the effectiveness was even greater than applying a fixed training condition and that the training effects transferred to untrained stimuli. In a similar study, Qian et al. (2018) confirm the benefit of individualizing training for each learner by implementing a computer-based training. They found out that the participants' segmental perception improved significantly, and the effects also generalized to novel voices, but not to untrained words. One particularly interesting finding from this line of research is that training with audio-visual input can produce more positive outcomes than audio-only input (Hardison, 2003).

The issue of variability in the speech signal which presents considerable difficulties on the part of the listening teacher and learner actually can be overcome by employing an effective training method with high speaker variability as discussed above. This effective perceptual training has been well-known in speech perception research for nearly three decades, yet very few L2 listening teachers have been informed of it partly because those findings were largely published in speech perception research and academic journals to which L2 teachers do not usually have access. The remaining issue, also a persistent one, is the research-practice gap in L2 pedagogy (Rahman, M.M. ; Pandian, 2016). Fortunately, with some support from available technologies, we can overcome this issue because this training method can be conveniently incorporated into L2 classrooms, for instance, with a web-based application as proposed by Thomson (2011).

In this present study, the issue of variability in speech perception will be addressed by employing high variability perceptual training. To maximize the effectiveness of the training method, this study will make use of audio-visual input and employ AI technology for automatic provision of varied speakers' contextualized input.

2.7 Technology and L2 listening

2.7.1 Digital affordances for L2 listening

Scattered elsewhere in this chapter has seen the role of technology in addressing various difficulties in learning and teaching L2 listening. Regarding the potential of technology for L2 listening, Hubbard (2010) proposed a taxonomy for digital affordances (see Table **2.**2.1). According to Norman (1988 as cited in Hubbard, 2017, p. 94;), affordances refer to "the perceived possibilities that a user has for such tools and applications". For example, a button showing caption texts may have different affordances depending on how learners use it for a certain purpose. They may use it for aiding comprehension by reading the texts instead of listening, or they may use it for aiding listening comprehension, or simply for learning new words.

Table 2.1 Mediating characteristics of digital devices and networks

Characteristic	Definition	Examples
Archiving and indexing	Making copies and providing location information for searches	Copying and indexing through tagging, titles, or content
Transferring	Moving digital information from one place to another	Webcasting, streaming, and downloading audio and video, as with YouTube
Linking	Associated one item with another	Clicking on words for associated definitions
Controlling time and time shifting	Controlling time; moving non-linearly through materials	Tracking or limiting time; anytime listening; jumping to desired portions of a listening text
Transforming	Changing the form of a listening text	Text to speech; speech to graphic form

(Hubbard, 2017, p.94)

Archiving and indexing help users to create copies and search available contents stored in the networks. This type of affordance has brought about a diverse, abundant, ever-growing supplies of listening materials on the internet. *Transferring* allows for media streaming and downloading for online and offline use, for instance

the features in websites like YouTube or SoundCloud. News media and podcasting are also made available by this affordance. This affordance of media streaming and downloading has made listening learning become anytime, anywhere. *Linking* helps connect a content with another content, for instance, linking a video with captions, or linking a word to its glosses. *Controlling time* is an affordance which allows listener/viewer to navigate nonlinearly to any part of the listening media as they wish. Most media players today provide a playback control bar for this purpose. Interestingly, this affordance if combined with linking can create useful media annotations, for example by linking a word to a video set to playback at a given point. Furthermore, if smartly combined with indexing, these affordances can produce applications allowing searches of such media annotations which can be used as effective learning tools. There are some websites offering these affordances such as Youglish.com and Playphrase.me. The affordance of *transforming* allows converting text to speech as in speech synthesizing software or converting speech to text as in speech recognition software. For example, many videos on YouTube are provided with automatic-generated captions if this function is enabled. In addition, many applications, for example Adobe Audition or Audacity, can allow users to modify quality of recordings like changing speech rate or applying low-pass filtering.

2.7.2 Typology of help options for L2 listening

With the abovementioned affordances and other newly emerging ones, technology can potentially assist with L2 listening. Such assistance can be provided by presenting computer-based help options for the student. Cardenas-Claros and Gruba (2013) have proposed a comprehensive framework, called *CoDe*, for conceptualizing (Co) and designing (De) help options for L2 listening in CALL environments. The framework consists of four types of help option: operational, regulatory, compensatory, and explanatory.

Operational help options are aimed at giving introductory instructions regarding how the software works, what features or help options are available and how to access them, what technical problems may occur and how to deal with them. Operational help options are characterized as technical support (Hubbard, 2017;

Romeo & Hubbard, 2010). In most CALL contexts, these include user manuals, help menus, frequently asked questions (FAQs), and tutorials.

Regulatory help options are used to direct learners both before and after performing a task. Pre-task regulatory help options instruct learners in approaching a task, for example, by suggesting learning tips or providing strategy training. Post-task regulatory help options are designed to assist with self-regulation in learning. These include, for example, explanatory feedback on learners' responses to a task or comments on learners' performance.

Compensatory help options are designed to provide input enhancement, i.e. making modification and salience for input, which are supposed to aid comprehension and learning (Chapelle, 2003). Modification involves adding sources to listening material, which compensate learners' breakdowns in comprehension. These modifications include adding visuals (videos or still pictures) to audio, texts (captions or transcripts) to audio, or translated texts (L1 captions or L1 transcripts) to audio. Salience involves making some parts of the input become more accessible to the learner through control and repetition. In CALL software for listening, these help options for increased salience are presented as speed control buttons, media controller bars or sliders, and volume controller buttons (Hubbard, 2017).

Explanatory help options are also aimed at increasing salience of input by providing extra information pertaining to the input rather than by interacting with the original input. These operations involve making learners aware of some parts of the input by linking certain parts of a transcript to glossaries, definitions, or entries in dictionary. Explanatory help options can also enrich input by giving external references such as culture notes, grammar explanations, concordancers, or Wikipedia entries.

In the present study, the typology presented above was adopted to operationalize the affordances and design various functionalities of the proposed learning system.

2.7.3 Key research areas in L2 listening and technology

As internet connection is increasingly affordable and convenient with wireless and 3G/4G technology, online streaming multimedia such as podcasts, videos, narrated animations have become useful sources for authentic and pedagogic listening

materials. Much research effort has been made to investigate the usefulness of multimedia in L2 listening instruction. One typical use of multimedia in L2 listening is videos. To date, a growing amount of research suggests that videos be the principal material for L2 listening instruction instead of audio-only texts (H.-S. Kim, 2015; Sarani et al., 2014; Shin, 1998; Wagner, 2010). According to Cross (2018), the rationales for promoting the use of videos in listening are threefold. First, viewing videos has an ecological validity in that it replicates many real-life listening encounters that a learner may have. This is because videos can provide a visual channel in conjunction with aural information, a feature which a learner is likely to experience in real-life listening events. Second, since viewing videos is often interesting and considered a pleasure activity, it has a motivational impact on the learner and thus encourages autonomous learning. Third, videos present multiple input modalities which facilitate comprehension according to the Dual-coding theory (Paivio, 1986) and the cognitive theory of multimedia learning (Mayer, 2009).

Research on the use of videos also examined the effects of captioning in videos. Past research has shown that captioning in videos benefits learners' listening comprehension and spoken word segmentation (Gowhary et al., 2015; Winke, P., Gass, S., & Sydorenko, 2010). A meta-analysis study by Montero Perez, Van den Noortgate, and Desmet (2013) shows that viewing videos with captions can aid listening comprehension and significantly influence vocabulary learning. Winke et al. (2013) also affirm in their study that captions had a large effect size on both comprehension and vocabulary learning as opposed to no captions. However, a reasonable concern about using captioned videos is that learners may progressively rely on reading the captions for comprehension rather than listening to the spoken text (Gruba, 2018). This happens especially when the listening material includes unfamiliar contents or contains many novel words (Gass, Winke, Isbell, & Ahn, 2019). Research has revealed that learners with different L2 proficiency levels may have variations in behavior patterns of using captions and thus benefit differently from captions (Lwo & Lin, 2012; Montero Perez et al., 2013; Neuman & Koskinen, 1992; Taylor, 2015; Vanderplank, 2016) and that captioning can be helpful in listening given that "the viewer's language level must be not too far from the difficulty level of the captioned material for captioning" (Gass et

al., 2019, p.87), otherwise the viewer may heavily depend on fewer input modalities, i.e. depending largely on reading the caption texts or ignoring the captions to focus on images.

In addressing individual differences in language proficiency levels, previous research also investigated the effects of partial captioning, i.e. keyword captioning as opposed to conventional, full captioning (Guillory, 1998; Montero Perez, M., Peters, E. and Desmet, 2014). Mirzaei, Meshgi, Akita, and Kawahara (2017, 2018) proposed an interesting method for generating automatic partial captions. Mirzaei et al (2017, 2018) utilized an automatic speech recognition (ASR) software to generate automatic captions for English-speaking video. ASR captions are then used to inform the potential listening difficulties of EFL learners. Meanwhile, the human-annotated captioning is automatically reduced to keyword captioning according to speech rate, word frequency, and word specificity. The selection of the three criteria rests on the premise that fast speech rate may lead to difficulties in L2 listening comprehension (Rost, 2005), that words frequently occurred in a certain context are important in aiding comprehension (Bloomfield et al., 2010; Schmitt & McCarthy, 1997), and that unknown terminologies and academic words can greatly influence comprehension (Révész & Brunfaut, 2013). Learners are tested on their vocabulary size and speech rate tolerance before viewing so that the system can adjust the keyword captioning and speech rate of the video on an individual basis. ASR captions are also compared with automatic keyword captions to enhance the quality of keyword captioning. The researchers conclude that there was no significant difference between effects of partial captioning and that of full captioning on the learners' listening comprehension. However, they noted that while the two caption techniques can be used interchangeably, partial captioning might be preferred as it assists comprehension in a way that it promotes more listening and less reading than full captioning (Mirzaei et al., 2017).

Another important area of listening research is using technology for input modification. Though speech rate is a perceived difficulty for many listeners, modifying speech rate to assist listening comprehension through play speed controls has had mixed results (Hubbard, 2017). A more fruitful technique is applying low-pass filtering to the input (He, 2014). This is an interesting technique originated from the verbotonal

approach (Guberina, 1972 as cited in He, 2014). The verbotonal approach is effective both in speech perception and pronunciation training (He, Sangarun, & Lian, 2015; He, 2014), however, not much has been known about its effects on listening comprehension development. Given its potential for speech perception learning, positive influences on listening development can be expected.

Recent advances in multimodal technology and a keen interest in edutainment have led to new and exciting research dimension investigating effects of 3D immersion on listening comprehension. The most cited technological tool in this research line has been Second Life. Second life is a 3D multiuser virtual environment for online interactions with other people using avatars (Lan, 2015). Research has indicated potential benefits of Second Life in developing L2 listening performance. Experimental studies such as Levak and Son (2016) and Lan & Liao (2017) have revealed that being immersed in 3D multiuser virtual environments like Second Life leads to improvement in L2 listening skills and increases learner's motivation. One of the most beneficial aspects of this 3D immersion environment is that it provides authentic contextual, situational information which assists schema formation and schema activation. As motivation is concerned, this type of virtual immersion can blend learning with entertainment, as a result, removes any negative emotions in learning and fostering independent learning.

Finally, as artificial intelligence or AI is becoming a global trend in almost every aspect of the modern life, researchers and practitioners are beginning to look for possible uses of AI software for listening instructions. Multimedia-based applications with the integration of AI such as YouTube, Netflix, TuneIn radio, Spotify Music have been used for self-learning of listening comprehension (Suryana et al., 2020). Other virtual personal assistant applications such as Siri or Alexa powered by AI speech recognition and speech generation have also been used for teaching both English speaking and listening (Dizon, 2020). It is worth noting that while applications such as YouTube or Netflix are powered by AI, those applications are not developed for teaching or learning listening comprehension, and thus the AI in those apps only serves to personalize the users' listening preferences and has nothing to do with comprehension. Likewise, application like Alexa might be used for teaching and

learning speaking or listening to some extent, however, they are not designed for those purposes, and thus they possess no features that support learning. As Dizon (2020) also admitted in his study about the use of Alexa in improving students' listening and speaking performance, the students' listening comprehension did not improve significantly after the training with the virtual personal assistant Alexa, which is understandable since the app is not designed for that purpose. Therefore, as stated earlier, intelligent CALL applications specifically developed for listening comprehension are still very rare and this area of research should be investigated more.

2.7.4 Multimedia learning

It was not until 1980s that most language classrooms largely relied on the use of tape recordings, usually those accompanying language textbooks. The rapid development of technology has completely changed the teaching and learning situation of L2 listening since audio tapes were replaced by digital recordings such as CDs and later MP3 files or video files. Digitalization of audio recordings, new inventions of technological devices, and the extensive coverage of the internet altogether have made listening to native speakers as convenient as never before. These days, L2 learners do not have to rely on fixed hours of the listening class or ready-made tape recordings included in the language textbooks for practicing L2 listening. YouTube, BBC News, other news media websites, and mobile applications have provided apparently unlimited sources of materials for L2 learners, especially for EFL/ESL learners, to listen anytime, anywhere. For this change, learning L2 listening can take place in an online environment, and very often, with multimedia input. Therefore, there is a need for a sound theory that situates L2 listening learning in the multimedia environment as it is today. In response to this need, the present study will draw mainly on the theory of Multimedia Learning (Mayer, 2009).

Mayer (2009) proposed a cognitive theory of Multimedia Learning which promotes the use of multimedia input for effective, deeper learning. He defines multimedia instruction as "the presentation of material using both words and pictures" (Mayer, 2009, p. 3). Words can be printed or spoken. Pictures can be static (e.g. photos, maps, charts) or dynamic (e.g. video, animation). The central claim of Multimedia

Learning is that “people learn better from words and pictures than from words alone” (Mayer, 2009, p. 4). In addition to the core principle of multimedia learning, Mayer also suggests other principles for designing effective multimedia instruction which resulted from nearly 100 experimental comparisons conducted over the course of 20 years by him and his colleagues. Table 2.2 summarizes the twelve principles for promoting effective multimedia learning.

Table 2.2 Principles of Multimedia Design

Principle	Content
Coherence	People learn better when extraneous words, pictures, and sounds are excluded rather than included.
Signaling	People learn better when cues that highlight the organization of the essential material are added.
Redundancy	People learn better from graphics and narration than from graphics, narration, and on-screen text.
Spatial Contiguity	People learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen.
Temporal Contiguity	People learn better when corresponding words and pictures are presented simultaneously rather than successively.
Segmenting	People learn better when a multimedia lesson is presented in user-paced segments rather than as a continuous unit.
Pre-training	People learn better from a multimedia lesson when they know the names and characteristics of the main concepts.
Modality	People learn better from graphics and narration than from animation and on-screen text.
Multimedia	People learn better from words and pictures than from words alone.
Personalization	People learn better from multimedia lessons when words are in conversational style rather than formal style.
Voice	People learn better when the narration in multimedia lessons is spoken in a friendly human voice rather than a machine voice.
Image	People do not necessarily learn better from a multimedia lesson when the speaker’s image is added to the screen.

(Mayer, 2009)

Mayer's Multimedia Learning theory is constructed on three fundamental tenets of cognitive science: the dual channel assumption, the limited capacity assumption, and the active processing assumption (see Table 2.3)

Table 2.3 Three assumptions about how the mind works in multimedia learning

Assumption	Definition
Dual channel	Humans possess separate information processing channels for verbal and visual material.
Limited capacity	There is only a limited amount of processing capacity available in the verbal and visual channels.
Active processing	Learning requires substantial cognitive processing in the verbal and visual channels.

(Mayer & Moreno, 2003, p.44)

First, the dual channel assumption maintains that human cognition has separate channels for processing visual/pictorial information and auditory/verbal information (Mayer & Moreno, 2003). This assumption resulted from Paivio's (1986) Dual-coding theory and Baddeley's (2012) model of Working Memory. Second, each of the channel has a limited processing capacity, meaning that only a limited amount of visual information or auditory information is processed by each channel according to the Cognitive Load theory (Chandler & Sweller, 1991) and Working Memory theory (Baddeley, 2012). Third, meaningful learning involves processing of both visual and verbal information. As Mayer and Moreno (2003, p. 44) put it,

> *"These processes include paying attention to the presented material, mentally organizing the presented material into a coherent structure, and integrating the presented material with existing knowledge."*

Multimedia learning is a sound theory for designing effective L2 listening instruction in the digital age because it is rooted in mounting research-based evidence and has a learner-centered approach (Mayer, 2009). Mayer argued that most instructional approaches involving the use of technology in the past failed because they had a technology-oriented approach. When motion picture was invented, for

instance, it was promised to replace the textbooks and to revolutionize the whole educational system, yet it failed to fulfill the promise (Mayer, 2009). The same disappointing histories can be observed from the invention of radios and that of television.

> *Although different technologies underlie film, radio, television, and computer-assisted instruction, they all produced the same cycle. First, they began with grand promises about how the technology would revolutionize education. Second, there was an initial rush to implement the cutting-edge technology in schools. Third, from the perspective of a few decades later it became clear that the hopes and expectations were largely unmet.*
> (Mayer, 2009)

Rather than asking what a certain technology can do for learners, he contends, we should ask what it can improve learning in relation to how the human mind works. Multimedia Learning theory is well fitted with theories of constructivism and principles of personalized learning. For these reasons as well as the theory's promising outcomes, the present study drew on the principles of multimedia learning as ways for creating more effective L2 listening learning experience in an online learning environment. Specifically, a multimedia-based learning environment was constructed, working as a center for learning materials and references.

2.8 Chapter summary and theoretical framework of the study

The chapter starts with a review of some theoretical positions of constructivism and argues for a radical position as proposed by von Glasersfeld. The point is to arrive at a working conceptualization of learning, which is necessary for proposing an effective instructional approach to L2 learning. The core premises of radical constructivism are as follows. First, knowledge is a person's cognitive structures and is subjectively constructed. Second, knowledge construction is guided by personal experiences, thus personal perceptions of the world are unique, and those perceptions are not directly received from the environment but constructed. Third, knowledge is constructed and adapted to achieve viability conforming to constraints of the situation where it arises.

From these premises, learning is a change in the learner's knowledge structures, which involves restructuring the learner's pre-existing cognitive structures and redirecting the exploration of new experiences.

The chapter then highlights the importance of personalization in addressing the problems arising from individual differences. One of the methods to personalize learning is to provide effective corrective feedback. The chapter also reviews some issues in corrective feedback provision and how these issues may be overcome. From previous research, some conclusions can be made. First, providing CF seems more beneficial than withholding it. Second, CF provision is more beneficial if presented explicitly. Third, although it is still inconclusive regarding the optimal timing of CF provision, immediate CF provision has more theoretical support than delayed one. Fourth, in effecting CF provision, feedback has to be finely-tuned to individual learner's needs, or in other words, it needs to meet a high degree of precision. Lastly, computer can provide opportunities for personalizing CF, however, the reliability of the technology behind CF generation has been doubted.

As the study is concerned with L2 listening, conceptualization and operationalization of the construct listening comprehension is discussed with reference to structure-building comprehension model and communicative language approach to L2 listening assessment. Listening comprehension is conceptualized as a process of constructing a mental structure of meaning of an attended speech by activating relevant perceptual units which are then enriched or suppressed in terms of the goodness of fit to the structure. This process is supported by the interplay between bottom-up processing and top-down processing, as described in the interactive model of listening. The construct is operationalized as the listener's ability to understand the target spoken language in a communicative event, which provides both purposes and contexts for understanding, by drawing on the listener's various types of knowledge and skills.

The chapter also draws attention to pedagogical issues affecting L2 listening development. First, comprehension approach to L2 listening instruction is considered ineffective because it puts too much effort into checking comprehension and fails to attend to the learner's problems in the listening process. Process-based approaches,

on the other hand, place more emphasis on developing cognitive processes in listening comprehension, i.e. top-down and bottom-up. However, a process-based approach is only effective provided that it does not overemphasize one type of processing, either top-down or bottom-up and that it informs appropriate remedial actions.

Among the many challenges of L2 listening, the issue of variability in the speech signal is discussed and solutions to this issue from past research in speech perception are also reviewed. In short, this issue can be overcome by introducing speaker variability in perceptual training and if the training provides visual information accompanying the aural input, the effectiveness may be boosted.

Finally, the chapter ends with the reviews of major technological affordances, a comprehensive typology of help options for L2 listening in CALL software, and related research on L2 listening. A cognitive theory of multimedia learning that provides principles for creating an effective technology-based learning environment is also discussed. The fundamental principle of multimedia learning is that people learn better if they are presented with words and pictures rather than words alone.

On the grounds of the reviewed theories, the theoretical framework of the study is constructed and briefly summarized as follows (see Figure 2.6). Radical constructivism explains what learning is and how it can be conditioned. Personalized learning provides principles for creating effective learning. Structure-building model explains how listening comprehension operates. The operationalization of listening comprehension enables the measurement of the construct in this study. Both conceptualization and operationalization of listening comprehension inform methodological approaches to L2 listening. Specifically, process-based approach to L2 listening and high-variability perceptual training are employed as the primary methods. Multimedia learning theory provides a theoretical account for the use of multimodal, multimedia input as the main learning materials in this study. Finally, technology comes to realize the pedagogical principles resulting from those theories which include the following features: multimodal and multimedia input, help options, personalized practice.

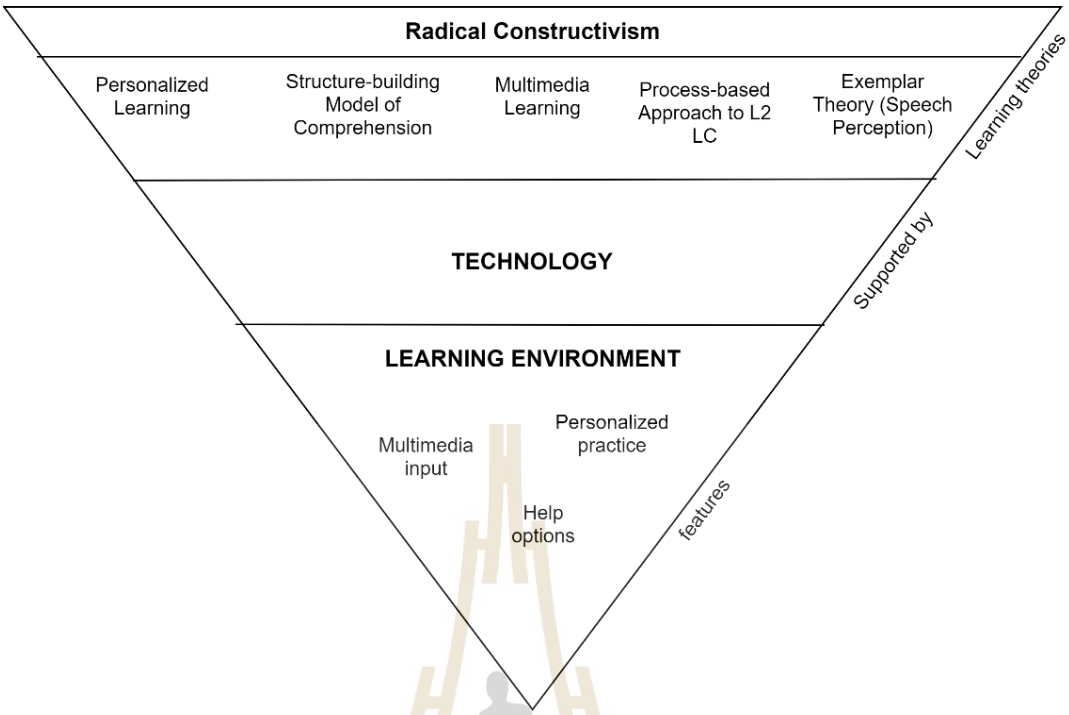


Figure 2.6 Theoretical framework of the study

CHAPTER 3
RESEARCH METHODOLOGY

This chapter explains the methodology of the present study. It provides information on and rationale for the methodological decisions in the research study, including, the subjects of the study, the design, the instruments, the procedures for conducting the research, data collection procedures, and analytical approaches to specific types of collected data. The chapter also discusses the pedagogic principles underlying the learning system derived from the theoretical framework as presented in Chapter 2, plus, describes how the system works. Finally, it reports on the results of the pilot study.

3.1 Research design

3.1.1 Mixed-methods approach

One of the main purposes of the present study was to investigate the effectiveness of an EFL listening training program using a web-based application, namely, Listening Hacked (as described in Section 3.4). The research was also interested in exploring the students' learning behaviors in an ICALL environment, particularly, specifically how the students approached a problem as it arose in the listening process and what influenced their learning behaviors. Also, the research was aimed at answering inquiries into how the students felt about the usefulness of the learning system.

To those ends, the research adopted a mixed-methods design in which both quantitative and qualitative data were gathered for seeking answers to the above research interests. Precisely, the research adopted a convergent mixed methods design (J. D. W. Creswell & Creswell, 2018). The pivotal methodological value of this design is that of allowing the researcher to draw on both quantitative data and qualitative data for a finer understanding of a phenomenon of interest from different perspectives (J. D. W. Creswell & Creswell, 2018). In this study, quantitative data and qualitative data

were collected and analyzed separately, and the results of which were compared for interpreting and answering relevant research questions.

First, to answer Research Question 1 regarding the effects of the training program using Listening Hacked on the students' EFL listening comprehension, the study drew on quantitative methods. In educational research, experimental designs usually involve using pretest and posttest in order to observe possible differences in a certain dependent variable provoked as a result of an intervention. Furthermore, to observe the potential causal relationships between variables, a Control Group is recommended in the design (J. W. Creswell, 2012), so that confounds are well controlled. Therefore, this study adopted an experimental, pretest-posttest, experimental-Control Group design. The independent variable in this study was the EFL listening training program using Listening Hacked. The dependent variable of interest was the students' EFL listening comprehension measured by their posttest scores and progress results.

Second, in answer to Research Question 2 pertaining to the students' learning behaviors and their reasons for exhibiting such behaviors on the learning system, activity logging (see Section 3.2.3) was used to collect the behavioral data. Additionally, qualitative data can give valuable insights into the students' reasons for having those behaviors as the nature of this inquiry is primarily exploratory. Therefore, qualitative methods for data collection were also employed for this research purpose. Specifically, the study utilized interviews for gathering data on this matter.

Finally, in addressing Research Question 3 as regards the students' opinions of the usefulness of Listening Hacked, a questionnaire was administered at the end of the training program. The questionnaire included both closed-ended and open-ended questions with the intention of collecting both quantitative data and qualitative data on the students' opinions. The two types of data collected could enrich understanding of the concerned research problem.

3.1.2 Participants

The population of interest in this research was Vietnamese undergraduates whose major of study was the English language. However, due to issues of accessibility, the sampling frame was reduced to current students in a faculty of foreign languages

at a university in Vietnam. The faculty of foreign languages has approximately 2500 undergraduates, with approximately 400 new English-major students annually.

This study employed the simple random sampling method in which the researcher approached all first-year, English major students by delivering brochures, sending out emails, posting on the Facebook page of the faculty of foreign languages, and contacting the students in their classes. The researcher randomly selected 100 students from the list of 151 students who registered to participate in the study. Then, the researcher randomly assigned the 100 selected participants into the Experimental group and Control group. It is noted that the final sample of the present study included only 53 participants since many students dropped out of the study at some stage. The study took place when schools resumed after a long lockdown due to Covid-19 pandemic. The students were thus overwhelmed with the pending assignments and tests, and they had to attend classes double the normal time. For this reason, many of the students decided to withdraw from the study to focus on their formal learning.

The rationale for recruiting first-year students in the research was that practically they would need more instructional support in the first year of study as compared to senior students in the faculty. Plus, sampling only first-year students potentially created a homogeneous group of participants in terms of their EFL listening proficiency, which was an advantage for making statistical inferences. All participants were asked to express their agreement in an informed consent form (See APPENDIX A).

3.1.2.1 The Experimental Group

Of the 53 participants, 30 students were in the Experimental Group. The other 23 students formed the Control Group. The participants in the study were 18 years old. They had at least 10 years of EFL formal learning since English is a compulsory subject from Grade Three in the national education system in Vietnam. Additionally, as first-year English-major students, they were attending a compulsory, 45-period course of EFL listening, namely Listening 1, at the time of the research study. However, the training program in this study was not a part of the listening course and did not fit in the semester time frame at the university.

The students in the Experimental Group learned EFL listening on the learning system (Listening Hacked) in 12 weeks. They were required to complete exactly three practice sessions per week on the learning system, i.e. approximately 2.25 hours per week (averagely 45 minutes per practice session).

3.1.2.2 The Control Group

The students in the Control Group were not given access to the learning system, however, they were required to do traditional listening exercises regularly which roughly equal the amount of practice that the students in the Experimental Group had to do; this means they did 5-6 listening exercises per week or 2.25 hours per week.

The listening exercises were posted weekly on an e-learning platform called Google Classroom. The exercises were taken from two reference books for EFL listening practice, starting from Intermediate level to Upper-intermediate level. The chosen books were Keynote Upper-intermediate (student's book and workbook) and Perspectives Level 3 (student's book and workbook). The decision for choosing the materials from those books was twofold. First, one of the two book series, i.e. Keynote series, has been used at the university and proven appropriate for the teaching situation at this university. Second, two listening instructors in the faculty of foreign languages confirmed with the researcher that the listening materials in Keynote Upper-intermediate were suitable for the current first-year, English-major students in the faculty. Perspectives Level 3 is another book series that bears a close resemblance to the Keynote series both in the format of the listening tasks and the level of task difficulty. In addition to the materials from the two mentioned books, more challenging exercises for quick learners were chosen from selective extracts of radio podcasts with comprehension questions. The selected exercises feature both English monologues and conversations accompanied by comprehension questions, including multiple choice questions, true/false statements, dictation, multiple matching, short answers, and summary completion. The length and number of the questions in each exercise vary, however, it normally takes an average student 20 minutes complete a normal exercise and 40 minutes to complete an advanced exercise.

The participants in the Control Group also did the pretest and posttest. Table 3.1 illustrates the design of the present study.

Table 3.1 The design of the study

Experimental Group (EG)	R	O_1	X	O_2	O_3	S	I
Control Group (CG)	R	O_1	C		O_3		

Notes:

R: Random assignment

O: Assessment of listening comprehension (O_1 Pretest; O_2 Progress tracking; O_3 Posttest)

C: Control (traditional listening exercises via Google Classroom)

X: The training program for EFL listening (Listening Hacked)

S: Questionnaire

I: Interview

3.2 Research Instruments

Hubbard (2006) recommends that the evaluation of CALL software should take student outcomes into account. He also notes various methods for gathering this information, i.e., through (a) observing student's learning progress, (b) tracking student's learning behaviors, (c) using a survey or questionnaires for knowing student's perceptions of the CALL software functionalities, (d) using pretesting and posttesting, and (e) having students reflect in a weekly journal writing.

In this study, the students' learning outcomes, i.e., students' listening comprehension, were measured by a pretest and a posttest. Additionally, the students' learning progress during the training was also observed and recorded by the learning system. For understanding the students' learning behaviors, semi-structured interviews were conducted. Originally, the research was planned to conduct interviews through instant messages; however, for some reasons discovered in the pilot study (see Section 3.7.3), the traditional, face-to-face interviews were employed. Finally, for measuring the students' perceived effectiveness of the learning system, an online questionnaire was administered at the end of the research.

3.2.1 Pretest and posttest

For the pretest and posttest, the Cambridge First Certificate (FCE) listening tests were utilized. The Cambridge FCE listening test is a test for B2 level of English

which assesses the candidates' ability to understand spoken speech, including news radio programs, talks, and daily conversations (*B2 First General*, 2021). The test consists of 4 parts with 30 questions in total and lasts about 40 minutes. Part One of the FCE test, featuring 8 short extracts from monologues, consists of 8 multiple-choice questions. Part Two is a monologue lasting about 3 minutes and accompanied by 10 sentence completion questions. Part Three features a series of 30-second monologues on a topic and 5 multiple matching questions. Part Four consists of a 3-4 minute conversation with 7 multiple-choice questions.

The reason for choosing FCE listening tests was twofold. First, the nature of the test is in line with the operationalization of listening comprehension in this study. Second, since FCE is a standardized language proficiency test, it has been systematically validated and acknowledged by the language testing community (Cambridge English, n.d.; Elliott, 2013; Elliott & Chismholm, 2015; Geranpayeh, 2004; Zeronis & Geranpayeh, 2015). Third, unlike IELTS tests, FCE listening tests are designed to assess English proficiency for general purposes and not specifically designed for academic purposes, which made it suitable for use in the present study where students' listening practice primarily involved viewing movies.

The researcher selected two Cambridge FCE listening tests, practice version from Cambridge published materials (see APPENDIX D and APPENDIX E). One of the tests was administered before the implementation of the training program. The pretest scores served as an indicator of the students' listening proficiency prior to the training. Another test was used as the posttest, administered at the end of the training program.

3.2.2 Progress tracking

Evidence of the students' listening improvement in this study could alternatively emanate from the students' cumulative practice results throughout the training program. Every time a student completed a listening practice session on Listening Hacked, a score was immediately given based on their task performance. A task performance score, calculated automatically by Listening Hacked, reflects how well a student performed a transcription task measured by the total number of words they successfully transcribed. These scores serving as the repeated measures of a student's listening performance can clearly depict the degree of development in EFL

listening comprehension that a student has made. In this study, these scores are referred to as Task Performance Scores (TPS).

To ensure the reliability of TPS, not all TPS were collected but only the TPS resulted from the three compulsory practice sessions which were carefully designed with the following criteria. First, the three practice sessions were compulsory to all students in the Experimental Group and were assigned in three designated timepoints in the research: Week 2, 7, and 12. The rationale for choosing the three timepoints was that Week 2 was the time when the students started learning on Listening Hacked; Week 7 and 12 were chosen to observe the students' learning progress with an equal interval of 5 weeks. Second, the sessions were taken from the first three episodes in the same movie series, namely Supergirl Season 1 (2015), therefore having the same baseline story and length. Third, the 10 chunks selected for each practice sessions satisfy the criteria: (1) its vocabulary is profiled at B2 level, (2) the speech rate ranges from 2.6 to 3.5 words per second, or from 160 to 210 words per minute, since the average speech rate in English is 180 words per minute (Szarkowska & Gerber-Morón, 2018), (3) each chunk consists of exactly 10 words, and (4) the words selected for transcribing are not proper names.

3.2.3 Activity logging

One of the objectives of the present study was to illuminate what behaviors and patterns of behaviors that students exhibited in a CALL environment and understand their reasons for having such learning behaviors as well. To this end, Listening Hacked was developed with a built-in function of logging all activities that the learner interacted with the learning system while doing their listening practice, for example, using the Repeat button, listening to a phrase again with more contexts, looking up a word in the dictionary, revising their previous submissions, and others. Those activity logs also include time reference. Information extracted from the activity logs can provide concrete evidence for the types of help options that the students used, how often they used each type, and the priority of each type of help options.

3.2.4 Semi-structured interviews

Activity logging may disclose frequent actions that the students did when learning in a CALL environment; however, if the goal is understanding the students'

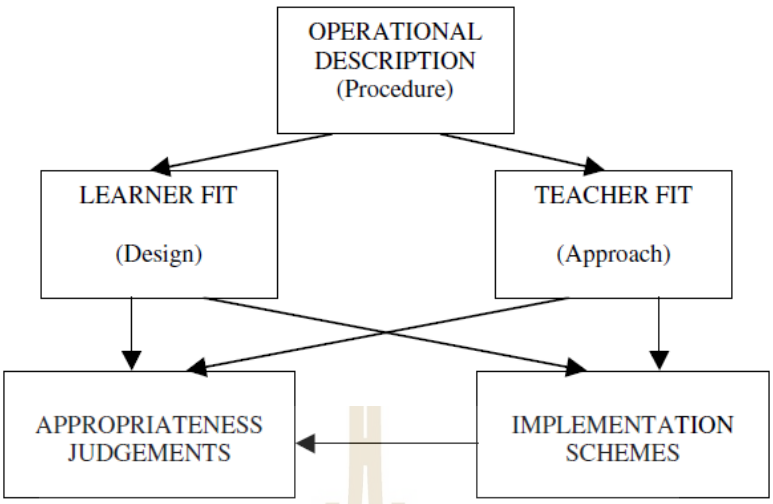
reasons for having such behaviors, activity logging alone is inadequate. Interviews are a potential research instrument for obtaining invaluable data on this matter. It is because interviews are regarded as a core research tool in qualitative research for gaining an understanding beneath the surface of a phenomenon from the students' perspective (K. Richards, 2009). Though interviews may not disclose precise instances of students' behaviors, they grant the possibility of understanding the driving force underneath the behaviors. Therefore, interview method was utilized in this study.

Essentially, the interviews sought the students' responses concerning the situations when they would use and prioritize specific help options in their learning on Listening Hacked and the reasons for doing so. For consistency of data collection, semi-structured interviews were used because this type of interviews works better than unstructured interviews when a study is concerned with different experiences and views about the same things, especially when most questions can be anticipated (K. Richards, 2003, 2009). For the questions used in the semi-structured interviews, please refer to APPENDIX C.

The sample for interviewing was 19 students from the Experimental Group. Originally, 20 students were randomly selected and invited for the interviews, however, one student could not make it to the interview in the end. The interviews in this study were conducted in Vietnamese and were audio-recorded.

3.2.5 Questionnaire

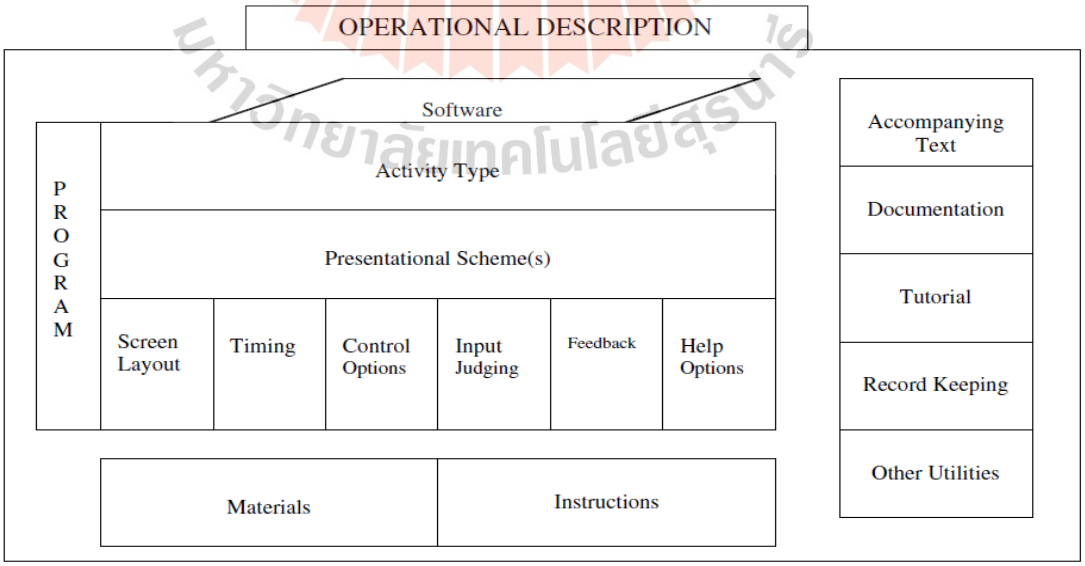
In this study, the students' opinions towards the usefulness of various features of the learning system were measured by a questionnaire with both closed-ended and open-ended questions. The questionnaire was constructed with reference to Hubbard's (Hubbard, 1988) framework for CALL courseware evaluation. In his framework, there are three main components: approach, design, and procedure (see Figure 3.1).



