

THE EFFECTS OF AN ARTIFICIAL INTELLIGENCE VOICE CHATBOT
ON ENGLISH SPEAKING SKILL OF VIETNAMESE
UNDERGRADUATE EFL STUDENTS



A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Master of Arts in English Language Studies
Suranaree University of Technology
Academic Year 2022

ผลของเซทบอทแบบใช้ปัญญาประดิษฐ์ที่ได้ตอบด้วยเสียงต่อทักษะ
การพูดภาษาอังกฤษของนักศึกษาเวียดนามระดับปริญญาตรี
ที่เรียนภาษาอังกฤษในฐานะภาษาต่างประเทศ



นายยูง หวอ ถิ่น ต้ม

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศิลปศาสตรมหาบัณฑิต
สาขาวิชาภาษาอังกฤษศึกษา
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2565

THE EFFECTS OF AN ARTIFICIAL INTELLIGENCE VOICE CHATBOT
ON ENGLISH SPEAKING SKILL OF VIETNAMESE
UNDERGRADUATE EFL STUDENTS

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master's Degree.

Thesis Examining Committee


.....
(Assoc. Prof. Dr. Pragasit Sitthitikul)

Chairperson


.....
(Dr. Suksan Suppasetsee)

Member (Thesis Advisor)


.....
(Asst. Prof. Dr. Jeffrey Dawala Wilang)

Member

มหาวิทยาลัยเทคโนโลยีสุรนารี


.....

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


.....

(Assoc. Prof. Dr. Anchalee Wannaruk)
Dean of Institute of Social Technology

ยุง หวอ ถั่น ต้ม: ผลของแชทบอทแบบใช้ปัญญาประดิษฐ์ที่โต้ตอบด้วยเสียงต่อทักษะ
การพูดภาษาอังกฤษของนักศึกษาเวียดนามระดับปริญญาตรีที่เรียนภาษาอังกฤษในฐานะ
ภาษาต่างประเทศ (THE EFFECTS OF AN ARTIFICIAL INTELLIGENCE VOICE
CHATBOT ON ENGLISH SPEAKING SKILL OF VIETNAMESE UNDERGRADUATE
EFL STUDENTS) อาจารย์ที่ปรึกษา: อาจารย์ ดร.สุขสรณ์ ศุภเศรษฐ์เสรี, 214 หน้า.

คำสำคัญ: ปัญญาประดิษฐ์/แชทบอทแบบใช้ปัญญาประดิษฐ์ที่โต้ตอบด้วยเสียง/ ทักษะการพูด
ภาษาอังกฤษ/ การสอนพูด

การศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อตรวจสอบผลของแชทบอทแบบใช้ปัญญาประดิษฐ์ที่โต้ตอบ
ด้วยเสียงต่อทักษะการพูดภาษาอังกฤษของนักศึกษาเวียดนามระดับปริญญาตรีที่เรียนภาษาอังกฤษใน
ฐานะภาษาต่างประเทศและสำรวจว่าปัญญาประดิษฐ์ที่โต้ตอบด้วยเสียงสามารถทำอะไรได้บ้างเพื่อ
ช่วยพัฒนาทักษะการพูดภาษาอังกฤษของนักศึกษา

นักศึกษาระดับปริญญาตรีที่ไม่ใช่วิชาเอกภาษาอังกฤษจำนวน 60 คนที่ลงทะเบียนเรียน
หลักสูตรภาษาอังกฤษ 2 ที่ Can Tho University ถูกสุ่มเลือกให้เข้าร่วมในการศึกษาครั้งนี้ ในบรรดา
ผู้เข้าร่วม 30 คนถูกจัดเป็นกลุ่มควบคุม และ 30 คนที่เหลือถูกเลือกสำหรับกลุ่มทดลอง การศึกษาใช้
ข้อมูลทั้งเชิงคุณภาพและเชิงปริมาณตลอดระยะเวลา 8 สัปดาห์ของการทดลอง เครื่องมือ
แบบทดสอบก่อนและหลังเรียนใช้เพื่อรวบรวมข้อมูลเกี่ยวกับผลของแชทบอทแบบใช้ปัญญาประดิษฐ์
ที่โต้ตอบด้วยเสียงต่อทักษะการพูดภาษาอังกฤษของนักศึกษาเวียดนามระดับปริญญาตรีที่เรียน
ภาษาอังกฤษในฐานะภาษาต่างประเทศ หลังการทดลอง แบบสอบถามที่มี 11 ข้อและการสัมภาษณ์
แบบกึ่งโครงสร้างที่มีคำถาม 7 ข้อถูกนำมาใช้เพื่อสำรวจความคิดเห็นของนักศึกษาเกี่ยวกับการใช้แช
ทบอทแบบใช้ปัญญาประดิษฐ์ที่โต้ตอบด้วยเสียงเพื่อเรียนรู้ทักษะการพูดภาษาอังกฤษ

ผลการทดสอบก่อนเรียนและหลังเรียนพบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่
 $p=0.000$ ($p<0.05$) ดังนั้น นักศึกษาจึงพัฒนาการพูดภาษาอังกฤษด้วยคะแนนเฉลี่ยของการทดสอบ
ก่อนเรียน ($M=12.83$, $SD=2.88$) และการทดสอบหลังเรียน ($M=13.65$, $SD=2.97$) นอกจากนี้
แบบสอบถามและผลการสัมภาษณ์แบบกึ่งโครงสร้างพบว่านักศึกษามีความคิดเห็นเชิงบวกเกี่ยวกับ
การใช้แชทบอทแบบใช้ปัญญาประดิษฐ์ที่โต้ตอบด้วยเสียงเพื่อเรียนรู้ทักษะการพูดภาษาอังกฤษ โดยมี
คะแนนเฉลี่ย $M=3.90$ และ $SD=0.77$ ความคิดเห็นในเชิงบวกของนักศึกษาอาจมาจากบรรยากาศ
การเรียนรู้ที่เป็นกันเองและปราศจากความเครียด การเรียนจากเสียงของเจ้าของภาษา เวลาที่ไม่จำกัด
ในการพูด และสามารถเรียนได้ซ้ำๆ และการที่สามารถพัฒนาทักษะการพูดภาษาอังกฤษของนักศึกษา
(ไวยากรณ์ การออกเสียง คำศัพท์ การหยุดและความลึกลับ) เมื่อเวลาที่นักศึกษาพูดคุย กับแชทบอท
แบบใช้ปัญญาประดิษฐ์ที่โต้ตอบด้วยเสียง ในห้องเรียนที่ใช้ภาษาอังกฤษ แชทบอทแบบใช้
ปัญญาประดิษฐ์ที่โต้ตอบด้วยเสียงสามารถทำให้นักศึกษามีบรรยากาศที่น่าตื่นเต้นและสนุกสนาน และ
ดึงดูดการมีส่วนร่วมในทักษะการพูดภาษาอังกฤษ นักศึกษาจะไม่ประหม่าในการพูดภาษาอังกฤษเมื่อ
ใช้ แชทบอทแบบใช้ปัญญาประดิษฐ์ที่โต้ตอบด้วยเสียง

สาขาวิชาภาษาต่างประเทศ
ปีการศึกษา 2565

ลายมือชื่อนักศึกษา Tam Th
ลายมือชื่ออาจารย์ที่ปรึกษา Su Phan S.

DUONG VO THANH TAM: THE EFFECTS OF AN ARTIFICIAL INTELLIGENCE VOICE CHATBOT ON ENGLISH SPEAKING SKILL OF VIETNAMESE UNDERGRADUATE EFL STUDENTS. THESIS ADVISOR: SUKSAN SUPPASETSEREE, PH.D., 214 PP.

Keyword: Artificial Intelligence/ AI Voice Chatbot/ English Speaking Skill/ Teaching Speaking

This study aims at investigating the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL students and exploring what AI voice chatbots can do to help improve their English-speaking skill.

There were 60 non-English major undergraduate students who enrolled in the English 2 course at Can Tho University randomly selected to participate in the study. Among the participants, 30 were grouped in the control group and the remaining 30 were selected for the experimental group. The study utilised both qualitative and quantitative data over a period of 8 weeks of the experiment. The pretest and posttest instruments were used to collect data on the effects of an AI voice chatbot on English speaking skill of Vietnamese undergraduate EFL students. After the experiment, a questionnaire with 11 items and a semi-structured interview with 7 guided questions were deployed to explore the students' opinions on using an AI voice chatbot to learn English speaking skill.