(Hubbard, 1988)

Figure 3.1 Evaluation Framework: Central Components

Approach essentially refers to the theoretical assumptions about language learning underlying the software. *Design* involves the realizations of those assumptions including learning goals, learning tasks, and learning materials. *Procedure* is the implementation of the realizations including performing tasks and activities in the software. Of the three components, procedure or *operational description* is highly relevant to the third research question in this study because this component involves evaluating the operations of the learning system (see Figure 3.2).



Hubbard, 1988

Figure 3.2 Operational description (Procedure)

There are both central and peripheral components in the operational description, however, the usefulness of the software remains in the central components. Activity type may cover various *presentational schemes*. For example, cloze, word scramble, and strip story are differently presented but belong to the same category of activity, i.e. text reconstruction (Hubbard, 1988). The presentational scheme can be further described in six features: screen layout, timing, control options, input judging, feedback, and help options. *Screen layout* refers to how the learning material is presented to the students which may include text size, arrangement of text, use of colors, etc. on the screen. In a multimedia environment, it may include quality of the videos or effects of animations. *Timing* refers to the function of setting time limit or monitoring time during a task. *Control* involves the function of choosing what material to learn or interacting nonlinearly with the material. For example, a listening CALL software may allow users to pause or replay or switch to any desired part of the recording. *Input judging* refers to the types of response which the software requires the users to supply for a given task. Common types of input are selecting options from multiple choice questions, typing short answers for prompted questions, submitting recording files. *Feedback* is the information generated by the software based on student input. *Feedback* is often presented in texts, though sometimes accompanied by pictures or videos, to pinpoint student's errors. *Help options* include a wide range of instructional support given to a student, however, they can be categorized as instructional reviews or hints. For example, the program may provide a link to a review of a given grammar point or simply provide clues on how to answer a grammar question. It is noted that the component *material* is treated as being independent of the presentational scheme in Hubbard's framework, however, if considering the learning material is video, then it can be part of the presentational scheme.

Based on the above-described framework, the questionnaire for use in the present research was constructed with reference to the components of operational description. The questionnaire, using online Google Form, gathered the participants' opinions of the usefulness of various features of Listening Hacked, including (a) using movies as the primary learning materials, (b) transcription tasks, (c) automated corrective feedback, (d) repeated listening, (e) high variability training, (f) performance

tracking and Listening Boosters, and (g) other supports available in Listening Hacked. In essence, the questionnaire asked participants to rate the level of usefulness for each feature, using a five-level Likert scale. There was also an open-ended question intended for each feature. These open-ended questions invited participants to explain their rating for a given feature. Finally, there was a question seeking participants' opinions of their overall experience with Listening Hacked, also followed by an open-ended question.

3.3 Research procedure

It is important to restate the primary purpose of the present study which was to develop an AI-powered, personalized system for learning EFL listening. Only when that purpose was achieved, the secondary purposes of investigating effectiveness and the students' opinions of the proposed system might be accomplished. For this reason, the present study was conducted in two phases (See Figure 3.3).

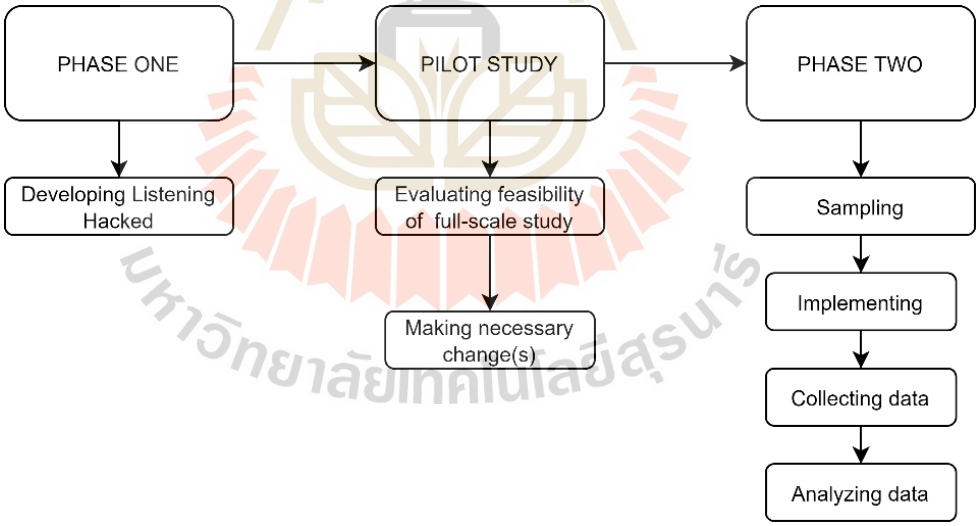


Figure 3.3 Research Procedure

Phase One was concerned with the process of generating the proposed learning system, namely Listening Hacked. This phase involved seven stages (See Figure 3.4): (1) constructing a theoretical base for the design of Listening Hacked, (2) defining the key functionalities of the software, (3) preparing the data for the software operation, (4) coding the software, (5) implementing the software, (6) testing and modifying its

functionalities, and (7) operating and maintaining the whole system. Based on the theoretical framework constructed in Chapter 2, stages (1), (2), and (3) are discussed in detail in Section 3.4. Stage (5) involved installing and configuring the software on a cloud server and uploading the prepared data to the server. Stage (6) was concerned with carrying out a pilot study to test all the functionalities and made necessary modifications to the learning system. In addition, during the whole course of the development process, the researcher worked with the software coder to test and modify any detected unsatisfactory functionalities. The major changes made to the final version were reported in the results of the pilot study (See Section 3.7). Stage (7) was to ensure the smooth operation of Listening Hacked by upgrading the server to full solid-state drive (SSD) storage and more powerful hardware.

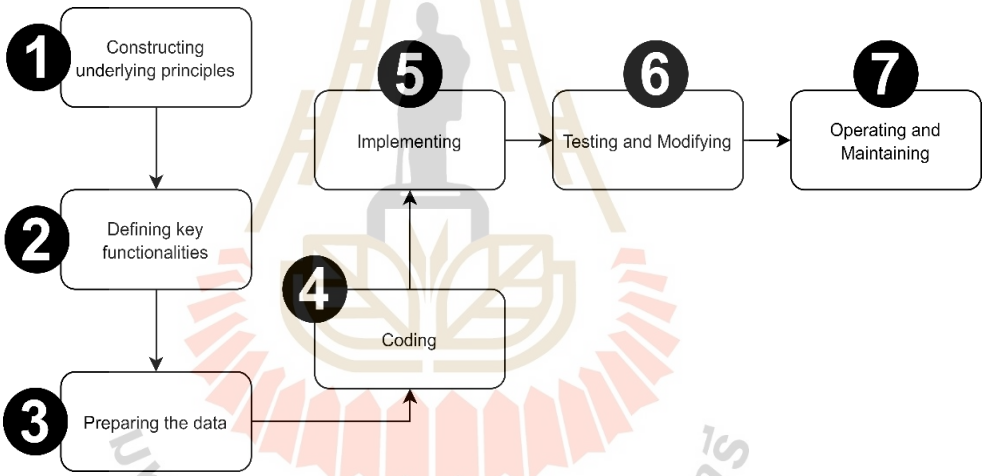


Figure 3.4 The process of developing Listening Hacked

Phase Two of the study was designed to investigate the effectiveness of the learning system, explore the students' learning behaviors, and survey the students' opinions of the system. In this phase, an experiment and social research methods were employed. Sections 3.1, 3.2, 3.4.2, 3.4.3 the research design, materials, and methods used in the second phase.

3.4 The learning system: Listening Hacked

3.4.1 Principles of the learning system

Based on the theoretical framework described in Chapter 2, following are the principles underpinning the design of the learning system, Listening Hacked.

(1) On the grounds that knowledge is actively constructed by the learner and it cannot be imposed intact on them, the system should *respect the listener's meaning-making process*. For instance, pre-listening activities, a popular method of introducing key vocabulary items and/or activating topic knowledge, were not employed in this system. The reason is that by presenting some vocabulary items to the listener in advance, the teacher also makes assumptions about what is important and what is less important as to understanding a certain listening text on behalf of the learner. This is a form of imposition of the teacher's knowledge and experiences on the learner. In fact, introducing specific keywords or phrases of the teacher's interest may create distractions as the listener may be searching for those specific lexical items while listening instead of attending to other important information in the text. Similarly, activating the listener's background knowledge on a topic of listening is useful, but if it is done before the listening event in a pre-listening activity, it may bring more complications into the listening process when the listener's predictions of the contents are inaccurate. Thus, while activation of background knowledge is valuable, it should be done as the listening event takes place, and by the listener as a result of the needs generated by the listener's meaning-making process. Please refer to 0 and 0 for fuller discussions on these issues.

(2) Given learning is an adaptive process and knowledge construction must have viability, feedback provision is critical in the learning process. As discussed in 0, the effectiveness of corrective feedback (CF) depends on several factors, including, but not restricted to, the explicitness of CF, the timing of CF provision, the degree of precision and personalization of CF, and the reliability of CF given that it is generated by a computer. Taking these issues into consideration, *feedback provision should be explicit, immediate, and personalized*, *by targeting specific and real problems of an individual listener*.

(3) Personalization is an effective solution to individual differences in language learning. The system, therefore, should *offer various help options or learning paths* to cater for different learners effectively.

(4) Listening is an interactive process between top-down and bottom-up processing, which is why *listening practice should promote the types of activity that can tap into those two primary processes of listening* as an attempt to reflect its operation in real-life situations.

(5) As suggested in the theory of Multimedia Learning (Mayer, 2009), learners benefit more when the learning materials are presented with multimodalities (e.g. both auditory channel and visual channel) multimedia (e.g. both pictures and written texts) than being restricted to a single medium or input modality. On that account, t*he system needs to provide contextualized and multimedia input for the listeners*.

(6) With the aim of addressing the challenge of variability in the speech signal as discussed in 0, listening training should *increase speaker variability in the input* rather than reduce it, to facilitate perceptual learning.

(7) Listening activities should aim for *qualified input*. According to Rost (2006), the quality of listening texts includes relevance, difficulty, and authenticity. Relevance requires *listening materials to have motivational effects* on the listener. In other words, the listener should have an authentic need to process and understand the listening texts. This is because real-life listening is characterized not only by the cognitive processes but can be affected by the listener's willingness to listen (Worthington & Bodie, 2018). Weaver is the first listening scholar to notice the affective dimension of listening in communication when proposing in his book that "willingness to listen is probably as important as capacity to listen" (1972, p. 8). To rephrase it, the choice that a listener makes whether to attend to the speech or to ignore it is a vital component and prerequisite of listening. Authenticity involves *using genuine target-language texts* whenever possible. As far as the goal of EFL listeners is to understand the language actually used by native speakers, they need to be introduced to that type of language use (Rost, 2006). Text difficulty refers to the cognitive load of a text imposing on the listening process. These include, but are not limited to, "length, speed, familiarity, information density, and text organization" (Rost, 2006, p. 50). Text

difficulty, in this regard, was signified to the learners by the researcher's suggested labels for the graded difficulty of certain learning materials. The learners, however, were free to consult these labels and decide what to learn.

3.4.2 Learning approaches and activities

3.4.2.1 Materials

The main materials for use in this learning system were movies. This was an attempt to realize principle (5) of multimedia, multimodal learning. Previous research confirms the positive impacts of audio-visual input or videos on listening comprehension. The rationale for using videos in EFL listening was threefold. First, videos can provide adequate contextual information about the situational settings within which the speech is used, or as Kelly (1969, p.253, cited in Gruba, 2018) put it: "the most valuable aspect of television, and one shared with film, was its power of linking linguistic behavior with the environment and cultural context which occasions it". Second, viewing videos assists with the activation of relevant schemata by means of non-verbal cues such as gestures, postures, facial expressions, or physical settings of the speech event (Guichon & McLornan, 2008; C.-D. Nguyen & Newton, 2018; Vandergrift, 2004). Past studies on the application of videos in listening comprehension, such as Shin (1998), Wagner (2007, 2010), Sueyoshi and Hardison (2005), reveal that students watching the video texts outperformed those listening to audio-only texts. This outperformance is attributed to the added visual source for schema activation. Third, watching movies instead of listening to audio-only texts may have a motivational effect on EFL students (Cross, 2018) as watching movies is commonly considered a form of relaxation, thus it also possibly promotes autonomous learning.

In this research, choosing movies as the primary materials in listening practice works as a solution to the issue of pre-listening activities. As discussed in principle (1), the learning system should respect the listener's meaning-making process by not imposing the teacher's experience or knowledge on the students and by not eliciting unnatural predictions for the listening texts. A movie by itself, especially some genres with story-telling features, can provide adequate information about the contexts, the settings, the characters or the speakers involved in any speech taking place in it. By viewing a movie, the listener can obtain correspondingly enough

contextual information without support from pre-listening activities. On top of this, it is this contextual information that can offer a natural source for schema activation. Sometimes, the listener might also need to consult a good few cultural notes to understand some details of a movie, but most of the time, everything is expected to be self-explanatory in movies. This was also an attempt to satisfy principle (5) of providing contextualized input.

It is noted that although speeches and utterances used in movies are scripted, they are supposed to be authentic in the sense that they occur in real communicative events, except that the events are in a movie. Therefore, it can satisfy principle (7). Additionally, unlike talks, a movie presents various characters as speakers, which increases the listener's exposure to several speakers; hence an increase in speaker variability in the input as principle (6) requires.

3.4.2.2 Task

The main task which the students had to do while learning on the system was to watch movies and occasionally transcribe some parts missing of caption texts. This technique for listening practice was adapted from what is originally known as paused transcription, proposed by Field (2008a, 2011). In the original version, students are asked to listen to an authentic audio-only text where some pauses are inserted at irregular intervals. Students are required to supply the last few words whenever they encounter a pause. The rationale for employing paused transcription is that this task can tap into both primary processes of listening comprehension, namely bottom-up process and top-down process. This type of listening task is purported to focus on aural decoding, i.e. attending to the speech signal to get perceptual units, while not removing the benefit of using contextual information from top-down processing (Sheppard & Butler, 2017). Field (2008d, pp. 16–17) claims that paused transcription "replicates a real-world one. Subjects listen to the recording with a view to following its meaning, and it is only when a pause occurs that they switch attention to word level," This task hence adheres to principle (4), which aims for the practice of primary processes in listening.

In the current study, the students watched movies instead of listening to audio-only recordings; however, the same rationale is still valid. In other

words, students watched a movie for understanding its global meaning and only switched attention to specific details whenever there was a pause indicating a need for transcribing certain missing caption texts. However, instead of having students listen one time only and asking them to supply the last few words that they can remember before the pause, the task in this study required the students to provide accurate transcription for a chunk immediately preceding a forced pause as to continue viewing the movie. The purpose of this was to eliminate the effects of memory on listening and instead to focus on comprehension.

3.4.3 Features

3.4.3.1 Help options

When performing a (modified) paused transcription task, it was anticipated that learners would encounter various listening problems. To assist with learning, the system provided three help options as described below.

a) Repeated listening

Because of issues related to working memory capacity, the listener may not remember precisely what they have heard or may fail to recall some words when they encounter a pause. For this reason, the learners were allowed to listen to the target chunk as many times as they need rather than being restricted to one-time listening as in the original paused transcription task. In real-life communication, the need for clarification or repetition is not unusual, and allowing repeated listening only assists individual learners with better comprehension (Dupuy, 1999; Krashen, 1996; Şendağ et al., 2018). In fact, the restriction of one-time listening may be only suitable for testing purposes, whereas if aiming for listening improvement, repeated listening can help the listener reduce the processing load and avoid memory problems since they can comfortably focus their attention to certain parts of the speech as they deem important. Listening again is not just repetition but each time the learners listen again to a text or a part of it, they may come up with different understandings. This feature particularly responds to principle (3) of the learning system, i.e. creating personalization as a solution to individual differences in learning. Note that the unit of listening practice in this study is sometimes referred to as a chunk which is operationalized as an intonation unit (or pause unit) consisting of phrases or clauses and uttered in a single

short burst of speech. In a movie, it regularly coincides with a speaker's complete utterance.

b) Corrective feedback

To maximize the effectiveness of repeated listening, the students should know exactly where to attend to; therefore, they need corrective feedback as stated in principle (2). The system utilized a built-in software to indicate any errors in the students' transcription input. The corrective feedback was given immediately after the students submitted their transcription texts. The mark-up software employed a string-matching technique that compared the student's submitted transcription with the pre-stored corresponding transcription. This technique can produce very high accuracy when capitalizations and punctuations are not considered as errors. The mark-up rules are as follows.

- Word(s) which actually appear(s) in the original caption text and is(are) displayed in their correct position and will be presented in blue.

- Word(s) which actually appear(s) in the original caption text and is(are) displayed in the wrong position will be presented in blue and underlined. The feedback will indicate the correct order of words only when the students have made at least three attempts to respond.

- Word(s) which do(es) not exist in the original text will be presented in red and strikethrough.

- Word(s) which exist(s) in the original text but is(are) not transcribed will be displayed with the symbol * in their correct position.

- Mismatched punctuation and capitalization are not counted as errors and are left untouched.

Table 3.2 shows an example of how the mark-up rules are applied to generate feedback for the students' submitted transcriptions.

Table 3.2 An illustration of the mark-up rules

Submission	Original text	Student's transcription	Feedback
1st time	What's going on here?	What's on going?	What's * on * going *?
2nd time	What's going on here?	What's on going there?	What's * on * going * ~~there~~?
3rd time	What's going on here?	What's on going here?	What's * on_3 * $going_2$ here?

Being immediate, detailed, personalized as such, feedback was hoped to help direct the listener's attention to where problems in listening prevailed and consequently increased the benefit of repeated listening by providing the listener opportunities to compare what they believed they have heard and what was actually present in the listening text.

c) High variability perceptual training

When students made a transcription error, this indicated that they were experiencing a perceptual problem. In conformity with principle (6) and as reviewed in 0, high variability training is an effective method for overcoming this issue. The app suggested personalized, high variability training for any perceptual problem being experienced in the transcription task. Specifically, the app performed the part-of-speech (POS) tagging and chunking for the utterance being transcribed. POS tagging is the process of labeling the lexical category, i.e., nouns, pronouns, verbs, adverbs, adjectives, conjunctions, prepositions, and other POS, of each word in the context of an utterance (D'Souza, 2018). Chunking involves segmenting and labeling sequences of POS tagged tokens, such as noun chunks, verb chunks, prepositional chunks, and the like (Bird et al., 2009). The results of the tagging and chunking operations inform the app whether or not the word being mistranscribed belonged to a chunk in the utterance. The app then searches in the entire movie data for several video supercuts containing the target chunk, otherwise, the target word, given that it does not belong to any chunk. The motivation behind the search for chunks rather than a single word is that chunks are formulaic and frequent sequences of language which the listener is likely to encounter, and accurate perception of which in speech can facilitate successful transfer to new listening situations.

The students were encouraged to watch all the suggested video supercuts, for the purpose of increasing exposure to various contextualized occurrences of the target word/chunk. Repeated listening to the target word/chunk in various contexts is actually a modified version of high variability training for a certain perceptual problem in listening. Note that the present study has developed a novel method for delivering high variability training in ways that while the conventional training is aimed at phonetic contrasts by having students expose to decontextualized

audio-only input, the type of perceptual training in this study was targeted at word and phrase level, and utilized audio-visual, contextualized materials. The modifications in the training were based on the findings of previous research on this matter, in which the inclusion of visual and contextualized information is expected to boost the effectiveness of the training (refer to 0 for a review of high variability perceptual training). In this study, these modifications included the use of videos as listening input and that of chunks as units for repeated exposure.

3.4.3.2 Follow-up practice

The focal intention of the system was to personalize learning for individual learners. While the three help options as described above can provide immediate support for learners in their listening practice, they still need more personalization in subsequent practice in order to effectively address certain listening problems that were not solved in previous practice. The app has two features to achieve this goal: progress tracking and Listening Boosters.

d) Progress tracking

Progress tracking allowed Listening Hacked to keep track of any word/phrase which was skipped during a paused transcription task. These skipped words were saved in a list and can be reviewed by the students on their personal page.

e) Listening Boosters

Listening Boosters was a follow-up practice targeting at students' personal decoding problems. Based on the list of skipped items, the app suggested several supercuts containing a random word/phrase from the list of skipped items. Students had to watch the suggested video clips and supply the target word/phrase which occurred in all of the video clips. If the student successfully completed the follow-up practice, the results would be updated in the list of skipped words/phrases. Specifically, successfully transcribed words/phrases were removed from the list. This type of practice involved a high degree of personalization; and for this reason, it was expected to increase listening success by utilizing high-variability training and targeting at personal decoding problems.

Apart from the above help options, of course, the students also needed to make necessary revisions to their transcriptions, consult the built-in dictionary for any unknown expressions, or search external sources for information. They may need to use captions for compensating their deficiencies in comprehension as well. Nevertheless, these help options were not meant to distract them from the main task but to assist with greater gains in listening comprehension.

Table 3.3 summarizes the main features and functionalities of the learning system, with categories for help options based on Cardenes-Claros and Gruba's (Cardenas-Claros & Gruba, 2013) framework for help options in CALL software (see Section 2.7.2 for detailed descriptions of the categories).

Table 3.3 Main features of Listening Hacked

Feature/functionality	Content	Category
Learning material	Movies	
Task	(modified) paused transcription	
Help options	Repeated listening	Compensatory
	Corrective feedback	Regulatory
	High variability perceptual training (with audio-visual input)	Explanatory
Other options	Reviewing and revising previous transcriptions	Regulatory
	Built-in dictionary	Explanatory
	Captions	Compensatory
Follow-up practice	Listening Boosters	Regulatory
Progress tracking	List of skipped words/phrases	Regulatory
Automatic grading	Task performance scores	

3.4.4 Main operations in Listening Hacked

The learning system in this study was designed as a web-based application and named Listening Hacked. A direct benefit of using a web-based application was that it required no installation on the part of the user, plus, it was compatible with most devices, supposed that the device can run web pages. For the purpose of tracking the students' progress, Listening Hacked required each student to log in the website

in order to use it Figure 3.5 and Figure 3.6 and are screenshots of the Log-in page and the Home page.

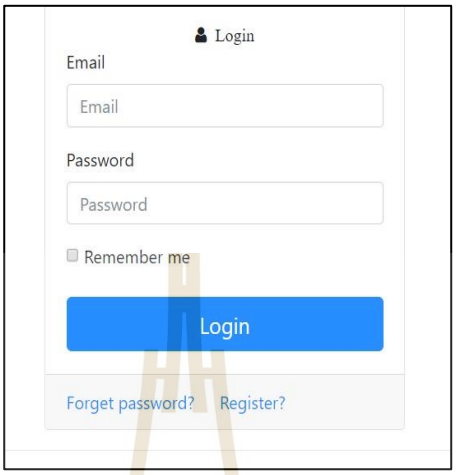


Figure 3.5 The Log-in page

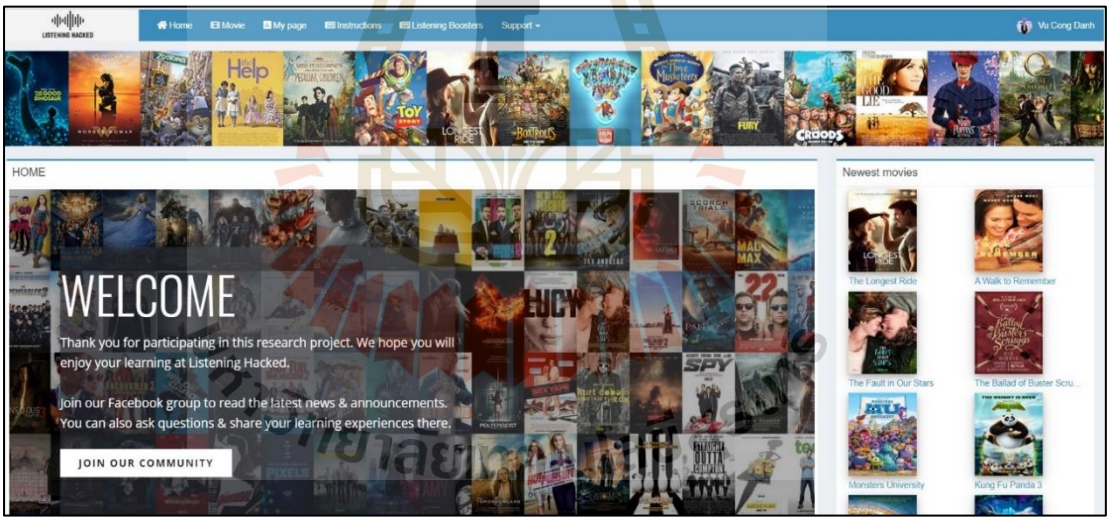


Figure 3.6 The Home page

To ensure that the students understood how to use Listening Hacked, a page dedicating for instructions was built up (see Figure 3.7 for some sections of this page).

Figure 3.7 The Instruction page

This page fits in the category of operational help options (Cardenas-Claros & Gruba, 2013).

Listening Hacked made available a plethora of movies of various genres and classified movies according to their most relevant genres, the information of which was obtained from the Internet Movie Database (https://www.imdb.com/). A movie normally lasts 90 to 120 minutes, thus, for convenience in practice, each movie was divided into three to four practice sessions, approximately 25 to 30 minutes per session. Students were free to choose a movie of their interest to view and practice, however, animated movies and documentaries were recommended at the outset, for the reason that these movies are generally slow-paced and the language is less idiomatic as opposed to other movie genres. It remained, however, the student's own decision whether to follow this recommendation. Figure 3.8 is a screenshot of the page for movie selection.

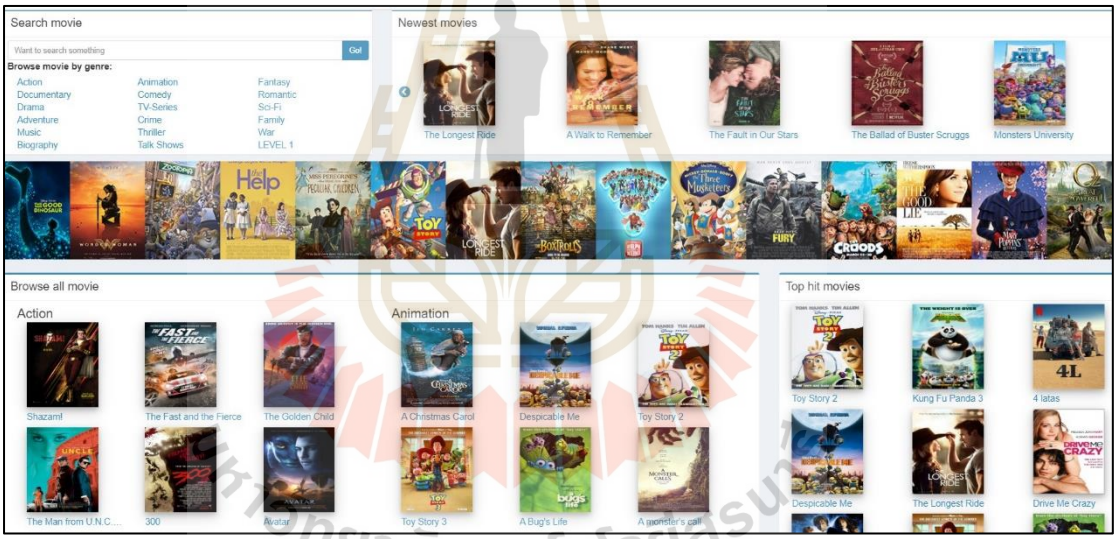


Figure 3.8 The Movie selection page

For a practice session, the app randomly selected 40 chunks for a transcription task and removed the caption texts of those chunks. The movie then can be played back with or without captions. Research on the effects of captioned videos has revealed that students comprehend the video contents more when captions accompany the videos, although whether or not comprehension gains are mainly attributed to reading caption texts is still under debate (Vandergrift & Cross, 2014). On

this account, the learning system in this study gave students complete control on caption viewing.

Whenever the movie playback reached a selected chunk, Listening Hacked would execute a forced pause and ask the student to transcribe the missing caption texts of the chunk immediately preceding the pause. The app did not allow the student to continue the playback unless the correct transcription was submitted, or unless they chose to skip the current chunk. However, the app allowed the students to replay the current chunk or any parts prior to the chunk with the Rewind options.

Once the student's input was submitted, the app compared it with the pre-stored transcript using the string-matching technique. If the student's transcription completely matched the original text, the app continued the movie playback. In case there was any mismatch, the app would perform three operations.

1) Displayed the corrective feedback as described in 0.
2) Allowed the student to replay the current chunk as many times as they wish.
3) Suggested a collection of supercuts containing the target word/phrase which was mistranscribed. This operation was triggered whenever the student clicked on the symbol * in the generated corrective feedback.

To perform the third operation, the app executed three processes in the following order: (1) tagging part-of-speech (POS) for each word in the utterance being transcribed, (2) chunking the tagged tokens, and (3) searching the target word/chunk in other videos. The app used a natural language processing toolkit, namely *NLTK* (version 3.4.5), for POS tagging and chunking (Bird et al., 2009). *NLTK* is an open-source code library for building NLP software in Python programming language (Yumusak et al., 2014). The natural language processing function of NLTK rests on 50 different datasets and allows the installation of text processing libraries such as WordNet, CoNLL 2000 Corpus, NPS Chat Corpus, VerbNet Lexicon, etc. (For more information, visit http://www.nltk.org/nltk_data/). The tagging process of Listening Hacked employed the trained POS tagger of NLTK which draws on the Penn Treebank corpus and its tagset (Santorini, 1990). Table 3.4 explains the tags used in this app with reference to the Penn Treebank tagset.

Table 3.4 The Penn Treebank tagset

Tag	Description	Example
CC	conjunction, coordinating	and, or, but
CD	cardinal number	five, three, 13%
DT	determiner	the, a, these
EX	existential there	there were six boys
IN	conjunction, subordinating or preposition	of, on, before, unless
JJ	adjective	nice, easy
MD	verb, modal auxillary	may, should
NN	noun, singular or mass	tiger, chair, laughter
PRP$	pronoun, possessive	my, your, our
RB	adverb	extremely, loudly, hard
TO	infinitival to	what to do?
VB	verb, base form	think
VBN	verb, past participle	a sunken ship
VBG	verb, gerund or present participle	thinking is fun
WDT	wh-determiner	which, whatever, whichever
WP	wh-pronoun, personal	what, who, whom
WRB	wh-adverb	where, when

For chunking the POS tagged tokens, the app exploited the chunking function of NLTK by applying regular expressions rules (regex rules). Regex rules are pre-defined patterns for identifying any chunk of interest which can be formulated from the POS tagged tokens. For example, a possible pattern for noun chunks is: an optional determiner (e.g. a, an, the), followed by an adjective (e.g. good, bad), and then followed by a head noun (e.g. dog, computer). This pattern can be expressed in regex rule as: DT + JJ + NN.

Table 3.5 details the regex rules for use in this app. Note that the optional components of the regex rules are not printed in bold.

Table 3.5 Regex rules for chunking

Chunk type	Regex rule	Example
Noun	DT + JJ + NN	A (DT) happy (JJ) father (NN)
	NN + IN + NN	Son (NN) of (IN) Zeus (NN)
	NN + IN + PP$	A (DT) friend (NN) of (IN) yours (PP$)
	PR + IN + PR	Nothing (PR) with (IN) it (PR)
	PR + IN + RB	Somewhere (PR) near (IN) here (RB)
	PR + TO + VB	Nothing (PR) to (TO) do (VB)
	WP + TO + VB	What (WP) to (TO) say (VB)
	PR + RB	Anyone (PR) else (RB)?
	PR + JJ	Nothing (PR) new (JJ)
	WP + PR/NN + VB	What (WP) you (PR) mean (VB)
Prepositional	IN + DT + NN	At (IN) the (DT) office (NN)
	IN + PR	By (IN) bus (NN)
		With (IN) her (PR)
Verb	MD + VB	I've (MD) paid (VB)
	MD + NN/PR + VB	Will (MD) you (PR) come (VB)
	EX + MD + VB	There (EX) is(VB) / there (EX) are (VB)
	MD + VB + NN/PR + RB	See (VB) you (PR) soon (RB)
	VB + IN	I'd (MD) like (VB) to (IN)
	VB + PR/NN + IN	Bring (VB) it (PR) back (IN)
	VB + PR/NN + TO + VB	Buy (VB) a (DT) book (NN) to (TO) read (VB)
	VB + IN + IN	Look (VB) down (IN) on (IN) employees (NN)
	MD + VB + PR/NN	love (VB) her (PR)
	VB + PR/ NN+ VB	let (VB) me (PR) know (VB)
	VB + PR/NN + VBG	Watch (VB) them (PR) running (VBG)
	MD + VB + IN + PR/NN	was (MD) shot (VB) with (IN) a (DT) gun (NN)
	VB + PR/NN + PR/NN	Elected (VB) him (PR) president (NN)
	VB + PR/ NN + RB	Leave (VB) me (PR) alone (RB)
	VB + PR/ NN + JJ	Make (VB) John (NN) crazy (JJ)
	VB + PR/NN + IN	Take (VB) care (NN) of (IN)
		Read (VB) it (PR) through (IN)
	VB + TO + VB	want (VB) to (TO) leave (VB)
	VB + VBG	enjoy (VB) jogging (VBG)
	VB + JJ	Be (VB) careful (JJ)
	VB + CC + VB	Come (VB) and (CC) see (VB)
	VB + PR + JJ	Say (VB) it (PR) right (JJ)
	VB + PR + NN	Take (VB) it (PR) an (DT) honor (NN)
	VB + PR/ PP + VBN	Make (VB) yourself (PP) respected (VBN)
Adjective	RB + JJ	very (RB) lovely (JJ)
	JJ + CC + JJ	back (JJ) or (CC) white (JJ)
	JJ + IN	Kind (JJ) to (IN) me (PR)
	JJ + IN + VBG	Responsible (JJ) for (IN) teaching (VBG)
Adverbial	RB + IN + DT + NN	Out (RB) of (IN) the (DT) blue (NN)
	RB + JJ/RB + IN	As (RB) good (JJ) as (IN)
	WRB + DT + NN/PR	When (WRB) the (DT) train (NN) leaves (VB)

After identifying the target chunk, the app performed a word-in-video search in the entire Listening Hacked's movie data and extracted all of the relevant parts from the movie data. Listening Hacked then presented those video clips in the area called *In Other Movies*, an area which was reserved and intended for high variability training. Figure 3.9 presents a screenshot of the main working space in Listening Hacked.

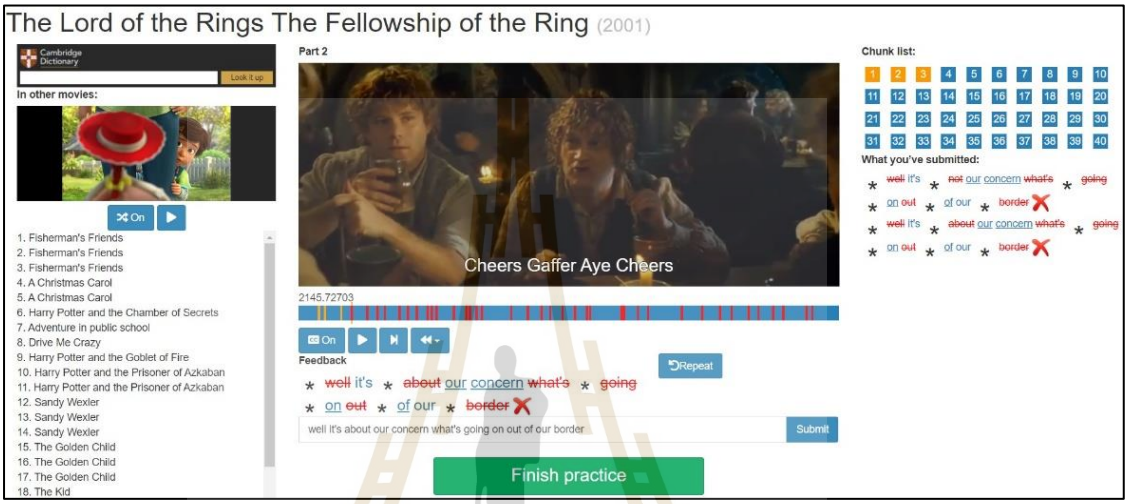


Figure 3.9 The main working space

After using up the available help options, the student needed to revise their previous transcription. They could also review their previous submissions during the revision. The student's revised transcription went through the same process as described above, starting from comparing it with the pre-stored transcript. However, after at least five attempts of failing to provide an accurate transcription, the student could choose to skip the current chunk to continue watching the movie. Any skipped word/phrase was saved in *the list of skipped items* for further practice. Additionally, students could look up any word in the integrated dictionary.

When the student completed a practice session, Listening Hacked automatically calculated and displayed the task performance score for the student. *Task performance scores* (TPS) were calculated by dividing the number of successfully transcribed words by the total number of words to be transcribed and then multiplied it by 10. For example, if the student successfully transcribes 190 words out of 400 words in a practice session, the TPS is 4.75 (out of 10).

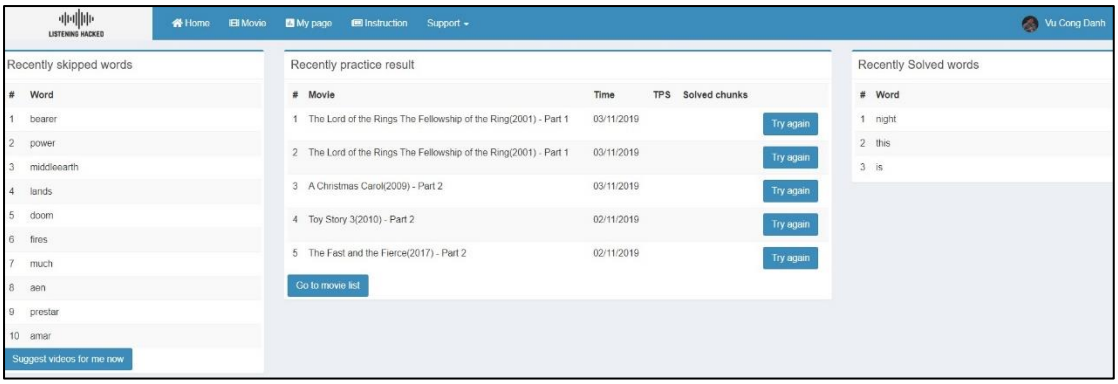


Figure 3.10 The personal page

Listening Hacked also had a personal page, called My Page (see Figure 3.10), for showing students the summaries of what they have done in the practice sessions including their practice results, a list of skipped words, and a list of recently repaired words/phrases.

By visiting the personal page, the students could have more personalized practice with Listening Boosters. *Listening Boosters* generated appropriate exercises for individual students based on the list of skipped items. Whenever students clicked on the request button '*Suggest videos for me now*', the app randomly selected a word/phrase from the list of skipped items as a keyword for performing a word-in-video search, identical to the technique used in the main practice sessions. The app then presented all the video clips containing the target word/phrase and requested the students to supply just the word/phrase occurring in all the presented video clips. The app also provided corrective feedback if the students' transcription was inaccurate as well as allowed the students to submit multiple revisions. When the students can provide an accurate answer, the app removed the target word/phrase from the list of skipped items. Figure 3.11 is a screenshot of the page featuring Listening Boosters.

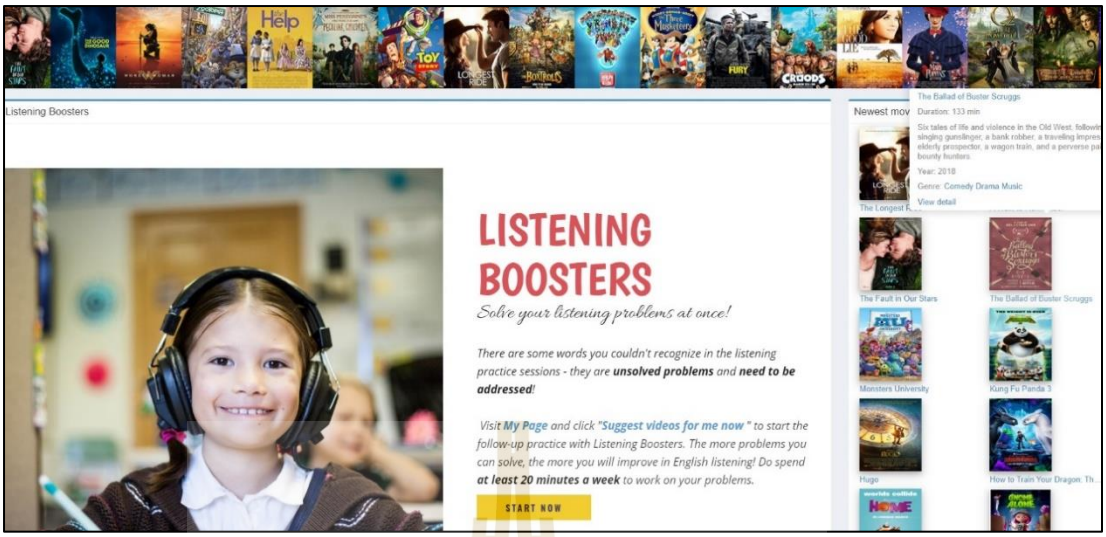


Figure 3.11 Listening Boosters

Figure 3.12 summarizes the main operations underlying Listening Hacked.

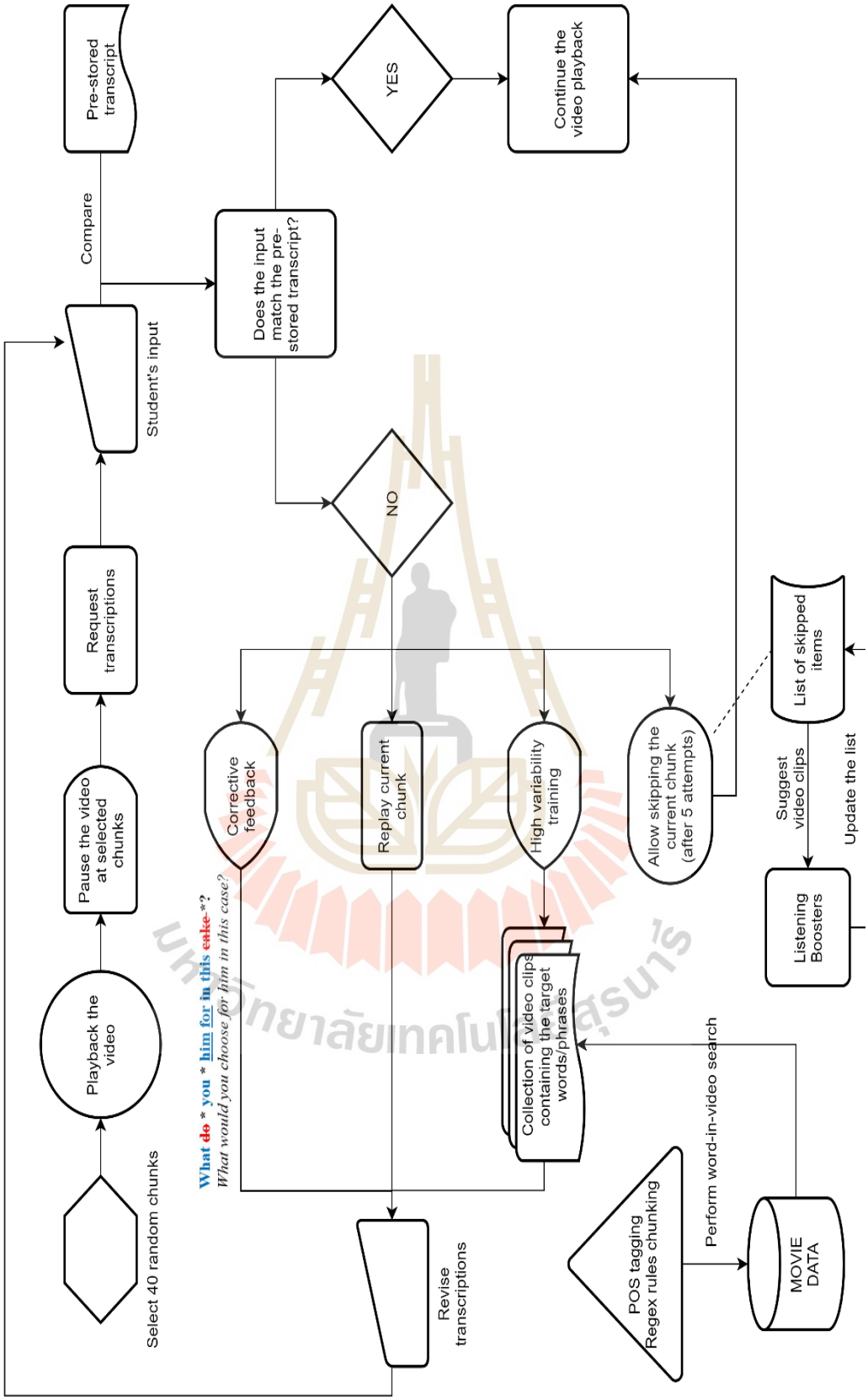


Figure 3.12 Main operations in Listening Hacked

3.5 Data collection procedures

The study was carried out in a fourteen-week time frame and did not follow the semester period at the university. There were separate procedures for the Experimental Group and the Control Group as described in Figure 3.13.

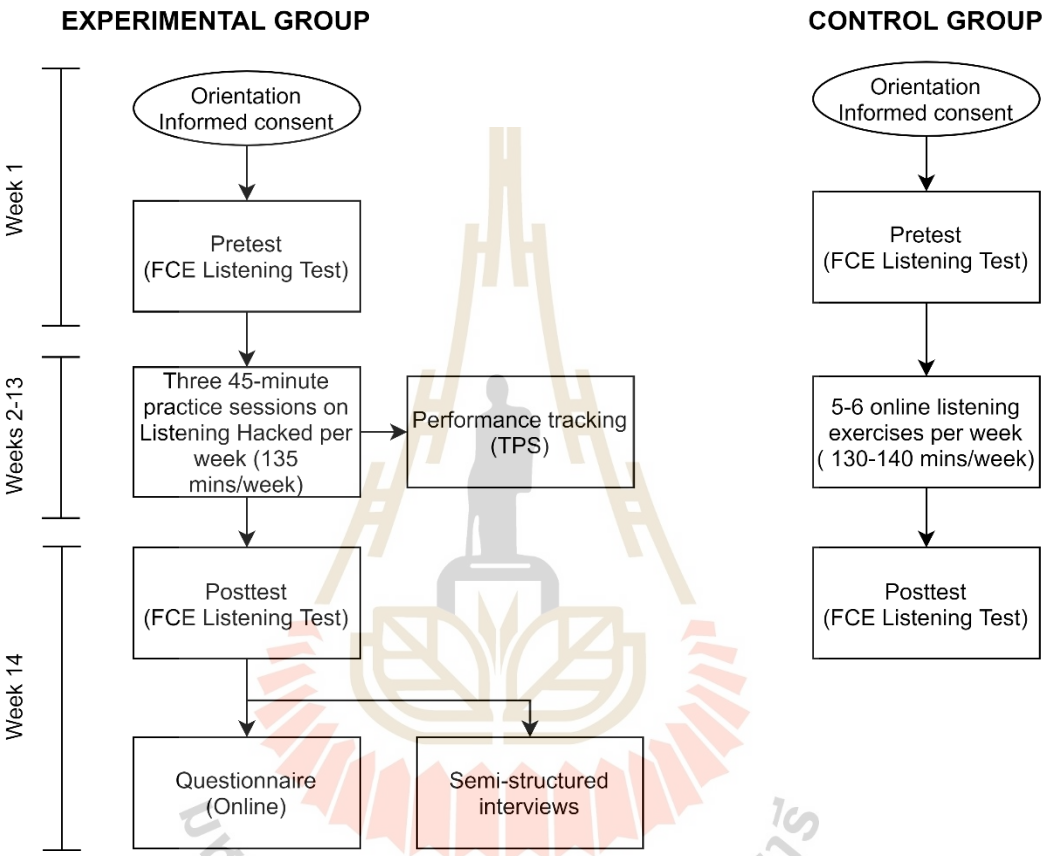


Figure 4.13 Summary of data collection procedures

After the process of sampling and group assignments as described in Section 3.1.2, the research assistant helped collect demographic information of the participants. Two orientations were scheduled in the first week by the research assistant, one for the Experimental Group and the other for the Control Group. The students in the Experimental Group were informed about the research purposes, how to use Listening Hacked, and how often they needed to use it. They were also informed that the results of the training program did not affect the results of their formal learning. The students were also instructed to sign up for an account and experiment with the features of

Listening Hacked. In the second orientation, the research assistant met the participants in the Control Group, also explaining the purposes of the research, and giving them guidelines on how to do the listening assignments online. After the orientation, the pretest was administered to all participants.

In the remaining weeks, the students in the Experimental Group started using Listening Hacked in their independent learning. They were required to complete three practice sessions per week and did so in a consecutive 12 weeks, meaning that each student had to complete 36 practice sessions by the end of the training period. The students were not required to complete a full movie but only a practice session of it. Moreover, they were advised to check the list of skipped items and solve those listening problems by doing suggested tasks in Listening Boosters. Upon completion of each listening practice session, the students received a task performance score. Information about these task performance scores was collected.

The students in the Control Group, on the other hand, had to do weekly listening exercises on the Google classroom. As described in Section 3.1.2, they did an equal amount of listening practice as compared to their counterparts, i.e. 5-6 listening exercises per week, or approximately 2.25 hours per week in 12 consecutive weeks. The exercises were listening to English audio recordings taken from English course books (Keynotes Upper-intermediate and Perspectives 3) and English podcasts accompanied with comprehension questions (multiple choice questions, true/false statements, multiple matching).

In the last week of the period (Week 14), the posttest was administered to both the Experimental Group and the Control Group. A questionnaire was sent to the participants in the Experimental Group in Week 14. Additionally, 19 participants in the Experimental Group were randomly selected and interviewed in Week 14.

Table 3.6 summarizes the collected data and the methods for obtaining them.

Table 3.6 Summary of data and methods for collection

No.	Data	Collection method	Time
1	*Pretest scores*	FCE Listening Test	Week 1
2	*Posttest scores*	FCE Listening Test	Week 14
3	*Task performance scores*	Progress tracking	Weeks 2-13
4	*Learning behaviors on Listening Hacked*	Activity logging	Week 2-13
5	*Students' opinions of Listening Hacked*	Online questionnaire with closed-ended and open-ended questions	Week 14
6	*Students' responses about their learning behaviors*	Instant messaging interviews	Week 14

3.6 Data analysis

3.6.1 Research Question 1

To answer the first research question, data were obtained from the pretest and the posttest. The first step involved marking the test papers using the answer key accompanying the FCE listening tests. Each correct response in the test was awarded 1.0, and the maximum score was 30. The researcher marked the test papers and a colleague helped to verify the reliability of the marking.

The second step was to apply statistical analysis to answer Research Question 1. First, Levene's test and independent samples t-test were employed to analyze the homogeneity of the students' pretest scores. These tests informed whether the students' EFL listening proficiency level significantly differed from that prior to the experiment. Second, paired samples t-test was applied on the pretest and posttest of each group in order to examine whether each group has made any changes in their listening scores. Third, independent samples t-test was run on the posttest scores to investigate the comparative effect of Listening Hacked on the students' EFL listening performance.

Regarding the task performance scores, only the TPS resulted from the three compulsory sessions were collected. The TPS were sorted according to the weeks which the compulsory sessions were assigned. The statistical procedures involved

calculating repeated measures of ANOVA to see whether the students have improved their TPS over the course of learning on Listening Hacked. Post-hoc tests were also employed to further explore where the differences in TPS occurred.

3.6.2 Research Question 2

Regarding Research Question 2 about the students' learning behaviors on Listening Hacked, the students' learning behavior data were retrieved from the app's database. The learning behavior data went through five analyses. The first analysis was to conduct frequency testing, i.e. to calculate the frequency and other descriptive statistics of the students' learning behaviors. This was done with Pivot function in Microsoft Excel and then Frequencies and Descriptive statistics in SPSS version 23.

The second analysis involved examining possible correlations between certain learning behaviors and EFL listening test performance by using Person correlation coefficient. Furthermore, multiple linear regression analysis was employed when any significant correlations were found.

The third analysis was to conduct a trend analysis of the students' learning behaviors over the observed period of 12 weeks. Repeated measures of ANOVA was used to examine whether there were significant changes in each behavior between the three months (each month consisted of 4 weeks in the observed period). Post-hoc test using Bonferroni correction was also employed to investigate where the changes reached the statistical significance level at .05.

The fourth analysis was to conduct a sequential pattern analysis to examine the common patterns of learning behaviors among the students. The analysis was done by calculating the Z-scores of the transitions between every two behaviors. Data were prepared and coded with Kutools extension in Microsoft Excel and analyzed for residuals table with GSEQ 5.1 (Quera et al., 2007).

The last analysis involved analyzing the qualitative data from the interviews and the open-ended responses in the questionnaire in order to investigate the students' reasons for exhibiting certain learning behaviors. Categorical content analysis procedures like the one used in Goh (2000) were adapted. The procedures were detailed as follows.

First, the researcher carefully read through all the students' responses to familiarize with the ideas expressed in those responses. Then, the researcher marked any descriptions explaining the reasons and situations that the students decided to use or not to use certain features of Listening Hacked.

After that, the researcher assigned the marked quotations with a code. For example, suppose a student wrote "*The feature of Listening Booster helps me practice what I'm having problems with, but sometimes it takes too long to load the videos*", the researcher would assign a code *Solving personal problems* to the phrase "*helps me practice what I'm having problems with".* Likewise, the clause "*sometimes it takes too long to load the videos"* would be applied with a code *Technical issues.* When a similar description is found, the existing code was applied to it and no new code was created. After the initial coding phase, the researcher grouped the codes into categories and reducing overlapping codes. Finally, the researcher looked for the relationships between the categories and refined the categories to bring out the themes.

3.6.3 Research Question 3

To answer the third research question, *what were the students' opinions about the usefulness of the learning system and its features?,* quantitative data from the students' questionnaire responses were analyzed for the central tendency and the variability of distribution regarding the students' opinions of each feature, plus the overall perceived usefulness. Statistical measures included the mean, standard deviation, and proportion of each rating (Urdan, 2010).

For a deeper understanding of the quantitative findings, qualitative data from the open-ended questions and the semi-structured interviews were analyzed using a similar approach as described above. However, the analysis was concentrated on any students' descriptions or mentions of positive or negative opinions about Listening Hacked and its features.

Note that the whole qualitative analysis procedures were performed in the original language of the students' responses, i.e., in Vietnamese, and only some quotations were selected and translated into English for reporting purposes. The qualitative analyses were conducted with ATLAS.ti version 9.

Table 3.7 Summary of methods for data collection and data analysis

Research question	Data collection method	Analytical method
1) What were the effects of Listening Hacked on Vietnamese English-major undergraduates' EFL listening comprehension?	Pretest (EG & CG)	Levene's test, Independent t-test
	Pretest vs Posttest (EG)	Paired-samples t-test
	Posttest (EG & CG)	Independent t-test
	Progress tracking (TPS)	Repeated measures of ANOVA, Post-hoc test
2) What happened when students encountered a listening problem in the learning system? What were the students' reasons for exhibiting such behaviors?	Activity logging (EG) Semi-structured interviews (EG)	Descriptive statistics Association testing Pearson correlation Multiple linear regression Repeated measures of ANOVA and Post-hoc test Categorical Content analysis
3) What were the students' opinions about the usefulness of the learning system and its features?	Questionnaire Semi-structured interviews	Descriptive statistics Categorical content analysis

3.7 The pilot study

3.7.1 Purposes of the pilot study

A pilot study is a small-scale study that is conducted to test the feasibility of the full-scale research project (Hassan et al., 2006). It is conducted prior to the main study to inspect the research protocol and the instruments to avoid any potential problems that may occur in the full study (Kraemer et al., 2006). For these reasons, the present pilot study was conducted with the following objectives.

1. To investigate the feasibility of conducting the full-scale study
2. To test the research instruments, including the features of activity logging and automatic grading, the appropriateness and comprehensibility of the items in the questionnaire, the appropriateness of data collected in the semi-structured interviews as well as using instant messaging (IM) as an interview technique
3. To identify any potential problems in the process of data analysis

4. Most importantly, to see whether the operation of Listening Hacked can fulfill the expectations of the study

3.7.2 Sampling, participants, and procedures

The pilot study used a convenient sample in which the researcher invited his former students to participate in the study for two weeks. The participants of the pilot study were 13 students, with ages from 21 to 26. At the end of the period, two students withdrew from the study. The remaining number of participants was 11.

The participants were instructed on how to learn with Listening Hacked and were required to use the app at least 4 hours in two consecutive weeks. After the course of two weeks, a questionnaire was sent to all of them. Finally, four of the participants were randomly selected to be interviewed via Facebook Messenger.

3.7.3 Results

3.7.3.1 Feasibility of conducting the full-scale study

From the sample in the pilot study, it can be concluded that the full-scale study was feasible. In general, the participants did not encounter any problems in terms of gaining access to the learning system and they expressed in the questionnaire high levels of satisfaction with their learning on Listening Hacked. The only problem preventing them from using Listening Hacked more frequently, as they reported, was that the period of conducting the pilot study coincided with the time when Vietnam's government ceased the lockdown measures for COVID-19, therefore, they were occupied with their job and schoolwork after the long lockdown period. Besides, there was only one participant reporting account registration and log-in problems.

3.7.3.2 The research instruments

Regarding the feature of activity logging, the system did a satisfactory work. The students' important interactions with the learning system were logged in the database as expected. However, one type of interactions i.e. when the students turned on or off the caption texts for movies was not logged. This type of interaction was important to understand whether the students used captions to assist their listening comprehension as a compensatory help option, thus, it was added to the list of logged interactions/ behaviors after the pilot study.

For automatic grading, some students reported that occasionally it misleadingly reflected their listening ability. They said that some words were mistranscribed because they were proper nouns or interjections which made them unreasonably difficult to spell correctly. Consequently, those mistranscribed words affected their task performance scores. In response to this issue, the transcription tasks in Listening Hacked no longer included proper nouns or interjections in the chunks to be transcribed.

Regarding the questionnaire, the results reveal that the data collected were useful but somehow inadequate for the purposes of the study. In addition to investigating the students' perceived effectiveness of the learning system, one important goal of the study was an understanding of how the learning system may assist with EFL listening learning. However, the fact that most informants merely reported on what and how certain features of Listening Hacked should be designed or improved was insufficient. Though not all of the informants, the majority of them tended to comment on the technical aspects without referring to the impact on their listening learning. For instance, a student commented on the use of captions in watching movies.

> *I think that the subtitles were carefully compiled and edited.*

Similarly, when being asked to explain how the transcription tasks affected their listening learning, another student pinpointed some technical issues in the transcription tasks and suggested a future change. It is also noted that the system had a bug reporting form for this purpose.

> *Very useful, but there are some flaws to be fixed. I have already submit the feedback form listing that. Besides, I think the feedback features for transcription should be more flexible.*

It was expected that the collected data should include information illuminating both the students' perceived usefulness of the learning system and the ways in which a certain feature of Listening Hacked assisted with the students' EFL listening learning.

For dealing with this issue, two solutions were proposed. First, the wording of the items in the questionnaire was rephrased so that it reminds informants of making references to their learning when commenting on each feature. Second, as observed in the data collected from the interviews, much information provided by the interviewees was either overlapping or supporting their responses in the questionnaire, therefore, the researcher decided to combine the two sources of data concerning the students' perceptions for answering both Research Question 2 and Research Question 3. The combination of the two data sources worked as a method for data triangulation, i.e. data corroboration.

Regarding the instant messaging interviews, from the four interviews conducted via Facebook Messenger, although this technique could be effective at collecting data as suggested by Fontes and O'Mahony (2008), Kaufmann and Peil (2019), Pearcea et al. (2014), the fact that the interviewees' responses were carefully revised and well-thought made their responses very condense. Condense responses are good but at the mercy of information loss. Another issue was that it took 70 minutes on average to complete an instant messaging interview which involved lots of typing amount and was also exhausting due to the students' effort in thinking, compiling expressions of ideas, and typing. For these reasons, especially to ensure the richness of data, the interview technique was replaced by the traditional face-to-face, oral interviews.

3.7.3.3 Potential problems in the process of data analysis

In general, the analysis of the qualitative data extracted from the questionnaire and interviews did not experience significant difficulties. There were no major challenges in analyzing the data obtained from Listening Hacked database either.

3.7.3.4 The operation of Listening Hacked

Since the whole study completely relied on the robust operation of the learning system, the most important goal of the pilot study was to explore any areas of weaknesses in the operation of Listening Hacked so that the app could be improved to ensure proper and smooth operation as expected in the full study.

Based on the observations of the students' learning behaviors on Listening Hacked and their feedback in bug reports, 11 major changes were made to

the latest version of Listening Hacked. Table 3.8 details the changelog of Listening Hacked after the pilot study and the reasons for those changes.

Table 3.8 Changelog after the pilot study

No.	Change	Reason
1	Reduced the number of chunks for transcribing from 40 to 10 per practice session	An original practice session (40 chunks) normally took a student 2.5-3 hours which was too long and exhausting.
2	Proper nouns i.e. names of people or places, ... were displayed and not selected for transcription	Some names of people and places were unfamiliar and caused difficulties in spelling.
3	The Rewind button now allowed students to listen to the current chunk, plus 2 seconds before and after that chunk	This change made this button significantly different from the Repeat button in that it allowed listening with more contexts. It also solved the technical issue of inaccurate synchronization between audio and captions.
4	Logged the students' actions of turning on/off caption texts	Student's decisions on using captions are important in understanding their learning.
5	Added an option of "*View All*" to display all items in the *List of recently skipped words* featured in the students' personal page	The original list only showed up to 10 recently skipped items while some students may need to review a full list.
6	Added a new page for *Listening Booster* and put it on the main menu	Some students did not notice Listening Booster and did not know how to use it.
7	Added a video on how to use Listening Hacked	Some students reported difficulties in reading long written instructions and preferred instruction videos.
8	Fix captions ● Removed irrelevant texts ● Removed texts for hearing impaired ● Removed HTML formatting tags ● Removed lyrical texts ● Split long lines of conversations ● Resynchronized captions with movies using Automatic Speech Recognition (ASR) technology ● Removed interjections (whoawhoo, whoowhee,...) and vocal sounds (aahh, hmmm, ...) ● Used standard contraction forms (u, ya → you, 'em → them, ...)	Most of the technical frustrations in the transcription tasks were due to irrelevant or non-standard caption texts, or inaccurate synchronization with the audios.

Table 3.8 Changelog after the pilot study (Cont.)

No.	Change	Reason
9	Reworked the mechanism of phrase search in other movies: When a phrase is searched and not found in other movies, the app will automatically search the whole movie data and display the five most frequent phrases containing the keyword (that causes difficulties in the transcription task). In case the phrase search returns no results, the app will perform single-word search, instead of sticking to that phrase	Because the movie data is limited, very often, the app reported "not found" results when performing a phrase search in other movies. While single word searching often returned many relevant results, exact phrase searching requires much bigger movie data to have similar results.
10	Phrase Search in other movies can now be used with successfully transcribed words	High variability perceptual training can be used with already recognizable words to enhance word recognition skills.
11	Enhanced the sound quality: Change the HTML5 video player so that the sound decoder was improved	Some students reported that the sound quality of many movies was too poor, causing great difficulties in listening comprehension.

3.8 Chapter summary

This chapter describes and discusses the methodological decisions and procedures for conducting the present study. Furthermore, it summarizes the main results of the development process of the proposed learning system (Listening Hacked), including how the principles of the learning system were formulated and transformed into certain features, and how those features were developed and implemented in the system. Finally, the chapter reports on some important changes resulting from the pilot study.

CHAPTER 4
LITERATURE REVIEW

This chapter provides the answers to the research questions posed in Chapter 1 by reporting and summarizing the results of the data analysis. Section 4.1 presents the results of the pretest and the posttest in response to Research Question 1. Section 4.2 discloses the students' learning behavior data with reference to Research Question 2. Finally, the results of the survey and interviews are reported in Section 4.3.

4.1 The effectiveness of Listening Hacked

One primary purpose of the present research was to investigate the effectiveness of the AI-powered, personalized learning system, namely Listening Hacked, constructed based on the proposed conceptual framework (see Section 2.8 for more details) on Vietnamese English-major undergraduates' EFL listening comprehension. In order to accomplish this, Research Question 1 was raised: *What were the effects of Listening Hacked on Vietnamese English-major undergraduates' EFL listening comprehension?*

To this end, an English listening pretest and a posttest were administered to both groups of participants, the Experimental Group and the Control Group. The analytical procedures for answering Research Question 1 involved comparing the pretest scores, comparing differences between the pretest and posttest scores within each group, and comparing the posttest performances between the two groups.

4.1.1 Pre-treatment levels of EFL listening proficiency

The first analysis in the procedures was to test the homogeneity of the two groups regarding the pretest scores which were indicative of their pre-treatment levels of EFL listening proficiency. Table 4.1 reports the results of the pretest of the two groups.

Table 4.1 Descriptive statistics of the Pretest

Experimental Group				Control Group			
N	Mean	Std. Deviation	Std. Error Mean	N	Mean	Std. Deviation	Std. Error Mean
30	9.2	4.45978	0.81424	23	7.2174	3.32971	0.69429

At first glance, it seems that there was a mean difference of roughly 2.0 in the prestest scores between the two groups. Specifically, the Experimental Group apparently outperformed the Control Group on the pretest by approximately 2.0. However, as this mean difference was observed in the sample; to determine whether this difference was statistically significant, Levene's test for equality of variances and independent samples t-test were employed. Table 4.2 summarizes the results of the Levene's test and the t-test on the pretest.

Table 4.2 Levene's test and independent t-test on the Pretest

Levene's Test for Equality of Variances		t-test for Equality of Means		
F	P-value	t	df	P-value (2-tailed)
1.697	0.198	1.783	51	0.08

The result of Levene's test shows that there was no statistically significant variance in the pretest between the two groups (F=1.697; p=0.198>.05). The results of the independent samples t-test also indicate that there was no statistically significant difference in the pretest between the two groups (t=1.783; p>.05). However, since the size of the Control Group (N=23) was under 30 which does not guarantee the central theorem for normal distribution, the bootstrap method for calculating the 95% confidence interval (95% CI) was employed to validate the t-test results (Brownlee, 2018; Kulesa et al., 2015).

4.3 summarizes the results of the Bootstrap procedures.

Table 4.3 Bootstrap results of independent t-test on the Pretest

Mean Difference	Bootstrap				
	Bias	Std. Error	P-value (2-tailed)	95% Confidence Interval	
				Lower	Upper
1.98261	-0.0228	1.05178	0.07	-0.1372	3.9578

Regarding the bootstrap results based on 1000 samples for independent t-test, the 95% Confidence Interval (CI) contains value 0 (Lower bound: -.137, Upper bound: 3.958), thus the mean difference (1.98) was statistically insignificant. This points to a similar result as in the independent samples t-test, and that means there was no statistically significant difference in the pretest scores between the two groups. In order words, the students in the Experimental Group and the Control had the same starting levels of EFL listening proficiency as their pretest scores were not statistically different.

4.1.2 Within-group changes in EFL listening performance

To see whether each group has made any progress in their EFL listening comprehension ability, the analyses then focused on the changes in their pretest and posttest scores within each group. Table 4.4 summarizes the descriptive statistics for the pretest and posttest of the two groups. Table 4.5 presents the results of the paired-samples t-test for each group.

Table 4.4 Descriptive statistics of the Pretest and the Posttest

	Experimental Group		Control Group	
	Pretest	Posttest	Pretest	Posttest
Mean	9.2	12.5	7.2174	8.2609
N	30	30	23	23
Std. Deviation	4.45978	6.05008	3.32971	4.24497
Std. Error Mean	0.81424	1.10459	0.69429	0.88514

Table 4.5 Paired samples t-test on the Pretest and Posttest

		Experimental Group	Control Group
		Mean difference (Posttest – Pretest)	Mean difference (Posttest – Pretest)
Paired Differences	Mean	3.3	1.04
	Std. Deviation	4.37	3.54
	Std. Error Mean	0.798	0.737
t		4.134	1.416
df		29	22
P-value (2-tailed)		.000	0.171

As shown in Table 4.5 , the Experimental Group improved their listening test scores by 3.30 on average (N=30; t=4.134; p<.01), with a relatively large effect size (d=.76).

Regarding the Control Group, the results of the paired samples t-test indicate that there was no statistically significant difference in the pretest and posttest scores (N=23; t=1.416; p>.05). Since there might be concerns over the small sample size of the Control Group, the bootstrap method was again calculated, based on 1000 resampling times. The bootstrap results (see Table 4.6) show that the 95% CI contains value 0 (Lower bound: .-348; Upper bound: 2.43), implying a statistical insignificance. This is in line with the t-test results.

Table 4.6 Bootstrap results of the paired samples t-test on the Control Group Pretest and Posttest

	Mean	Bootstrap				
		Bias	Std. Error	P-value (2-tailed)	95% Confidence Interval	
					Lower	Upper
Posttest – Pretest (Control Group)	1.04348	0.00783	0.72159	0.173	-0.34783	2.43478

In addition to the pretest and posttest, another source of listening performance was the task performance scores (TPS) which the students in the Experimental Group received upon completing compulsory practice sessions in the three timepoints: Weeks 2, 7, and 12. Table 4.7 summarizes the results of the statistical analyses on the TPS.

Table 4.7 Mauchly's test of sphericity and repeated measures of ANOVA

Mauchly's W	Approx. Chi-Square	df	Sig.	
0.876	3.71	2	0.156	
Source		df	F	P-value
Week	Sphericity-assumed	2	22.373	.000
Error (Week)	Sphericity-assumed	58		

The Mauchly's test of sphericity indicates that the dataset does not violate the assumption of sphericity ($\chi 2(2)=.876$; $p>.05$). The repeated measures of ANOVA with sphericity assumed reveals that the differences in the TPS between the three timepoints were statistically significant ($F(2, 58)=22.373$; $p<.001$). The post-hoc tests using Bonferroni correction (see Table 4.8) were also conducted to investigate further where the differences occurred. The results show that the differences were statistically significant between Week 2 and Week 7 (mean difference=-1.507; $p<.001$), and between Week 2 and Week 12 (mean difference=-1.885; $p<.001$). There was a difference of -.378 between Week 7 and Week 12, however, it did not reach the statistical significance level at .05. In short, the results suggest that the students' TPS increased significantly over the observed timepoints.

Table 4.8 Post-hoc tests for pair-wise comparisons

Week (I)	Week (J)	Mean Difference (I-J)	Std. Error	P-value
2	7	-1.507	.254	.000
	12	-1.885	.334	.000
7	12	-.378	.289	.607

To put it concisely, the students in the Experimental Group improved their EFL listening comprehension performance significantly, whereas the students in the Control Group did not. Interestingly, 30.43% of the students in the Control Group performed worse in the posttest than the pretest. This proportion was lower, that is 23.33%, in the Experimental Group (see Table 4.9).

Table 4.9 Changes between the pretest and posttest

	Experimental Group (N=30)		Control Group (N=23)	
	Number of cases	Percent	Number of cases	Percent
Posttest – Pretest <0	7	23.33	7	30.43
Posttest – Pretest =0	1	3.33	3	13.04
Posttest – Pretest >0	22	73.33	13	56.52

4.1.3 Post-treatment levels of EFL listening performance

To see the effects of Listening Hacked on the students' EFL listening performance, independent samples t-test was conducted on the students' posttest between the Experimental and the Control Group. Table 4.10 reports the results of the independent t-test on the students' posttest scores.

Table 4.10 Independent t-test on the Posttest

Levene's Test for Equality of Variances		t-test for Equality of Means				
F	P-value	t	df	P-value (2-tailed)	Mean Difference	Std. Error Difference
3.087	0.085	2.861	51	0.006	4.23913	1.48183

The result of the Levene's test reveals that there was statistically equal variance in the posttest scores in the two groups (F=3.087; p>.005). From the results of the independent t-test, it could be concluded that the Experimental Group outperformed the Control Group on the Posttest by 4.24 on average (t=2.861; p<.05).

Because equal variance was assumed in the data according to the Levene's test result, the assumption of normal distribution was not a concern in this analytical procedure; however, to corroborate the results of the t-test, the bootstrap method

was performed for calculating the 95% CI. The results of bootstrap for independent samples t-test (see Table 4.11) show that the 95% CI does not contain value zero (Lower bound 1.60; Upper bound 6.82) which indicates a similar result as compared to the t-test results.

Table 4.11 Bootstrap results of the independent t-test on the Posttest

Mean Difference	Bootstrap				
	Bias	Std. Error	P-value (2-tailed)	95% Confidence Interval	
				Lower	Upper
4.23913	0.04926	1.36715	0.005	1.59506	6.81685

In conclusion, the students in the Experimental Group significantly improved in their listening comprehension, and they outperformed their counterparts in the Control Group on the posttest.

4.2 The students' learning behaviors on Listening Hacked

Research Question 2, *What happened when students encountered a listening comprehension problem on Listening Hacked? What were the students' reasons for exhibiting such behaviors?* was mainly concerned with exploring the students' learning behaviors, the patterns of learning behaviors on the learning system, and the possible reasons for exhibiting those behaviors. Accordingly, there were two parts in the analyses in relation to Research Question 2. Firstly, the analyses involved (a) calculating the frequency and distribution of the learning behaviors (see Section 4.2.1) (b) examining whether there were any potential correlations between certain behaviors and the students' listening achievements (see Section 4.2.2), (c) investigating whether there were any changes in the behaviors over the period of 12 weeks of the experiment (see Section 4.2.3), and (d) identifying the sequential patterns of behaviors (see Section 4.2.4). Secondly, the qualitative data obtained from the interviews and the open-ended responses in the questionnaire were analyzed to understand the students' reasons for showing certain learning behaviors.

4.2.1 Frequency of the learning behaviors

In total, there were 11 types of learning behaviors recorded in the system database. These behaviors were coded as shown in Table 4.12. Note that the category "Total_Repeat" in Figure 4.1 refers to the total number of times that the students used all types of repetition, including Repeat, Rewind, Rewind 1, Rewind 2, Rewind 5, and Rewind 10. The category "Total_Interaction" refers to the total number of interactions that the students made over the period of 12 weeks.

Table 4.12 Codes for students' learning behaviors

Code	Behavior
DICTIONARY	Use the built-in dictionary
FEEDBACK	Read the automatically generated corrective feedback
OFFCAP	Turn off English captions
ONCAP	Turn on English captions
PHRASE SEARCH	Perform a phrase search by clicking on an asterisk in the feedback or by double-clicking on a word in the feedback
REPEAT	View/ listen again to the current chunk being transcribed
REWIND	View/ listen again to the current chunk, plus 2 seconds before and 2 seconds after that chunk
REWIND 1	Rewind the movie playback 1 minute from the current chunk
REWIND 2	Rewind the movie playback 2 minutes from the current chunk
REWIND 5	Rewind the movie playback 5 minutes from the current chunk
REWIND 10	Rewind the movie playback 10 minutes from the current chunk

Table 4.13 Frequency of students' learning behaviors

	Frequency	Mean	Std. Deviation	Percentage
DICTIONARY	274	9.13	8.13	0.23
FEEDBACK	50452	1681.73	598.99	43.12
OFFCAP	476	15.87	11.3	0.41
ONCAP	1519	50.63	24.02	1.3
PHRASESEARCH	17988	599.6	570.86	15.37
REPEAT	26365	878.83	800.98	22.53
REWIND	19275	642.5	558.06	16.47
REWIND 1	569	18.97	20.85	0.49
REWIND 10	15	0.5	1.01	0.01
REWIND 2	42	1.4	1.94	0.04
REWIND 5	21	0.7	0.99	0.02
TOTAL	116996	3899.87	1437.52	100

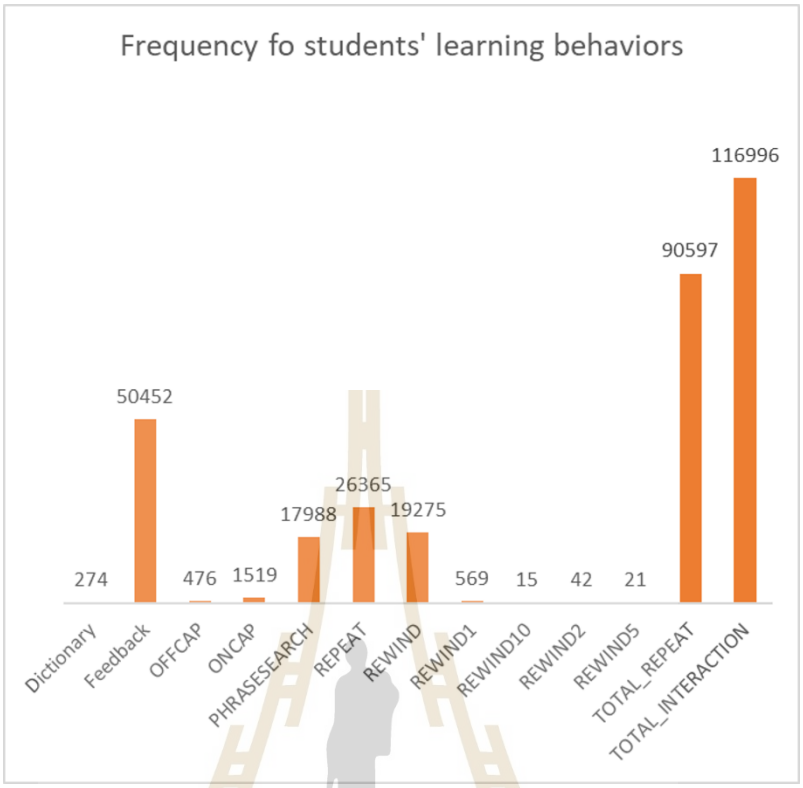


Figure 4.1 Frequency of students' learning behaviors

As seen in Table 4.13 and Figure 4.1, the most frequent behavior when the students encountered a listening problem was to consult the feedback (M=1681.73), which accounted for 43.12% of the total actions that students did. This was followed by using the Repeat function (M=878.83; 22.53%) and the Rewind function (M=642.50; 16.47%). Among the most frequent behaviors was to perform a phrase search (M=599.6), which contributed to 15.37% of the students' total interactions.