The results of the pretest and posttest showed that there were significant differences at $p=0.000$ ($p<0.05$). Accordingly, the students improved their English speaking with average scores on the pretest ($M=12.83$, $SD=2.88$) and posttest ($M=13.65$, $SD=2.97$). Furthermore, the questionnaire and the semi-structured interview results revealed that students had positive opinions on using an AI voice chatbot to learn English speaking skills with the average score of $M=3.90$ and $SD=0.77$. Their positive opinion might come from the friendly and stress-free atmosphere, the native voice, the unlimited time of speaking and repetition, and the improvement in the students' English speaking skills (grammar, pronunciation, vocabulary, pauses and hesitations) when they talked with it. In English-speaking classrooms, the AI voice chatbot can bring the students an exciting and joyful atmosphere and attract their engagement in English speaking skills. The students will not become nervous in speaking English when using the AI voice chatbot.

School of Foreign Languages
Academic Year 2022

Student's Signature Tam Th
Advisor's Signature Suksan S.

ACKNOWLEDGEMENTS

My master dissertation would never have been accomplished without priceless support and strong encouragement of many individuals, to whom I would like to truly express my gratitude. I would like to take this valuable opportunity to respectfully express my appreciation from the bottom of my heart to all the valued persons who have guided me and accompanied with me thorough the journey of learning at the School of Foreign Languages of Suranaree University of Technology (SUT).

First and foremost, I would like to express my truly gratitude to the Chair of School of Foreign Languages, Dr. Butsakorn Yodkamlue, for granting me a “Vithedbundit” Scholarship to pursue my graduate program.

I would like to express my inclusive gratitude to my thesis advisor, Dr. Suksan Suppasetsee, for his restless support and precious guidance to me thorough the way we explore the new knowledge. He has spent his countless and priceless hours to help me with investigating my book, giving comments, and supporting me restlessly. Honestly, my dissertation would not have been accomplished without my dear advisor who always has meaningful and careful advice.

My strong gratitude and truly thanks go to all of thesis committee members: Assoc. Prof. Dr. Pragasit Sitthitikul and Asst. Prof. Dr. Jeffrey Wilang, who worked professionally in giving their priceless comments and meaningful feedback to build up and develop the content of my dissertation. Their valuable advice and support come a great appreciation to me.

My sincere appreciation goes to all lecturers in the English Language Studies program of School of Foreign Languages who played a significant role in giving me the knowledge and guiding me to become more analytical, critical, and intellectual in research activities during my course work.

My truly gratitude and special thanks go to all the experts who assisted me in developing a variety of research instruments to ensure the most trustworthy and reliable value. That is my deep appreciation to Dr. Suksan Suppasetsee, M.A. Dao Minh Trung, and M.A. Huynh Huu Linh who spent their priceless time in assessing and

validating my research instruments. Their professional and insightful feedbacks throughout the entire process are greatly appreciative.

My truly thanks go to all examiners who contributed their precious time in giving scores to the pre-speaking and post-speaking tests during the experimental study. That is my special thanks to Ajarn Dao Minh Trung, Ajarn Huynh Thi Ngoc Kieu, Ajarn Kim Hoai Tan, Ajarn Nguyen Hieu Thao, M.A. Thomas Kirst, and Mr. Kevin Nadeau. Their professionals and carefulness in evaluating the students' speaking performance thorough the speaking tests come to my strong gratefulness. Especially, my great gratitude is expressed to Dr. Gregory Brown for his dedicative time of proofreading the writing to make it elaborative and reader-friendly.

Besides, I would like to express my sincere thanks to all my classmates at SUT who always offered me strong support and encouraged me during my journey to knowledge. They are always very cooperative, supportive, and heartfelt to me whenever I ask for their favor.

I am deeply indebted to my former instructors: Ph.D. Thai Cong Dan and M.A. Dao Minh Trung, who dedicated their valuable time to write recommendation letters for me so that I could successfully get in SUT and become part of School of Foreign Languages to pursue my master degree until now.

Last but not least, I would like to express my respectful and heartfelt thanks to my mother, my mother-in-law, my beloved wife Ngan, and my dearest son Kien for their strong support and long-lasting endurance during my master journey. Without my family, my accomplishment would be meaningless.

Tam Duong

TABLE OF CONTENTS

	Page
ABSTRACT (THAI)	I
ABSTRACT (ENGLISH)	II
ACKNOWLEDGEMENTS	III
TABLE OF CONTENTS	V
LIST OF TABLES	IX
LIST OF FIGURES	X
LIST OF ABBREVIATIONS	XI
CHAPTER	
1 INTRODUCTION	1
1.1 Background and Context of the Study	1
1.2 Statement of the Problems.....	4
1.3 Rationale of the Study.....	7
1.4 Research Objectives.....	8
1.5 Research Questions.....	8
1.6 Significance of the Study.....	8
1.7 Scope of the Study	10
1.8 Operational Definitions of Key Terms.....	10
1.9 Summary.....	12
2 LITERATURE REVIEW	13
2.1 English Speaking Skill	13
2.2 The Potential Use of Technology to Improve English Speaking Skill	15
2.3 Artificial Intelligence (AI) Chatbots	16
2.3.1 Characteristics of AI Chatbots.....	16
2.3.2 Operational Structures of AI Chatbots	17
2.3.3 The Difference Between an AI Text Chatbots and an AI Voice Chatbots ..	32
2.4 Previous Studies about AI Voice Chatbots in EFL Contexts	33
2.5 How Vietnamese Undergraduate EFL Students Learn English Speaking in Traditional Classrooms.....	36
2.5.1 Traditional English Speaking Classrooms in Vietnam.....	37
2.5.2 Difficulties of Vietnamese Undergraduate EFL Students in Achieving English Speaking Skill in the Traditional Classrooms	38

TABLE OF CONTENTS (Continued)

	Page
2.6 Technology-Enhanced Language Learning (TELL)	42
2.6.1 The TELL Model	42
2.6.2 Learning Theories behind TELL	44
2.6.2.1 Behaviorism	44
2.6.2.2 Constructivism	45
2.6.2.3 Social Constructivism	45
2.6.2.4 Connectivism	46
2.7 Conceptual Framework for Developing the Lessons	48
2.7.1 Pre-Speaking	48
2.7.2 While-Speaking	49
2.7.3 Post-Speaking	49
2.8 Summary	49
3 RESEARCH METHODOLOGY	51
3.1 Research Context	51
3.2 Research Design	51
3.3 Participants	52
3.4 Instruments	54
3.4.1 Research Instruments	54
3.4.1.1 The Pre-speaking and post-speaking Tests	54
3.4.1.2 The Questionnaire	55
3.4.1.3 The Semi-Structured Interview	57
3.4.2 Instructional Instruments	58
3.4.2.1 The AI Voice Chatbot	58
3.4.2.2 The Learning Website for The Experimental Group	59
3.4.2.3 The Speaking Rubric	60
3.4.2.4 The Lesson Plans	61
3.4.2.4.1 The Lesson Plan for the Control Group	61
3.4.2.4.2 The Lesson Plans for The Experimental Group	62
3.5 Data Collection Procedure	64
3.5.1 Quantitative Data Collection	64
3.5.2 Qualitative Data Collection	67
3.6 Data Analysis	67
3.6.1 Quantitative Data Analysis	68

TABLE OF CONTENTS (Continued)

	Page
3.6.2 Qualitative Data Analysis.....	68
3.7 Trustworthiness and Validity	69
3.7.1 Trustworthiness	69
3.7.2 Validity	70
3.8 Pilot Study.....	70
3.8.1 General Process.....	71
3.8.2 Specific Process.....	71
3.8.3 Pilot Results	72
3.8.3.1 Quantitative Data Analysis Results.....	72
3.8.3.2 Qualitative Data Analysis Results	73
3.8.4 Problems Found	75
3.9 Summary	76
4 RESULTS AND DISCUSSION.....	77
4.1 Results.....	77
4.1.1 The Effects of an AI Voice Chatbot on English Speaking Skill of Vietnamese Undergraduate EFL Students.....	77
4.1.2 The Students’ Opinions on Using an AI Voice Chatbot	79
4.1.2.1 Questionnaire Responses.....	79
4.1.2.2 Semi-Structured Interview Responses.....	81
4.2 Discussion	85
4.2.1 The Effects of an AI Voice Chatbot on English Speaking Skill of Vietnamese Undergraduate EFL Students.....	85
4.2.1.1 The Students’ Improvement in Speaking Accuracy and Fluency	85
4.2.1.2 The Differences in English Speaking Skill of the Students in the Control Group and in the Experimental Group	86
4.2.2 The Students’ Opinions on Using an AI Voice Chatbot	87
4.3 Summary	89
5 CONCLUSION AND RECOMMENDATIONS	90
5.1 Summary of the Major Findings.....	90
5.2 Implications of the Study.....	91
5.2.1 Educational Implications for Students	91
5.2.2 Pedagogical Implications	91