4.2.2 Correlations between the students' learning behaviors and their listening performance

For there were differences in the frequency of each behavior, the analyses continued to examine whether there were any potential correlations between each learning behavior and the students' listening performance on the posttest. Correlational analyses were conducted with Pearson correlation coefficient and summarized in Table 4.14

Table 4.14 Correlations between students' posttest and certain learning behaviors

Learning Behavior	Posttest	
	r	*p-value (2-tailed)*
DICTIONARY	-0.091	0.632
FEEDBACK	-0.196	0.298
OFFCAP	0.065	0.735
ONCAP	0.026	0.892
PHRASE SEARCH	-.495**	0.005
REPEAT	0.022	0.907
REWIND	0.01	0.957
REWIND1	.513**	0.004
REWIND2	0.109	0.568
REWIND5	0.061	0.75
REWIND10	0.257	0.17

The results of the correlational analyses show a moderate negative correlation between the students' behavior of performing phrase searches and their listening posttest scores (r=-.495; p<.05; N=30). There was also a moderate positive correlation between the students' behavior of rewinding 1 minute and their posttest scores (r=.513; p<.05; N=30). Remarkedly, while other behaviors exhibited high frequencies as reported in

4.13, their correlations with the posttest scores did not reach the significance level at .05, and only two of the observed behaviors were correlated with the students' posttest scores.

Furthermore, since there were correlations between those variables, a multiple linear regression analysis was conducted to predict the students' listening posttest scores from the two behaviors: performing phrase searches and rewinding 1 minute. Table 4.15 presents the results of the multiple regression analyses.

Table 4.15 Multiple Regression Model

Model summary			
R	R^2	Adjusted R^2	Std. Error of the Estimate
0.678	0.46	0.42	4.60943

ANOVA					
	Sum of Squares	df	Mean Square	F	P-value
Regression	487.835	2	243.917	11.48	.000
Residual	573.665	27	21.247		
Total	1061.5	29			

Coefficients					
	Unstandardized Coefficients		Standardized Coefficients	t	P-value
	B	Std. Error	Beta		
(Constant)	12.77	1.513		8.44	.000
PHRASESEARCH	-0.005	0.002	-0.446	-3.133	.004
REWIND 1 minute	0.135	0.041	0.466	3.272	.003

The results suggest that the two types of behavior statistically predicted the students' posttest scores ($F(2,27)=11.48$; $p<.001$; $R^2=.46$). Moreover, each of the two observed behaviors contributed significantly to the prediction of the posttest scores ($p<.05$). Specifically, the results of the coefficient analysis imply that for every additional time performing a phrase search, a student's listening posttest score would decrease by 0.005 point ($B=-.005$; $p<.05$). To put it another way, a student's posttest score would improve by 0.005 point for every single drop of using Phrase Search. On the contrary, the coefficient of the predictor Rewind 1 minute was 0.135, meaning that for every time a student performed this action, their listening posttest score was supposed to increase by 0.135 ($B=.135$; $p<.05$).

4.2.3 Changes in the students' learning behaviors on Listening Hacked

To investigate the changes in the students' learning behaviors on Listening Hacked, trend analyses were conducted by calculating the repeated measures of ANOVA for each behavior in every month. Note that each month consists of four consecutive weeks in the experiment timeframe of 12 weeks.

4.2.3.1 Looking up a word/phrase in the built-in dictionary (DICTIONARY)

The Mauchly's test of sphericity (see Table 4.16) indicates that the assumption of sphericity was violated (χ2(2)=20.608; p<.05). As the data violated the assumption of sphericity, Greenhouse-Geisser correction was used to interpret the results of the ANOVA with repeated measures.

Table 4.16 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.479	20.608	2	**.000**

Source		df	F	P-value
Month	Greenhouse-Geisser	1.315	9.694	**.002**
Error (Month)	Greenhouse-Geisser	38.133		

The repeated measures of ANOVA with a Greenhouse-Geisser correction reveals that the mean frequency of using the built-in dictionary differed significantly between the observed months (F(1.315, 38.133)=9.694; p<.05). In addition, the post-hoc tests using Bonferroni correction indicate that students reduced using the built-in dictionary between Month 1 and Month 2 (-4.233; p<.05). Their dictionary usage remained unchanged till the end of the research period.

Table 4.17 Post-hoc tests for pair-wise comparisons

Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	4.233*	1.185	.004
	3	4.233*	1.379	.014
2	3	0.000	.625	1.000

4.2.3.2 Reading corrective feedback (FEEDBACK)

As seen in Table 4.18, the data did not violate the assumption of sphericity (χ2(2)=.504; p>.05). The results of the repeated measures of ANOVA indicate a change in the mean frequency of reading feedback between the three months (F(2,58)=9.604); p<.001).

Table 4.18 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.982	.504	2	.777

Source		df	F	P-value
Month	Sphericity Assumed	2	9.604	.000
Error (Month)	Sphericity Assumed	58		

As shown in Table 4.19, there was a decline of using feedback in the first two months of the period; however, this change was statistically insignificant (p>.05). The real change occurred between Month 2 and Month 3 when there was a drop by 133.900 (p<.05). Overall, the observed period saw a downward trend in the students' behavior of reading corrective feedback (p<.001). That is to say, they tended to read feedback less than they used to in the first month.

Table 4.19 Post-hoc tests for pair-wise comparisons

Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	93.333	55.391	.308
	3	227.233*	49.621	.000
2	3	133.900*	51.188	.042

4.2.3.3 Turning off English captions (OFFCAP)

As shown in Table 4.20, the repeated measures of ANOVA with sphericity assumed reveals a significant difference in the students' behavior of turning off the English captions between the time points (F(2,58)=26.823; p<.001). A closer look at the results of the post-hoc analyses (see Table 4.21) further indicates that there was a gradual decline in the behavior of toggling off the English captions over the period, and this fall was statistically significant between Month 1 and Month 2 (p<.001) and between Month 1 and Month 3 (p<.001). In short, turning off English captions steadily became less frequent among the students' learning behaviors over the months.

Table 4.20 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.886	3.400	2	.183

Source		df	F	P-value
Month	Sphericity Assumed	2	26.823	.000
Error(Month)	Sphericity Assumed	58		

Table4.21 Post-hoc tests for pair-wise comparisons

Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	6.000	1.084	.000
	3	8.133	1.329	.000
2	3	2.133	1.018	.135

4.2.3.4 Turning on English captions (ONCAP)

The results of the repeated measures of ANOVA with sphericity assumed (see Table 4.22) suggest that there was a statistically significant difference in the students' behavior of turning on English captions in their learning ($F(2,58)=17.596$; $p<.001$). The post-hoc tests as shown in Table 4.23 highlight that the difference occurred between Month 1 and Month 2 when there was a mean decrease by 9.33 ($p<.001$). Overall, students reduced the behavior of turning on English captions while viewing movies on Listening Hacked over the period ($p<.001$).

Table 4.22 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.974	.726	2	.696

Source		df	F	P-value
Month	Sphericity Assumed	2	17.596	.000
Error (Month)	Sphericity Assumed	58		

Table 4.23 Post-hoc tests for pair-wise comparisons

Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	9.333*	2.048	.000
	3	11.633*	2.228	.000
2	3	2.300	1.945	.740

4.2.3.5 Performing a phrase search (PHRASE SEARCH)

As presented in Table 4.24, the data of performing phrase searches in the three months violated the assumption of sphericity ($\chi 2(2)=20.572$; $p<.05$), accordingly, Greenhouse-Geisser correction was used to interpret the repeated measures of ANOVA. The results of ANOVA determine that the students' mean usage of Phrase Search differed significantly between the timepoints ($F(1.315, 38.149)=6.258$; $p<.05$). The post-hoc tests using Bonferroni correction (see Table 4.25) show that the students steadily reduced the number of phrase searches performed between Month 1 and Month 2 (-88.167) and between Month 2 and Month 3 (-29.267), however, these changes were statistically insignificant. The decline in using Phrase Search only reached the significant level between Month 1 and Month 3 (-117.433; $p<.05$).

Table 4.24 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.480	20.572	2	.000

Source		df	F	P-value
Month	Greenhouse-Geisser	1.315	6.258	.011
Error (Month)	Greenhouse-Geisser	38.149		

Table 4.25 Post-hoc tests for pair-wise comparisons

Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	88.167	43.140	.150
	3	117.433*	36.513	.010
2	3	29.267	19.710	.445

4.2.3.6 Replaying chunks (REPEAT)

As shown in Table 4.26, the Mauchly's test implies a violation of sphericity in the dataset of REPEAT behavior ($\chi 2(2)=6.88$; $p<.05$), thus, the Greenhouse-Geisser correction should be used for interpreting the repeated measures of ANOVA. The repeated measures of ANOVA results with Greenhouse-Geisser correction indicate a statistically significant difference in the students' behavior of using the Repeat button between the time points ($F(1.642, 47.625)=11.247$; $p<.001$). The post-hoc tests (see Table 4.27) further reveal that students declined using the Repeat function in their practice sessions between Month 1 and Month 2 ($p=.001$). They continued to decline using it between Month 2 and Month 3 while this decline was statistically insignificant ($p>.05$). Overall, the students steadily decreased replaying chunks in their listening practice over the period (-217.467; $p<.05$).

Table 4.26 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.782	6.880	2	.032

Source		df	F	P-value
Month	Greenhouse-Geisser	1.642	11.247	.000
Error(Month)	Greenhouse-Geisser	47.625		

Table 4.27 Post-hoc tests for pair-wise comparisons

Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	151.000	37.914	.001
	3	217.467	56.448	.002
2	3	66.467	44.737	.444

4.2.3.7 Replaying chunks with more contexts (REWIND)

The trend of replaying a chunk with 2 seconds before and 2 seconds after it (Rewind function) was similar to that of using the Repeat function. As reported in Table 4.28, the repeated measures of ANOVA with sphericity assumed determines that the students' usage of Rewind function differed significantly between the

timepoints (F(2, 58)=6.152; p<0.5). As the post-hoc tests show, the students continually decreased using the Rewind function between Month 1 and Month 2 (-85) and between Month 2 and Month 3 (-40.3), although these changes were not statistically significant. The decline was, however, statistically significant between Month 1 and Month 3 (-125.3; p<.05), meaning that the students slowly reduced using the Rewind function over the whole period.

Table 4.28 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.944	1.619	2	.445

Source		df	F	P-value
Month	Sphericity Assumed	2	6.152	.004
Error(Month)	Sphericity Assumed	58		

Table 4.29 Post-hoc tests for pair-wise comparisons

Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	85.000	38.882	.111
	3	125.300*	38.245	.008
2	3	40.300	31.879	.649

4.2.3.8 Replaying the movie 1 minute before the current chunk (REWIND 1)

The repeated measures of ANOVA with Greenhouse-Geisser correction (see Table 4.30) indicates a significant difference in the students' behavior of rewinding the movie for 1 minute from the current chunk being transcribed (F(1.045, 30.308)=13.209; p=.001). The post-hoc analyses (see Table 4.31) further reveal that there was a decrease in using the Rewind 1 minute function between Month 1 and Month 2 (-10.1; p<.05), between Month 2 and Month 3 (-1.533; p<.05), and between Month 1 and Month 3 (-11.633; p<.05). The data generally show that the students significantly decreased using the Rewind 1 function over the months.

Table 4.30 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.086	68.596	2	.000

Source		df	F	P-value
Month	Greenhouse-Geisser	1.045	13.209	.001
Error(Month)	Greenhouse-Geisser	30.308		

Table 4.31 Post-hoc tests for pair-wise comparisons

Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	10.100*	2.866	.004
	3	11.633*	3.100	.002
2	3	1.533*	.569	.035

4.2.3.9 Replaying the movie 2 minutes before the current chunk (REWIND 2)

The behavior of rewinding the movie playback for 2 minutes before the chunks being transcribed declined significantly as demonstrated in the results of repeated measures of ANOVA with Greenhouse-Geisser correction. The results indicate a statistically significant change in the students' use of this function ($F(1.126, 32.645)=9.033$; $p<.05$). The post-hoc analyses as shown in Table 4.33 suggest that the change was a decline of this behavior between Month 1 and Month 2 (-1.2; $p<.05$). There was a slight increase of rewinding 2 minutes between Month 2 and Month 3 (+0.1), but this increase was statistically insignificant ($p>.05$). Overall, the students tended to reduce using the Rewind 2 minutes function over the period (-1.1; $p<.05$) and this was a statistically significant downward trend.

Table 4.32 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.223	41.975	2	.000

Source		df	F	P-value
Month	Greenhouse-Geisser	1.126	9.033	.004
Error(Month)	Greenhouse-Geisser	32.645		

Table 4.33 Post-hoc tests for pair-wise comparisons

Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	1.200	.363	.008
	3	1.100	.388	.025
2	3	-.100	.111	1.000

4.2.3.10 Replaying the movie 5 minutes before the current chunk (REWIND 5)

The results from repeated measures of ANOVA with Greenhouse-Geisser correction (see Table 4.34) point to a significant difference in the students' behavior of rewinding the movie playback for 5 minutes before the chunk ($F(1.17, 33.943)=6.226$; $p<.05$). The post-hoc tests (see Table 4.35) show that the change occurred between Month 1 and Month 2 when students decreased using this function by 0.533 ($p<.05$). However, the students slightly increased using Rewind 5 minutes by 0.67 on average between Month 2 and Month 3 although this was not statistically significant ($p>.05$). On the whole, the behavior of using Rewind 5 minutes did not change significantly at the end of the period ($p>.05$).

Table 4.34 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.291	34.539	2	.000

Source		df	F	P-value
Month	Greenhouse-Geisser	1.170	6.226	.014
Error(Month)	Greenhouse-Geisser	33.943		

Table 4.35 Post-hoc tests for pair-wise comparisons

Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	.533	.190	.027
	3	.467	.202	.084
2	3	-.067	.067	.977

4.2.3.11 Replaying the movie 10 minutes before the current chunk (REWIND 10)

The repeated measures of ANOVA with Greenhouse-Geisser correction (see Table 4.36) detects a significant difference in the students' behavior of replaying the movie 10 minutes before the chunk ($F(1.046, 30.348)=5.725$; $p<.05$). The post-hoc analyses (see Table 4.37), however, reveal that the differences between the timepoints were not statistically significant ($p>.05$). Thus, it can be said that the students' behavior of using the Rewind 10 minutes function did not change much over the period.

Table 4.36 Mauchly's test of sphericity and repeated measures of ANOVA

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	P-value
Month	.089	67.794	2	.000

Source		df	F	P-value
Month	Greenhouse-Geisser	1.046	5.725	.022
Error(Month)	Greenhouse-Geisser	30.348		

Table 4.37 Post-hoc tests for pair-wise comparisons

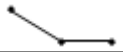
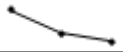
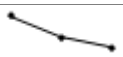
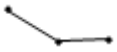
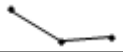
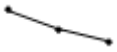
Month (I)	Month (J)	Mean Difference (I-J)	Std. Error	P-value
1	2	.433	.190	.090
	3	.467	.184	.051
2	3	.033	.033	.977

4.2.3.12 Summary

From the above analyses in Section 4.2.3, it is clear that most of the students' learning behaviors changed over the period and that the students tended to use less support on Listening Hacked. This was evident by the significant drops in every learning behavior between the 12-week period with exceptions of rewinding the movie playback for 5 and 10 minutes. Table 4.38 illustrates the trends of the students' learning behaviors in the observed timepoints. As shown in the table,

most of the behaviors experienced a downward trend between the timepoints. This means that the students gradually used less help options on Listening Hacked to accomplish the listening tasks while their TPS and EFL listening performance improved as pointed out in Section 4.12.

Table 4.38 Changes in learning behaviors

Behavior	Mean Month 1	Mean Month 2	Mean Month 3	Trend
DICTIONARY	5.87	1.63	1.63	
FEEDBACK	667.43	574.10	440.20	
OFFCAP	10.00	4.00	1.87	
ONCAP	23.87	14.53	12.23	
PHRASESEARCH	268.40	180.23	150.97	
REPEAT	415.77	264.77	198.30	
REWIND	284.27	199.27	158.97	
REWIND 1	13.57	3.47	1.93	
REWIND 2	1.23	0.03	0.13	
REWIND 5*	0.57	0.03	0.10	
REWIND 10*	0.47	0.03	0.00	
Total	1691.43	1242.10	966.33	

4.2.4 Patterns of learning behaviors on Listening Hacked

To identify the students' patterns of learning behaviors in this study, behavioral sequence analysis was conducted with GSEQ 5.1 (Quera et al., 2007). The results of the sequential pattern analysis generated the adjusted residuals table of the students' behavioral transitions as shown in Table 4.39. The rows represent the starting behaviors, and the columns represent the subsequent behaviors. Statistically significant transitions were indicated by a Z-score greater than 1.96 and printed in bold. In total, there were 40 significant sequences of behaviors ($Z>1.96$ and $p<.05$).

Table 4.39 Adjusted residual table

Starting behavior	DICTIONARY	FEEDBACK	OFFCAP	ONCAP	PHRASE SEARCH	REPEAT	REWIND	REWIND 1	REWIND 2	REWIND 5	REWIND 10
DICTIONARY	0	-5.01	0.3	**4.46**	0.23	-1.18	**9.44**	**2.41**	-0.22	-0.18	-0.14
P-value	1	<.01	0.76	<.01	0.82	0.24	<.01	0.02	0.83	0.86	0.89
FEEDBACK	-3.98	0	-13.0	-8.66	-7.3	**18.95**	-6.07	-8.06	-3.82	-3.41	-1.99
P-value	<.01	1	<.01	<.01	<.01	<.01	<.01	<.01	<.01	<.01	0.05
OFFCAP	0.3	-11.9	0	**74.13**	-4.62	-6.57	**8.46**	**6.55**	-0.34	-0.28	-0.22
P-value	0.76	<.01	1	<.01	<.01	<.01	<.01	<.01	0.74	0.78	0.83
ONCAP	**3.8**	-16.8	**100.7**	0	-8.55	-11.9	**32.83**	**11.13**	**2.8**	-0.49	-0.39
P-value	<.01	<.01	<.01	1	<.01	<.01	<.01	<.01	0.01	0.62	0.7
PHRASE SEARCH	**3.55**	-2.97	-5.12	-2.66	0	-2.87	**11.43**	-1.49	-0.5	-1.49	-0.16
P-value	<.01	<.01	<.01	0.01	1	<.01	<.01	0.14	0.61	0.14	0.87
REPEAT	0.95	**2.98**	-6.29	-7.41	**13.64**	0	-15.2	0.42	-1.83	-0.63	-1.54
P-value	0.34	<.01	<.01	<.01	<.01	1.0	<.01	0.67	0.07	0.53	0.12
REWIND	1.49	**11.31**	-0.34	**9.16**	**4.5**	-29.9	0	**10.36**	**2.34**	1.37	0.63
P-value	0.14	<.01	0.73	<.01	<.01	<.01	1	<.01	0.02	0.17	0.53
REWIND 1	**2.41**	-10.4	**14.5**	**16.91**	-2.44	-4.97	**17.51**	0	**25.1**	**20.32**	**17.24**
P-value	0.02	<.01	<.01	<.01	0.01	<.01	<.01	1	<.01	<.01	<.01
REWIND 2	**4.35**	-2.65	-0.34	**2.8**	-1.81	-1.83	0.5	**19.44**	0	**70.72**	-0.06
P-value	<.01	0.01	0.74	0.01	0.07	0.07	0.62	<.01	1	<.01	0.96
REWIND 5	-0.18	-4.12	-0.28	-0.49	-0.7	-1.29	**4.36**	**13.45**	**42.41**	0	**43.68**
P-value	0.86	<.01	0.78	0.62	0.49	0.2	<.01	<.01	<.01	1	<.01
REWIND 10	**6.65**	-3.51	-0.23	**4.66**	-0.24	-1.59	6	**3.97**	-0.06	-0.05	0
P-value	<.01	<.01	0.82	<.01	0.81	0.11	<.01	<.01	0.95	0.96	1

For convenience in visualizing the students' learning behavior patterns, the behaviors of rewinding 1 minute or greater were coded as **Rewind 1+** (which includes Rewind 1, Rewind 2, Rewind 5, and Rewind 10), and regarded as a single group of behaviors in the visual map. The patterns were identified and illustrated in Figure 4.2. The arrows in the visual map signify the direction of the sequence and the numbers on each line indicate the Z-values of the sequence.

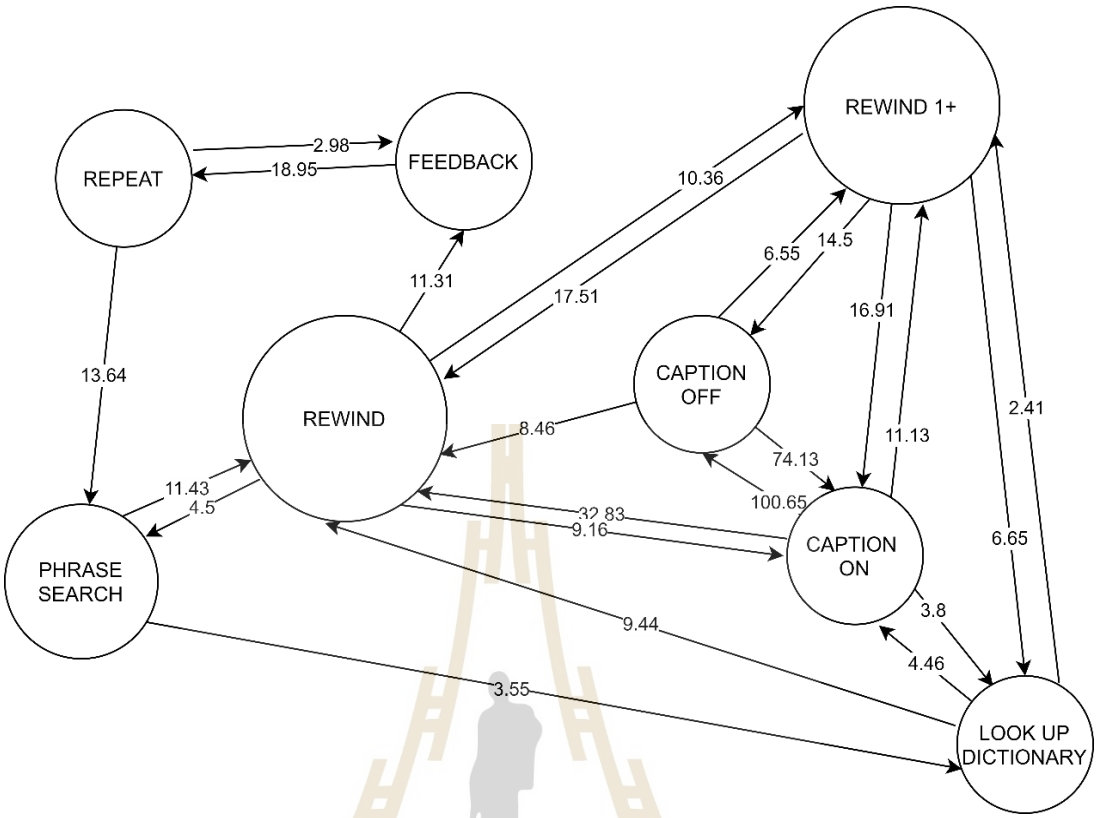


Figure 4.2 Patterns of learning behaviors

As can be seen in Figure 4.2, some behavior sequences were bilateral, such as REPEAT⇆ FEEDBACK, PHRASE SEARCH⇆REWIND, REWIND⇆REWIND 1+, or ONCAP⇆OFFCAP. These two-direction sequences mean that the students often performed the pairs of actions back and forth. For example, the two-direction sequence REPEAT⇆FEEDBACK indicates that the students first read the feedback, then replayed the chunk (Z=18.95). Alternatively, they may consult another feedback (by resubmitting a new transcription) after replaying and rehearing the chunk; however, it was less likely to happen as the probability of the sequence REPEAT→FEEDBACK (Z=2.98) was weaker than the other sequence REPEAT→PHRASE SEARCH (Z=13.64). To put it simple, the above map of learning behavior patterns suggests that after reading feedback, the students often replayed and listened again to the chunk, then they might perform a phrase search rather than immediately submitting a new transcription. Following the trails of probability, the most prominent sequential patterns of learning behaviors, excluding two-direction sequences, were identified as follows.

(1) FEEDBACK→ REPEAT→ PHRASE SEARCH→ REWIND→ FEEDBACK

(2) REWIND 1+→ REWIND→ FEEDBACK

(3) OFFCAP→ ONCAP→ REWIND→ FEEDBACK

(4) ONCAP→ OFFCAP→ REWIND→ FEEDBACK

(5) DICTIONARY→ REWIND→ FEEDBACK

Note that all of the above patterns ended with FEEDBACK, the behavior which involved resubmitting a (revised) transcription and reading another feedback. It is reasonable because the paused transcription tasks required the students to provide a correct transcription as to continue viewing the movies, thus, the ending behavior in most sequences was the one which can help them evaluate their submitted transcriptions and proceed with the movie playback. This was the function of the Submit button which was recorded in the database as FEEDBACK (since every time the students submitted a transcription, the app would generate feedback accordingly). The identified patterns above can be interpreted as closed sequences or loops which always ended with the act of resubmitting and receiving feedback.

4.2.5 Students' reasons for the observed learning behaviors

This section of the chapter is mainly concerned with reporting the reasons underlying the students' learning behaviors that have been explored so far. Data from the open-ended responses in the questionnaire and the semi-structured interviews were the sources of the results. For the sake of clarity, the reasons were categorized and presented according to the corresponding learning behavior. To illustrate each category of the reasons for certain behaviors, quotations extracted from the students' responses were presented. The quotations were selected based on its brevity and representativeness of the information being illustrated.

4.2.5.1 DICTIONARY

Students reported two primary reasons for using the built-in dictionary: searching for information and developing their vocabulary knowledge.

f) Searching for information

One of the students' most common reasons for using the built-in dictionary was to search for reference information, for example, information about word meaning, usage, pronunciation, and word class. Many students reported that they

only searched the dictionary when encountering unfamiliar lexical items in the movies. The following excerpt illustrate this.

> *Excerpt 1: For dictionary, I think when I didn't know the meaning of a word, or I wasn't sure about its meaning, I can type the word in [the search box] to look up that word while watching the movie. I think it would give me more knowledge about the word and its usage in various contexts.*

g) *Expanding linguistic knowledge*

Coupled with information searching, the students also reported that they used the dictionary for expanding their linguistic knowledge. One student maintained that looking up new words in the dictionary would enlarge her vocabulary and benefit her listening comprehension accordingly.

> *Excerpt 2: The built-in dictionary would help me know more words. I can learn vocabulary while learning listening, and I think it would be useful in my learning.*

Some students said that they used the dictionary in their learning since it improved their vocabulary retention. Two students mentioned it in the excerpt below.

> *Excerpt 3: When I encountered an unknown word, I would search for its meaning, so that I can remember that word longer.*

4.2.5.2 FEEDBACK

Feedback was the most common behavior out of the 11 types of learning behaviors on Listening Hacked. Two reasons motivated the students to consult the feedback were evaluating their comprehension and regulating their attention in listening.

a) *Monitoring and evaluating comprehension*

The results show that the most frequent reason for consulting the corrective feedback was to *monitor and evaluate comprehension*. This reason was reported 67 times in the questionnaire and in the interview data. The students

mentioned that they read the feedback in order to assess whether their submitted transcriptions were correct or not, and whether their comprehension of the speech was accurate. The comment below demonstrates this point.

> *Excerpt 4: The feedback showed me where I was wrong, and it also indicated how I have improved in listening. When receiving less and less feedback, I understood that my listening has improved.*

b) Regulating attention in listening

Another important reason for consulting the feedback mentioned by several students was that it helped them regulate their attention in listening. This involved knowing the correct parts of the transcriptions and allocating more listening attention to the parts where problems persisted. The except below illustrates this point.

> *Excerpt 5: (When reading feedback) I could attend to that certain word. For example, if I have recognized the first three words in the chunk, so when I listened again to that chunk, I knew the fourth word was the one I should listen, and so I would have enough attention and preparation for my listening.*

4.2.5.3 ONCAP

Regarding the use of English captions, the students reported four main reasons, i.e., evaluating their comprehension, compensating for their incompetency in word recognition, compensating for lacking unfamiliar contents, and expanding their linguistic knowledge.

a) Evaluate comprehension

The first reason was to evaluate their comprehension of some parts in the movies. This mostly involved using the captions for checking the accuracy of word recognition. Note that the use of English captions was only possible in any parts of the movies where transcriptions were not required. In the excerpt below, a student explained how she used the English captions for assessing her word recognition attempts.

Excerpt 6: I wanted to know what they actually said, and whether I heard accurately. It was the case when I could hear they said something, so I turned on the captions to check whether I was right.

b) <u>Compensating for incompetency in word recognition</u>

Alternatively, some students reported that they used captions for compensating for their incompetency in word recognition. It was the case when they encountered unfamiliar or unknown words and needed the captions for figuring out the words used in the movies.

Excerpt 7: The captions helped me a lot because there were some words I listened but I couldn't figure out what words they were. So when I turned on the captions, it could help me know exactly what the words were.

Sometimes aspects of connected speech such as assimilations or elisions were the causes of the students' incompetency in word recognition, and they stimulated the students' use of captions.

Excerpt 8: When I couldn't follow what words were spoken because native speakers usually link sounds together, captions helped me catch up with the movie and know what the words were.

Like features of connected speech, fast speech or unfamiliar accents can be the causes of the students' failures in word recognition. In these situations, the students likely chose to turn on the captions as they commented below.

Excerpt 9: For some movies containing difficult words or fast conversations, I would watch them with captions, so I could understand quite well the contents of the movies.

c) <u>Compensating for lack of comprehension of the movie contents</u>

The students also reported using captions for compensating for their lack of understanding of movie contents, especially when the contents were unfamiliar.

Excerpt 10: I mainly used captions for watching action movies, but not for animated movies, because I think action movies contain many high level words which are difficult for listening.

d) Expanding linguistic knowledge

Apart from the above compensation purposes, the students also reported using captions for expanding their linguistic knowledge, notably vocabulary and pronunciation. In the excerpt below, a student maintained that he used captions for learning new vocabulary and pronunciation of some interested words.

Excerpt 11: For me, using captions helped increase my vocabulary range and it is also good for my writing because many action movies, for example, or documentaries, contain lots of advanced words.

4.2.5.4 OFFCAP

Contrasting with using captions, several students chose to turn off the captions in their learning. The reason underlying this behavior was that they wanted to *direct their attention* in the listening process and avoided reading the caption texts. In the excerpts below, the students explained that they wanted to focus on the process of meaning making in listening.

Excerpt 12: In my opinion, when using the captions, we only focus on reading the texts but not on listening and understanding what the characters are saying. I think if we focus on what the people are saying, we can understand the whole story. But if we read, we only recognize how some words are pronounced, and that does not help our listening at all.

Captions were at times thought as a distractor which may hinder the listening process as a student explained in the excerpt below.

Excerpt 13: I turned it off because I didn't want to be dependent on the captions... when my eyes were always on the captions... and I couldn't focus on listening at all. I

wanted to understand the contents of what they were saying in the movies.

4.2.5.5 PHRASE SEARCH

Performing phrase searches was a common behavior and accounted for 15.37% of the total students' interactions with Listening Hacked. The students reported three major reasons for this behavior: finding support for word recognition, compensating for unclear speech signal, and expanding their linguistic knowledge.

a) Seeking assistance in word recognition

The most frequently mentioned reason for conducting phrase searches was to seek assistance in word recognition. The students often used Phrase Search when they wanted to have more speaker and context variability which may help them in word recognition. This happened when they experienced failures in word recognition after several attempts of repeated listening. The following excerpt demonstrates this reason.

Excerpt 14: In my opinion, this is my favorite function. Because in many movies the characters spoke in ways that I couldn't get anything at all, even when I tried listening again. In those contexts, the characters intentionally spoke very fast, and it was impossible to get it. But in other contexts in other movies, other characters may speak slowly. I found it very useful, and it helped me recognize the words which I missed earlier.

Likewise, when the students faced problems in word recognition due to fast speech or unfamiliar accents, they were likely to use Phrase Search to get speaker and context variability. The following excerpt contains the students' description of this purpose for conducting phrase searches on Listening Hacked.

Excerpt 15: I think it (using phrase search) provided me with various accents, various contexts, various ways of speaking... sometimes in other movies they may speak the same utterance a little more slowly or a little faster. So,

if we can listen with different speeds of delivery, we can recognize words more easily.

Interestingly, one student reported a quite unexpected reason for her use of Phrase Search, i.e. to help with prediction. Nevertheless, this reason is also related to seeking variability of contexts for word recognition.

Excerpt 16: Sometimes when I couldn't hear a word, I clicked on the asterisk, then it took me to a different movie, and I can guess the meaning of that word in that movie... then I can guess that word.

b) Compensating for unclear speech

Another common reason for using Phrase Search was to compensate for unclear speeches. This was the case when the actors' voices were too soft, or when there was much background noise or overlapping voices.

Excerpt 17: Sometimes the actor's pronunciation in the present movie wasn't clear or there was background noise, viewing in other movies would make word recognition easier.

c) Expanding linguistic knowledge

In addition to the support for word recognition, the students also mentioned that they used Phrase Search to expand their knowledge of words and their usage in context. In the excerpt below, the student described that she used Phrase Search in order to learn common English phrases and collocations.

Excerpt 18: It helped me know which words often go together... the words often go together which I usually couldn't recognize.

Similarly, another student insisted that using Phrase Search can help her learn word usage in context.

Excerpt 19: A very good thing was that the missed word can be listened again in other movies, the same word but in different positions. It would help me know more about the usage of a word.

4.2.5.6 REPEAT

Repeat is the act of replaying the chunk being transcribed in order to listen again to that chunk. It was found the most frequent learning behavior of most students. Overall, the students reported six main reasons for this behavior: rehearing the missed information, compensating for limited working memory capacity, compensating for unclear speech, seeking support for word recognition, regulating attention in listening, and developing their linguistic knowledge.

a) Rehearing the missed information

Most students said that they repeated the chunk simply because they had missed some information in the chunk in the first-time listening and thus wanted to rehear it.

> *Excerpt 20: There were some parts which were very difficult to listen...which I couldn't get it for the first time listening. So I used Repeat button 2-3 times, and then I could recognize the words being said.*

Some students believed that repeated listening could facilitate the process of making sense of the attended speech. In the excerpt below, one student explained this idea.

> *Excerpt 21: When listening again and again to a part, I could easily focus on and understand what the character wanted to say. This was impossible when I could listen for it once only.*

b) Compensating for limited working memory capacity

Another reason for the students' use of the Repeat function was to compensate for their limited capacity of working memory. The students told us that the Repeat function was used when they had to transcribe long utterances; that was when they wanted to repeat the chunk several times to get the whole chunk by allocating their attention to different parts of the chunk once at a time.

> *Excerpt 22: There were some long utterances I couldn't get all the words if I just listened for one time. But the*

more times I used Repeat, the more I could recognize those words.

c) Compensating for unclear speech

Conversations in movies generally contain background noise, sound effects, music, and overlapping voices. Those are common elements in most movies that may produce unclear speeches for the listener. Therefore, some students said that they needed to repeat some chunks due to the vagueness in the speech signal. Here is a student's mention of this reason.

Excerpt 23: For example, there was a scene in a movie when the character was speaking from a very far distance... so the voice was very soft...so if using Repeat I could pay more attention to that voice and I could get it.

d) Seeking support for word recognition

When having problems completing the transcription tasks, mostly due to problems in word recognition, the students first turned to the Repeat function in seeking support for their word recognition. One student described this in her interview response as follows.

Excerpt 24: It (Repeat) helped me listen again to the words I missed, or the words I couldn't follow because they were spoken too fast, or because they were hissing sounds or silent sounds spoken very fast in speech...that I couldn't hear it. So if I could repeat it many times, I could hear it and I felt that it got into my ears more easily.

Similarly, when encountering contractions or features of connected speech such as linking sounds, elisions, assimilations, the students usually chose to repeat the chunks in order to recognize the confusing words or sounds.

Excerpt 25: I think I listened to an utterance over and over for many times to recognize words like contractions, and new vocabulary, and native speaker's pronunciation of those words.

Some students claimed that repeated listening helped them get used to unfamiliar accents, and consequently assisted them with word recognition. The following excerpt illustrates this point.

> *Excerpt 26: Actually, I didn't always succeed with only one time listening, instead I listened again and again for many times. I think the nature of repetition would be beneficial to me...when I listened repeatedly for many times, I get used to the pronunciation and the way people speak.*

e) Regulating attention

As mentioned earlier, the students sometimes wanted to allocate their attention in listening due to the limited capacity of working memory. However, they may need to allocate their listening attention to different parts of the chunk even when the utterances were not long. This was the case when they realized some problems in their transcriptions, and thus they only wanted to focus their attention to those areas instead of the whole utterance. A student described this situation in the following excerpt.

> *Excerpt 27: I think it would be better if I just focused more on the parts which I couldn't hear them well, and it also saved my time... so I chose to use it (Repeat).*

f) Expanding linguistic knowledge

Occasionally, the students would replay the chunks for several times in order to learn vocabulary or pronunciation. Some students maintained that by listening repeatedly to an utterance, they can imitate and acquire the native speaker's intonation and pronunciation.

> *Excerpt 28: It helped me know how the words should be pronounced, and it helped me practice imitating the actors' voices appropriately.*

Some students also reported using the Repeat function for learning vocabulary and remembering the words longer.

Excerpt 29: Listening again helped my brain remember the vocabulary, the pronunciation better, and also the intonations in different contexts.

4.2.5.7 REWIND

The Rewind function allowed the students to replay the chunk being transcribed with an addition of 2 seconds before and after that chunk. To some extents, Rewind was like the Repeat function in the way that it allowed repeated listening of the current chunk. Therefore, the students who used Rewind also reported similar reasons for using it like those who used Repeat. However, they noted one important additional reason for this behavior that was to *seek contextual information for listening comprehension*. In the following excerpt, a student explained why he decided to use Rewind instead of repeating the chunk.