TABLE OF CONTENTS (Continued)

	Page
5.2.3 Research Implications	92
5.3 Limitations and Recommendations for Future Studies	93
5.4 Summary	94
REFERENCES	95
APPENDICES	101
APPENDIX A: THE SPEAKING TEST PACK.....	102
APPENDIX B: THE QUESTIONNAIRE	107
APPENDIX C: THE SEMI-STRUCTURED INTERVIEW.....	111
APPENDIX D: EXAMPLES OF LESSON PLANS FOR CONTROL GROUP.....	113
APPENDIX E: EXAMPLES OF LESSON PLANS FOR EXPERIMENTAL GROUP	121
APPENDIX F: THE SPEAKING ASSESSMENT RUBRIC	132
APPENDIX G: THE CRITERIA FOR ENGLISH SPEAKING EXAMINERS.....	133
APPENDIX H: THE IOC VALUE FOR THE QUESTIONNAIRE ITEMS.....	134
APPENDIX I: THE IOC VALUE FOR THE INTERVIEW QUESTIONS	138
APPENDIX J: THE IOC VALUE FOR THE SPEAKING TEST PACK.....	141
APPENDIX K: THE IOC VALUE FOR THE SPEAKING ASSESSMENT RUBRIC	146
APPENDIX L: THE RATING SCALE FOR THE LESSON PLANS	151
APPENDIX M: THE EXPERTS' EVALUATION ON THE INSTRUMENTS	154
APPENDIX N: INSTRUMENTAL EVALUATION RESULTS.....	165
APPENDIX O: THE PILOT STUDY RESULTS	169
APPENDIX P: INTERVIEW QUESTIONS FOR THE MAIN STUDY	204
APPENDIX Q: THE INFORMED CONSENT FORMS.....	212
CURRICULUM VITAE	214

LIST OF TABLES

Table	Page
2.1. TELL Model	43
3.1. The Standard Scale for Evaluating English Proficiency	53
3.2. The Interpretation of Points in the Questionnaire Part 2	57
3.3. Summary of Research Instruments and Data Analysis	67
4.1. Results of the Paired sample T-tests	78
4.2. The Students' Personal Information	80
4.3. Students' Opinions on Using the AI Voice Chatbot	80
N1. Validity of The Questionnaire Items	165
N2. Validity of The Semi-Structured Interview Questions	165
N3. Validity of The English Speaking Test Pack	166
N4. Validity of The Speaking Assessment Rubric	167
N5. Validity of The Lesson Plans	168
O1. Changes in The Speaking Test Questions	169
O2. The Mean Scores of the Participants in the Speaking Tests in the Pilot Study ...	180
O3. The Result of the Paired Samples T-Test in the Pilot Study	181
O4. The Number of Males and Females in the Pilot Study	181
O5. The Range of Ages of the Participants in the Pilot Study	182
O6. The Majors of the Participants in the Pilot Study	182
O7. The Participants' Year of Study in The Pilot Study	183
O8. The Participants' English Placement Score in the Pilot Study	183
O9. The Participants' Technological Experience in The Pilot Study	183
O10. The participants' Responses to Question 7 in the Pilot Study	184
O11. The participants' Responses to Question 8 in the Pilot Study	184
O12. The participants' Responses to Question 9 in the Pilot Study	185
O13. The participants' Responses to Question 10 in the Pilot Study	185
O14. The participants' Responses to Question 11 in the Pilot Study	186
O15. The Participants' Responses to Interview Question 1 in The Pilot Study	187
O16. The Participants' Responses to Interview Question 2 in The Pilot Study	189
O17. The Participants' Responses to Interview Question 3 in The Pilot Study	191
O18. The Participants' Responses to Interview Question 4 in The Pilot Study	194
O19. The Participants' Responses to Interview Question 5 in The Pilot Study	196
O20. The Participants' Responses to Interview Question 6 in The Pilot Study	199
O21. The Participants' Responses to Interview Question 7 in The Pilot Study	201

LIST OF FIGURES

Figure	Page
2.1 Chatbot System Processes.....	18
2.2 System Architecture	20
2.3 The Ontology Model	21
2.4 The Reference Framework.....	22
2.5 The eXtensible Markup Language (XML).....	24
2.6 Example of RiveScript Code	25
2.7 Example of an Intent	27
2.8 Example of Entities.....	28
2.9 General Chatbot Structure.....	30
2.10 An AI Text Chatbot and an AI Voice Chatbot	33
O1-O4 Participant’s Chat Sessions with Andy English Bot in the Pilot Study	169
O5-O6 Participant’s Chat Sessions with Peter in the Pilot Study	175
O7-O10 Participant’s Warm-Up Sessions with Elsa in the Pilot Study	176

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AIML	Artificial Intelligence Markup Language
CALL	Computer-Assisted Language Learning
CLT	Communicative Language Teaching
CMC	Computer-Mediated Communication
EAP	English for Academic Purposes
EFL	English as a Foreign Language
ESL	English as a Second Language
ESP	English for Specific Purposes
LDA	Latent Dirichlet Allocation Algorithm
NLG	Natural Language Generation
NLP	Natural Language Processing
NLU	Natural Language Understanding
PCs	Personal Computers
TELL	Technology-Enhanced Language Learning
WP	Word Processing
XML	eXtensible Markup Language
ZPD	Zone of Proximal Development

CHAPTER 1

INTRODUCTION

This chapter provides the background and context of the study. It includes a statement of the problem, and rationale of the study. Research objectives, research questions, and the significance of the study are then introduced. Finally, operational definitions of key terms used in this study are defined, with a discussion of the scope and limitations of the study.

1.1 Background and Context of the Study

In English language education, four skills - listening, speaking, reading, and writing have been identified of highest importance (Brown, 2000). Among these skills, speaking is the most vital skill to be mastered by learners as it is the manner that language learners measure how much they understand a language, and demonstrate how effectively they communicate with other people (Nazara, 2011). Additionally, the proficient speaking of English can provide students with many potential advantages. One of the advantages is the advancement in language acquisition. Nergis (2020) mentioned that the students with low English proficiency capacity have difficulties focusing on both the form and meaning of the language, which led them to speak slowly with frequent mistakes. Therefore, it can be inferred that if students wish to speak English effectively, they need to improve both fluency and accuracy so that they can speak English with higher ability and less mistakes. Once the students fluency and accuracy are improved, their English speaking skill would likely advance.

Another advantage of proficiency in English is that the students can obtain greater academic achievement. Genesee, Lindholm-Leary, Saunders, and Christian (2006) claimed that the communicative skills and competence allow a student to become successful in academic aspects as proficiency requires the acquisition of vocabulary, grammar, and semantics of a language. It also provides learners with practice actually using the language for communicating with other speakers of that language, from which it is inferred that the student can obtain higher academic success.

A final advantage is that the students can receive better jobs and higher likelihood of promotion of the course of their lifetime. Tuyen and Loan (2019)

reported that the growth of national economy depends on students' competence in spoken English in reference to international markets. Therefore, it is predictable that if EFL students are proficient in spoken English, they will receive more chances to have their job of choice and career advancement. As a result, it is valuable to further discuss the current status of how English is taught and learned in Vietnam in general, and how English speaking is important to Vietnamese undergraduate EFL students. It is therefore necessary to find the most appropriate intervention and facilitation to support students' progress.

English teaching in Vietnam has moved forward to communicative-focused instruction as it is an international language. Most of research studies have been conducted to improve communicative competences of the EFL students in English classrooms (Nguyen & Nguyen, 2016). As a result, English speaking becomes an important skill which the EFL students need to achieve (Quyen, Nga, & Nguyen, 2018). Interestingly, the Vietnamese government has driven the country's advancement through nationwide computer investment, and has encouraged teachers to utilize technology in their teaching, to enhance traditional teaching methods. This trend has improved the EFL students' positive attitudes toward their studies because they tend to believe that they can be proficient in English when they utilize technologies (Thao et al., 2019).

Within this context, previous studies worked on investigating the effects of using technology in teaching communication skills to the Vietnamese undergraduate EFL students. Duc (2017) applied a computer-based model in helping the students to assess their speaking performance and become more aware of their learning. Quyen and Loi (2018) worked on improving the students' English speaking skill by assigning the students to watch and listen to videos spoken by native speakers. Anh and Nhu (2021) employed visual aids to help the students improve English speaking skill by reducing the stress of the speaking environment. All studies mentioned previously reported positive findings and statistically significant differences from their experimental interventions. Therefore, it can be reasonably concluded that using technologies could significantly help Vietnamese undergraduate EFL students improve their English speaking skill.

There are a variety of technologies used for educational purposes such as digital readers and tablets, 3D printing, virtual reality (VR), gamification, cloud technology, artificial intelligence (AI), and mobile technology (Fulton, 2019). Among the technologies for education, artificial intelligence (AI) is one of the most influential technologies and it has been widely popularized in education (Cheng, Sun, & Zarifis,

2020). In Vietnam, AI technologies are suggested for automatic grading of learners exercises, improving their awareness in studies and assessment, helping them to correct grammatical mistakes, and improving the students pronunciation and knowledge of preposition usage.

Nga (2022) suggested that using AI technologies such as ELSA Speak could help students improve pronunciation in speaking. Grammarly could help to fix grammatical mistakes in writing. Turnitin was broadly used by Vietnamese universities to check for students possible plagiarism. Nghi, Phuc, and Thang (2019) reported that using AI technologies like chatbots with conversational activities in the English classrooms could help students learn new grammar knowledge in terms of preposition usage. AI technologies are considered to be one of the effective tools in foreign language education as they can provide flexible, interactive, and learner-center learning which gives them significant potential in EFL contexts, facilitating the enhancement of oral communication (El Shazly, 2020). Studies on use of AI technologies to improve Vietnamese undergraduate EFL students English speaking skill have not been published to date, which inspired the basis for this study.

Having emerged in recent years, AI technologies are used in the fields of robotics, computer vision, speech recognition, and natural language processing (Kaplan, 2016). Nowadays, speech recognition and natural language processing are two main technologies used in developing voice chatbots in order to create the bots which can communicate with a user like a human, called AI voice chatbots. Thanks to the AI technologies, AI voice chatbots can identify questions from an end user and respond with appropriate answers quickly and explicitly based on the assistance of AI algorithms. This technology allows chatbots to be considered a type of e-tutors (Colace et al., 2018).

According to Kim (2016), AI chatbots in general are used as communication tools and are particularly divided into AI text chatbots and AI voice chatbots. AI voice chatbots have been continuously developed with artificial intelligence algorithms and voice recognition technology and can have a voice chat with users, especially EFL students, and help them improve their English speaking skills. Furthermore, Kim (2018) mentioned that the convenience of AI voice chatbots in usage was their rapid and free functions with online access, which allows the students to use them flexibly and freely. Kim, Cha, and Kim's (2019) finding revealed that AI voice chatbots could speak like a native speaker and conduct humanlike conversations with users by with the simulation of an AI-based communication environment, by also imitating human speech patterns, users were convinced that they were talking to a real person.

Based on the findings mentioned previously, it can be predicted that AI voice chatbots may be useful for the Vietnamese undergraduate EFL students because AI voice chatbots were reported to help students concentrate on grammar knowledge, improve English speaking, and provide native-like and humanlike conversations. These findings suggest potential solutions to the improve challenges found in improving Vietnamese undergraduate EFL students English speaking skill.

1.2 Statement of the Problems

Previous research studies defined the success of English speaking based on how accurately and fluently a speaker uses the spoken language during a conversation. Actually, Dao (2017) mentioned that the clear communication in a foreign language is achieved when speakers are able to have listeners understand what they mean during a conversation. This achievement requires speakers to be accurate and fluent when they speak the language. Brown (2000) also stated that accuracy and fluency distinctively demonstrate learners language competence. Manurung (2015) further mentioned that fluency and accuracy are the components of English speaking which need to be achieved. Therefore, it can be understood that accuracy and fluency are inseparable elements which determine the success of learners English speaking. However, many researchers on the field of English speaking in Vietnam found that Vietnamese undergraduate EFL students could not speak English successfully due to their insufficient skill, their limited exposure to English speaking environments, and their unwillingness to speak in English classrooms.

Quyen, Nga, and Nguyen (2018) makes reference to the students English insufficiency within the limitations of using correct vocabulary, grammatical structures, and sentences. Therefore, these challenges may hinder the students in their efforts to speak English successfully. Actually, English speaking was found to be the most challenging skill for Vietnamese undergraduate EFL students to achieve successfully, reflected in their low scores for the speaking tests compared to listening, reading, and writing. They could not engage in long discussions without code switching to the Vietnamese language, and they made grammatical and pronunciation mistakes when speaking English (Dao, 2017). This situation was also clarified by Thao and Nguyet (2019) that EFL learners witnessed a high frequency of difficulties in developing language because they lacked grammatical and vocabulary knowledge. In addition, Quyen and Loi (2018) reported that due to the students lack of English proficiency, they did not feel confident speaking in English classrooms. However, it is worth asking why the students had insufficient English speaking

proficiency. To answer for this question, we may argue that apart from the difficulties of language itself, as previously mentioned, the limitation of exposure to English is also likely the primary contribution to the students' lack of proficiency.

In fact, this problem has also been reported by some previous studies. Wang (2014) found that English speaking was greatly ignored in EFL classrooms because most of teachers tended to speak a lot in class instead of giving their students chances to speak English. As a result, the students were unable to develop their spoken English, which led to their insufficient speaking competence. In addition, instructional approaches have not always been updated (Chen & Hwang, 2019). The challenge for English speaking is also due to the lack of practice and exposure that the students from non-English speaking countries witness, as they only get English input and practice from language classes, schools, colleges, or universities (Baek & Lee, 2018). Vietnamese undergraduate EFL students also have experienced a similar situation. Quyen and Loi (2018) mentioned that the students in English classrooms were given limited opportunity for practicing English speaking due to insufficient meeting time in English courses. Sharing similar context, Quyen et al. (2018) pointed out the limitation of an English speaking environment caused by the frequent use of the mother tongue in English speaking classrooms and the design of English teaching syllabus and curriculum which did not provide adequate speaking activities for the students. Students need to frequently practice English speaking and see how the language is used in a social setting to become comfortable and fluent (Nguyen & Tran, 2015; Thao & Nguyet, 2019). Therefore, this is another argument that the limitation of English exposure is one of the problems preventing Vietnamese undergraduate EFL students from achieving English speaking skills.

Another problem was found to contribute to the insufficient English speaking skill of Vietnamese undergraduate EFL students, which is the students' unwillingness to speak English. According to Tuyen and Loan (2019), the students' lack of willingness to speak may cause unsuccessful language production. However, the reasons for the unwillingness of the students in speaking were caused by several factors, among which, the degree of the students' motivation was primary. This observation was also reported by Tram (2020), that the students found no motive expressing their ideas in English. Trinh and Pham (2021) further supported this position by confirming that sometimes the students chose not to say anything whenever they felt uncertain with what they were supposed to say, often because they had no ideas about the specific topics they were given. Therefore, we may assume that Vietnamese undergraduate EFL students' unwillingness in English

speaking may also relate to their insufficiency in knowledge of the English language, which prevents them from speaking English successfully. This assumption may be right because we know that without learning vocabulary and grammatical structures, Vietnamese undergraduate EFL students may not know how to compose words and sentences to produce the spoken language so that the hearer can understand. As a result, they may be afraid of failure in English speaking, which makes them decide to remain silent.