Excerpt 30: In some cases, if I only listened to the current chunk and they (the actors) spoke too fast, I couldn't understand it. So, if I could listen before and after the chunk, I could understand what they were saying before it and after it… then I could guess what they were saying in the chunk… It gave me more context and information.

4.2.5.8 REWIND 1+

This group of behaviors involves replaying the movie 1 minute, 2 minutes, 5 minutes, or 10 minutes from the current chunk being transcribed. The students did not usually do this because it took time, however, generally when they decided to rewind 1 minute or more, they had the same reasons as using the standard Rewind. The only difference was that they wanted to get even more contextual information for listening. In the following excerpts, a student described the situations when she would prefer using Rewind 1+.

Excerpt 31: It depends on the situations. When I can listen well enough, I just need to use Repeat. But when I couldn't get anything at all, I would replay the movie step by step to understand the situation better… that would help me understand the utterance better.

4.3 The students' opinions of Listening Hacked

In addition to the investigation into the effectiveness of Listening Hacked, the present research was also concerned with the students' opinions of the usefulness of the app and its various features. For this purpose, Research Question 3 was posed as follows: *What were the students' opinions of the usefulness of Listening Hacked and its features in the development of their EFL listening comprehension?*

In answering this research question, the data from the questionnaire with five-level Likert scale were obtained and analyzed with SPSS version 23. The qualitative data from the open-ended items and the interviews were also utilized to explicate the rating results. The results regarding the students' opinions will be presented separately in two sections, one reporting their opinions about each feature (Section 4.3.1) and the other reporting their overall opinions about the app Listening Hacked (Section 4.3.2).

4.3.1 Students' opinions about the features of Listening Hacked

Table 4.40 reports the results of the students' ratings for the surveyed features of Listening Hacked. The 5-level scale was interpreted as follows.

1.0 means the feature was thought to be completely useless; 2.0 means the feature was useless; 3.0 means the students were unsure about the usefulness of the feature; 4.0 means the feature was useful; 5.0 means the feature was very useful.

Note that the data yielded from the questionnaire were ordinal and categorical in nature; therefore, together with reporting the means, the results were also reported in percentages to indicate frequencies.

Table 4.40 Students' opinions about features of Listening Hacked

Feature	Mean	Completely useless (%)	Useless (%)	Unsure (%)	Useful (%)	Very Useful (%)
Listening Boosters	4.03	0.00	0.00	40.00	16.67	43.33
List of skipped items	4.03	3.33	6.67	13.33	36.67	40.00
Built-in dictionary	3.67	0.00	6.67	40.00	33.33	20.00
Phrase search	4.53	0.00	3.33	10.00	16.67	70.00
Repeated listening	4.73	0.00	0.00	0.00	26.67	73.33
Feedback	4.10	0.00	3.33	20.00	40.00	36.67
Automatic grading	3.87	0.00	3.33	30.00	43.33	23.33
Paused transcription tasks	4.40	0.00	0.00	6.67	46.67	46.67
English captions	4.57	0.00	0.00	3.33	36.67	60.00
Viewing movies	4.30	0.00	0.00	3.33	63.33	33.33

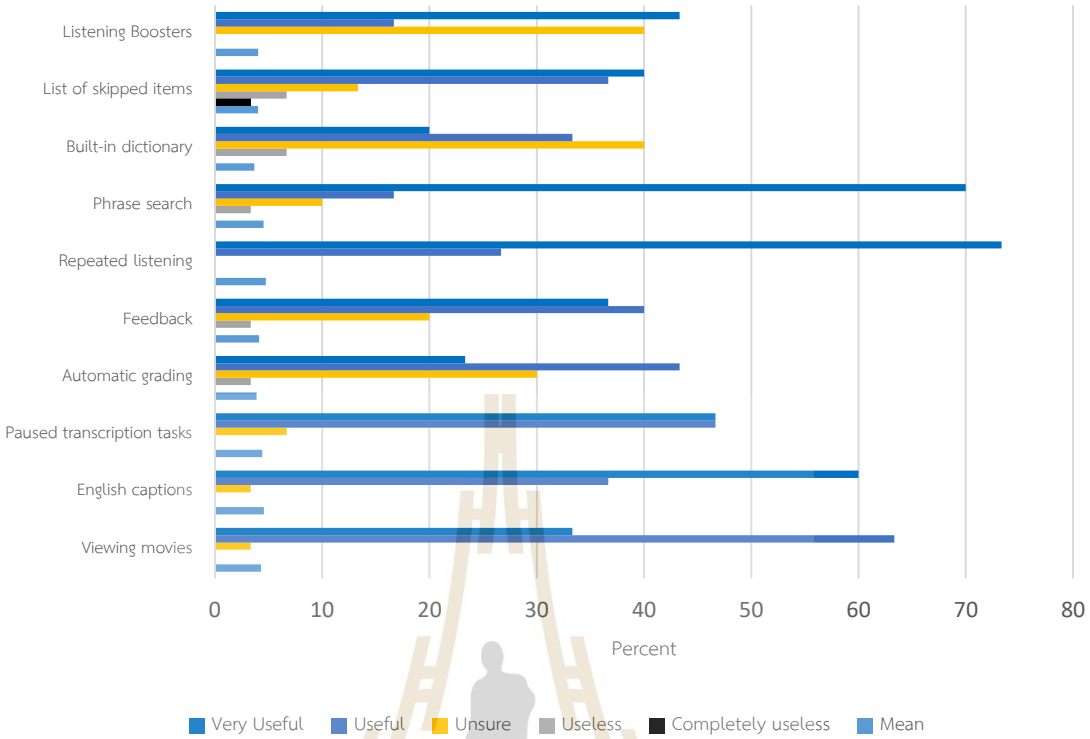


Figure 4.3 Students' opinions about features of Listening Hacked

4.3.1.1 Repeated listening

As can be seen in Table 4.40 and Figure 4.3, 8 out of 10 surveyed features received positive opinions from the students (M>4.0; N=30). Among those features, ***repeated listening*** was considered the most useful feature. 73.33% of the informants believed that this feature was *very useful* for their listening improvement; 26.67% of them said that it was *useful* for their listening improvement, and none had any doubts about its usefulness. A student described its importance in EFL listening learning as follows.

> *Excerpt 32: In my viewpoint, Repeat is very important because learning English needs many repetitions, or we will forget everything. And I think repetition will engrave things into my brain, my memory, and so it is good for listening. Actually it is also good for other skills.*

4.3.1.2 Viewing movies

In the second place, 96.66% of the informants said ***viewing movies*** helped them in EFL listening and only one student was unsure about its usefulness. Most of them agreed that movies helped them learn while entertaining, and that viewing movies was good for their listening skills as well as expanding linguistic knowledge. Here is a student's comment on this.

> *Excerpt 33: This feature helped improve my listening a lot. From the one who needed captions for watching movies, now I can watch movies without any captions.*

4.3.1.3 Turning on/off English captions

The feature that allowed the students to ***turn on and off English captions*** received similar results to that of viewing movies, with 96.67% of the students believing it was useful or very useful, and only 3.33% expressing uncertainty. In addition to the reasons for using or not using English captions as mentioned previously in Section 4.2.5, the students also emphasized its strength in personalizing their learning experience as shown in the excerpt below.

> *Excerpt 34: Being able to turn on or hide captions helped me adjust with my listening ability.*

4.3.1.4 Paused transcription tasks

Regarding the ***paused transcription tasks*** used in all practice sessions on Listening Hacked, 93.33% of the students said the tasks were useful (46.67%) or very useful (46.67%) for their learning; only 6.67% (2 out of 30 informants) said that they were not sure about its usefulness. In general, the students reported two main benefits of paused transcription tasks: (a) it created learning opportunities and (b) it helped them focus on listening comprehension. Firstly, the students believed that paused transcription tasks helped them notice the input and then learned something from it. Without doing the tasks, they may not notice and learn anything. Here are some students' mentions of this benefit.

> *Excerpt 35: Chunks are an important element for learners to practice listening while viewing movies. Without having*

chunks, the learners will easily miss out utterances and new vocabulary which they should learn.

Excerpt 36: In my opinion, when the transcription tasks extract 10 chunks from a movie session, it can evaluate better how accurately I could listen to an utterance than just listening to what was in the movie. Especially, those chunks also helped me practice word recognition skills.

Another benefit of paused transcription tasks mentioned by several students was that the tasks helped them focus on listening comprehension and the process of meaning-making.

Excerpt 37: When there are chunks, we need to focus more on viewing the movies because we do not just view the movies the way we like, but we need to view the movies in order to understand the contents so that we can do the next chunks.

Some students, on the other hand, reported some frustrations towards the operations of paused transcription tasks, primarily related to technical issues. In a paused transcription task, the app automatically selected 10 chunks from a movie session and asked the students to provide the correct transcriptions of those selected chunks. The app was programmed to eliminate interjection words from the required transcriptions of any selected chunk; however, the list of English interjections as programmed in the app was no way near exhaustive. Thus, there were some chunks selected with strange interjections and that made the students' tasks problematic.

Excerpt 38: The selection of the chunks was not reasonable, although I knew they were randomly selected by AI. For example, words like "yeah", "well", "no no"... are not qualified enough as chunks that the listeners must listen.

4.3.1.5 Phrase search

In the fourth place, ***Phrase search*** was thought to be very useful or useful by 70% and 16.67% of the students, respectively. Most of the students' positive opinions about this feature were in line with their reasons for using it, as presented in

Section 4.2.5. Here is an additional comment in which a student explained how Phrase Search assisted in his learning.

> *Excerpt 39: I think it can help quite a lot because with it I can listen to an utterance in different situations and different intonations, such as US and UK.*

However, there were 17.86% of the students who were either unsure about its usefulness or thought that Phrase Search was useless to some extents. The main cause of this attitude might be due to the limited resources of Listening Hacked which affected the operation and efficiency of Phrase Search. A student said in the interview that Phrase Search sometimes could not return any results, simply because no other movies in the database contained the searched words or phrases.

> *Excerpt 40: One drawback of Phrase Search was that it depends on the movie data, so there were some words which were not available in other movies.*

4.3.1.6 Built-in dictionary

On the other side of the coin, the ***built-in dictionary*** received the least positive opinions from the students, with only 53.33% of the informants believing it was useful or very useful for their listening learning. A large proportion of the students (40%) were unsure about its usefulness, probably because they thought that the pop-up results from the dictionary interfered with their learning. The excerpts below illustrate this point.

> *Excerpt 41: When using the dictionary, a new window would appear. That could interfere with viewing the movie and transcribing the chunk.*
>
> *Excerpt 42: I didn't usually use the dictionary but after a few times using it, I found it a bit time-consuming to access the main page of the dictionary.*

Nevertheless, well over half of the informants still held positive opinions towards the built-in dictionary. Here is a student's mention of the usefulness of this feature.

Excerpt 43: I found the dictionary very useful. When I searched for the word meaning, the dictionary also gave me international phonetic transcriptions, both US and UK. The definitions were written with simple English words so that I could understand. Besides, there were also word usage, and synonyms, antonyms, verbs, adjectives of that word.

4.3.1.7 Listening Boosters

For ***Listening Boosters***, a feature designed for personalized follow-up listening practice, 60% of the students thought that it was useful or very useful, but a relatively significant proportion 40% of them expressed uncertainty about its usefulness. In the two following excerpts, one illustrates the students' positive opinions about this feature and the other contains some doubts on its effectiveness.

Excerpt 44: The feature for following-up practice helped me listening again several times to the words that I missed in the movies, helped me fix my pronunciation mistakes, and helped me remember vocabulary more effectively.

Excerpt 45: I used it for a few times, but then I found it quite ineffective, so I stopped using it.

4.3.1.8 List of skipped items

The ***List of skipped items*** received generally positive opinions from the students (76.67%), however, it was also the feature receiving the highest rate of negative opinions among all features, with 3.33% saying it was completely useless, 6.67% saying it was useless, and 13.33% being unsure about its usefulness. Those who believed it were useful mentioned the benefit of regulating and keeping track of their learning.

Excerpt 46: With this feature, I could review and give attention to the words, and take notes for every time I listened.

For those who had negative opinions about the feature List of skipped items, it is probably because they only had trivial mistakes in the transcription

tasks, therefore their list of skipped items possibly contained unimportant words; as a result, they did not benefit much from the information in the list. A student explained why he had negative opinions about this feature as follows.

> *Excerpt 47: I think this feature was normal because it did not work well for me. Most of my skipped items were onomatopoeic words and simple vocabulary which I made mistakes in transcriptions due to my carelessness.*

4.3.1.9 Automatic grading

Regarding ***Automatic grading***, 66.67% of the students thought that it was useful or very useful for their learning. The students expressed two main positive opinions about this feature: tracking their learning progress and creating motivation in learning.

> *Excerpt 48: Automatic grading let me know the result of the session I just finished to know whether I did well on it or not, and whether I made any progress. Especially, it helped motivate my listening.*

However, 3.33% of them said automatic grading was useless, and 30% were skeptical about its benefits. One student explained that the feature of grading may give her motivation in learning, but it may also demotivate her when the grades did not satisfy her expectations.

> *Excerpt 49: The grading basically affected the listener's emotions. If the score was high, it motivated me to do other practice sessions; but if it was low, I may get frustrated a little bit.*

Another student asserted that automatic grading did not assess her listening ability accurately.

> *Excerpt 50: Since the grading was based on 10 chunks of the session, it couldn't assess the whole listening process.*

4.3.1.10 Corrective feedback

Regarding the ***corrective feedback*** which was automatically generated by Listening Hacked, 40% of the students rated it as useful, and 36.67%

thought that it was very useful. Most of the students' positive opinions about this feature point to their reasons for using it, as mentioned earlier in Section 4.2.5. Here are some students' comments on the usefulness of corrective feedback.

Excerpt 51: It helped me know the words that I misheard and saved time on correcting those mistakes.

While the majority of students found feedback useful for their learning, a few students were not happy with it. In the following excerpt, a student described a technical issue of this feature which made it ineffective. This happened when the chunk to be transcribed had words that are identical in pronunciation, as in homonymy or polysemy.

Excerpt 52: If the chunk had two identical words, the feedback only showed the correct order of the first word in the chunk. The number of identical words and their correct order in the chunk were unknown.

Another student commented that the corrective feedback was not useful because it lacked details; unfortunately, she could not recall what kind of details should be added.

Excerpt 53: Since the feedback was not detailed enough, it did not help much.

4.3.2 Students' overall opinions about Listening Hacked

Table 4.41 and Figure 4.4 present the students' overall opinion of Listening Hacked in improving their EFL listening comprehension.

Table 4.41 Students' overall opinion of Listening Hacked

	Mean	Very dissatisfied (%)	Dissatisfied (%)	Neutral (%)	Satisfied (%)	Very satisfied (%)
Overall opinion	4.27	0	0	6.67	60.00	33.33

The results show that 93.33% of the students participated in the Experimental Group expressed high levels of satisfaction towards Listening Hacked; only 6.67% (2 students) felt uninterested in using this app. The qualitative data further

reveal the reasons for their opinions, which were classified as either positive opinions or negative opinions.

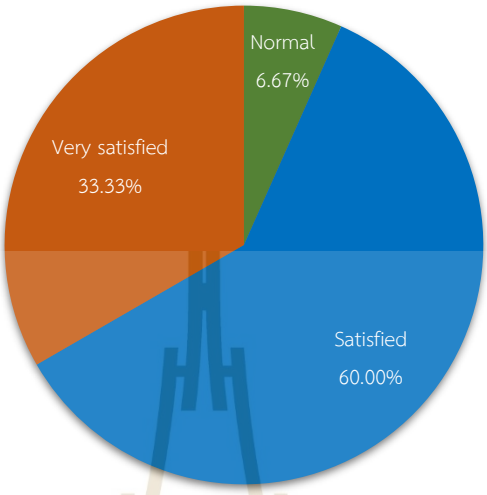


Figure 4.4 Students' overall opinion of Listening Hacked

4.3.2.1 Positive opinions about Listening Hacked

The students' positive opinions about Listening Hacked can be summarized in four categories, as shown in Figure 4.5.

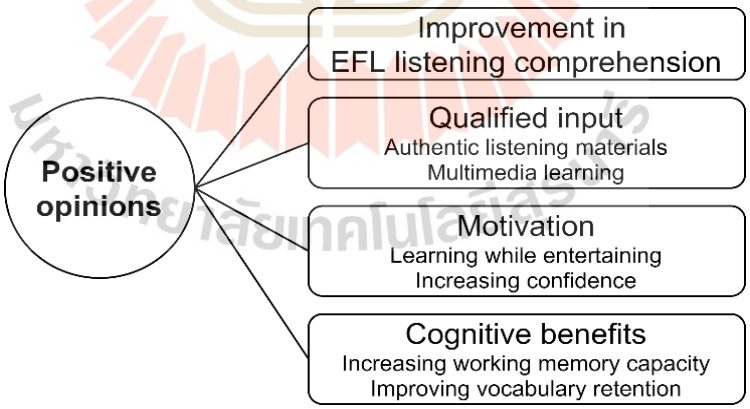


Figure 4.5 Students' positive opinions about Listening Hacked

a) Improvement in EFL listening comprehension

Most students believed that learning on Listening Hacked has contributed to their improvement in EFL listening comprehension. Below are some

students' comments on the effectiveness of Listening Hacked on their EFL listening development.

> *Excerpt 54: At the beginning, my level of listening was just a bit above the average, but after the course, I think that my level has improved a lot.*

b) Qualified input

The use of movies for EFL listening also received positive opinions from most students. They believed that movies were a source of quality, authentic input for their listening practice. Moreover, movies are a type of multimedia input which made their learning more effective. The following excerpt demonstrates the point.

> *Excerpt 55: I think learning English through movies is very useful and effective. We can listen to what native speakers use in real life, recognize words and structures that they use, and how they pronounce the words. And regarding learning while entertaining, the app Listening Hacked has done a good job.*

c) Motivation

They also said that their learning on Listening Hacked was motivated since the main learning activities were viewing movies which is a form of entertainment. Here is a student's comment on the motivational aspect of learning listening on Listening Hacked.

> *Excerpt 56: I really like watching movies and learning English. So it was wonderful when I could watch movies while learning English. I have learned phrases used in daily conversations that I have never learned in textbooks, such as: So comfy, It's on me, etc. Although I might not have progressed substantially, I have improved since the beginning of the project.*

In addition, one student asserted that she has become more confident in listening thanks to her learning on Listening Hacked.

Excerpt 57: This course really helped improve my listening skills a lot. I feel a little more confident in listening now.

d) Cognitive benefits

Some cognitive benefits resulting from the learning on Listening Hacked were also reported. The students maintained that they could remember new words more effectively and longer thanks to various features of Listening Hacked. These benefits have been mentioned in Excerpts 4, 5, 44, 45 (see Section 4.2.5). Additionally, one student affirmed that he could process and comprehend longer utterances which he could not prior to learning on Listening Hacked.

Excerpt *58: My listening comprehension has improved somehow. At the beginning, I could not understand too long utterances, but now I can understand some.*

4.3.2.2 Negative opinions about Listening Hacked

Despite having mostly positive opinions about the app, the students also expressed some concerns about the app which are summarized and presented in Figure 4.6.

a) Technical issues

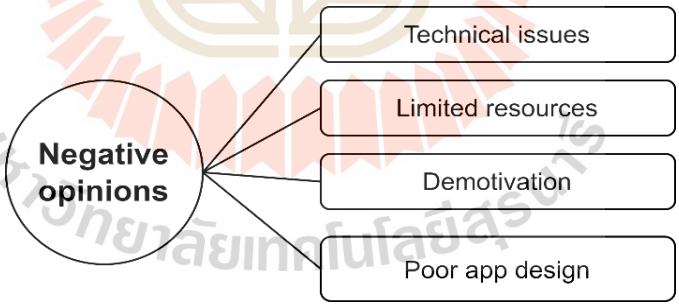


Figure 4.6 Students' negative opinions about Listening Hacked

Most of the students' frustrations when using Listening Hacked involved certain technical issues (14 mentions). One of them was the mechanism of chunk selection. As mentioned earlier, this operation was automated and consulted a pre-defined list of unnecessary words to determine whether a word in a selected chunk should be displayed or not. The list was created by the researcher with his best effort

to include as many as possible all the English interjection words, onomatopoeic words, proper names so that those words could not interfere with the transcription tasks. Nevertheless, since the list was not exhaustive, there were cases when unnecessary words were selected for the transcription tasks, and that made the transcription tasks become unreasonably difficult.

> *Excerpt 59: There were some onomatopoeic words that the characters often say, such as ooh, aah, whoa, etc. There were several chunks with two or three of those words; I mistranscribed those and couldn't know how they should be properly transcribed, so I had to give up the points.*

Another reported technical problem was the errors in the original transcript texts. The app generated feedback according to the matching mechanism based on the students' submitted transcription and the original transcript texts, thus, errors in the original transcript texts such as typos or alternative spellings of certain words would result in inaccurate feedback. Additionally, the app was reported to malfunction or crash when not running on Microsoft Edge or Google Chrome web browsers, plus the server was not responsive during high usage time, but these problems did not happen frequently.

> *Excerpt 60: But there were some system failures, for example, the website was inaccessible, or the feedback was inaccurate due to some words typed without spacing, wrong spellings, …*

b) <u>Limited resources</u>

Another source of negative opinions came from the limited quantity of the movies in the app's movie data. A few students informed us that they expected the movie data would be larger and more varied in genres. At the end of the research project, the movie data contained 257 movie titles, with 932 movie sessions; that approximately equals to 466 hours of viewing. Below is a student's mention of this expectation.

Excerpt 61: Regarding the movies, the number of movies on Listening Hacked was not anything really substantial. I hope that Listening Hacked will update many more movies, more genres, for example TV shows or TV series…

Indeed, the limited quantity of movies may also affect the operation of other features, e.g. Phrase Search and Listening Boosters. As already shown in Excerpt 66, the limited movie data could result in the unfruitful operation of Phrase Search.

c) Demotivation

Opposing to the entertaining aspects of Listening Hacked, some students reported difficulties in doing the transcription tasks when they encountered difficult chunks. A student explained that too difficult chunks could demotivate her learning.

Excerpt 62: Since my starting level of listening was not good, there were chunks that I listened again and again but still failed to get them; so I sometimes got frustrated and just wanted to skip them.

d) Poor app design

The students also complained about the design of the app, notably the design of the built-in dictionary. They maintained that while the dictionary was necessary and useful, the display of the search results was not user-friendly and interfered with viewing the movies. This was understandable because the dictionary was programed to return search results in a pop-up window. Accordingly, in order to view the results, the students had to switch from the main working window to the result window. The following excerpt contains a student's explanation for not using the dictionary.

Excerpt 63: When using the dictionary, a new window would appear. That could interfere with viewing the movie and transcribing the chunk.

In addition to the design of the dictionary, some students also mentioned that they were unsatisfied with the small size of the movie player screen. They preferred watching the movies in full-screen mode.

> *Excerpt 64: Regarding the screen of the movie playback, it was quite small (probably covering only 1/9 of the laptop screen), this was quite annoying and made the viewers unsatisfied, and so could affect the learning. It would be wonderful if full-screen feature was possible.*

4.4 Chapter summary

The chapter reports on the results of the present research by presenting the data and analyses in response to the three research questions posed in Chapter 1. In summary, the results show that the students who learned on Listening Hacked have improved their EFL listening ability and outperformed their counterparts who did not use the app in their learning. The results of behavioral analyses also reveal important learning patterns which were closely related to the students' reasons for using or not using certain features of the app. Finally, the results from the questionnaire and interview data suggest that the students preferred learning on Listening Hacked and that they were satisfied with the proper assistance for their learning from the various features of the app. The next chapter will discuss the interpretations and implications of those findings.

CHAPTER 5
DISCUSSIONS

This chapter discusses possible interpretations of the results presented in Chapter 4. It starts with the discussions about the effectiveness of the proposed learning system. The students' learning behaviors on Listening Hacked and their opinions expressed in the questionnaire and interviews are also discussed to enrich our understandings of the processes and strategies involved in the learning process on a technology-based, autonomous learning environment.

5.1 Improvement in EFL listening comprehension

The present research set out to investigate the effects of the proposed learning system, namely Listening Hacked, on the students' EFL listening comprehension. The results show that after 12 weeks learning on Listening Hacked, the students in the Experimental Group improved their EFL listening performance significantly, with a relatively large effect size (d=.76), while the students in the Control Group who learned EFL listening with the traditional method did not show evidence of improvement. In fact, there were 30.43% of the students in the Control Group deteriorating in their posttest performance. The improvement and outperformance in EFL listening of the Experimental Group suggest that the app Listening Hacked was effective in assisting the students in their listening learning. The effectiveness of Listening Hacked could be possibly attributed to the students' development in perceptual processing, learner autonomy, and increased motivation in learning EFL listening.

5.1.1 Improvement in perceptual processing

Previous studies show that improvement in perceptual processing could lead to better comprehension in L2 listening (Field, 2003; Jia & Hew, 2019; Leonard, 2019; Matthews & O'Toole, 2015). The present study seems to suggest similar

conclusions. In this study, the students' improved performance in paused transcription tasks as evidenced by the significant increases in the task performance scores (TPS) implies that their word recognition ability has improved and thus led to better comprehension in EFL listening. Field (2003) asserts that many problems in listening, including those related to higher level processes, are in fact rooted in problems of perceptual decoding. Without adequate recognizable perceptual information, top-down information would become useless, and the speech would become unintelligible. In a study examining learners' problems in L2 listening, Goh (2000) found that several problems are related to the learners' inability to recognize words. For example, they could not recall the meanings of words in speech although they said that those words were familiar to them. This happens because the learners could not establish the link between the spoken and the written forms of the words. This link is established through the same linguistic representations to which both spoken and written forms need to contact in speech recognition (Johnsrude & Buchsbaum, 2017). While the connection between the written forms and their corresponding representations in the brain is normally robust, that between the spoken forms and the linguistic representations may not always be easily established since the spoken form of a word have multiple variants. This explains why one may find some words familiar in some cases and the same words unrecognizable in speech in other cases. Regarding this aspect, the paused transcription tasks in this study directly addressed this issue. The tasks provided the students with opportunities to practice establishing the link between written and spoken forms of words and overcoming the problems of variability in the speech signal. As a result, the students' perceptual decoding ability improved and their listening comprehension followed.

Attention problems in listening could also stem from problems in word recognition. For example, some learners in Goh's (2000) study said that they missed the next part of the text when spending time thinking about the meanings of certain words. Due to the limited capacity of working memory (Baddeley, 1992, 2012), more attention allotted to perceptual processing means that less attentional resources are

spared for other processes, and that would also affect perceptual processing of subsequent speech signal. According to the view that working memory is a dual function system consisting of processing and storage (Daneman, 1991; Finardi, 2006), when processing function does not take much space, more space is available for storage, and the limited capacity of working memory is well used to process and store more incoming information. In other words, it could be said that the study helped to improve the students' working memory by reducing the negative impacts of the trade-off between perceptual processing efficacy and allocation of attentional resources. Therefore, that the students in the present study improved their perceptual processing or word recognition ability could solve their attention problems in listening and positively affect their general EFL listening ability is sensible.

5.1.2 Development in learner autonomy

The improvement in EFL listening comprehension could also be explained by the development in learner autonomy. While autonomy was not measured explicitly in this study, it is an obvious part of learning on Listening Hacked which is an autonomous learning system. Unlike the Control Group studying with the traditional method, the students in the Experimental Group had to decide almost everything in their learning, including choosing the appropriate listening materials, developing a set of strategies for using help options, and working on their own listening problems. Those decisions and learning behaviors were definitely contributing to their autonomy in learning as the construct learner autonomy is defined as the ability to take responsibility for one's learning (Holec, 1981). Research shows that learner autonomy is actually correlated with EFL proficiency (Dafei, 2007) and EFL listening comprehension ability (Safari & Tabatabaei, 2016). Therefore, the students' development in learner autonomy could also lead to better learning, hence better EFL listening ability.

5.1.3 Increased motivation in L2 listening learning

Another possible explanation for the students' improvement in EFL listening comprehension is that they were motivated in learning. In the L2 literature, there is a

consensus that motivation can affect L2 learning (Zoltán Dörnyei, 1998) and L2 listening comprehension (Chon & Shin, 2019; Goctu, 2017; Mohammad Jafari, 2010). Thanks to the entertaining aspect of movie viewing, the students informed in the questionnaire and the interviews that they were motivated to learn on Listening Hacked because they enjoyed viewing movies while improving their listening skills as well as expanding their linguistic knowledge. This is considered as the intrinsic motivation which encourages people to do things "for its inherent satisfactions rather than for some separable consequence" (Ryan & Deci, 2000, p. 56). Likewise, the students said that they were also motivated in learning as their performance in the practice sessions improved, which was indicated by the increases in TPS. This latter type of motivation is more external than the former, with the increases in TPS being the rewards. All in all, with increased motivation, both intrinsic and extrinsic sources, learning on Listening Hacked was likely to happen effectively and that would accordingly contribute to the students' listening skill development.

5.2 The slowing down of listening development

The repeated measures of ANOVA on the task performance scores (TPS) over the two time periods in the experiment reveals some interesting facts: the TPS in Week 7 increased 15% compared to that in Week 2, while the TPS in Week 12 was only 3.8% higher than that Week 7. This indicates that the students' improvement in EFL listening ability was slowed down with time. One possibility is that the learning system has its own limits, and hence it might not be capable of helping the students grow further once they reached the threshold. This could happen when they used out the available help options in the system and none of those could bring necessary support for them to overcome certain listening problems. In this case, it is important that the students' performance data should be obtained and analyzed so that new functionalities can be devised accordingly.

Another possible explanation is that higher levels of listening proficiency may require longer time of practice than lower levels. That is to say, an increase of 2.0 TPS from 2.0 to 4.0 could be less effortful and take much less time than that of 2.0 TPS from 8.0 to 10 (the maximum TPS). According to a Cambridge University Press report, to achieve A1 level in the CEFR, an adult learner at beginner's level would need to spend 90 to 100 hours on guided learning, whereas to reach C1 level from B2 level, they should spend 200 to 300 hours of learning, and an additional of 300 to 400 hours for achieving C2 level (Knight, 2018). This means that the speed of improvement did not slow down, however, higher levels of listening proficiency encompass more requirements, both in linguistic knowledge and processing skills, and thus they simply demanded more time. If it was the case, the implication here is that the learning system should be frequently updated to provide adequate and various learning resources for extensive listening practice in order to ensure the students' continuing development.

5.3 Students' learning behaviors on Listening Hacked

Along with the investigation into the effectiveness of Listening Hacked, this study has an explanatory orientation in that it explored the behaviors which the students exhibited. In fact, it was the students' behaviors of using the features and help options of the learning system that determined their gain in EFL listening ability. The decisions and reasons for those behaviors are considered as the strategies in this study. Therefore, this study also examined the strategies which the students employed in using the app's affordances and resources as to conveniently regulate their learning. This section is devoted to exploring those behaviors and strategies in a few dimensions to offer insights into the processes involved in the learning on Listening Hacked which subsequently led to their EFL listening improvement.

5.3.1 Strategies for using corrective feedback

The frequencies of the students' learning behaviors on the learning system indicate both levels of engagement and degrees of trust that the students gave to

certain features or help options of the learning system in assisting with their learning. The frequencies, however, are not indicative of the isolated effects of certain behaviors on the gain in listening comprehension. Put another way, the behavior with the highest frequency does not necessarily point to a correspondingly high contribution to the students' improvement in EFL listening comprehension. Having said that, the frequencies of the learning behaviors when being considered together with other aspects can describe the possible roles of particular behavior in their learning.

5.3.1.1 Frequency of using corrective feedback

The results show that reading the corrective feedback automatically generated by the learning system was the most frequent behavior, with a total of 50452 times in 12 weeks of learning on Listening Hacked. That was approximately 140.14 times per student per week. This behavior alone made up 43.12% of all students' interactions throughout the whole experiment period. These figures indicate that reading feedback contributed to a large amount of learning on the system. The findings accord with the observations in previous studies, which show that students usually made considerable use of feedback in their learning (M.S. Cárdenas-Claros, 2020).

5.3.1.2 Regulating attention

The high frequency of feedback usage can be explained by the important role of feedback in learning which has been extensively discussed in the general L2 literature, but still rarely touched on in L2 listening research. Nevertheless, we still have some empirical evidence that corrective feedback can improve the learner's listening performance such as Lee et al. (2012), Pérez-Segura et al. (2020), and Yan (2012). The results of the present study further expand the findings of those studies by providing insights into this relationship. In the following paragraphs, we discuss the different roles of feedback in L2 listening learning by demonstrating how the students made use of it.

Pujolá's (2002) and Cardenas-Claros' (2011) studies report that most students used feedback for assessment purposes rather than for reflection or learning

opportunities. The present study confirms that the students did use feedback for assessing their comprehension and task completion, however, they also used feedback for regulating their attention in listening learning, which has not been described in past L2 listening research.

Corrective feedback has been known to facilitate the comprehension process (S. P. Lee et al., 2012). It helps the learners notice their problems in this process of meaning construction. The listening process involves the interactions of the two-way processing of aural information: top-down and bottom-up (Matthews & O'Toole, 2015; Siegel & Siegel, 2015; Vandergrift et al., 2012). When the listener encounters a problem in only one direction of information processing, the whole comprehension process is affected. For example, a listener may have false interpretations and incorrect word recognition due to lacking understanding of the contexts in which the utterance is spoken (LaScotte, 2015). Alternatively, when listening to unfamiliar, regional accents a listener may face greater challenges in word recognition even though they know the contexts of the speech (Adank et al., 2009; Floccia et al., 2006; Van Engen & Peelle, 2014). Additionally, the listener can have problems in comprehension when they encounter ambiguous words, or when they misunderstand the contexts. In those cases, receiving corrective feedback in time is useful and effective in improving listening comprehension because it can assist them with eliminating ambiguity and wrong hypotheses (Nassaji & Swain, 2000) and reallocating their attentional resources to where necessary (Hattie & Timperley, 2007), and hence enhancing the listening process. In the following excerpt from the interviews, a student described exactly just those benefits.

> *Excerpt 65: It (reading the feedback) let me know which words in the chunk I transcribed correctly, and whether I put the words in the correct order. With that, I didn't have to make any guesses, but I knew exactly which I mistranscribed in the chunk, which I got it wrong in the transcriptions, and which I should listen again in that chunk.*

It is noted that the benefits of feedback only come with precision (A.-P. Lian & Sangarun, 2017). Generalized feedback cannot lead to effective learning as it does not effectively draw out the learner's awareness to errors. As seen in Excerpt 7, the student can eliminate confusion and redirect her attention only when she was aware of her own listening problems. Therefore, the efficiency of feedback depends on whether it is personalized, and precisely points out the learner's errors. When feedback lacks precision, its effectiveness is questionable, as a student commented in Excerpt 53.

5.3.2 Strategies for using repeated listening

Another behavior that is worth noting is replaying and rehearing the chunks being transcribed. This behavior had the second highest frequency and accounted for 22% of all interactions in the learning system. Throughout the period of 12 weeks, the students pressed the Repeat button 26365 times as to rehear the chunks, and that does not include 1927 times pressing the Rewind button. While these figures do not directly point to the students' listening improvement, they reveal the students' authentic needs for repetitions in L2 listening learning. The finding is different from that of Sendag et al (2018) whose study found that students were demotivated with repeated listening activities. In their study, the students were required to listen repeatedly to a number of listening texts for 3, 7, 15 times which caused the students' boredom. Note that in the present study, the students who learned EFL listening on Listening Hacked were not required to listen repeatedly to any spoken texts; it was their decisions to listen again to what they felt necessary. This explains why repeated listening in this study did not cause demotivation or boredom as in previous studies.

5.3.2.1 Constructing meaning

In fact, the need for repeated listening is natural, both in listening learning and real communications. Speech contains several factors that can bring challenges to the process of comprehending foreign speech. They could be noise (Lecumberri & Cooke, 2006; Yang et al., 2017), accented speech (Ferguson et al., 2010; Gass & Varonis, 1984), or aspects of connected speech (Kuo, 2012; S. Lee, 2013; Rosa,

2002; Wong et al., 2019) among others. Sometimes the minimal and favored approach to overcoming those challenges in listening comprehension is simply rehearing what is missed. The results of this study have confirmed the value of this strategy. Most students in this study first looked to repetitions when they noticed their problems in listening, either by reading the feedback or by themselves.

Past research posits that repeated listening/viewing can improve listening comprehension (Dupuy, 1999; Krashen, 1996; Şendağ et al., 2018; Skouteris & Kelly, 2006). The findings of this study offer possible explanations for this relationship. One explanation might be that for every attempt of repetition, the listener may attend to and notice different parts of the texts and that could lead to various ways of constructing the appropriate meaning. This process has been mentioned earlier as the accommodation process (Bodner, 1986) in which the listener tries to find a way to construct the meaning so that it can gradually fit in the mental structure of meaning (Gernsbacher, 1990). A student mentioned this point in Excerpt 21.

5.3.2.2 Recapturing the lost information

With reference to the model of listening comprehension presented in Section 2.4.1, every attempt of repeated listening could contribute to better perceptual or/and conceptual processing of the speech. It is because repeated listening gives the listener additional opportunities to retake the lost information in the speech signal as normally seen in reading comprehension. As the students reported in the interviews, most information was lost due to fast speech, unclear speech signal, or low working memory capacity. These challenges might be overcome simply by rehearing. When the perceptual information or the conceptual information is well obtained, the whole process of listening comprehension would also be enhanced due to the interactive nature of the listening process. Excerpt 20 and Excerpt 24 are two examples in which the students mentioned the possibilities of recapturing the lost information in listening with the support of repeated listening.

In short, the findings of the present study suggest that repeated listening plays an important role in the process of meaning construction in L2 listening.

It is the first type of support that the students would want to use in order to solve their problems in listening. The observations from the study further indicate that the values of repeated listening much rely on whether the students are capable of controlling what, how, and when to listen again. In doing so, repeated listening can assist with L2 listening comprehension without causing unnecessary and redundant efforts on the part of the listener.

5.3.3 Strategies for using the rewind functions

5.3.3.1 Retaking contextual information

In this study, the learning system provided two types of repeated listening: exact repetitions and repetitions with additional contexts. The exact repetitions were provided through the Repeat button which allowed students to listen again to the exact chunks being transcribed while the Rewind button and other options for rewinding longer than 1 minute allowed students to rehear the parts before the chunks. It is important to note that there is a conceptual difference between the two types of repetitions though they may seem alike. The exact repetitions are merely about repeated listening as discussed in the previous section, while the main purpose of the Rewind button and other longer rewinding options (coded as Rewind and Rewind 1+) is to assist the listener with retaking the contextual information. Therefore, the students' decisions of rewinding reflect a need for retaking the contexts required for comprehending the texts, as a student mentioned in Excerpt 30.