In summary, it can be implied that the students' insufficient English speaking skill, their limited exposure to English speaking environments, and their unwillingness to speak in English classrooms have an intertwined relationship with each other. Vietnamese undergraduate EFL students' lack of language proficiency may be due to their limited opportunities for exposure to English speaking, which prevents them from acquiring English speaking skill effectively and cause them to be unwilling or fearful regarding speaking English. Hence, using technologies like virtual reality (VR), gamification, cloud technology, mobile technology, or artificial intelligence voice chatbots (Fulton, 2019) may provide a helpful solution. These technologies provide a new environment for the students to have exposure to English language, which improves their English proficiency and encourages them to learn English more successfully. Although previous research studies found the problems in promoting Vietnamese undergraduate EFL students' improvement in English speaking skills, those studies did not consider any forms of technology, such as AI voice chatbots, into consideration for solving the problems. For example, Dao (2017) suggested the involvement of both teachers and students in addressing speaking problems of the students and minimizing the hesitation behind their speaking. Also, Nguyen and Tran (2015) and Quyen et al. (2018) proposed the application of speaking activities, the enrichment of speaking environment for the students to practice English speaking, and the increase in the students' awareness of independent learning styles when studying English speaking. Based on the research studies reviewed previously, no technological solutions like AI voice chatbots were suggested to help Vietnamese undergraduate EFL students overcome their difficulties in achieving English speaking skills. Therefore, this study is focused upon finding the effects of an artificial intelligence voice chatbot on English speaking skills of Vietnamese undergraduate EFL students. It will assess the ability to fill the gap between knowledge and experience because AI chatbots were reported to be effective for EFL students in terms of English speaking due to their particular positive features such as learner-center environment, human-like interaction, and authentic input/output (El Shazly, 2020;

Kim, 2018; Kim et al., 2019). Within the scope of this research, the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL students will be investigated to explore whether an AI voice chatbot can improve the target students speaking skill by enhancing their speaking fluency and accuracy.

1.3 Rationale of the Study

Based on the background and context along with the problems mentioned previously, an investigation on the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL students may provide a unique solution. Therefore, this study is support three key benefits in terms of language exposure, curriculum, and speaking environment. First, an AI voice chatbot will create more chances for the EFL students to become exposed to English speaking. This is important because language exposure, according to Krashen (1982), is the key for learners to achieve higher language proficiency. This is also expected that the EFL students will be able to speak English with more fluency and accuracy when they achieve higher language proficiency. In fact, Kim's (2016) study explored the possibility that chatbots could bring up many opportunities for foreign language learners to practice a target language. Second, this study expects that an AI voice chatbot will be integrated into the teacher s curriculum for teaching English speaking in Vietnamese universities because curriculum and learning environment are among the key factors affecting the students success in English speaking apart from teachers, students, and methods (Dao, 2017). Therefore, the integration of an AI voice chatbot into English speaking s curriculum is predicted to allow Vietnamese undergraduate EFL students to practice English speaking. Then by evaluating themselves, they will have feedback to improve autonomous practice of English speaking. Practically, Fryer and Carpenter (2006) suggested that students could proceed with self-practice and evaluation by doing an exercise of chatting with the chatbots, recording the transcript and then examining it to self-evaluate their speaking performance. Third, a new English speaking environment in association with an AI voice chatbot is hoped to solidly motivate Vietnamese undergraduate EFL students to speak English more frequently and confidently because AI chatbots can be used flexibly in terms of time and place in addition to communicating authentically with the learner (Kim et al., 2019). Within this perspective, Haristiani (2019) reported that AI chatbots motivated students in speaking English because the

students could use the chatbots at anytime, anywhere, and speak more confidently with the chatbots than speaking while facing teachers.

This study will need to gather data to find the significant differences in the English speaking skill between Vietnamese undergraduate EFL students who study in the traditional classrooms compared to those who study by using an AI voice chatbot which was designed for improving the students' English speaking skill. This allows us to investigate the effects of an artificial intelligence voice chatbot on the English speaking skill of Vietnamese undergraduate EFL students to assess potential effects an AI voice chatbot has on Vietnamese undergraduate EFL students in improving English speaking skill. In addition, exploring Vietnamese undergraduate EFL students' opinions on using the AI voice chatbots for improving English speaking skill is also an important aspect to assess. Therefore, this study comes up with two research objectives.

1.4 Research Objectives

This study has two objectives:

1. To investigate the effects of an AI voice chatbot on the English speaking skills of Vietnamese undergraduate EFL students.
2. To explore Vietnamese undergraduate EFL students' opinions on interacting with an AI voice chatbot.

1.5 Research Questions

On the basis of the objectives, this study is focused upon the following questions:

1. What are the effects of an AI voice chatbot on English speaking skill acquisition of Vietnamese undergraduate EFL students?
2. What are the opinions of Vietnamese undergraduate EFL students about using an AI voice chatbot?

1.6 Significance of the Study

This study is conducted to investigate the effects of an AI voice chatbot on the English speaking skills of Vietnamese undergraduate EFL students. Therefore, it is expected that with its findings, this study will enhance learning experiences, provide for pedagogical improvement, and add to the research data on this topic.

Regarding learning enhancement, this study is expected to contribute an option for students to motivate them to speak English in classroom environment as Kim (2016) reported that the AI chatbots helped to enhance the students' positive attitude to motivate the students to speak in a relaxed atmosphere. Also, the introduction of an AI voice chatbot in this study will provide learners with more learning options. In fact, Nghi, Phuc, and Thang (2019) found that when class meetings are combined with the chatbots, they elevate students' performance when they learn the lessons of the course book. In addition, through the investigation of Vietnamese undergraduate EFL students' opinions about using an AI voice chatbot to improve English speaking skills, this study may propose vital adjustments for Vietnamese undergraduate EFL students to enhance their ability to self-practice. Technological skills may also improve, because Thao et al. (2019) found that the integration of technology-enhanced language learning (TELL) tools into an English language learning (ELL) curriculum can increase the students' autonomy and centeredness.

For pedagogical enhancement, teachers may benefit from using an AI voice chatbot as a teaching assistant for teaching English speaking because Colace et al. (2018) mentioned that AI voice chatbots can play the role of e-tutors thanks to the association of natural language processing (NLP) algorithms and the ontology model which can help chatbots understand user words to successfully communicate with users. Moreover, teachers may also need to integrate an AI voice chatbot into their English speaking curriculum to enrich their instructional resources. El Shazly (2020) suggested that the integration of AI applications into English classrooms can optimize learning experiences.

For research data, researchers on technology-enhanced language learning (TELL) may benefit from this study as it will explore and report possible effects of an AI voice chatbot on English speaking skill of Vietnamese undergraduate EFL students. The findings in this study may become useful clues for future researchers to continue investigating further effects of an AI voice chatbot on enhancing receptive skills and other productive skills of Vietnamese undergraduate EFL students when they learn English. This perspective is also supported with the point of view from Kim et al. (2019) that the effects of the use of AI voice chatbots should be shifted to the investigation of the four basic language skills: listening, reading, speaking, and writing.

1.7 Scope of the Study

Based on the significance previously mentioned, this study is narrowed to the investigation of the effects of an AI voice chatbot on English speaking skill of Vietnamese undergraduate EFL students. Previous studies proved statistical significance even though they were conducted with small and medium-scale sample sizes (Han, 2020; Kim, 2016, 2017, 2018; Kim et al., 2019; Nghi et al., 2019). This study will conduct its research with an AI voice chatbot only, as AI voice chatbots in general were reported to be able to communicate through not only text messages but also voice messages (Han, 2020). For the effects of an AI voice chatbot on English speaking skill of Vietnamese undergraduate EFL students, this study will mainly delve into investigating the effects of an AI voice chatbot on Vietnamese undergraduate EFL students' English speaking skills in terms of fluency and accuracy when they speak English, because these two components were reported by previous researchers to be the central tasks to the success of EFL learners in English speaking.

1.8 Operational Definitions of Key Terms

The artificial intelligence voice chatbot or the AI voice chatbot in this study is a program integrated with artificial intelligence algorithms and can talk with users through the text-to-speech and voice recognition processes operating under internet connection. According to Kaplan (2016) and Colace et al. (2018), an AI voice chatbot is a kind of program or app which works under artificial intelligence algorithms such as machine learning combined with natural language processing operated with internet connection and can have a chat with users via a mobile device. In this study, Andy English Bot, called Andy, is an AI voice chatbot used thorough the experiment as it was reported by Kim et al. (2019) that it could chat with users in form of text chat and voice chat and can help users learn and practice English conversation, vocabulary, and grammar.

The Vietnamese undergraduate EFL students in the study means the non-English majored undergraduate students whose English level is equivalent to A2 level in CEFR standard, classified by a TOEIC Bridge test. Those students are in the first year or second year and learn English 2 which is one of the three modules of General English taught by Faculty of Foreign Languages at Can Tho University, Vietnam. They come from various schools and colleges in Can Tho University such as College of Natural Sciences, School of Education, College of Technology, College of Agriculture.. Accordingly, Vietnamese undergraduate EFL students are re ed by those

who major in non-English academic programs such as Chemistry Education, Agriculture, Primary Education, and Computer Networking.

Traditional classrooms mentioned in this study refers to the English classrooms in Can Tho University where students study on textbooks and follow teachers guidance in the courses. Traditional classrooms are conducted in a face-to-face manner and English language skills are taught by a teacher using oral instruction, handouts, or electronic media such as CDs, MP3 files, or presentation slides throughout the lesson.

AI-aided classrooms suggested as a term in this study implies the English classrooms in Can Tho University in which an AI voice chatbot is integrated into the teaching curriculum. The AI-aided classrooms are also conducted in a face-to-face manner but the students learn English speaking through a website and practice English speaking with the integrated AI voice chatbot while the teacher only provides technological support to the students.

English speaking skill in this study refers to the students competence of demonstrating accurate and fluent spoken language to communicate during a conversation in English as a foreign language, from which the hearer can understand the speaker's message. This key term is based on Dao (2017) mentioning that English speaking skill is achieved when the speaker can successfully convey his meaning to the listener during a conversation by speaking the language fluently and accurately.

Fluency is defined in this study as one of the distinctive criteria demonstrating learners' proficient speaking skill. Fluency is the flowing or the natural language produced by the EFL students when they speak English, demonstrated through speaker's appropriate hesitations and pauses. Brown and Lee (2015) explained that fluency in speaking is best achieved when the speech flow is performed smoothly. In addition, Wang (2014) determined hesitations and pauses in speaking as the main obstructions to speaking fluency.

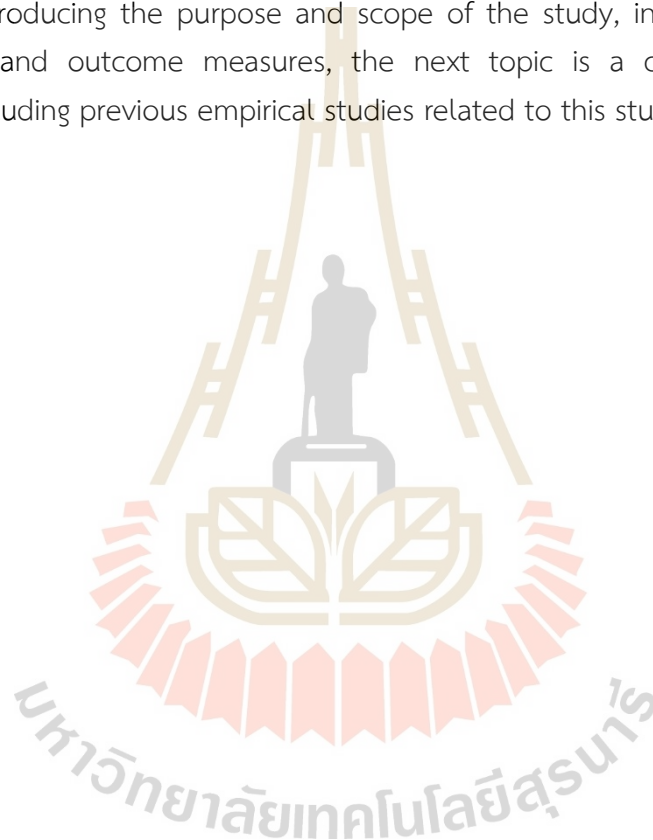
Accuracy in this study is defined as the EFL students' correct use of pronunciation, vocabulary, and grammar when they speak English. According to Brown and Lee (2015), accuracy focuses on the accurate use of phonology, grammar, morphosyntax, and discourse in the spoken output. Wang (2014) also stated that speaking accuracy involves the correct use of pronunciation, vocabulary, and grammar in English speaking. Together with fluency accuracy is seen in this study as another distinctive criteria to demonstrate learners' proficiency in speaking skills.

Opinions mentioned in this study refers to the Vietnamese undergraduate students' opinions on using the AI voice chatbot for learning English speaking skill,

classified into four themes: practice process, fluency, accuracy, and feeling. This term was composed by this study to address four main aspects of the investigation referring to (1) how long and how frequently the students practiced English speaking with the AI voice chatbot, (2) to what degree the students improved their English speaking fluency, (3) in what manner their English speaking accuracy was improved, and (4) how the students felt after using the AI voice chatbot.

1.9 Summary

After introducing the purpose and scope of the study, including the intending instruments and outcome measures, the next topic is a critical review of the literature including previous empirical studies related to this study topic.



CHAPTER 2

LITERATURE REVIEW

The literature review firstly delves into English speaking skill to explore how fluency and accuracy are closely intertwined with each other in evaluating the success of learners' English speaking skill based on previous research study's findings. Secondly, a review on the potential use of technology to improve English speaking skill will be discussed. Thirdly, a thorough review of AI chatbots including definitions, characteristics, operational structures, and types of AI chatbots will be discussed. This is followed by the review on AI voice chatbot applications in the EFL contexts to discover their potentials in enhancing English language skills in EFL contexts, in particular, in learning English speaking as a foreign language. Next, a review of how Vietnamese undergraduate EFL students learn English speaking in traditional classrooms will be conducted to discover the current standard of the learning process which may have caused difficulties for the students achieving English speaking skills successfully. This will allow introduction of this research which will propose an additional method to help the target students overcome the difficulties. Finally, this chapter will conduct a literature review on technology-enhanced language learning (TELL) which holds the theoretical framework of this research.

2.1 English Speaking Skill

Brown and Lee (2015) determined that reading, listening, writing, and speaking were the four most important skills for EFL learners to achieve in English language education. Among the skills, Nazara (2011) stated that speaking is the most essential skill required to be developed by EFL students in order to communicate effectively. Speaking skill is a skill set of the students regarding producing words, language, or expressing ideas (Socheath, 2018). Therefore, speaking is determined as a productive skill in a language. According to Brown and Lee (2015), learners should learn speaking skill with a message-oriented approach in which the focus is on speakers' ability to communicate in different contexts so that listeners can understand. Brown (2000) stated that accuracy and fluency are important goals to pursue in spoken language, and they determine the success of English speaking. Hedge (2000) proposed accuracy

and fluency to be the criteria for evaluating speaking. Accordingly, an L2 speaker is determined to be a competent L2 speaker if he or she at first can use English accurately in terms of pronunciation, grammar, lexis, with appropriate use in specific contexts. Second, he or she should speak English fluently. Manurung (2015) suggested that successful speaking skill is aimed at the mastering of speaking fluency, accuracy, and comprehensibility. According to Brown and Lee (2015), learners fluency and accuracy demonstrate correct comprehension in English speaking and infer language acquisition. Therefore, it can be understood that observing the learners fluency and accuracy in English speaking can reflect learners English speaking skill. However, accuracy and fluency are two connected aspects because learners require both for successful speaking skills. Actually, Brown (2000) stated that when learners pay too much attention to fluency in speaking, they try to speak quickly and smoothly, and then they tend to use wrong grammar, pronunciation, and inappropriate vocabulary and vice versa. In other words, the faster they try to speak, the more mistakes are found during their speech. Therefore, it is a recommendation that teachers focus their teaching approach to a message-oriented mode so that learners can initially use the language, and then progress to using it properly.

Previous researchers have different views on speaking accuracy and fluency. Wang (2014) determined accuracy in English speaking as the correct use of pronunciation, vocabulary, and grammar. Accuracy from Brown and Lee's (2015) point of view, is the accurate use of voice and articulation, grammar, and phonology. Dao (2017) related speaking accuracy with the learners' correct use of grammar and pronunciation. According to Socheath (2018), speaking skill requires learners to have the knowledge of vocabulary, parts of speech, pronunciation, expression, sentence structures, and tense. Therefore, this suggestion implies speaking accuracy in terms of vocabulary, pronunciation, and grammar (elements of speech, pronunciation, expression, sentence structures, and tense).

Among the research studies reviewed, the work of Wang (2014) and Socheath (2018) provides a more comprehensive discussion of speaking accuracy when compared with other research studies reviewed. Other authors, such as Brown (2015) and Dao (2017) relate speaking accuracy with only the correct use of grammar and pronunciation. Fluent English speaking is explained by Walker and White (2013) as the ability of L2 speakers to maintain the smooth flow of the speech without hesitations and in a coherent manner, from which they will maintain the thought flow. Within this perspective, it can be seen that the main purpose of speaking is to

negotiate meaning of the messages between the speaker and the hearer of the language, while linguistic features are to support the process of meaning negotiation.