5.3.3.2 Seeking alternative solutions other than bottom-up processing

The results of the study show that the behavior of rewinding had a relatively high frequency like that of repeated listening. In total, the students used the Rewind function 19275 times for seeking more contextual information, that accounted for 16.47% of all interactions. However, other longer rewinding options, which are rewinding 1, 2, 5, 10 minutes before the current chunks, constituted a very small proportion of 0.56% of all interactions. This suggests that the students rarely felt the need for rewinding longer than 2 seconds before the current chunks. Additionally, among the longer rewinding options, rewinding 1 minute had the highest frequency

and the frequencies of other options decreased as the rewinding time was longer. Some students said that they found rewinding 1 minute (or longer than 1 minute) time-consuming and they believed that the standard Rewind function was helpful enough.

In this study, a significant positive correlation between the students' behavior of rewinding 1 minute and their listening improvement was found. In addition, the multiple linear regression analysis shows that this behavior statistically predicted the students' gains in listening comprehension. Thus, it is a predictor of the students' listening improvement. Interestingly, while repeated listening had much higher frequency than that of rewinding, it was not statistically correlated with the students' listening improvement. One possibility is that the students usually looked for repetitions, and only when that failed, they would seek solutions from obtaining more contextual clues. In fact, the learning behavior data also discloses a common interaction sequence that gives support to this action scheme. As presented in 0, one of the most common behavior patterns of students' learning on Listening Hacked was FEEDBACK→ REPEAT→ PHRASE SEARCH→ REWIND→ FEEDBACK. According to this pattern, the students usually started addressing their listening problems with repeated listening (REPEAT). If repeated listening failed to help them, they would try Phrase Search function, and then rewinding the chunks for more contexts.

Both the frequency and the above behavior pattern suggest that it was apparently a bottom-up processing priority in solving listening problems and that when bottom-up processing did not help, the students would seek an alternative solution from revisiting top-down information, and they often succeeded with it. One possible explanation might come from the role of top-down information in listening. According to the listening comprehension model of structure-building, the listening comprehension process involves activating, enriching, and suppressing perceptual units from the speech signal (Gernsbacher, 1990). Problems in one of these three processes would cause inefficiency in building a coherent mental structure of the speech. On the contrary, when multiple perceptual units are activated simultaneously, efficiency in suppressing or enriching the best-fit activated units for the mental structure could bring

significant advantages in the listening comprehension process (Gernsbacher & Faust, 1991). It follows that when top-down information is properly obtained, the activated schemata in the conceptual source could serve as a mental framework for the suppression and enrichment of bottom-up information (S. Graham & Macaro, 2008).

5.3.3.3 Predicting the speech message

In the literature, the benefits of contextual or top-down information are often discussed in relation to prediction, inference, or metacognitive strategies (Macaro et al., 2007). Consistent with the literature, the students in this research also reported that rewinding provided them with more contextual clues that allowed prediction and inference of the speech message, and thereby assisted them with word recognition and comprehension in general. Here is a student's mention of this benefit.

Excerpt 66: *When I used Rewind 1 minute, 3 minutes, I could listen to understand the situation as to predict the meaning of the current utterance.*

5.3.4 Strategies for performing phrase searches

High variability phonetic training has been employed in speech perception learning, specifically in improving segmental perception, word recognition, and suprasegmental perception (Barriuso & Hayes-Harb, 2018; Clopper & Pisoni, 2004; Lively et al., 1993). In the context of listening comprehension, perceptual learning involves enhancing the efficiency of activating the best-fit perceptual units in the speech signal. It highlights the importance of data-driven processing in the listening process. The integration of this method in the function Phrase Search was aimed at providing the students with effective perceptual learning, and subsequently enhancing their bottom-up processing as to improve their overall listening comprehension ability. In the present study, we explored how the students made use of Phrase Search in addressing perceptual problems.

The results show that the behavior of performing phrase searches constituted 15.37% of all interactions, with a total of 17988 searches, or an average of 599.6 searches performed by a student in 12 weeks. These results could be interpreted in a few ways. First, these figures indicate that this behavior was quite common in the

students' learning, meaning that it took a significant proportion of the students' autonomous learning. Second, the high frequency of using Phrase Search might stem from the students having many perceptual problems in learning and their needs for solving those problems. Third, it might reflect the students' positive experience using Phrase Search in addressing those problems. These interpretations are backed up by the students' perceived usefulness of Phrase Search in which 86.67% of them said the function was *useful* or *very useful* in their listening learning. Furthermore, the students reported three main purposes of using Phrase Search.

5.3.4.1 Seeking support for word recognition

The students said that using Phrase Search could provide them with additional support for word recognition. This suggests that high variability perceptual training could be an alternative approach to addressing word recognition problems. To date, the common approach to word recognition is repeated listening (Dupuy, 1999; Ellis & Le, 2016; Mordaunt & Olson, 2010; Şendağ et al., 2018). In other words, if the listener cannot recognize the words in speech, the most common and straightforward way is to listen again. While repeated listening may work, there are many cases when it does not help at all and may cause demotivation, as a student mentioned in the following excerpt.

Excerpt *67*: *...in many movies the characters spoke in ways that I could not catch it at all, even when I tried listening again. In those contexts, the characters intentionally spoke very fast, and it was impossible to get it. But in other contexts in other movies, other characters may speak slowly. I found it very useful, and it helped me recognize the words that I missed earlier.*

In cases when repetitions do not help, perceptual learning with speaker and context variability might be helpful. The students' interaction data in the present study illustrates this. The sequential pattern of learning behaviors FEEDBACK→ REPEAT→ PHRASE SEARCH→ REWIND→ FEEDBACK demonstrates that the students often tried high variability perceptual learning (PHRASE SEARCH) after being unsuccessful with repeated listening (REPEAT). The findings can be explained with

reference to the exemplar model of speech perception (Keith Johnson, 1997). In the exemplar model, a unit of perception is stored as exemplars and recognized through the shared features of those exemplars. In high variability perceptual training, the listener's exposure to sources of variability including speaker and context variability allows the intake of additional exemplars of certain perceptual units. As more exemplars added, the representations of those units become robust, and with that, perception of varied pronunciations is enhanced (Barcroft & Sommers, 2005; Clopper & Pisoni, 2004). This explains why high variability perceptual learning could improve bottom-up processing in this study.

5.3.4.2 Dealing with adverse listening conditions

The students purported that Phrase Search could be more effective than repeated listening, particularly in adverse listening conditions such as noise, unclear speeches, or overlapping voices. Previous research has suggested that variability in non-phonetic features such as amplitude does not affect the listener's perception while variability in phonetic features such as quality of speech, speech rate, or speaker variability does (Bradlow et al., 1999; Perrachione et al., 2011; Sommers & Barcroft, 2007). That perhaps explains why students resorted to Phrase Search in the hope of finding finer speech signal which could possibly ease their perceptual processing, as a student described in Except 17.

5.3.4.3 Learning vocabulary in contexts

The students reported an interesting purpose of using Phrase Search, somehow irrelevant to perceptual decoding, that is to learn vocabulary in contexts. Research shows that exposure to a variety of contexts in which certain lexical items are used could make vocabulary acquisition happen (Rapaport, 2005; Sonbul, 2012). In this study, the students used Phrase Search, not only for the intended purpose of perceptual learning but for acquiring vocabulary from various contexts as well. It is difficult to evaluate the isolated effect of learning vocabulary in this way on the students' overall listening comprehension; however, given this opportunity for expanding vocabulary, the effect on listening comprehension could not be of any less

positive since improved vocabulary knowledge can contribute to better listening comprehension (Noreillie et al., 2018; Stæhr, 2009; F. Teng, 2014, 2016). On the other hand, research has indicated that the knowledge of frequently co-occurring lexical items is essential in L2 listening comprehension (Conklin & Schmitt, 2012; L. Li et al., 2020), however, one needs a certain number of encounters in order to acquire such word knowledge (Peters et al., 2016; van Zeeland & Schmitt, 2013). The use of Phase Search in this way allows repeated exposure to various occurrences of a lexical item, thus, could be useful in vocabulary acquisition and meaning recall.

One interesting finding of the present study is that the students' frequency of performing phrase searches negatively correlated with their listening improvement. The regression analysis also indicates a corresponding result that the students' EFL listening proficiency would improve when the frequency of using Phrase search decreased. One possible explanation is that the students' Phrase Search usage was related to their problems in word recognition, that is, the more problems they had, the more phrase searches would be expected. As their word recognition ability improved, the students' reliance on Phrase Search for word recognition became less frequent. Put another way, they would need less support from perceptual training when their bottom-up processing capacity has improved, hence the resulting negative correspondence between the frequency of this behavior and listening proficiency. This study still holds that using Phrase Search is beneficial to building up bottom-up processing and word recognition skills; the observed negative correlation between frequency and listening improvement was just a reflection of the changes in the students' need for perceptual learning support as they progressed in listening comprehension.

5.3.5 Strategies for using English captions

The use of captioning in L2 listening comprehension has received much attention in the literature despite its effectiveness in L2 listening being inconclusive (Leveridge & Yang, 2014; F. Teng, 2019). Much research on captioning and L2 listening has focused on three aspects: the impacts of captioning on listening comprehension

(Hayati & Mohmedi, 2011; Mitterer & McQueen, 2009) and on L2 acquisition (d'Ydewalle & Pavakanun, 1997; Montero Perez, M., Peters, E. and Desmet, 2014), the instructional approach to using captions in L2 listening (Hayati & Mohmedi, 2011; Leveridge & Yang, 2014), and the processes involved in L2 listening comprehension with and without captions (Chai & Erlam, 2008; Winke, P., Gass, S., & Sydorenko, 2013). The discussions of the findings will be centered on the last aspect regarding the processes and strategies involved in the use of captioning in L2 listening learning. Some pedagogic considerations will also be discussed.

Past research has indicated that choosing the timing regarding display or removal of captions in teaching and learning L2 listening is already a dilemma (Hayati & Mohmedi, 2011). As Leveridge and Yang (2013) note, both early and late removal of caption texts can bring negative influence on the learner. They contend that removing captions too early can demotivate the learner while removing captions too late can interfere with the listening process. To address this issue, a few options have been proposed: using self-reports to consult the learner's perceptions of their reliance on captions, or using an instrument for assessing their reliance on captions such as in Leveridge & Yang (2014) study. In the present study, we propose an alternative approach that is autonomy-oriented: let the students decide when and how to use the captions by themselves.

Regarding the frequency of using captions, the students decided to turn on the captions 1519 times throughout the observed period. Additionally, they intentionally turned off the captions 476 times. Note that by default Listening Hacked turned off caption viewing in all practice sessions, and in total the students have taken 1468 practice sessions including both complete and incomplete ones. From those data, it seems that the students tended to view the movies without captions rather than with captions. They explained in the interviews that viewing without captions can help them focus on listening and avoid reliance on reading the caption texts.

> Excerpt *68: I think turning off captions would help me focus on listening more. Turning on captions made me focus on reading comprehension.*

This finding is somehow incongruent with some previous studies examining the effects of caption viewing on comprehension such as Markham et al. (2001) or Montero Perez et al. (2014). The results of those studies suggest that caption viewing leads to better comprehension than non-caption viewing. However, the students in the present study did not seem to value the idea of using captions in the whole time viewing the movies. They showed a tendency to viewing the movies without captions and believing it was more beneficial to their listening comprehension, and they only switched on caption when they felt the need for it. This indicates a benefit of the autonomous approach in which the learner is responsible for deciding when captioning should be used. The decisions are learner-made and based on specific cases at the time. Appropriate decisions could bring about advantages to the listening process. Inappropriate decisions, on the other hand, can be assessed online and rewound, which may lead to better results as compared to the self-report or pre-assessment methods. In fact, it was common that the students in this study often switched between caption viewing and non-caption viewing as observed in the sequential patterns of students' learning behavior (Patterns 3 and 4 in Section 4.2.4). Therefore, the need for toggling the captions is dynamic and only individual learners can decide what works best for them.

(3) OFFCAP → ONCAP → REWIND → FEEDBACK

(4) ONCAP → OFFCAP → REWIND → FEEDBACK

In addition to the need for non-caption viewing, the students also explained four main reasons for using captions as they watched the movies: to evaluate their comprehension, to compensate for their lack of comprehension or incompetency in word recognition, and to learn new vocabulary. These findings are in line with the results of many previous studies which highlight the benefits of using captions in L2 learning, including, improving comprehension of video viewing (Chung, 1999; Isabel

Borras & Robert C, 1994; Markham et al., 2001), enhancing written word recognition (Sydorenko, 2010), and developing L2 vocabulary knowledge (H.-C. Huang & Eskey, 1999; Majuddin et al., 2021; Neuman & Koskinen, 1992; Peters & Webb, 2018). It is worth noting that the decisions of using captions in movie viewing are not completely contradictory to that of not using captions. They simply differ in modality but serve the same purpose, that is to better understand the movie contexts. Listening Hacked only allows the students to toggle captions in non-selected chunks, in order words, the decisions of turning on or off captions were only applicable to the movie parts preceding the chunks being transcribed. Therefore, whether the students decided to view movies with or without captions was essentially related to the need for understanding the movie contents or the contexts prior to the chunks being transcribed. In this sense, focusing on the aural and visual signal as in non-caption viewing, or using textual support from caption viewing might be different in modality of input but similar in the goal of obtaining necessary top-down information for comprehension of the target chunks.

5.3.6 Strategies for using dictionary

In the present study, vocabulary was not taught explicitly as in traditional approaches to L2 listening instruction. The students were expected to acquire vocabulary knowledge through extensive movies viewing and consultations with a built-in dictionary. The results show that the built-in dictionary had a very low level of engagement; its frequency of usage was 274 times in total, taking only 0.23% of all learning behaviors. This result is consistent with the students' perceived usefulness of the built-in dictionary in which 6.67% of the students said it was useless and 40% reported doubtfulness. A further exploration in the interviews reveals that most students believed that the dictionary was necessary for their learning; however, they found its design not user-friendly enough. Despite having low level of usage and mixed opinions about the usefulness, the students confirmed that they did make use of dictionaries, but mostly external ones such as the online Oxford Learner's Dictionaries or Google Translate. Therefore, the behavior data in this study may not accurately

reflect the students' behavior of using the dictionary on Listening Hacked. Nevertheless, the need for it in L2 listening learning is a real one since it is related to vocabulary knowledge.

Several studies have suggested that vocabulary knowledge plays a considerable role in developing L2 listening comprehension. Matthews (2018), Wang and Treffers-Dallerb (2017), L. Li et al. (2020) maintain that L2 vocabulary knowledge, referring to both vocabulary size and depth, is a predictor of L2 listening proficiency. The students in this study reported three ways that they employed in order to expand their vocabulary knowledge in support for EFL listening comprehension. First, they consulted dictionaries, the built-in or external ones, for unfamiliar lexical items when viewing the movies or reading the captions. The students' main purposes of looking up dictionaries were concerned with word meanings, word usage, and spellings. Second, they made use of the Repeat function to learn pronunciation of certain words. The main purpose, thus, was to enhance spoken word recognition. Some students also maintained that this strategy helped with their vocabulary retention. Third, they reported using the Phrase Search function to learn word usage in various contexts, especially words that frequently occur together, as already discussed in Section 4.2.5.

5.3.7 The overall picture

Previous studies such as Pujolà (2002) and Cardenas-Claros (2020) have explored how students interact with help options in their L2 listening learning. These studies have identified several strategies that students employ when facing different listening problems. The discussions on the students' strategies for using specific features of the app Listening Hacked thus far also point to similar findings, however, a big picture of how the students put all the help options in use has not been clearly described. Based on the analyses of the students' learning patterns on Listening Hacked (see Section 4.2.4) and the discussions thus far, an overall pattern of the students' strategies for using the learning system resources was formulated as shown in Figure 5.1

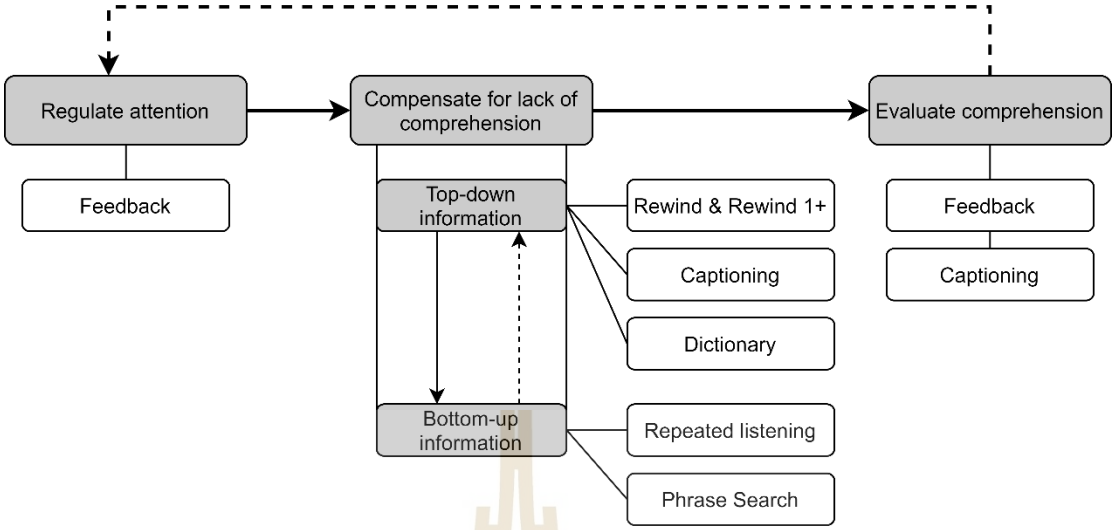


Figure 5.1 The overall pattern of learning on Listening Hacked

The figure illustrates the sequential pattern of the decisions that the students made in their autonomous learning on Listening Hacked, and together with that, the help options being used to execute those decisions. The pattern starts with the students' decision to regulate their attention in listening. Put another way, the students first decided to consult corrective feedback in order to identify any mismatches in their understanding and the expected understanding of the spoken speech. The confrontation with those mismatches would trigger the redirection of attentional focus in listening. The next stage involves using other help options to compensate for their lack of comprehension as they have noticed earlier through the feedback. At this stage, the students were likely to prioritize addressing perceptual problems through repeated listening, using the Repeat function, or through high variability perceptual learning, using the Phrase Search function. However, when the support for bottom-up processing could not solve the problems, they would turn to other help options to recapture the relevant contexts since obtaining more accurate top-down information is also a way to assist with bottom-up processing. These usually involved using the Rewind functions, turning on or off the caption texts, or searching for the meanings of unfamiliar content words in the dictionary. As discussed earlier, the Rewind functions differ from the Repeat function in that they are instrumental in

retaking the lost contextual information in movie viewing. Additionally, the acts of turning on or turning off captions while differ in the modality and focus of learning, share the same purpose since caption texts were only available for movie parts with non-selected chunks. The last stage in the sequence is the decision of evaluating comprehension. The evaluation process was done through reading the generated corrective feedback; however, sometimes when feedback was unavailable as for non-selected chunks, English captions were used instead. Note that the sequential pattern may continue as a loop if the evaluation of comprehension still showed unsolved problems, otherwise it ended when the evaluation showed no remaining problems.

From the overall pattern of learning in such a CALL application for EFL listening comprehension as Listening Hacked, it can be seen that regulation of attention in listening, strategies for compensating for lack of comprehension, and evaluation of comprehension are all important components in the learning process. The confrontation and noticing of listening problems in the regulation of attention stage are instrumental in making learning happen; the supports for recapturing the lost perceptual and contextual information are contributing to the process of meaning construction in the listening process; the evaluation of comprehension through feedback makes the learning process complete and effective. What follows is that to enhance the learning process, the learning system should provide necessary support and affordances that can facilitate each of the said learning stages.

Be that as it may, a few studies have cautioned that the presence of help options is not always a merit to L2 listening comprehension as they may cause dependence and interference in learning, especially the use of captioning or subtitling (Diao et al., 2007; Stewart & Pertusa, 2004). However, the present study found that the students tended to use less help options as they progressed in learning. As the results have shown, nine out of eleven types of help options had significant downward trends in the overall usage patterns. This line of findings, when being considered with the students' reported usefulness of the help options and their listening improvement, implies that the help options were well designed and have achieved the goal of

assisting the learners in developing their listening skills. This also explains why the frequencies of using the help options were high at the beginning and gradually decreasing as they have fulfilled their objectives. The results, thus, may be interpreted in the way that the students' reliance on help options was real but until they improved their listening skills, then the need for those supports were no longer relevant. For that reason, it can be concluded that the use of help options is not detrimental to the students' learning provided that they are properly designed to assist specific stages in the learning process as mentioned earlier and that they are under the students' control.

5.4 Chapter summary

This chapter provides discussions on the findings of the study. It first explains the possible reasons for the students' improvement in EFL listening comprehension. These include the students' development in word recognition ability, learner autonomy, and increased motivation in learning. The chapter then discusses the students' behaviors and strategies for using the affordances on the learning system as to support their independent learning. Lastly, the chapter presents an overall pattern of behaviors and strategies employed in the learning process on Listening Hacked.

CHAPTER 6
CONCLUSION

This chapter concludes the present study. Firstly, it summarizes the objectives, the procedures, and the main findings of the study. Secondly, it discusses the implications of the study. Thirdly, it points out some limitations of the study and recommends possible directions for future research.

6.1 Summary of the study

This study set out to investigate the effects of the proposed AI-based, personalized learning system Listening Hacked on Vietnamese undergraduates' EFL listening comprehension. Another aim was to explore and understand more about the learners' learning behaviors in a technology-based, autonomous learning environment as such. With those aims, the study was designed with mixed methods in which both quantitative and qualitative data were collected and analyzed. Fifty-three first-year, English-major students at a university in Vietnam participated in the research project; 30 of them were randomly assigned in the Experimental Group and 23 in the Control Group. The experiment employed the pretest-posttest design and last 12 weeks. The system recorded all the students' interactions with the app in the database which allowed the analyses of the students' learning behaviors. Furthermore, the students in the Experimental Group were surveyed via a questionnaire which was aimed to document their perceived usefulness of the learning system. Additionally, 19 students in the Experimental Group were interviewed to enhance the interpretations of their surveyed perceptions and learning behaviors.

The most obvious finding emerged from this study is that the students in the Experimental Group have improved their EFL listening proficiency significantly and that they outperformed their counterparts in the Control Group on the EFL listening posttest. These results have illuminated the positive effects of the app Listening Hacked on the students' EFL listening comprehension.

The analyses of the students' learning behaviors in the learning system further reveal that the students employed different strategies for using the help options on Listening Hacked as to address various listening problems arisen in their learning and that those strategies followed a pattern which involved the students' need for regulating attention in listening, the need for compensating for lost information or lack of comprehension, and the need for evaluating their comprehension. The analyses of learning behavior frequencies and patterns also highlight the importance of certain help options in the students' EFL listening learning, including the provision of corrective feedback, different types of repeated listening, high variability perceptual learning with audio-visual materials, and captioning.

Regarding the results of the qualitative analyses from the questionnaire and interviews, the students expressed high levels of satisfaction with the app and its features. They also believed that the app assisted them well in developing their EFL listening skills, motivating them to learn EFL listening, and enhancing their cognitive ability such as working memory capacity and vocabulary recall. On the contrary, some students expressed concerns about the technical issues they experienced when using the app, including the limited learning resources and the design of the app. Nonetheless, they maintained that those concerns were not too serious, and the app has brought them far more positive experiences in EFL listening learning.

6.2 Implications of the study

Based on the findings and the discussions presented thus far, some important implications of the study are discussed below.

6.2.1 Theoretical implications

L2 listening researchers insist that the traditional method of some kind is usually less effective than the newer methods; for instance, considering the product approach with the process approach to L2 listening instruction (Field, 2008c; Porter-Szucs, 2018; Vandergrift & Goh, 2012). Despite the ineffectiveness of the so-called traditional method, it still exists in some forms in most language classrooms in Vietnam because the application and implementation of any innovative methods will take time and effort, and some are even impractical to implement in certain contexts. This study

demonstrates that with a principled framework, the development and implementation of technology-based solutions like Listening Hacked are feasible and could produce encouraging results. The framework rests on the theoretical foundations presented in Chapter 2 and offers seven principles for designing technology-based learning system for L2 listening as illustrated in Figure 6.1.

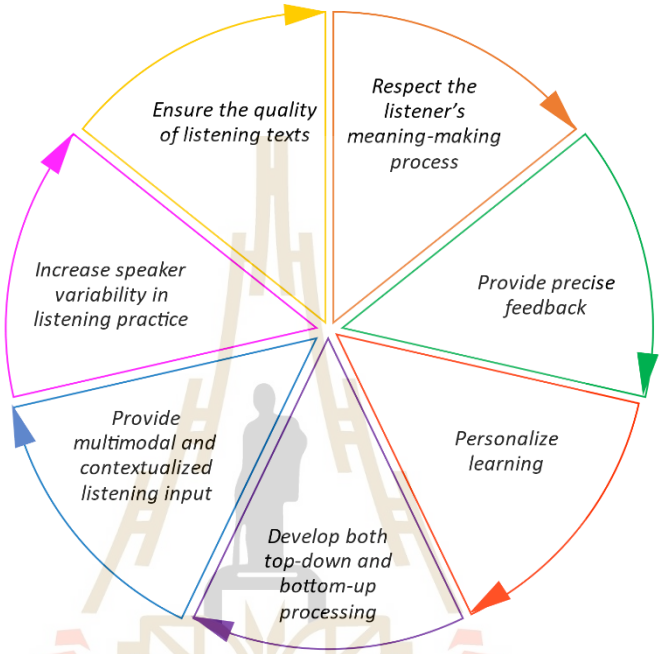


Figure 6.1 Seven principles for designing CALL applications for L2 listening

Descriptions of the principles could be found in Section 3.4.1, however, it should be noted that those principles can be applied in the development process in a flexible way. For instance, the principle requiring the system to provide tasks for practicing both top-down and bottom-up processing was applied in Listening Hacked in the form of paused transcription tasks; however, any other tasks that can achieve this purpose should also be considered in the development process. Similarly, while movies in Listening Hacked can serve as multimodal, contextualized input for listening practice, other types of input such as audiobooks with illustrations, or presentations with voice-over could also be qualified, depending on the focus of the practice. Therefore, the applications or realizations of those principles in the development of any technology-based learning systems for L2 listening comprehension are open to the researchers' creativity and interests, but at the same time, the principles formulate the vital components of the system to ensure the desired effectiveness of the system.

For this reason, these principles can be used as a guidance for designing ICALL software effectively rather than detailed recipes for replicating the learning system in this study.

6.2.2 Pedagogic implications

6.2.2.1 Using technology in personalizing learning experience

The present study shows that personalization could be created with the help of technology and AI which possess the affordances for supplying and enhancing listening input, and the capabilities of efficiently personalizing students' learning experience.

Firstly, the results of this study suggest that technology can create personalization in learning. The construction of the multimedia database in this study illustrates how technology can allow us to create a large collection of listening materials that suits every learner in the way that learners are offered more learning choices and can decide what fits their learning preferences and experiences. These conceivable virtues of technology can solve the matter of input inappropriacy caused by the conventional classrooms in which "the whole class" are obliged to study the same materials. The multimedia database can be used as the main listening materials as well as the reference library like the Listening Boosters feature.

Secondly, the results of the study highlight the capabilities of technology in input enhancement (Cardenas-Claros & Gruba, 2013). These capacities are associated with media playback controls (e.g. play/pause/rewind buttons) and multimedia materials (e.g. videos, captions) (Hubbard, 2017). Regarding the ephemeral nature of listening, using media players with playback controls for listening practice can help learners pace their own learning so that perceptual processing overload is efficiently addressed on an individual basis. Media controller bars and sliders, for example, or the Repeat/Rewind buttons can help make authentic materials more accessible in ways that students can rewind the speech to a certain part and listen again to it as many times as they wish. Note that in order to optimize the usefulness of sliders, we should place certain restrictions in their usage, that is, they can allow students to rewind freely, but not to fast-forward or skip forward the listening texts. Alternatively, deficiency in comprehension can be tackled when learners are given options for using multimedia such as videos (H.-S. Kim, 2015) or captioned videos (Gass,

Winke, Isbell, & Ahn, 2019; Winke, P., Gass, S., & Sydorenko, 2010) in their listening practice.

6.2.2.2 Creating teacherless classrooms and developing leaner autonomy

As the results indicated, the students in the present study not only improved in their EFL listening but reported increased motivation in EFL learning as well. An implication of this is the possibility that technology-based learning systems like Listening Hacked can make L2 listening learning effective without the presence of the teacher while, at the same time, are capable of solving the paradox between the large-sized language classroom and the teacher's inability to cater for individual learners' needs. In the future, when more ICALL apps are developed, online teacherless classrooms might be a practical solution to L2 listening learning. As this study took place during the COVID-19 pandemic, it further demonstrates that autonomous learning systems as such may be an indispensable part of L2 listening instruction in the future and of L2 pedagogy in general. In fact, the data collection process in this study was carried out completely online and there were no physical teacher-student interactions during the learning process due to the social distancing policies in Vietnam. The students' learning activities were chosen and completed by the students themselves, yet the learning outcomes were satisfying. From this point, we can be confident that online autonomous learning systems with AI integration like the one in this study will be soon capable of developing learner autonomy and solving the administrative issues in teaching and learning L2 listening.

6.2.2.3 Using multimedia in L2 listening learning

The use of videos in L2 listening instruction is not new and has been supported by research (Shin, 1998; Sueyoshi & Hardison, 2005; Wagner, 2010). Overall, this study supports the use of videos in L2 listening. Had the multimedia listening materials used in this study been replaced with audio-only materials, the results of the study would not have been very positive. Videos are not only assisting in the comprehension process but also creating motivational effects on the learners. The study therefore suggests that videos be used as the primary listening materials as in this study, or as feedback or references in L2 listening teaching and learning.

6.2.2.4 Focusing on both form and meaning in L2 listening comprehension

While bottom-up processing might be prioritized in L2 listening learning, this study insists that working on both bottom-up and top-down processing is important in developing L2 listening comprehension. Traditional listening drills for building up bottom-up skills such as dictation or listening cloze can be useful, however, those drills primarily draw the learner attention to form recognition and often neglect meaning. The method used in this study for training bottom-up processing skills, i.e. paused transcription tasks, has proven useful in drawing the students' attention to bottom-up information in the listening process without leaving the benefit of contextual information. Additionally, this study provided different types of repeated listening which can assist the learner in dealing with different types of cognitive processing. Exact repetitions (e.g. the Repeat function) can assist with bottom-up processing while longer repetitions (e.g. variations of the Rewind function) can assist with clarifying uncertain contextual information. One implication of the study is that when paused transcription is used with repeated listening, it can serve as an intensive listening practice which helps the learner work on bottom-up skills without losing the focus on contextual information and global meaning.

6.2.2.5 Addressing perceptual problems in L2 listening

As a matter of fact, many students find listening to a foreign language difficult because they cannot recognize the spoken words in connected speech. There are several teaching methods for enhancing word recognition skills used in L2 listening class such as dictation, listening cloze, or error identification (Rost, 2011). This study has demonstrated that high variability perceptual learning is not only effective in learning phonetic contrasts but useful in training word recognition and phrasal recognition as well. This is a promising alternative way for addressing perceptual problems in L2 listening. Furthermore, the use of audio-visual materials has shown added values as compared to the traditional audio-only materials in high variability phonetic training. The study also implies that perceptual training should be personalized and not followed a general set of common perceptual problems. That means the learner should only work on their own perceptual problems as they feel necessary.

6.2.2.6 Personalizing caption-viewing

In the conventional L2 listening class, the teacher decides when to show the transcripts of the listening material or decides in advance whether the students can use the transcripts/captions for certain listening practice. The decisions certainly may not suit every learner's need. In fact, there are controversies over the use of transcripts/captions in L2 listening learning as discussed in the previous chapters. This study has presented a simple way that is to allow the learner to decide the timing of using captions. This approach implies personalization and autonomy in learning. The findings of the present study show that some students did not use captions at all in their learning, while some reported many intentional uses of captions, especially when they needed to recapture the contexts of the movies. The fact that the students exhibited different patterns of using captions in their listening implies that the one-size-fits-all approach does not work, and that rather than asking the question of when captions should be used, we should let the learner decide it themselves.

6.2.2.7 Teacher training programs

This study has demonstrated a successful implementation of a technology-based, autonomous learning system for EFL listening. However, its implication is not necessarily limited to L2 instructions but also extends to teacher training programs. As can be seen in this study, the development and maintenance of such technology-based systems inevitably require the teacher to have some IT knowledge, and more importantly, expertise in developing, selecting, and providing appropriate learning resources. In other words, the job of a traditional teacher who is used to interacting with the learners and delivering instructions now should involve much work in material development and course design. Unfortunately, most teacher training programs in Vietnam currently do not have much concern over these issues, instead they normally concentrate on teaching methods, classroom management, lesson planning, to name a few. Another reality in Vietnam is that EFL course designers are usually those who do not teach, and that EFL teachers are often incapable of designing new courses or materials, but reuse the available materials provided in commercial EFL coursebooks. With the considerable benefits of technology-based, autonomous learning systems like Listening Hacked, such platforms are likely to

expand in use soon, yet pre-service teachers have not been properly trained and prepared for this trend. Therefore, it is recommended that teacher training programs should give emphasis to knowledge courses and practice modules for designing, preparing appropriate learning resources for use in technology-based learning platforms. Furthermore, some basic IT knowledge and skills such as how to create, edit, and publish multimedia materials, how to create and manage databases, are also of importance in future teacher training programs, when the role of teacher gradually shifts from a knowledge giver to a learning resources provider who will be responsible for the effective operation of a technology-based learning platform.

6.3 Limitations of the study and recommendations for future research

Although the study has followed rigorous research procedures and produced valuable insights and implications for researching and teaching L2 listening, it cannot avoid certain limitations.

Firstly, the study had a relatively small sample size and only took its samples from only one university. In addition, the participants in this study were English-major undergraduates. While this sampling technique is useful for producing a homogeneous group and advantageous to statistical analyses, generalizations to other contexts and non-English major students might be cautious. This study could be replicated using a larger sample size and include a variety of participants whose levels of EFL listening proficiency might be different.

Secondly, as the present study only investigated the effects of the app in 12 weeks, further work could assess the long-term effects on the students' EFL listening comprehension. This will help to explore the limits of the app in developing students' listening ability. On the other hand, as the students' listening development in this study was slowed down in the last 5 weeks of the experiment, further research could also investigate the optimal time of using the app as to maximize its effectiveness in the shortest time.

Thirdly, since the present study only examined the overall effects of Listening Hacked on the students' EFL listening comprehension, the effects of the separate help options and instructional methods embedded in the design of the app have not been

quantified. We have gained knowledge of how each help option or method may assist the students in their listening learning and we have known that they did contribute to some extent to the students' improvement in EFL listening, but obviously the isolated effects of each have not been measured yet. Therefore, the question of to what extent a certain help option or a pedagogic decision underlying the design of Listening Hacked influenced the students' EFL listening comprehension remains unknown. It is also difficult to conclude whether the use of automated corrective feedback, for example, was contributing to the students' listening improvement more or less than that of Phrase Search. Therefore, further work needs to be done to determine the effects of separate methods employed in this study on L2 listening comprehension by isolating the effects of each method or help option. A possible way is to include more groups of participants in the research design, and each group will be allowed to use only one method or help option in their learning. Such investigations will allow the comparisons of certain help options in assisting L2 listening learning.

Fourthly, due to the experimental nature of the research, the design of the app had to impose much control over the students' capability of using some features of the app, which reduced some affordances of the app and create trouble when using it. For example, in order to collect data on the students' behaviors of using the Repeat/Rewind functions, the app was programed to provide predefined Repeat/Rewind options such as exact repetition of the chunk, 1-minute rewind, 2-minute rewind, etc. and thus it did not provide a flexible playback controller or slider as in most online multimedia players. Some students complained that this caused impatience and frustration in their learning, and possibly affected the learning outcomes, hence affecting the measurement of the dependent variable in this study. Moreover, some technical issues in the operations of the app were another source of weakness in this study which could also have affected the measurement of the effects. While the movie data in the app contained 257 movies with 266 hours of continuous viewing, some participants reported that it was still insufficient. The inadequate quantity of movies in the database may not only restrict the students' choices of learning materials but also could directly affect the operation and quality of some features of the app such as Phrase Search or Listening Boosters. Other technical

problems in feedback generation or chunk selection mechanisms could also interfere with the students' learning. These above issues pose an important practical implication for future investigations into measuring the effects of technology-based solutions like Listening Hacked. The challenge now is to collect only important data without restricting any affordances. More importantly, greater efforts are needed to ensure smooth operation of the proposed software and to minimize the technical issues occurring during experiments.

Finally, this research made an initial step in promoting the use of AI and verifying the effectiveness of ICALL applications for L2 listening comprehension, notably in optimizing and personalizing the provision of corrective feedback and follow-up listening practice. This could be a fruitful area for future work. Further research is strongly recommended to expand the applications of AI in L2 listening instructions such as in personalizing listening materials for individual learners. For instance, the CALL software can collect data on the learner's L2 listening performance and listening topics of interest, then automatically suggests appropriate listening contents for them based on the performance and listening preference data. Alternatively, AI may also be employed to assist the learner in using captions judiciously in listening learning. For example, the software may collect and analyze the data on the learner's errors, then it can use a prediction model to optimize the display and timing of captions in listening learning for individual learners like the partial and synchronized captioning approach proposed by Mizaei et al (2017, 2018). This line of research is a necessary step in establishing the right direction which will move forward the field of L2 listening teaching and learning.

6.4 Closing remarks

This thesis has taken a journey in exploring the potential of using AI-powered learning systems in personalizing and assisting the learning of EFL listening comprehension. Much has been illuminated about the effectiveness of such a learning platform, the students' learning behaviors and their strategies for using the affordances on the system, and their opinions of the system. All in all, the results of this study indicate that the development and implementation of technology-based solutions,

especially with the use of AI in CALL applications, for L2 listening learning and teaching are really encouraging. Much work still needs to be done in research and development in order to produce and implement more successful ICALL applications for L2 listening, however, researchers and practitioners can consider this study as a guideline and a source of inspiration for taking further steps in creating comprehensive and practical solutions to L2 listening learning and teaching.

REFERENCES

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of Familiar and Unfamiliar Native Accents Under Adverse Listening Conditions. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 520–529. https://doi.org/10.1037/a0013552

Aljohani, M. (2017). Principles of "Constructivism" in Foreign Language Teaching. *Journal of Literature and Art Studies*, *7*(1), 97–107. https://doi.org/10.17265/2159-5836/2017.01.013

Anderson, N. J. (2002). The Role of Metacognition in Second Language Teaching and Learning. ERIC Digest, 2–5.

Attali, Y., Bridgeman, B., & Trapani, C. (2010). Performance of a generic approach in automated essay scoring. *The Journal of Technology, Learning and Assessment*, *10*(3), 1–17.