This author concludes that accuracy and fluency can be considered the main principles of successful English speaking skills. The next section will provide a review of the potential use of technology to improve students' English speaking skill.

2.2 The Potential Use of Technology to Improve English Speaking Skill

In our daily life, we tend to communicate with each other through the applications on the internet such as Zoom, LINE, Facebook Messenger, and Telegram or through text messages. In those cases, we use written language to communicate. For example, we tend to write or type "hey", "ur", "lol", or "555" in our messages to simulate speaking and expressing our emotions. Regarding this phenomena, Walker and White (2013) stated that there were no clues of spoken language found in such types of communication as we could not see any body language or intonation to confirm it was a real speech, but the use of written language could still provide a variety of opportunities fostering learners ability to practice a second language such as the speech transcribed from movies and plays. Along with written language, spoken language could be found in the communication between people through webinar platforms such as Skype, Zoom, Google Meet or in the communication between humans and the AI chatbot applications such as Google Dialogflow, Andy English Bot, Siri, ELIZA, Duolingo, or SpeechAce.

These examples are the specific to platform and situation. Through the webinar platform, we can see everyone's body language and hear their voice and intonation. For the AI chatbot applications, although the visibility of body language cannot be observed, we can still hear the AI speaker's voice and intonation. Walker and White (2013) suggested that rehearsing spoken language before having a conversation in real life is a way to help students familiarize themselves with real life speaking situations. This suggestion requires AI voice chatbots which can help with this process when they can perform repeatable question-answer processes (Kim, 2017). They can speak with students about the same topic tirelessly and attentively (Fryer and Carpenter, 2006). It is necessary to delve into a discussion about AI voice chatbots and their applications for learning English in EFL contexts and to provide details of how this technology can help with enhancing EFL students' language proficiency.

2.3 Artificial Intelligence (AI) Chatbots

The term “artificial intelligence (AI) chatbots” is well known to researchers in this area. Fryer and Carpenter (2006) reported that chatbots have witnessed a long history of their development since the first appearing version, called ELIZA in the early 1960s. In defining an AI chatbot, Haristiani (2019) stated that an AI chatbot is a computer program which is able to carry out conversations through audio or text. When it is integrated with artificial intelligence, it has the ability to update the knowledge from previous conversations. Kim et al. (2019) also describe AI chatbots as agents which engage in a conversation through providing their users with the natural language interface to simulate a human-like interaction with the users if integrated with artificial intelligence. Adamopoulou and Moussiades (2020) related AI chatbots with artificial conversation entities (i.e. the agents which are able to interact with users, the intelligent bots, such as digital assistants such as Siri, Cortana, and Alexa).

2.3.1 Characteristics of AI Chatbots

Previous studies determined the characteristics of AI chatbots in terms of repeatability, flexibility, productivity, and innovation. With repeatability, Fryer and Carpenter (2006) confirmed that an AI chatbot could discuss the same material with students repeatedly and endlessly as it did not feel bored or tired as human being might. Moreover, when speaking with an AI voice chatbot, the students could often repeat or rephrase an utterance to negotiate meaning (Kim, 2017). Regarding flexibility, an AI chatbot can perform both text and speech conversations with the students (Fryer and Carpenter, 2006). Importantly, an AI chatbot is available on the internet and can be downloaded onto smartphones at no cost (Kim, 2017; Kim, 2018). Ahmad, Che, and Zainal (2018) also mentioned that an AI chatbot can receive and respond to human inquiry 24 hours a day. In terms of productivity, Ahmad et al. (2018) and Adamopoulou and Moussiades (2020) mentioned that the best benefit of an AI chatbot is that it can serve a great number of users at the same time and reach a broad audience on a messenger system. Finally, an AI chatbot is innovative because it is continuously updated with artificial intelligence and voice recognition technology (Kim, 2016).

Moreover, Colace et al. (2018) discussed using an AI chatbot as the most innovative solution erasing the distance between technology and education as it could provide the students with interactive learning experiences, enhance skills, and provide a personalized learning environment. Kim (2018) concluded that an AI chatbot could perform human-like English conversation with the students as native speakers, giving the students authentic language input. Based on the characteristics

described, many studies were conducted to investigate the potentials of AI chatbots on language learning and found that AI chatbots could help with students communication skills, provide the students with negotiation of meaning, and increase the students motivation and interest in learning. In the next section, this study reviews the operational structures of AI chatbots to explore how they are constructed and operated.

2.3.2 Operational Structures of AI Chatbots

Previous researchers described the structure of AI chatbots in different ways. However, the components used for operating an AI chatbot were described quite similarly from one researcher to another; and the AI chatbot in their descriptions seems to serve the same purpose. Different operational structures of AI chatbots proposed by previous studies are reviewed below.

Ahmad et al. (2018) conducted a study to investigate the structure of AI chatbots in terms of their working system and structuring techniques. The researchers used Chat.io, Collect.chat, and Cleverbot as the samples for their investigation. The results showed that the AI chatbots worked under a chain of text processing which runs in the background and are supported by a variety of techniques such as Artificial Intelligence Markup Language (AIML), Pattern Matching, Language Tricks, Chatscript, Parsing, Database, and Markov Chain. It is worth examining two specific parts which are the text processing and the techniques associated in the process. Figure 2.1 shows the chatbot system processes adopted from Ahmad et al. s (2018) s work.

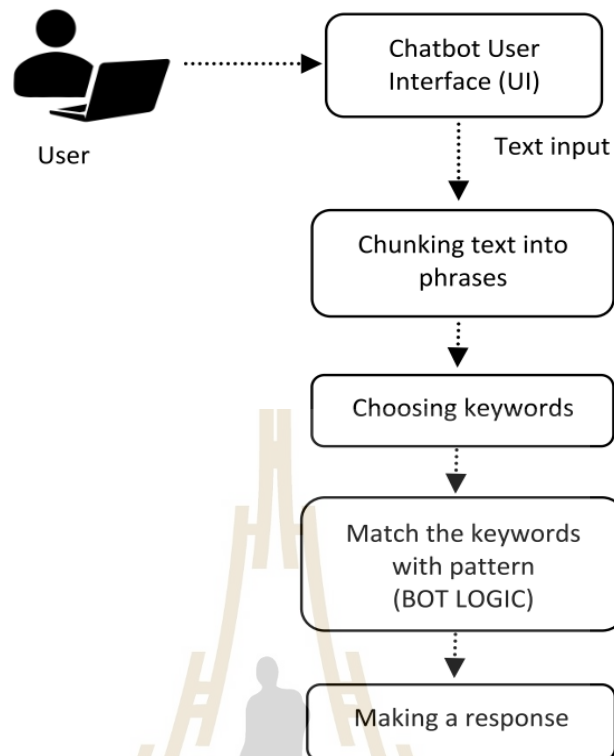


Figure 2.1 Chatbot System Processes (Ahmad et al., 2018, p. 9)

The text processing starts when a user inputs text into the user interface. At that time, the system will receive and analyze the user's text by separating it into small words and phrases to treat those items as the keywords and then matching them with the text pattern stored in the database before processing the output in form of text response.

With regard of the techniques used by an AI chatbot, the researchers regarded the AIML as a core unit which administers the process of modeling the conversation into a response process, followed by other supportive techniques. First, the Pattern Matching generates appropriate responses from matching the user's questions with the existing conversational patterns. Second, the Learning Tricks enriches the AI chatbot's knowledge base with additional words, phrases, or sentences to provide more persuasive responses to a user. Third, the Chatscript, including the written scripts coded by developers, provides syntax elements for the system so that an AI chatbot can give plausible responses even if it fails in matching the user's text with the existing conversational pattern. Fourth, the Parsing uses natural language and computer language to analyze whether a text contains any semantic or further information. This technique refers to the Natural Language Processing which was defined by Liddy (2001) as the computerized approach used to

analyze texts based on both a set of theories and a set of technologies. Fifth, the Markov Chain technique builds the consequential responses in order to make better sense. Finally, the database stores the conversational data including previous keywords and patterns so that an AI chatbot can remember and perform more successful matches when giving responses to a user. One thing to be noticed in this study is that AI chatbots at that time could communicate with users via text input and output.