Azizinezhad, M., & Hashemi, M. (2013). A Look at the Status of Computer Assisted Language Learning and its Applications. *Procedia - Social and Behavioral Sciences*, *93*, 121–124. https://doi.org/10.1016/j.sbspro.2013.09.163

B2 First General. (2021). https://www.cambridgeenglish.org/exams-and-tests/first/exam-format/

Baddeley, A. (1992). Working Memory: The Interface between Memory and Cognition. *Journal of Cognitive Neuroscience*, *4*(3), 281–288. https://doi.org/10.1162/jocn. 1992.4.3.281

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Barcroft, J., & Sommers, M. S. (2005). Effects of acoustic variability on second language vocabulary learning. *Studies in Second Language Acquisition*, *27*(3), 387–414. https://doi.org/10.1017/S0272263105050175

Barriuso, T. A., & Hayes-harb, R. (2018). High Variability Phonetic Training as a Bridge

From Research to Practice. *The CATESOL Journal*, *30*(1), 177–194.

Barriuso, T. A., & Hayes-Harb, R. (2018). High Variability Phonetic Training as a Bridge From Research to Practice. In *The CATESOL Journal* (Vol. 30).

Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python - Analyzing Text with the Natural Language Toolkit* (1st ed.). O'Reilly Media. http://www.nltk.org/book_1ed/

Bitchener, J. (2012). Written Corrective Feedback for L2 Development: Current Knowledge and Future Research. *TESOL Quarterly*, *46*(4), 855–860. https://doi.org/10.1002/tesq.62

Block, D. (2000). Revisiting the gap between SLA researchers and language teachers. *Links & Letters*, *7*, 129–143.

Bloomfield, A., Wayland, S. C., Rhoades, E., Linck, J., & Ross, S. (2010). What makes listening difficult? Factors affecting second language listening comprehension. In *Maryland University College Park* (Issue July 2016). University of Maryland, Center for Advanced Study of Language. file:///C:/Users/HOME/Downloads /CASL_study_FINAL_Lit_Rev (4).pdf

Bodner, G. M. (1986). CONSTRUCTIVISM: A THEORY OF KNOWLEDGE. *Journal of Chemical Education*, *63*, 873–878.

Botarleanu, R.-M., Dascalu, M., Sirbu, M.-D., Crossley, S. A., & Trausan-Matu, S. (2019). *ReadME – Generating Personalized Feedback for Essay Writing Using the ReaderBench Framework*. 133–145. https://doi.org/10.1007/978-3-319-92022-1_12

Bouillon, P., Cervini, C., & Rayner, M. (2016). Translation and technology: The case of translation games for language learning. In *The Routledge Handbook of Language Learning and Technology* (pp. 536–549). https://doi.org/10.4324/9781315657899

Bowles, M. (2005). *Effects of verbalization condition and type of feedback on L2 development in a CALL task*. Georgetown University.

Bozorgian, H. (2014). The Role of Metacognition in the Development of EFL Learners'

Listening Skill. *International Journal of Listening*, *28*(3), 149–161. https://doi.org/10.1080/10904018.2013.861303

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception and Psychophysics*, *61*(2), 206–219. https://doi.org/10.3758/BF03206883

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English / r / and / l /: IV. Some effects of perceptual learning on speech production . *The Journal of the Acoustical Society of America*, *101*(4), 2299–2310. https://doi.org/10.1121/1.418276

Brau, B. (n.d.). Constructivism. In R. Kimmons (Ed.), *Navigating Education Research*. https://educationresearch.pressbooks.com/chapter/constructivism/

Brown, H. D. (2007). *Principles of language learning and teaching* (5th ed.). Pearson Education.

Brownlee, J. (2018). A Gentle Introduction to the Bootstrap Method. *Machine Learning Mastery*. https://machinelearningmastery.com/a-gentle-introduction-to-the-bootstrap-method/

Brunfaut, T., & Rukthong, A. (2018). Integrative Testing of Listening. *The TESOL Encyclopedia of English Language Teaching*, 1–7. https://doi.org/10.1002/9781118784235.eelt0622

Buck, G. (2001). *Assessing listening*. Cambridge University Press.

Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The Criterion online writing system. *AI Magazine*, *25*(3), 27–36.

Cambridge English. (n.d.). *Quality and accountability*. Retrieved July 28, 2020, from www.cambridgeenglish.org/research-and-validation/quality-and-accountability/

Cardenas-Claros, M. S., & Gruba, P. A. (2013). Decoding the “CoDe”: a framework for conceptualizing and designing help options in computer-based second language listening. *ReCALL*, *25*(2), 250–271. https://doi.org/10.1017/S0958344013000049

Cárdenas-Claros, M.S. (2020). Spontaneous links between help option use and input features that hinder second language listening comprehension. *System*, *93*. https://doi.org/10.1016/j.system.2020.102308

Cárdenas-Claros, Monica S., & Gruba, P. A. (2009). Help Options in CALL: A Systematic Review. *CALICO Journal*, *27*(1), 69–90. https://doi.org/10.11139/cj.27.1.69-90

Cárdenas-Claros, Mónica Stella. (2011). *A preliminary framework of help options in computer-based second language listening* [The University of Melbourne]. https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=cat00006a&AN=melb.b4168873&scope=site%0Ahttp://repository.unimelb.edu.au/10187/10042

Cárdenas-Claros, Mónica Stella, & Gruba, P. A. (2014). Listeners' interactions with help options in CALL. *Computer Assisted Language Learning*, *27*(3), 228–245. https://doi.org/10.1080/09588221.2012.724425

Chai, J., & Erlam, R. (2008). The Effect and the Influence of the Use of Video and Captions on Second Language Learning. *New Zealand Studies in Applied Linguistics*, *14*(2), 25–44.

Chamot, A. U., O'Malley, J. M., & Chamot, A. U. (1990). Learning strategies in second language acquisition. In *Cambridge, CUP*. Cambridge University Press. https://doi.org/10.2307/329782

Chamot, A. U., O'Malley, J. M., & Chamot, A. U. (1996). The Cognitive Academic Language Learning Approach: A Model for Linguistically Diverse Classrooms. *The Elementary School Journal*, *96*(3), 259–273. https://doi.org/10.1086/461827

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*, 293–332.

Chang, A. C.-S., & Read, J. (2006). The Effects of Listening Support on the Listening Performance of EFL Learners. *TESOL Quarterly*, *40*(2), 375–397.

Chapelle, C. A. (2003). *English Language Learning and Technology: Lectures on Applied Linguistics in the Age of Information and Communication Technology*. John Benjamins.

Chapelle, C. A., Cotos, E., & Lee, J. (2015). Validity arguments for diagnostic assessment using automated writing evaluation. *Language Testing*, *32*(3), 385–405. https://doi.org/10.1177/0265532214565386

Chen, C. M., & Chung, C. J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers and Education*, *51*(2), 624–645. https://doi.org/10.1016/j.compedu.2007.06.011

Chen, C. M., & Li, Y. L. (2010). Personalised context-aware ubiquitous learning system for supporting effective english vocabulary learning. *Interactive Learning Environments*, *18*(4), 341–364. https://doi.org/10.1080/10494820802602329

Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). Twenty Years of Personalized Language learning: Topic Modeling and Knowledge Mapping. *Educational Technology and Society*, *24*(1), 205–222.

Chon, Y. V., & Shin, T. (2019). Profile of second language learners' metacognitive awareness and academic motivation for successful listening: A latent class analysis. *Learning and Individual Differences*, *70*, 62–75. https://doi.org/10.1016/j.lindif.2019.01.007

Chung, J. M. (1999). The effects of using video texts supported with advance organizers and captions on Chinese college students' listening comprehension: An empirical study. *Foreign Language Annals*, *32*(3), 295–308. https://doi.org/10.1111/j.1944-9720.1999.tb01342.x

Clancy, Ú., & Murray, L. (2016). CALL tools for listening and speaking. In *The Routledge Handbook of Language Learning and Technology* (pp. 491–508). https://doi.org/10.4324/9781315657899

Clopper, C. G., & Pisoni, D. B. (2004). Effects of talker variability on perceptual learning of dialects. *Language and Speech*, *47*(3), 207–239. https://doi.org/10.1177/00238309040470030101

Cole, P. (1992). Constructivism revisited: A search for common ground. *Educational Technology*, *33*(2), 27–34.

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, *32*, 45–61. https://doi.org/10.1017/S0267190512000074

Cooper, C. (2002). *Individual differences* (2nd ed.). Arnold.

Creswell, J. D. W., & Creswell, J. D. W. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (Fifth). SAGE Publications.

Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (Fourth). Pearson Education.

Cross, J. (2018). Video in Listening. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–6). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118784235.eelt0606

D'Souza, J. (2018). *Learning POS Tagging & Chunking in NLP*. https://medium.com/greyatom/learning-pos-tagging-chunking-in-nlp-85f7f811a8cb

d'Ydewalle, G., & Pavakanun, U. (1997). Could Enjoying a Movie Lead to Language Acquisition? *New Horizons in Media Psychology*, 145–155. https://doi.org/10.1007/978-3-663-10899-3_10

Dafei, D. (2007). An exploration of the relationship between learner autonomy and English proficiency. *Asian EFL Journal*, *24*(4), 1–23. http://www.asian-efl-journal.com/pta_Nov_07_dd.pdf

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, *20*(6), 445–464. https://doi.org/10.1007/BF01067637

De Raad, B. (2000). Differential psychology. In *Encyclopedia of psychology* (pp. 41–44). American Psychological Association and Oxford University Press.

DeKeyser, R. (2012). Age effects in second language learning. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 442–460). Routledge.

Diao, Y., Chandler, P., & Sweller, J. (2007). The effect of written text on comprehension of spoken English as a foreign language. *American Journal of Psychology*, *120*(2), 237–261. https://doi.org/10.2307/20445397

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning and Technology*, *24*(1), 16–26. https://doi.org/10.125/44705

Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Lawrence Erlbaum Associates Publishers.

Dörnyei, Zoltán. (1998). Motivation in second and foreign language learning. *Language Teaching*, *31*(3), 117–135. https://doi.org/10.1017/S026144480001315X

Dudzik, D. L., & Nguyen, Q. T. N. (2015). Vietnam: Building English Competency in Preparation for ASEAN 2015. In R. Stroupe & K. Kimura (Eds.), *ASEAN Integration and the Role of English Language Teaching* (pp. 41–71). IELTS.

Duffy, T. M., & Jonassen, D. H. (1992). Constructivism: New implications for instructional technology. In T. H. Duffy & D. H. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. 1–16). Lawrence Erlbaum Associates.

Duong, T. M., & Chau, N. T. (2019). English listening comprehension problems perceived by English majors at The Saigon International University. In T. T. Dang (Ed.), *Proceedings of International Conference on Language Teaching and Learning Today 2019: Autonomy and Motivation for Language Learning in the Interconnected World* (pp. 209–222). Vietnam National University-Ho Chi Minh City Press.

Dupuy, B. C. (1999). Narrow Listening: An alternative way to develop and enhance listening comprehension in students of French as a foreign language. *System*, *27*(3), 351–361. https://doi.org/10.1016/S0346-251X(99)00030-5

El-Ebyary, K., & Windeatt, S. (2010). The impact of computer-based feedback on students' written work. *International Journal of English Studies*, *10*(2), 121–142.

Elliott, M. (2013). Test taker characteristics. In A. Geranpayeh & L. Taylor (Eds.), *Examining Listening: Research and Practice in Second Language Listening (Studies in Language Testing, volume 35),* (pp. 36–76). Cambridge University Press.

Elliott, M., & Chismholm, A. (2015). Revising FCE and CAE Listening tests. *Research Notes*, *62*, 21–23.

Ellis, R. (2010). Second language acquisition, teacher education and language pedagogy. *Language Teaching*, *43*(2), 182–201. https://doi.org/10.1017/S0261444809990139

Ellis, R. (2017). Oral corrective feedback in L2 classrooms: What We Know so Far. In H. Nassaji & E. Kartchava (Eds.), *Corrective Feedback in Second Language Teaching and Learning: Research, Theory, Applications, Implications* (pp. 3–18). Routledge.

Ellis, R., & Le, C. (2016). The Effects of Inference-Training and Text Repetition on Chinese Learners' Incidental Vocabulary Acquisition While Listening. *Chinese Journal of Applied Linguistics*, *39*(2). https://doi.org/10.1515/cjal-2016-0009

Enyedy, N. (2014). Personalized Instruction: New Interest, Old Rhetoric, Limited Results, and the Need for a New Direction for Computer-Mediated Learning. *National Education Policy Center*, *I*, 22.

Erlam, R. M. (2008). What Do You Researchers Know about Language Teaching? Bridging the Gap Between SLA Research and Language Pedagogy. *Innovation in Language Learning and Teaching*, *2*(3), 253. https://doi.org/10.2167/illt072b.0

Fang, L., Tuan, L. A., Hui, S. C., & Wu, L. (2018). Personalized question recommendation for English grammar learning. *Expert Systems*, *35*(2). https://doi.org/10.1111/exsy.12244

Ferguson, S. H., Jongman, A., Sereno, J. A., & Keum, K. A. (2010). Intelligibility of foreign-accented speech for older adults with and without hearing loss. *Journal of the American Academy of Audiology*, *21*(3), 153–162. https://doi.org/10.3766/jaaa.21.3.3

Field, J. (1999). Bottom-up'and 'top-down. *ELT Journal*, *53*(4), 338–339.

Field, J. (2003). Promoting perception: Lexical segmentation in L2 listening. *ELT Journal*, *57*(4), 325–334. https://doi.org/10.1093/elt/57.4.325

Field, J. (2004). An insight into listeners' problems: Too much bottom-up or too much top-down? *System*, *32*(3), 363–377. https://doi.org/10.1016/j.system.2004.05.002

Field, J. (2008a). Bricks or Mortar: Which Parts of the Input Does a Second Language Listener Rely On? *TESOL Quarterly*, *42*(3), 411–432. https://doi.org/10.1002/j.1545-7249.2008.tb00139.x

Field, J. (2008b). Guest editor's introduction Emergent and divergent: A view of second language listening research. *System*, *36*, 2–9.

Field, J. (2008c). *Listening in the language classroom*. Cambridge University Press. https://doi.org/10.1093/elt/ccq026

Field, J. (2008d). *The L2 Listener : Type or Individual ?* (Working Papers in English and Applied Linguistics in Honour of Gillian Brown).

Field, J. (2009). *Listening in the Language Classroom*. Cambridge University Press. https://doi.org/10.1017/cbo9780511575945.017

Field, J. (2011). Into the mind of the academic listener. *Journal of English for Academic Purposes*, *10*(1), 102–112.

Finardi, K. (2006). Working memory capacity and speech production in L2: evidences from a picture description task. *Rev. Est. Ling.*, *14*(1), 231–260.

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1276–1293. https://doi.org/10.1037/0096-1523.32.5.1276

Flowerdew, J., & Miller, L. (2005). *Second language listening: Theory and practice*. Cambridge University Press.

Flowerdew, John, & Miller, L. (1992). Student Perceptions, Problems and Strategies in Second Language Lecture Comprehension. *RELC Journal*, *23*(2), 60–80. https://doi.org/10.1177/003368829202300205

Fontes, T. O., & O'Mahony, M. (2008). In-depth interviewing by Instant Messaging. *Social Research Update*, *53*, 1–4.

Gass, Susan, & Varonis, E. M. (1984). the Effect of Familiarity on the Comprehensibility of Nonnative Speech. *Language Learning*, *34*(1), 65–87. https://doi.org/10.1111/j.1467-1770.1984.tb00996.x

Gass, Susan, Winke, P., Isbell, D. R., & Ahn, J. (2019). How captions help people learn languages: A working-memory, eye-tracking study. *Language Learning and Technology*, *23*(2), 84–104. https://doi.org/https://doi.org/10125/44684

Gass, Sussan. (1997). *Input, interaction, and the second language learner*. Erlbaum.

Geranpayeh, A. (2004). Reability in First Ceritificate in English objective papers. *Research Notes*, *15*, 21–23.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Lawrence Erlbaum Associates, Inc.

Gernsbacher, M. A. (1991). Cognitive Processes and Mechanisms in Language Comprehension: The Structure Building Framework. *Psychology of Learning and Motivation - Advances in Research and Theory*, *27*(C), 217–263. https://doi.org/10.1016/S0079-7421(08)60125-5

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes*, *23*(3), 265–304. https://doi.org/10.1080/01638539709544994

Gernsbacher, M. A., & Faust, M. E. (1991). The Mechanism of Suppression: A Component of General Comprehension Skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(2), 245–262. https://doi.org/10.1037/0278-7393.17.2.245

Gilman, R. A., & Moody, R. L. M. (1984). What Practitioners Say About Listening: Research Implications for the Classroom. *Foreign Language Annals*, *17*(4), 331–334. https://doi.org/10.1111/j.1944-9720.1984.tb03236.x

Goctu, R. (2017). The Effects of Motivation on Listening Skills of ELT Students in Georgia (IBSU Case). *Smart Moves Journal Ijellh*, *4*(5), 15. https://doi.org/10.24113/ijellh.v4i5.1342

Goh, C. (2016). Cognition, Metacognition, and L2 Listening. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (3rd ed., pp. 214–228). Routledge.

Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, *28*(1), 55–75. https://doi.org/10.1016/

S0346-251X(99)00060-3

Gowhary, H., Pourhalashi, Z., Jamalinesari, A., & Azizifar, A. (2015). Investigating The Effect Of Video Captioning On Iranian EFL Learners' Listening Comprehension. *Procedia - Social and Behavioral Sciences*, *192*, 205–212.

Graham, S., & Macaro, E. (2008). Strategy instruction in listening for low-intermediate learners of French. *Language Learning*, *58*(4), 7470783.

Graham, Suzanne. (2006). Listening comprehension: The learners' perspective. *System*, *34*(2), 165–182. https://doi.org/10.1016/j.system.2005.11.001

Grgurović, M., & Hegelheimer, V. (2007). Help options and multimedia listening: Students' use of subtitles and the transcript. *Language Learning and Technology*, *11*(1), 45–66.

Gruba, P. (2018). Technology for Teaching Listening. In *The TESOL Encyclopedia of English Language Teaching*. John Wiley & Sons, Inc.

Guichon, N., & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. *System*, *36*(1), 85–93. https://doi.org/10.1016/j.system.2007.11.005

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

Han, Z. (2008). Fine-tuning Corrective Feedback. *FOREIGN LANGUAGE ANNALS*, *34*(6), 582–599.

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, *24*(4), 495–522. https://doi.org/10.1017/S0142716403000250

Haristiani, N., Danuwijaya, A. A., Rifai, M. M., & Sarila, H. (2019). Gengobot: A chatbot-based grammar application on mobile instant messaging as language learning medium. *Journal of Engineering Science and Technology*, *14*(6), 3158–3173.

Hasan, A. S. (2000). Learners' perceptions of listening comprehension problems. *Language, Culture and Curriculum*, *13*(2), 137–153. https://doi.org/10.1080/07908310008666595

Hassan, Z. A., Schattner, P., & Mazza, D. (2006). Doing A Pilot Study: Why Is It Essential? *Malaysian Family Physician : The Official Journal of the Academy of Family Physicians of Malaysia*, *1*(2–3), 70–73. http://www.ncbi.nlm.nih.gov/pubmed/27570591%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4453116

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112.

Hayati, A., & Mohmedi, F. (2011). The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology*, *42*(1), 181–192. https://doi.org/10.1111/j.1467-8535.2009.01004.x

He, B., Sangarun, P., & Lian, A.-P. (2015). Improving the English pronunciation of Chinese EFL university students through the integration of CALL and verbotonalism. In J. Colpaert, A. Aerts, M. Oberhofer, & M. G.-C. Plana (Eds.), *Seventeenth International CALL Research Conference: Task Design and CALL* (pp. 276–285). University of Antwerpen. http://wwwa.fundacio.urv.cat/congressos/public/usr_docs/call_2015_conference_proceedings.pdf

He, Bi. (2014). *Improving the English pronunciation of Chinese EFL learners through the integration of CALL and verbotonalism*. Suranaree University of Technology.

Heift, T. (2012). Intelligent Computer-Assisted Language Learning. In *The Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd. https://doi.org/10.1002/9781405198431.wbeal0548

Heift, T., & Hegelheimer, V. (2017). Computer-assisted corrective feedback and language learning. In H. Nassaji & E. Kartchava (Eds.), *Corrective Feedback in Second Language Teaching and Learning: Research, Theory, Applications, Implications* (pp. 51–65). Routledge.

Heift, T., & Rimrott, A. (2008). Learner responses to corrective feedback for spelling errors in CALL. *System*, *36*(2), 196–213. https://doi.org/10.1016/j.system.2007.09.007

Hendry, G. D. (1996). Constructivism and educational practice. *Australian Journal*

OfEducation, *40*(1), 19–45. https://doi.org/10.1177/000494419604000103

Hoang, V. V. (2010). The current situations and issues of teaching of English in Vietnam. *Ritsumeikan Studies in Language and Culture*, *22(1)*, 7–18. http://r-cube.ritsumei.ac.jp/bitstream/%0A10367/4129/1/LCS_22_1pp7-18_HOANG.pdf

Holec, H. (1981). *Autonomy in Foreign Language Learning*. Pergamon.

Hooshyar, D., Yousefi, M., Wang, M., & Lim, H. (2018). A data-driven procedural-content-generation approach for educational games. *Journal of Computer Assisted Learning*, *34*(6), 731–739. https://doi.org/10.1111/jcal.12280

Hsu, C. K., Hwang, G. J., & Chang, C. K. (2013). A personalized recommendation-based mobile learning approach to improving the reading performance of EFL students. *Computers and Education*, *63*, 327–336. https://doi.org/10.1016/j.compedu.2012.12.004

Huang, H.-C., & Eskey, D. E. (1999). The Effects of Closed-Captioned Television on the Listening Comprehension of Intermediate English as a Second Language (ESL) Students. *Journal of Educational Technology Systems*, *28*(1), 75–96. https://doi.org/10.2190/rg06-lywb-216y-r27g

Huang, Y. M., Huang, Y. M., Huang, S. H., & Lin, Y. T. (2012). A ubiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use. *Computers and Education*, *58*(1), 273–282. https://doi.org/10.1016/j.compedu.2011.08.008

Hubbard, P. (1988). An Integrated Framework for CALL Courseware Evaluation. *CALICO Journal*, *6*(2), 51–72.

Hubbard, P. (2006). Evaluating CALL Software. In L. Ducate & N. Arnold (Eds.), *Calling on CALL: From Theory and Research to New Directions in Foreign Language Teaching* (pp. 1–26). CALICO.

Hubbard, P. (2010). Listening to Learn: New Opportunities in an Online World. In C. Ward (Ed.), *RELC Anthology #51. The Impact of Technology on Language Learning and Teaching: What, How and Why*. RELC Publications.

Hubbard, P. (2017). Technologies for Teaching and Learning L2 Listening. In *The Handbook of Technology and Second Language Teaching and Learning* (pp. 93–106). John Wiley & Sons, Inc.

Isabel Borras, & Robert C, L. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *The Modern Language Journal*, *78*.

James, C. (2006). Validating a computerized scoring system for assessing writing and placing students in composition courses. *Assessing Writing*, *11*(3), 167–178.

Jia, C., & Hew, K. F. T. (2019). Supporting lower-level processes in EFL listening: the effect on learners' listening proficiency of a dictation program supported by a mobile instant messaging app. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2019.1671462

Johnson, K. (2008). *An introduction to foreign language learning and teaching* (2nd ed.). Routledge.

Johnson, Keith. (1997). Speech perception without speaker normalization: An exemplar model. In Keith Johnson & J. W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 145–165). Academic Press.

Johnsrude, I. S., & Buchsbaum, B. R. (2017). Representation of speech. In M. G. Gaskell & J. Mirković (Eds.), *Speech perception and spoken word recognition*. Routledge.

Kang, E., & Han, Z. (2015). The Efficacy of Written Corrective Feedback in Improving L2 Written Accuracy: A Meta-Analysis. *The Modern Language Journal*, *99*(1), 1–18. https://doi.org/10.1111/modl.12189

Karagiorgi, Y., & Symeou, L. (2005). Translating Constructivism into Instructional Design: Potential and Limitations. *Educational Technology & Society*, *8*(1), 17–27.

Kaufmann, K., & Peil, C. (2019). The mobile instant messaging interview (MIMI): Using WhatsApp to enhance self-reporting and explore media usage in situ. *Mobile Media & Communication*, 205015791985239. https://doi.org/10.1177/2050157919852392

Kim, H.-S. (2015). Using Authentic Videos to Improve EFL Students' Listening Comprehension. *International Journal of Contents*, *11*(4), 15–24.

Kim, J. H. (2003). Issues of Corrective Feedback in Second Language Acquisition. *Working Papers in TESOL & Applied Linguistics, 4*(2), 1–24.

Knight, B. (2018). How long does it take to learn a foreign language. In *Experience Better Learning* (Issue january 2018). Cambridge University Press.

Kraemer, H. C., Mintz, J., Noda, A., Tinklenberg, J., & Yesavage, J. A. (2006). Caution regarding the use of pilot studies to guide power calculations for study proposals. *Archives of General Psychiatry*, *63*(5), 484–489. https://doi.org/10.1001/archpsyc.63.5.484

Krashen, S. D. (1996). The case for narrow listening. *System*, *24*(1), 97–100.

Kulesa, A., Krzywinski, M., Blainey, P., & Altman, N. (2015). Points of Significance: Sampling distributions and the bootstrap. *Nature Methods*, *12*(6), 477–478. https://doi.org/10.1038/nmeth.3414

Kuo, F. (2012). Factors Affecting Chinese EFL Learners ' Spoken Word Recognition. *NCUE Journal of Humanities*, *6*, 1–14.

Lado, R. (1961). *Language Testing: the construction and use of foreign language tests*. Longman.

Lan, Y.-J. (2015). Contextual EFL learning in a 3D virtual environment. *Language Learning & Technology*, *19*(2), 16–31.

Lan, Y.-J., & Liao, C.-Y. (2017). The effects of 3D immersion on CSL students' listening comprehension. *Innovation in Language Learning and Teaching*, *12*(1), 35–46. https://doi.org/10.1080/17501229.2018.1418242

LaScotte, D. K. (2015). Listening in context: incorporating strategy instruction in L2 listening practice. *MinneTESOL Journal*, *31*(1).

Lecumberri, M. L. G., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *The Journal of the Acoustical Society of America*, *119*(4), 2445–2454. https://doi.org/10.1121/1.2180210

Lee, A. H., & Lyster, R. (2015). The effects of corrective feedback on instructed L2

speech perception. *Studies in Second Language Acquisition*, *38*(1), 35–64. https://doi.org/10.1017/S0272263115000194

Lee, A. H., & Lyster, R. (2016). Effects of Different Types of Corrective Feedback on Receptive Skills in a Second Language: A Speech Perception Training Study. *Language Learning*, *66*(4), 809–833. https://doi.org/10.1111/lang.12167

Lee, S. (2013). Korean University Students' Listening Comprehension and Learning Effects of Connected Speech Processes in English. *Korean Journal of Applied Linguistics*, *29*(3), 133. https://doi.org/10.17154/kjal.2013.09.29.3.133

Lee, S. P., Su, H. K., & Lee, S. Da. (2012). Effects of computer-based immediate feedback on foreign language listening comprehension and test-associated anxiety. *Perceptual and Motor Skills*, *114*(3), 995–1006. https://doi.org/10.2466/28.11.21.PMS.114.3.995-1006

Leonard, K. R. (2019). Examining the relationship between decoding and comprehension in L2 listening. *System*, *87*. https://doi.org/10.1016/j.system.2019.102150

Leong, C. X. R., Price, J. M., Pitchford, N. J., & van Heuven, W. J. B. (2018). High variability phonetic training in adaptive adverse conditions is rapid, effective, and sustained. *PLOS ONE*, *13*(10), e0204888. https://doi.org/10.1371/journal.pone.0204888

Levak, N., & Son, J. (2016). Facilitating second language learners' listening comprehension with Second Life and Skype. *ReCALL*, *29*(2), 200–218. https://doi.org/10.1017/S0958344016000215

Leveridge, A. N., & Yang, J. C. (2013). Testing learner reliance on caption supports in second language listening comprehension multimedia environments. *ReCALL*, *25*(2), 199–214. https://doi.org/10.1017/S0958344013000074

Leveridge, A. N., & Yang, J. C. (2014). Learner perceptions of reliance on captions in EFL multimedia listening comprehension. *Computer Assisted Language Learning*, *27*(6), 545–559. https://doi.org/10.1080/09588221.2013.776968

Li, L., Xu, H., & Zhang, X. (2020). Exploring the role of phraseological knowledge and

syntactic knowledge in L2 listening comprehension. *Lingua*, *248*. https://doi.org/10.1016/j.lingua.2020.102957

Li, S. (2010). The Effectiveness of Corrective Feedback in SLA: A Meta-Analysis. *Language Learning*, *60*(2), 309–365. https://doi.org/10.1111/j.1467-9922.2010.00561.x

Li, S. (2017). Student and teacher beliefs and attitudes about oral corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *Corrective Feedback in Second Language Teaching and Learning: Research, Theory, Applications, Implications* (pp. 143–157). Routledge.

Li, S., Zhu, Y., & Ellis, R. (2016). The Effects of the Timing of Corrective Feedback on the Acquisition of a New Linguistic Structure. *The Modern Language Journal*, *100*(1), 276–295. https://doi.org/10.1111/modl.12315

Li, Z., Link, S., Ma, H., Yang, H., & Hegelheimer, V. (2014). The role of automated writing evaluation holistic scores in the ESL classroom. *System*, *44*(1), 66–78. https://doi.org/10.1016/j.system.2014.02.007

Lian, A.-P., & Sangarun, P. (2017). Precision Language Education: A Glimpse Into a Possible Future. *GEMA Online® Journal of Language Studies*, *17*(4), 1–15. https://doi.org/https://doi.org/10.17576/gema-2017-1704-01

Lian, A. P. (2014). ON-DEMAND GENERATION OF INDIVIDUALISED LANGUAGE LEARNING LESSONS. *Journal of Science Ho Chi Minh City Open University*, *9*(1), 25–38.

Littlejohn, A. (2008). Digging deeper: learners' disposition and strategy use. In *Strategies in Language Learning and Teaching*. SEAMEO Regional Language Centre.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, *94*(3), 1242–1255. https://doi.org/10.1121/1.408177

Lo, J. J., Yeh, S. W., & Sung, C. S. (2013). Learning paragraph structure with online annotations: An interactive approach to enhancing EFL reading comprehension. *System*, *41*(2), 413–427. https://doi.org/10.1016/j.system.

2013.03.003

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese Listeners To Identify English /R/ And /1/: A First Report. *Journal of the Acoustical Society of America*, *89*(2), 874–886. https://doi.org/10.1121/1.1894649

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Long, Michael H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12*, 251–286.

Lwo, L., & Lin, M. (2012). The effects of captions in teenagers' multimedia L2 learning. *ReCALL*, *24*(2), 188–208.

Lyddon, W. J., & McLaughlin, J. T. (1992). Constructivist psychology: A heuristic framework. *Journal of Mind and Behaviour*, *13*(1), 89–108.

Lyster, R., & Saito, K. (2010). ORAL FEEDBACK IN CLASSROOM SLA. *Studies in Second Language Acquisition1*, *32*, 265–302. https://doi.org/10.1017/S0272263109990520

Lyster, Roy. (2018). Roles for Corrective Feedback in Second Language Instruction. In *The Encyclopedia of Applied Linguistics* (pp. 1–6). Wiley. https://doi.org/10.1002/9781405198431.wbeal1028.pub2

Macaro, E., Graham, S., & Vanderplank, R. (2007). A review of listening strategies: focus on sources of knowledge and on success. *Language Learner Strategies*, 165–186.

Majuddin, E., Siyanova-Chanturia, A., & Boers, F. (2021). Incidental acquisition of multiword expressions through audiovisual materials. *Studies in Second Language Acquisition*. https://doi.org/10.1017/S0272263121000036

Markham, P. L., Peter, L. A., & McCarthy, T. J. (2001). The effects of native language vs. target language captions on foreign language students' DVD video comprehension. *Foreign Language Annals*, *34*(5), 439–445. https://doi.org/10.1111/j.1944-9720.2001.tb02083.x

Mascolo, M. F., & Fischer, K. W. (2005). Constructivist theories. In *Cambridge*

Encyclopedia of Child Development (pp. 49–63). Cambridge University Press.

Matthews, J. (2018). Vocabulary for listening: Emerging evidence for high and mid-frequency vocabulary knowledge. *System*, *72*, 23–36. https://doi.org/10.1016/j.system.2017.10.005

Matthews, J., & O'Toole, J. M. (2015). Investigating an innovative computer application to improve L2 word recognition from speech. *Computer Assisted Language Learning*, *28*(4), 364–382. https://doi.org/10.1080/09588221.2013.864315

Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). Cambridge University Press.

Mayer, R. E., & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, *38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

McCarthy, J. (2007). *What is artificial intelligence*. http://www-formal.stanford.edu/jmc/whatisai/whatisai.html

McDonough, S. (2006). Learner strategies: an interview with Steven McDonough. *ELT Journal*, *60*(1), 63–70.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746--748.

McNamara, D. S., & Magliano, J. (2009). Chapter 9 Toward a Comprehensive Model of Comprehension. In B. H. Ross (Ed.), *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 51, pp. 297–384). https://doi.org/10.1016/S0079-7421(09)51009-2

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. https://doi.org/10.1017/S0958344017000039

Mirzaei, M. S., Meshgi, K., & Kawahara, T. (2018). Exploiting automatic speech recognition errors to enhance partial and synchronized caption for facilitating second language listening. *Computer Speech and Language*, *49*, 17–36. https://doi.org/10.1016/j.csl.2017.11.001

Mitterer, H., & McQueen, J. M. (2009). Foreign subtitles help but native-language

subtitles harm foreign speech perception. *PLoS ONE*, *4*(11). https://doi.org/10.1371/journal.pone.0007785

Mohammad Jafari, S. (2010). On the relationship between listening comprehension motivation and listening comprehension among Iranian EFL learners. *Indonesian Journal of English Language Teaching*, *6*(2), 89–106.

Montero Perez, M., Peters, E. and Desmet, P. (2014). Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension. *ReCALL*, *26*(1), 21–43.

Montero Perez, M., Leuven, K. U., Peters, E., & Clarebout, G. (2014). Effects of Captioning on Video Comprehension and Incidental Vocabulary Learning. *Language Learning and Technology*, *18*(1), 118–141.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, *41*(3), 720–739. https://doi.org/10.1016/j.system.2013.07.013

Mordaunt, O. G., & Olson, D. W. (2010). Listen, listen, listen and listen: Building a comprehension corpus and making it comprehensible. *Educational Studies*, *36*(3), 249–258. https://doi.org/10.1080/03055690903172068

Murray, M. C., & Pérez, J. (2015). Informing and performing: A study comparing adaptive learning to traditional learning. *Informing Science*, *18*(1), 111–125. https://doi.org/10.28945/2165

Nassaji, H. (2012). The relationship between SLA research and language pedagogy: Teachers' perspectives. *Language Teaching Research*, *16*(3), 337–365. https://doi.org/10.1177/1362168812436903

Nassaji, H., & Kartchava, E. (2017). *Corrective Feedback in Second Language Teaching and Learning: Research, Theory, Applications, Implications* (H. Nassaji & E. Kartchava (Eds.)). Routledge.

Nassaji, H., & Swain, M. (2000). A Vygotskian Perspective on Corrective Feedback in L2: The Effect of Random Versus Negotiated Help on the Learning of English Articles. *Language Awareness*, *9*(1), 34–51. https://doi.org/10.1080/09658

410008667135

Neisser, U. (1978). Perceiving, Anticipating, and Imagining. *Perception and Cognition Issues in the Foundations of Psychology*, *9*, 89–105.

Neisser, U. (2014). *Cognitive psychology* (Classic Ed). Psychology Press.

Neuman, S., & Koskinen, P. (1992). Captioned television as 'comprehensible input': Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, *27*(1), 95–106.

Nguyen, C.-D., & Newton, J. (2018). Schemata in Listening Comprehension. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–7). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118784235.eelt0592

Nguyen, H., & Abbott, M. L. (2016). Promoting Process-Oriented Listening Instruction in the ESL Classroom. *TESL CANADA JOURNAL*, *34*(11), 72–86. https://doi.org/http://dx.doi.org/1018806/tesl.v34i1.1254

Nguyen, N. Q., & Thai, C. D. (2018). Listening comprehension: First year English-major students' perceptions and problems. *Can Tho University Journal of Science*, *54(2)*, 75. https://doi.org/10.22144/ctu.jen.2018.010

Noreillie, A.-S., Kestemont, B., Heylen, K., Desmet, P., & Peters, E. (2018). Vocabulary knowledge and listening comprehension at an intermediate level in English and French as foreign languages. *ITL - International Journal of Applied Linguistics*, *169*(1), 212–231. https://doi.org/10.1075/itl.00013.nor

Norman, D. A. (1988). *The Psychology of Everyday Things*. Basic Books.

Nunan, D. (2002). Listening in language learning. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 238–241). Cambridge University Press.

Oxford, R. L. (1990). *Language Learning Strategies: What Every Teacher Should Know*. Newbury House.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.

Pearce, G., Thøgersen-Ntoumani, C., & Duda, J. L. (2014). The development of

synchronous text-based instant messaging as an online interviewing tool. *International Journal of Social Research Methodology*, *17*(6), 677–692. https://doi.org/10.1080/13645579.2013.827819

Pereira, A. B., Souza Jr, G. N., Monteiro, D. C., Barros, E. S., Costa, H. P., Nascimento, P. a, Marques, L. B., De Souza, D. G., Salgado, F. M., & Bessa, R. Q. (2012). A AIED Game to help children with learning disabilities in literacy in the Portuguese language. *2012 Brazilian Symposium on Games and Digital Entertainment*, 134–143. http://sbgames.org/sbgames2012/proceedings/papers/computacao/comp-full_17.pdf

Pérez-Segura, J. J., Sánchez Ruiz, R., González-Calero, J. A., & Cózar-Gutiérrez, R. (2020). The effect of personalized feedback on listening and reading skills in the learning of EFL. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2019.1705354

Perrachione, T. K., Lee, J., Ha, L. Y. Y., & Wong, P. C. M. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America*, *130*(1), 461–472. https://doi.org/10.1121/1.3593366

Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, *63*, 134–148. https://doi.org/10.1016/j.system.2016.10.002

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing l2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551–577. https://doi.org/10.1017/S0272263117000407

Petersen, K. (2010). *Implicit corrective feedback in computer-guided interaction. Does mode matter?* Georgetown University.

Phillips, M. D., Lowe, M. J., Lurito, J. T., Dzemidzic, M., & Mathews, V. P. (2001). Temporal lobe activation demonstrates sex-based differences during passive listening. *Radiology*, *220*(1), 202–207. https://doi.org/10.1148/radiology.220.1.r01jl34202

Porter-Szucs, I. (2018). Process Approach Versus Product Approach. *The TESOL*

Encyclopedia of English Language Teaching, 1–7. https://doi.org/10.1002/9781118784235.eelt0612

Price, D., Smith, J. K., & Berg, D. A. G. (2017). Personalised feedback and annotated exemplars in the writing classroom: An experimental study in situ. *Assessment Matters*, *11*, 122–144. https://doi.org/10.18296/am.0027

Pujolà, J. T. (2002). CALLing for help: Researching language learning strategies using help facilities in a web-based multimedia program. *ReCALL*, *14*(2), 235–262. https://doi.org/10.1017/S0958344002000423

Qian, M., Chukharev-Hudilainen, E., & Levis, J. (2018). A system for adaptive high-variability segmental perceptual training: Implementation, effectiveness, transfer. *Language Learning & Technology*, *22*(1), 69–96.

Quera, V., Bakeman, R., & Gnisci, A. (2007). Observer agreement for event sequences: Methods and software for sequence alignment and reliability estimates. *Behavior Research Methods*, *39*(1), 39–49. https://doi.org/10.3758/BF03192842

Quinn, P. (2014). *Delayed versus immediate corrective feedback on orally produced passive errors in English* [University of Toronto]. http://hdl.handle.net/1807/65728

Quinn, P. G., & Nakata, T. (2017). The timing of oral corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *Corrective Feedback in Second Language Teaching and Learning: Research, Theory, Applications, Implications* (pp. 35–47). Routledge.

Rahman, M.M. ; Pandian, A. (2016). The Gap between Research and Practice in the Field of SLA: The Rationale behind It and Bridging the Gap. *Journal of Applied Linguistics and Language Research*. http://www.jallr.com/index.php/JALLR/article/view/294

Ranta, L., & Lyster, R. (2007). A cognitive approach to improving immersion students' oral language abilities: The awareness–practice–feedback sequence. In R. DeKeyser (Ed.), *Practicing for second language use: Perspectives from applied linguistics and cognitive psychology* (pp. 141–160). Cambridge University Press.

Rapaport, W. J. (2005). *In Defense of Contextual Vocabulary Acquisition*. 396–409.

https://doi.org/10.1007/11508373_30

Reinders, H., & Hubbard, P. (2013). CALL and learner autonomy: Aff ordances and constraints. In M. Thomas, H. Reinders, & M. Warschauer (Eds.), *Contemporary Computer-Assisted Language Learning* (pp. 359–375). Bloomsbury Publishing.

Renandya, W. A., & Farrell, T. S. C. (2011). "Teacher, the tape is too fast!" Extensive listening in ELT. *ELT Journal*, *65*(1), 52–59. https://doi.org/10.1093/elt/ccq015

Révész, A., & Brunfaut, T. (2013). Text characteristics of task input and difficulty in second language listening comprehension. *Studies in Second Language Acquisition*, *35*(1), 31–65.

Richards, Jack. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly*, *17*(2), 219–240.

Richards, JC. (2008). *Teaching listening and speaking: From theory to practice*. Cambridge University Press. http://www.learningdomain.com/MEdHOME/Topic6.Theory.Richards-Teaching-Listening-Speaking2.pdf

Richards, K. (2003). Qualitative inquiry in TESOL. In *Qualitative Inquiry in TESOL* (1st ed.). Palgrave Macmillan. https://doi.org/10.1057/9780230505056

Richards, K. (2009). Interviews. In J. Heigham & R. A. Croker (Eds.), *Qualitative Research in Applied Linguistics: A Practical Introduction* (pp. 182–199). Palgrave Macmillan.

Rivens Mompean, A., & Guichon, N. (2009). Assessing the use of aids for a computer-mediated task: taking notes while listening. *Jaltcall*, *5*(2), 45–60. http://nicolas.guichon.pagesperso-orange.fr/_media/bv000019.lkdoc.rivens guichon2009.pdf

Roach, P. (2009). *English Phonetics and Phonology: A practical course* (Fourth). Cambridge University Press.

Robinson, J. (2019). *Phonological variation across the UK*. https://www.bl.uk/british-accents-and-dialects/articles/phonological-variation-across-the-uk

Robots, M. (2019). *What is AI? Some classic definitions*. https://medium.com/moral-robots/what-is-ai-some-classic-definitions-a85d6acc4cb8

Romeo, K., & Hubbard, P. (2010). Pervasive CALL learner training for improving listening proficiency. In M. Levy, F. Blin, C. B. Siskin, & O. Takeuchi (Eds.), *WorldCALL: International Perspectives on Computer-Assisted Language Learning* (pp. 215–229). Routledge. https://doi.org/10.4324/9780203831762

Rosa, M. (2002). Don'cha know? A survey of ESL teachers' perspectives on reduced forms instruction. *Second Language Studies*, *21*(1), 49–78.

Rost, M. (1990). *Listening in language learning*. Longman.

Rost, M. (2005). L2 listening. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 503–527). Erlbaum.

Rost, M. (2006). Areas of research that influence L2 listening instruction. *Current Trends in the Development and Teaching of the Four Language Skills*, 47–74. https://doi.org/Book_Doi 10.1515/9783110197778

Rost, M. (2011). *Teaching and Researching Listening* (3rd ed.). Routledge.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Safari, S., & Tabatabaei, O. (2016). Relationship between Autonomy and Listening Comprehension Ability among Iranian EFL Learners. *International Journal of Foreign Language Teaching and Research*, *4*(13), 33–46.

Sakai, H. (2018). Working Memory in Listening. *The TESOL Encyclopedia of English Language Teaching*, 1–6. https://doi.org/10.1002/9781118784235.eelt0579

Santorini, B. (1990). Part-of-Speech Tagging Guidelines for the Penn Treebank Project (3rd Revision). In *University of Pennsylvania 3rd Revision 2nd Printing* (Vol. 53, Issue MS-CIS-90-47). https://doi.org/10.1017/CBO9781107415324.004

Sarani, A., Behtash, E. Z., & Arani, S. M. N. (2014). The Effect of Video-Based Tasks in Listening Comprehension of Iranian Pre-intermediate EFL Learners. *Gist Education and Learning Research Journal*, *8*, 29–47.

Sasayama, S. (2018). Learner Characteristics, Individual Learner Differences, and Learner Role. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–6).

John Wiley & Sons, Inc. https://doi.org/10.1002/9781118784235.eelt0214

Sato, M., Sata, M., & Sato, M. (2017). Oral peer corrective feedback: Multiple Theoretical Perspectives. In H. Nassaji & E. Kartchava (Eds.), *Corrective Feedback in Second Language Teaching and Learning: Research, Theory, Applications, Implications* (pp. 19–34). Routledge.

Schmitt, N., & McCarthy, M. (1997). *Vocabulary: Description, acquisition and pedagogy* (Mitt (Ed.)). Cambridge University Press.

Scovel, T. (1998). *Psycholinguistics* (First Edit). Oxford University Press.

Şendağ, S., Gedik, N., & Toker, S. (2018). Impact of repetitive listening, listening-aid and podcast length on EFL podcast listening. *Computers and Education*, *125*. https://doi.org/10.1016/j.compedu.2018.06.019

Sheppard, B., & Butler, B. (2017). Insights Into Student Listening From Paused Transcription. *The CATESOL Journal*, *29*(2), 81–108.

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

Siegel, J., & Siegel, A. (2015). Getting to the bottom of L2 listening instruction: Making a case for bottom-up activities. *Studies in Second Language Learning and Teaching*, *5*(4), 637–662. https://doi.org/10.14746/ssllt.2015.5.4.6

Skouteris, H., & Kelly, L. (2006). Repeated-Viewing and Co-Viewing of an Animated Video: An examination of factors that impact on young children's comprehension of video content. *Australasian Journal of Early Childhood*, *31*(3), 22–30. https://doi.org/10.1177/183693910603100305

Sommers, M. S., & Barcroft, J. (2007). An integrated account of the effects of acoustic variability in first language and second language: Evidence from amplitude, fundamental frequency, and speaking rate variability. *Applied Psycholinguistics*, *28*(2), 231–249. https://doi.org/10.1017/s0142716407070129

Sonbul, S. (2012). Approaches to Second Language Vocabulary Teaching. *The Encyclopedia of Applied Linguistics*. https://doi.org/10.1002/9781405198431.wbeal0034

Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in english as a foreign language. *Studies in Second Language Acquisition*, *31*(4), 577–607. https://doi.org/10.1017/S0272263109990039

Stewart, M. A., & Pertusa, I. (2004). Gains to Language Learners from Viewing Target Language Closed-Captioned Films. *Foreign Language Annals*, *37*(3), 438–442. https://doi.org/10.1111/j.1944-9720.2004.tb02701.x

Stockwell, G. (2016). Mobile language learning. In *The Routledge Handbook of Language Learning and Technology* (pp. 296–307). https://doi.org/10.4324/9781315657899

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(4), 661–699. https://doi.org/10.1111/j.0023-8333.2005.00320.x

Suryana, I., Asrianto, A., & Murwantono, D. (2020). Artificial Intelligence To Master English Listening Skills for Non-English Major Students. *Journal of Languages and Language Teaching*, *8*(1), 48. https://doi.org/10.33394/jollt.v8i1.2221

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning and Technology*, *14*(2), 50–73. http://go.galegroup.com/ps/i.do?id=GALE%7CA232946131&v=2.1&u=s1185784&it=r&p=AONE&sw=w LA - English

Szarkowska, A., & Gerber-Morón, O. (2018). Viewers can keep up with fast subtitles: Evidence from eye movements. *PLoS ONE*, *13*(6). https://doi.org/10.1371/journal.pone.0199331

Taylor, G. (2015). Perceived processing strategies of students watching captioned video. *Foreign Language Annals*, *38*(3), 422–427.

Teng, C.-H., & Lin, H.-C. (2015). A Study of Metacognitive Listening Instruction to EFL Young Learners. *Critical Reflections on Foreign Language Education: Globalization and Local Interventions*.

Teng, F. (2014). Assessing the depth and breadth of vocabulary knowledge with listening comprehension. *Passa*, *48*(2), 29–56.

Teng, F. (2016). An In-Depth Investigation into the Relationship between Vocabulary

Knowledge and Academic Listening Comprehension. *Tesl-Ej*, *20*(2).

Teng, F. (2019). Maximizing the potential of captions for primary school ESL students' comprehension of English-language videos. *Computer Assisted Language Learning*, *32*(7), 665–691. https://doi.org/10.1080/09588221.2018.1532912

Thomson, R. I. (2011). Computer Assisted Pronunciation Training: Targeting Second Language Vowel Perception Improves Pronunciation. *CALICO Journal*, *28*(3), 744–765. https://doi.org/10.11139/cj.28.3.744-765

Tiippana, K. (2014). What is the McGurk effect? *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00725

Tran, T. Q., & Duong, T. M. (2020). Insights into listening comprehension problems: A case study in Vietnam. *Pasaa*, *59*, 77–100.

Trinh, V. H. (n.d.). *DIFFICULTIES AND STRATEGIES IN LISTENING COMPREHENSION*. Retrieved October 8, 2019, from https://lhu.edu.vn/139/662/Difficulties-And-Strategies-In-Listening-Comprehension-Trinh-Vinh-Hien-03av4.Html

Tschichold, C., & Schulze, M. (2016). Intelligent CALL and written language. *The Routledge Handbook of Language Learning and Technology*, 522–535. https://doi.org/10.4324/9781315657899

Urdan, T. C. (2010). *Statistics in plain English*. Routledge.

Van Engen, K. J., & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers in Human Neuroscience*, *8*(AUG). https://doi.org/10.3389/fnhum.2014.00577

van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: Adimensions approach. *System*, *41*(3), 609–624. https://doi.org/10.1016/j.system.2013.07.012

Vandergrift†, L., & Cross, J. (2018). Cognitive Listening Strategies. *The TESOL Encyclopedia of English Language Teaching*, 1–6. https://doi.org/10.1002/9781118784235.eelt0582

Vandergrift, L. (2002). *Listening: theory and practice in modern foreign language competence*. Good Practice Guide. https://www.llas.ac.uk/resources/gpg/67#toc_1

Vandergrift, L. (2004). LISTENING TO LEARN OR LEARNING TO LISTEN? *Annual Review of Applied Linguistics*, 3–25. https://doi.org/10.1017/S0267190504000017

Vandergrift, L. (2011). Second language listening. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (p. 455). Routledge.

Vandergrift, L., & Cross, J. (2014). Captioned video: How much listening is really going on? *Contact*, *40*(40), 31–33.

Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and Learning Second Language Listening*. *64*(5), 153–156.

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

Vandergrift, L., Tafaghodtari, M. H., Vandergrift†, L., & Tafaghodtari, M. H. (2010). Teaching L2 Learners How to Listen Does Make a Difference: An Empirical Study. *Language Learning*, *60*(2), 470–497. https://doi.org/10.1111/j.1467-9922.2009.00559.x

Vanderplank, R. (2016). Captioned Media in Foreign Language Learning and Teaching. In *Captioned Media in Foreign Language Learning and Teaching*. Palgrave Macmillan. https://doi.org/10.1057/978-1-137-50045-8

Vatz, K., Tare, M., Jackson, S. R., & Doughty, C. J. (2013). Aptitude-treatment interaction studies in second language acquisition: Findings and methodology. In G. Granena & M. H. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 273–292). John Benjamins.

von Glasersfeld, E. (1982). An Interpretation of Piaget's Constructivism. *Revue Internationale de Philosophie*, *36*(4), 612–635.

von Glasersfeld, E. (1991). An Exposition of Constructivism: Why Some Like it Radical. In *Facets of Systems Science* (pp. 229–238). Springer. https://doi.org/https://doi.org/10.1007/978-1-4899-0718-9_14

von Glasersfeld, E. (2007). *Key works in radical constructivism* (M. Larochelle (Ed.)). Sense Publishers.

von Glasersfeld, E., & Cobb, P. (1983). Knowledge as Environmental Fit. *Man-*

Environment Systems, *13*(5), 216–224.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wagner, E. (2007). ARE THEY WATCHING? TEST-TAKER VIEWING BEHAVIOR DURING AN L2 VIDEO LISTENING TEST. *Language Learning & Technology*, *11*(1), 67–86.

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

Wang, C., & Seneff, S. (2007). Automatic assessment of student translations for foreign language tutoring. *NAACL/HLT 2007*.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System*, *32*(4), 539–552. https://doi.org/10.1016/j.system.2004.09.011

Wang, Y., & Treffers-Daller, J. (2017). Explaining listening comprehension among L2 learners of English: The contribution of general language proficiency, vocabulary knowledge and metacognitive awareness. *System*, *65*, 139–150. https://doi.org/10.1016/j.system.2016.12.013

Weaver, C. H. (1972). *Human listening: Processes and behavior*. The Bobbs-Merrill Company.

Wilson, J. (2008). *How to teach listening*. Pearson Education Limited.

Wilson, J. (2018a). Macro Listening Skills. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–6). John Wiley & Sons, Inc.

Wilson, J. (2018b). Micro Listening Skills. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–6). John Wiley & Sons, Inc.

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

Winke, P., Gass, S., & Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *The Modern Language Journal*, *97*(1), 254–275.

Wolvin, A. D. (2009). Listening, Understanding, and Misunderstanding. In W. F. Eadie

(Ed.), *21st Century Communication: A Reference Handbook 21st century communication: A reference handbook* (pp. 137–146). SAGE Publications, Inc. https://doi.org/10.4135/9781412964005.n16

Wolvin, A. D. (2018). Listening Processes. *The TESOL Encyclopedia of English Language Teaching*, 1–7. https://doi.org/10.1002/9781118784235.eelt0569

Wolvin, A. D., Coakley, C. G., Wovin, A. D., Coakley, C. G., Wolvin, A. D., & Coakley, C. G. (2000). Listening Education in the 21st Century. *International Journal of Listening*, *14*(1), 143–152. https://doi.org/10.1080/10904018.2000.10499040

Wong, S. W. L., Dealey, J., Leung, V. W. H., & Mok, P. P. K. (2019). Production of English connected speech processes: an assessment of Cantonese ESL learners' difficulties obtaining native-like speech. *The Language Learning Journal*, 1–16. https://doi.org/10.1080/09571736.2019.1642372

Worthington, D. L., & Bodie, G. (2018). *The sourcebook of listening research : methodology and measures*. John Wiley & Sons, Inc.

Worthington, D. L., & Fitch-Hauser, M. E. (2018). *Listening: Processes, Functions, and Competency*. Routledge.

Wu, T. T., Huang, Y. M., Chao, H. C., & Park, J. H. (2014). Personlized English reading sequencing based on learning portfolio analysis. *Information Sciences*, *257*, 248–263. https://doi.org/10.1016/j.ins.2011.07.021

Xie, H., Zou, D., Lau, R. Y. K., Wang, F. L., & Wong, T. L. (2016). Generating Incidental Word-Learning Tasks via Topic-Based and Load-Based Profiles. *IEEE Multimedia*, *23*(1), 60–70. https://doi.org/10.1109/MMUL.2015.91

Yan, R. (2012). *Improving English Listening Self-efficacy Of Chinese University Students: Influences of Learning Strategy Training with Feedback on Strategy Use and Performance*. Durham University.

Yang, X., Jiang, M., & Zhao, Y. (2017). Effects of noise on english listening comprehension among chinese college students with different learning styles. *Frontiers in Psychology*, *8*(OCT). https://doi.org/10.3389/fpsyg.2017.01764

Yarnall, L., Means, B., & Wetzel, T. (2016). Lessons Learned From Early Implementations

of Adaptive Courseware. *SRI Education*, *April*, 53.

Yumusak, S., Dogdu, E., & Kodaz, H. (2014). Tagging Accuracy Analysis on Part-of-Speech Taggers. *Journal of Computer and Communications*, *02*(04), 157–162. https://doi.org/10.4236/jcc.2014.24021

Zeng, Y. (2014). Investigating the effects of metacognitive instruction on Chinese EFL learners' listening performance. *International Journal of Innovation in English Language Teaching and Research*, *3*(2), 139–158.

Zeronis, R., & Geranpayeh, A. (2015). Continuity and innovation: Updating FCE and CAE. *Research Notes*, *62*, 3–5.

APPENDICES

APPENDIX A

Suranaree University of Technology Institutional Ethics Committee	**Information Sheet for Participants and Informed consent Form**

This informed consent form is for students in the Faculty of Foreign Languages, Ton Duc Thang University who are invited to participate in the research, titled "*Developing an AI-powered, personalized learning system for EFL/ESL listening comprehension*".

Principal Investigator

Name: Cong Danh Vu

Email: martinkongzanh@gmail.com

Co-investigators

Name: Professor Dr. Andrew Lian

Email: mlapl1@gmail.com

Name: Dr. Peerasak Siriyothin

Email: peerasak@sut.ac.th

Organization

School of Foreign Languages, Institute of Social Technology, Suranaree University of Technology

Sponsor

Suranaree University of Technology

Project

Developing an AI-powered, personalized learning system for EFL/ESL listening comprehension (Version 1.0)

This Informed Consent Form has two parts:

• Information Sheet (to share information about the study with you)

• Certificate of Consent (for signature if you choose to participate)

You will be given a copy of the full Informed Consent Form

Part I: Information Sheet

1. Introduction

I am Cong Danh Vu, a lecturer in the Faculty of Foreign Languages, Ton Duc Thang University. I am doing research on improving English listening comprehension which is considered a challenging skill for many Vietnamese students. I am going to give you information and invite you to be part of this research. Please read the following information carefully. Please ask the researcher if there is anything unclear or if you need more information.

2. Purpose of the research

Listening plays an important role in language learning, however, many students are still struggling with this skill. The main purpose of the study is to develop and evaluate an online, self-learning system for English listening comprehension. The study also seeks understanding about Vietnamese undergraduate students' learning behaviors and their opinions of the features of the learning system, which will pave the way for future development.

3. Type of Research Intervention

This study will involve your participation in a training course on an online self-learning system, two (paper-based) testing sessions, a questionnaire, and an interview.

4. Participant Selection

You are invited to take part in this study because you are a first-year student whose study major is English language and is willing to be a participant.

5. Voluntary Participation

Your participation in this research is entirely voluntary. It is your choice whether to participate or not. The choice that you make will have no bearing on your study or on any study-related evaluations. You may change your mind later and stop participating even if you agreed earlier.

6. Procedures

Data will be collected through the learning activities on the system, the pre/post-course tests, questionnaire, and interviews. The table below provides an overall timeline of the study.

Activity	*Duration*	*Date*	*Place*
1. Learning activities	2.5 hours/week	Week 2-14 of the study	Online
2. Pre-course test	40 minutes	Week 1	TDTU
3. Post-course test	40 minutes	Week 15	TDTU
4. Questionnaire	20 minutes	Week 15	Online
5. Interview (audio-recorded)	1 hour	Week 15	Up to your preference

7. Duration

The study takes place over 15 weeks in total. During that time, we will meet you three times; one at the beginning of the study for conducting the orientation and pre-course test, another at the end of the study for the post-course test, and the last time for an interview with you The interview will be done in person or via video call, depending on the situation of COVID-19 in Vietnam.

8. Risks

There are no foreseeable risks for each procedure to be used in this study. You may decline to answer any or all questions and you may terminate your involvement at any time if you choose.

9. Benefits

It is hoped that the training course in this study will help you improve your English listening comprehension. Moreover, the information obtained from this study will help researchers and practitioners to develop better technology-based learning systems for English listening skill in the future which will benefit many learners of English.

10. Reimbursements

You will be provided with an account to access the learning system for self-learning when the study is completed for five years.

11. Confidentiality

All of the information collected in this study will be confidential and will only be used for research purposes. Your identity will be anonymous. Every effort will be made by the researcher to preserve your confidentiality, including the following:

- Assigning code names for participants that will be used on all research notes and documents
- Keeping interview transcripts and any other identifying participant information in a locked file cabinet in the personal possession of the researcher
- Storing electronic data on password-protected personal Google drive, accessible only to the researcher

Participant data will be kept confidential except in cases where the researcher is legally obligate to report specific incidents.

12. Sharing the Results

The results of this study will be published in the researcher's doctoral dissertation and future publications. You can read the dissertation through the link to the university's library after it is accepted. You will not be identified in any report/publication.

13. Right to Refuse or Withdraw

Your participation in this study is voluntary. It is up to you to decide whether or not to take part in this study. If you decide to take part in this study, you will be asked to sign a consent form. After you sign the consent form, you are still free to withdraw at any time and without giving a reason. Withdrawing from the study will not affect your study at Ton Duc Thang University. If you withdraw from the study before the data collection is completed, your data will be returned to you or destroyed.

14. Who to Contact

Name: Cong Danh Vu (principal investigator)

Address (Vietnam): Faculty of Foreign Languages, Ton Duc Thang University

Phone (Vietnam): 035 2435 121

Email: martinkongzanh@gmail.com or vucongdanh@tdtu.edu.vn

This proposal has been reviewed and approved by Ethics Committee for Researches Involving Human Subjects, Suranaree University of Technology, which is a committee whose task it is to make sure that research participants are protected from harm. If you wish to find about more about the EC, contact Ethics Committee Officer, Institute of Research and Development, Suranaree University of Technology Tel. 044-224757.

Part II: Certificate of Consent

I have read the foregoing information. I have had the opportunity to ask questions about it and any questions I have asked have been answered to my satisfaction. I consent voluntarily to be a participant in this study

Print Name of Participant_____

Signature of Participant _____

Date _____

Day/month/year

Statement by the researcher

I have accurately given /read out the information sheet to the participant, and to the best of my

ability made sure that the participant understands what will be done. I confirm that the participant was given an opportunity to ask questions about the study, and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily.

A copy of this ICF has been provided to the participant.

Print Name of Researcher _____

Signature of Researcher _____

Date _____

Day/month/year

APPENDIX B

QUESTIONNAIRE: STUDENTS' OPINIONS OF LISTENING HACKED

(Bảng câu hỏi: Ý kiến của người học về Listening Hacked)

Thank you for agreeing to be a respondent of this questionnaire examining your opinions of the effectiveness of the training. The information you provide is valuable for understanding how the training has helped you improve your English listening comprehension skill. This questionnaire should take approximately 20 minutes to complete. Your responses will be kept in confidentiality.

(Cảm ơn bạn đã đồng ý trả lời bảng câu hỏi của chúng tôi về ý kiến của bạn đối với sự hiệu quả của Khóa học nghe này. Các thông tin mà bạn cung cấp rất có giá trị để chúng tôi biết được Khóa học này đã giúp cải thiện kỹ năng nghe hiểu tiếng Anh của bạn như thế nào. Bảng khảo sát này tốn khoảng 20 phút để hoàn tất. Các thông tin phản hồi của bạn sẽ được bảo mật.)

You may respond to the questions in English or in Vietnamese.

(Bạn có thể trả lời các câu hỏi bằng tiếng Anh hoặc tiếng Việt.)

I. Personal Information *(Phần thông tin cá nhân)*

Your Listening Hacked User ID *(Tài khoản Listening Hacked)*: .....

II. Questions *(Phần câu hỏi)*

Question 1:

Please rate the usefulness of the following features in Listening Hacked for developing your English listening skills.

(Vui lòng đánh giá mức độ hữu ích của các chức năng có trong Listening Hacked sau đây cho việc phát triển kỹ năng nghe hiểu của bạn).

1	2	3	4	5
Completely useless	**Useless**	**Unsure**	**Useful**	**Very useful**
(Hoàn toàn vô ích)	***(Vô ích)***	***(Không rõ)***	***(Hữu ích)***	***(Rất hữu ích)***

No.	Feature	1	2	3	4	5
1	Viewing movies for improving English listening *(Chức năng xem phim để nâng cao kỹ năng nghe tiếng Anh)*	☐	☐	☐	☐	☐
Please explain your rating ***(Vui lòng giải thích việc đánh giá của bạn)***		*Please write your response here*				
2	English captions *(Chức năng phụ đề tiếng Anh)*	☐	☐	☐	☐	☐
Please explain your rating ***(Vui lòng giải thích việc đánh giá của bạn)***		*Please write your response here*				
3	Transcription tasks *(Các bài luyện tập nghe chép)*	☐	☐	☐	☐	☐
Please explain your rating ***(Vui lòng giải thích việc đánh giá của bạn)***		*Please write your response here*				
4	Automatic grading (Transcription task score) *(Chức năng chấm điểm tự động cho bài luyện nghe chép)*	☐	☐	☐	☐	☐
Please explain your rating ***(Vui lòng giải thích việc đánh giá của bạn)***		*Please write your response here*				
5	Automated feedback *(Chức năng nhận xét tự động)*	☐	☐	☐	☐	☐
Please explain your rating ***(Vui lòng giải thích việc đánh giá của bạn)***		*Please write your response here*				
6	Listening/viewing again to a certain part in the movies *(Chức năng nghe/xem lại một đoạn nào đó trong phim)*	☐	☐	☐	☐	☐

1	2	3	4	5
Completely useless	Useless	Unsure	Useful	Very useful
(Hoàn toàn vô ích)	*(Vô ích)*	*(Không rõ)*	*(Hữu ích)*	*(Rất hữu ích)*

No.	Feature	1	2	3	4	5
Please explain your rating *(Vui lòng giải thích việc đánh giá của bạn)*		*Please write your response here*				
7	In Other Movies *(Chức năng mục In Other Movies)*	☐	☐	☐	☐	☐
Please explain your rating *(Vui lòng giải thích việc đánh giá của bạn)*		*Please write your response here*				
8	Built-in dictionary *(Từ điển tích hợp)*	☐	☐	☐	☐	☐
Please explain your rating *(Vui lòng giải thích việc đánh giá của bạn)*		*Please write your response here*				
9	List of skipped words/phrases *(Danh sách từ/cụm từ đã bỏ qua trong lúc xem phim)*	☐	☐	☐	☐	☐
Please explain your rating *(Vui lòng giải thích việc đánh giá của bạn)*		*Please write your response here*				
10	Listening Boosters *(Chức năng mục Listening Boosters)*	☐	☐	☐	☐	☐
Please explain your rating *(Vui lòng giải thích việc đánh giá của bạn)*		*Please write your response here*				

Question 2:

What is your ***overall*** level of satisfaction with Listening Hacked in developing your English listening skill?

*(Mức độ hài lòng **chung** của bạn đối với việc sử dụng Listening Hacked cho việc phát triển kỹ năng nghe tiếng Anh của bạn như thế nào?)*

Very dissatisfied *(Rất không hài lòng)*	Dissatisfied *(Không hài lòng)*	Normal *(Bình thường)*	Satisfied *(Hài lòng)*	Very satisfied *(Rất hài lòng)*
☐	☐	☐	☐	☐
Please explain your rating ***(Vui lòng giải thích việc đánh giá của bạn)***			*Please write your response here*	

Question 3:

What other comments do you have on the English listening skill training with Listening Hacked?

(Bạn còn có những ý kiến nào khác về khóa học kỹ năng nghe tiếng Anh thông qua việc sử dụng Listening Hacked không?)

Please write your response here

-----Thank you for completing the questionnaire-----

(Cảm ơn bạn đã hoàn tất bảng câu hỏi)

APPENDIX C

GUIDING QUESTIONS FOR SEMI-STRUCTURED INTERVIEWS

1. How did you use feature A in your learning on Listening Hacked? *(Bạn đã sử dụng chức năng A vào việc học trên Listening Hacked như thế nào?)*

   List of features: Repeat (How did you use the options in the Repeat function?), Caption, Dictionary, Phrase search, Feedback, List of submitted transcription (submit)
2. How important was feature (A) in your listening learning? *(Chức năng A quan trọng như thế nào đối với việc học nghe của bạn?)*
3. According to the database, it seems that you rarely used the feature (A). Why didn't you use it frequently? *(Theo cơ sở dữ liệu, bạn có vẻ ít khi sử dụng chức năng A. Vì sao bạn không sử dụng chức năng này thường xuyên?)*

APPENDIX D

PRETEST

PART 1

You will hear people talking in eight different situations.

For questions **1-8,** choose the best answer **(A, B** or **C).**

---

1. You hear a woman talking on the radio about an actor.

What does the woman say about him?

A. His acting has improved over the years.

B. The media often criticise him unfairly.

C. He gets fewer film roles than he deserves.

2. You hear a hairstylist talking about her career.

She prefers working in the TV industry because she

A. feels that her contribution is valued.

B. is able to express her opinions freely.

C. thrives on the creative challenge the work presents.

3. You hear a comedian called Geoff Knight talking on the radio about his profession. What does Geoff like his act to contain?

A. stories that give people a surprise

B. things that everybody can relate to

C. material that nobody has used before

4. You hear a conversation between a customer and a coffee shop employee.

What is the employee doing?

A. waiting for a colleague's help

B. excusing a colleague's inefficiency

C. criticising a colleague's attitude

5. You hear a man telling a friend about an art exhibition.

What does he say about it?

A. It was well attended.

B. The lighting was effective.

C. The catalogue was worth buying.

6. You overhear a man ringing a sports shop.

Why is he calling?

A. to report an incident in the shop

B. to make a special order

C. to follow up an earlier query

7. You hear a man telling a friend about his work.

How does the man feel about his work?

A. resentment of his colleague's success

B. regret at the changes that have taken place

C. frustration at his lack of progress

8. You hear two people talking about a country walk they're doing.

What do they agree about?

A. It's much too long to complete.

B. The path is very difficult to follow.

C. They've chosen the wrong day to do it.

PART 2

You will hear a presentation given by a university student called Megan Rawlings about a forest survival course she went on in Australia. For questions **9-18**, complete the sentences with a word or short phrase.

Survival in the forest

It was Megan's (**9**) ..... who told her about the survival course.

Megan particularly appreciated the course leader John's use of (**10**) ..... at stressful moments.

Megan said the assistant's knowledge of (**11**) ..... was very useful during the course.

Megan was worried that her (**12**) ..... would be a problem in doing some of the tasks.

John emphasised that when it comes to safety, (**13**) ..... is the most dangerous reaction.

Megan's teammates were grateful for the (**14**) ..... which she'd brought with her.

Megan learned how to make a (**15**) ..... from the material found in the forest.

Megan and her group were told they should only use water from the (**16**) ..... for drinking.

Megan found that making a (**17**) ..... was hard for her.

Megan was surprised to find that the skill of (**18**) ..... benefited her.

PART 3

You will hear five short extracts in which people talk about a problem they had in their first few weeks in a new job. For questions **19-23**, choose what problem (**A-H**) each speaker says they had. Use the letters only once. There are three extra letters which you do not need to use.

A. I made an embarrassing comment.

B. I didn't get on with my colleagues.

C. I took on too much work.

D. I didn't get enough support.

E. I found the work too challenging.

F. I was over-confident.

G. l wasn't very punctual.

H. I was treated unreasonably.

Speaker 1		**19**
Speaker 2		**20**
Speaker 3		**21**
Speaker 4		**22**
Speaker 5		**23**

PART 4

You will hear an interview with an international concert pianist called Karen Hong.

For questions **24-30**, choose the best answer (**A, B** or **C**).

---

24. Why does Karen keep practising pieces of music she knows well?

A. to keep her confidence levels high

B. to warm up before playing difficult new pieces

C. to make small improvements to her performance of them

25. What does Karen say about her mother?

A. She still tries to have an influence over Karen.

B. She shows her emotions much more than Karen's father.

C. She could have been a competent pianist herself.

26. Karen says that after winning a big competition, she began

A. to lose interest in music.

B. to take offence easily.

C. to doubt her talent.

27. Karen's decision to take a break from performing allowed her to

A. spend a lot of time on her own.

B. regain full physical health.

C. put a new management team in place.

28. When she was performing on television regularly, Karen enjoyed the idea that

A. she was bringing people from different countries closer together.

B. she was improving people's mood and energy levels.

C. she was taking classical music to new places and people.

29. What does Karen say about pop music?

A. It is suitable for people of all ages.

B. It makes little impression on her.

C. It affects teenagers' behaviour in different ways.

30. Karen believes that when dealing with young children who play music

A. praise should only be given where it is justified.

B pushing them too hard will demotivate them.

C it's a mistake to make them nervous about the end result.

THE END OF THE TEST

APPENDIX E

POSTTEST

PART 1

You will hear people talking in eight different situations.

For questions **1-8,** choose the best answer **(A, B** or **C).**

---

1. You hear a man talking about collecting old coins.

What pleases him most about his hobby?

A. the satisfaction of aiming for a complete collection

B. the idea that someone has used the coins in the past

C. the thrill of searching for unusual coins for his collection

2. You hear a woman talking about playing the piano.

What does she say about learning to play the piano?

A. It's important to find the right teacher.

B. Everyone can play well if they try.

C. It requires more discipline than other instruments.

3. You *overhear* a man and a woman talking in an art gallery about a boy's paintings.

What do they agree about the paintings?

A. They show remarkable artistic maturity.

B. The gallery is asking too much money for them.

C. They probably weren't painted by the boy.

4. You hear two students talking about a university chemistry lecturer.

What do they agree about the lecturer?

A. She is good at explaining difficult concepts in lectures.

B. She is tolerant towards students who hand work in late.

C. She manages to make students feel enthusiastic about her subject.

5. You hear a woman talking to a work colleague about moving abroad for a new job.

What does the woman feel disappointed about?

A. the inflexible attitude to the start date

B. the lack of job security involved

C. the relatively low status of the work

6. You hear two friends talking about a job interview.

How does the woman feel now?

A. surprised that the interview went well

B. pleased to have impressed the interviewers

C. relieved that she wasn't asked any difficult questions

7. You hear part of a radio programme.

What is the woman talking about?

A. an environment group

B. a nature course for school children

C. a new walking route in the countryside

8. You hear a woman talking to her brother about his hair.

What is she doing?

A. admitting she cut his hair badly

B. teasing him about his haircut

C. suggesting he grow his hair longer

PART 2

You will hear a man called David Briggs giving a talk about his work as a volunteer on a turtle conservation programme in Western Australia. For questions **9-18**, complete the sentences with a word or short phrase.

Working on a turtle conservation programme

David first found out about the turtle programme from his (**9**) .....

David chose to work at the (**10**) ..... site because its location was more convenient.

David thinks his interest in (**11**) ..... helped him to get a place on the programme.

David was surprised to find that the ability to (**12**) ..... wasn't considered necessary.

Apart from the cost of (**13**) ..... everything essential was provided by the organisers.

David's shifts took place during the (**14**) ..... when the turtles could be checked on the beach.

David felt it was particularly important to be (**15**) ..... when handling the turtles.

Unlike his fellow volunteers, David found the (**16**) ..... didn't bother him.

David said that tiredness could lead to a loss of (**17**) ..... among the volunteers when they were collecting data.

David uses the name (**18**) ..... to refer to the most experienced volunteers.

PART 3

You will hear five short extracts in which writers give advice about writing comedy scripts for television. For questions **19-23**, choose which piece of advice (**A-H**) each speaker gives. Use the letters only once. There are three extra letters which you do not need to use.

A. Write about people who amuse you.

B. Team up with another writer.

C. Develop your characters well.

D. Rewrite your whole script several times.

E. Study comedy you like.

F. Listen to what other people say about your work.

G. Find your own way as a writer.

H. Let the audience in on the joke quickly.

Speaker 1		**19**
Speaker 2		**20**
Speaker 3		**21**
Speaker 4		**22**
Speaker 5		**23**

PART 4

You will hear an interview with a woman called Maya Gardi, whose daily life and business are based on waste-free principles. For questions **24-30**, choose the best answer (**A, B** or **C**).

---

24. What did Maya find most difficult when she started shopping in a waste-free way?

A. having to take more time over it

B. having to avoid things in plastic containers

C. having to remember to take her own bags

25. Maya decided to adopt a completely waste-free lifestyle when she

A. saw an article online about plastic rubbish.

B. noticed the bins outside her block of flats.

C. visited her local waste facility.

26. How did Maya's parents react to her decision to live waste-free?

A. They were worried that she would regret it.

B. They did not believe that she really meant it.

C. They did not think that she was likely to succeed.

27. How have Maya's cooking and eating habits changed?

A. She uses leftover food creatively.

B. She cooks more often for her friends.

C. She has developed her own cooking skills.

28. What does Maya say about socialising?

A. She sometimes has to forget her principles.

B. She doesn't worry about what people think of her.

C. She carefully chooses which events she attends.

29. What does Maya say about her new business?

A. She has an advantage when it comes to marketing.

B. Sales are increasing faster than expected.

C. She is expanding into a related sector.

30. How did Maya feel about the radio work she did recently?

A. nervous about taking part at the last minute

B. pleased to have the chance to explain her views

C. surprised that she was asked by a reporter

THE END OF THE TEST

CURRICULUM VITAE

Cong Danh Vu received his Bachelor's degree in English Language Teaching from Ho Chi Minh City University of Education, Vietnam. He obtained his Master's degree in Applied Linguistics from Curtin University, Australia. He started working as an English teacher at a language school in Ho Chi Minh city, Vietnam. He also taught English programs at Hoang Hoa Tham junior high school, Vietnam. He is currently working at the Faculty of Foreign Languages, Ton Duc Thang University, Vietnam. He had substantial experience in teaching English language skills, theoretical linguistics, translation, and American culture courses.

In 2017, he started his Ph.D. study in the School of Foreign Languages, Institute of Social Technology, Suranaree University of Technology, Thailand. His research interests include computer-assisted language learning, EFL listening comprehension, and corpus linguistics.