Colace et al. (2018) administered a study investigating an AI chatbot system designed for supporting educational purposes through the question-answer communicating process. Pandora, Forensic Challenger (TFC), IBM Bluemix, and Facebook Messenger Bot GUI were taken as samples of the AI chatbot system which were used as a case study and experiment to study an operational framework for designing an AI chatbot. The results of the study showed that when an AI chatbot was integrated with the Natural Language Processing (NLP) algorithms and an ontologies domain, it could detect user's messages and respond appropriately. When compared with the work of Ahmad et al. (2018), the writers of this study found that Colace et al.'s (2018) work described further development of an AI chatbot design when he engaged conversational flow into the evaluation to perform successful conversation and handle user requests. Conversational flow is defined by Koudenburg, Postmes, and Gordijn (2014) as *the extent to which a conversation is experienced as smooth, efficient and mutually engaging.* The reason why the conversational flow is seen to be a further step toward the development of an AI chatbot relates to the complexity of its structure. In Colace et al.'s (2018) system architecture shown in Figure 2.2, an AI chatbot did not work under the simple text processing based on the techniques described by Ahmad et al. (2018) any longer. Instead, it worked under the conversational flow design which engaged four main operational items: Front-End, Back-Office, Knowledge Base Module, and E-Learning BOT Module.

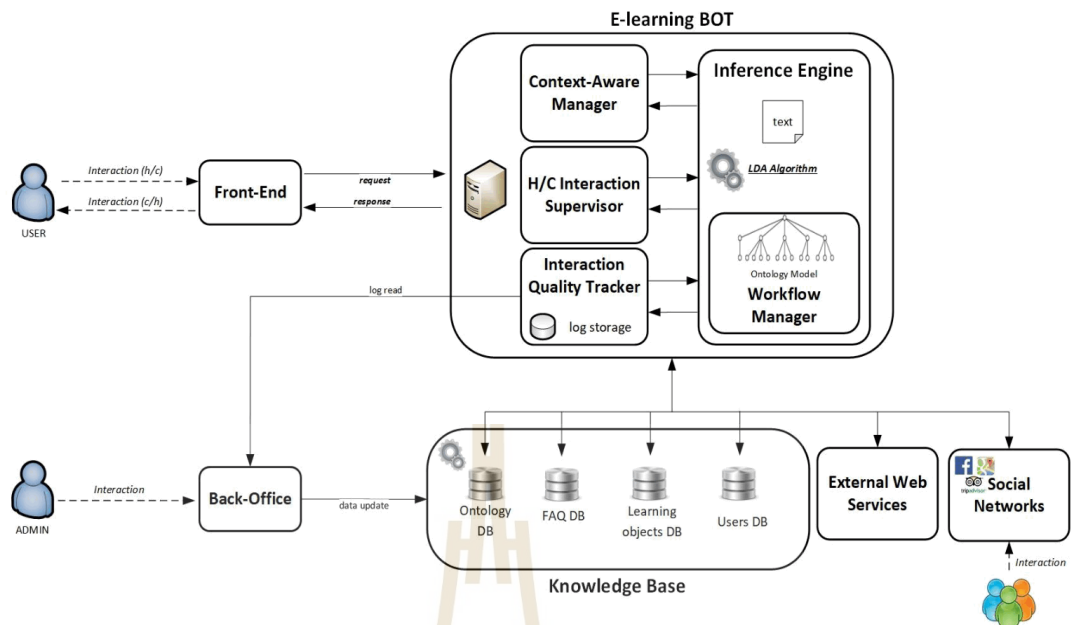


Figure 2.2 System Architecture (Colace et al., 2018, p. 531)

The Front-End is to provide a user-friendly interface. The Back-Office works in the background to manage the operations and is invisible to users. The Knowledge Base Module manages the database and processes the data. Finally, the E-Learn BOT Module was the core engine containing four smaller units: Interaction Quality Tracker, Human-Computer Integration Supervisor, Context-Aware Information Manager, and Inference Engine. Specifically, the Interaction Quality Tracker unit is used to monitor interactions between an AI chatbot and its users. The Human-Computer Integration Supervisor unit supervises dialogue, identifies ambiguous questions, and analyses the need for community support in case the system failed to answer the questions. The Context-Aware Information Manager unit drives dialogue based on context; and finally, the Inference Engine unit releases accurate responses to the user through the Latent Dirichlet Allocation Algorithm (LDA) and the database of the Workflow Manager. The LDA and the Workflow Manager are the important parts of an AI chatbot as they can manage and adopt the definition of an ontology or of pre-existing ontologies. They can then look up the ontology through word analysis based on the conversation to select appropriate sentences for an AI chatbot to provide correct answers to user's questions.

As previously mentioned, Colace et al.'s (2018) work involved two important items which helped an AI chatbot realize a user's question to respond correctly. Those items are the Natural Language Processing algorithms and the ontologies domain. Therefore, a further step is needed to examine Colace et al.'s

(2018) model of ontology and reference framework to see how the Natural Language Processing algorithms and ontologies domain work within the AI chatbot system. In Colace et al. s (2018) Ontology Model shown in Figure 2.3, the ontology model is built up under five elements: Topic of Study, User, Course, Lesson, and Learning Object.

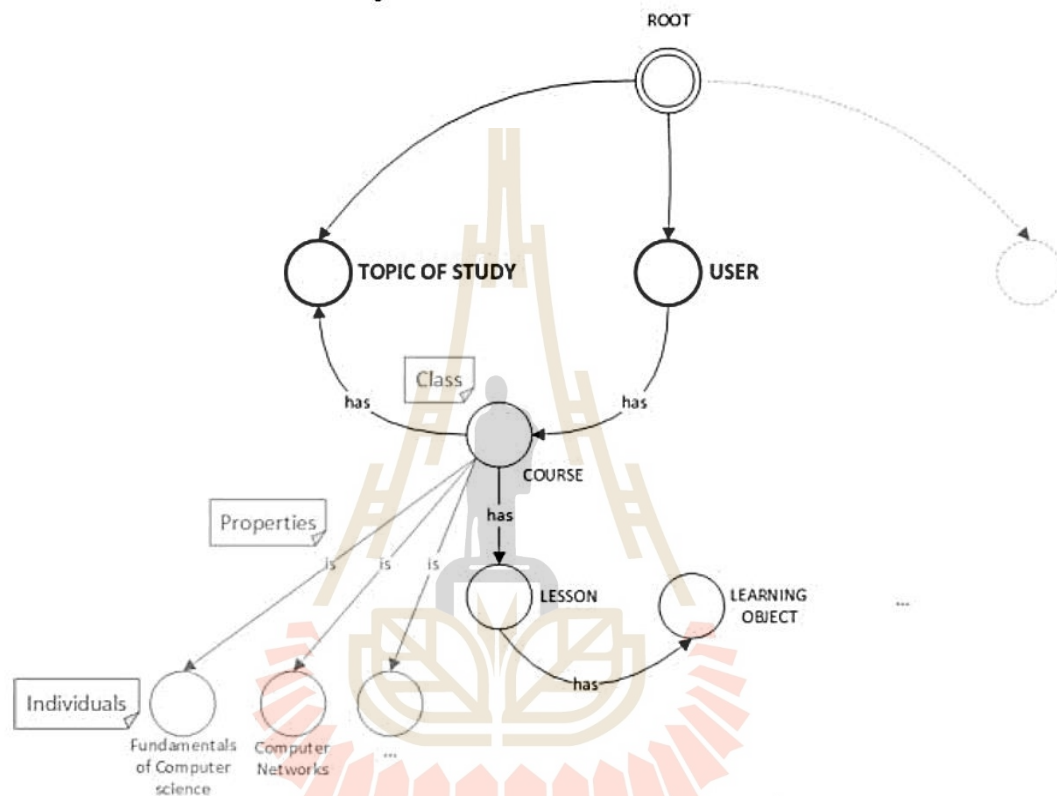


Figure 2.3. The Ontology Model (Colace et al., 2018, p. 531)

The first element refers to the knowledge of a set of subjects belonging to specific fields of study such as English Language Studies, Computer Sciences, or the like, reflected by the second element which is the User. The User contains both students and teachers who contribute their knowledge of a specific field of study into the ontology system. Such the knowledge is constructed from the remaining elements which are the Course (the smaller part or the topic of the study field such as English speaking, English listening, or the like), the Lesson (the specific module of the course), and the Learning Object (the digital or web-based resource supporting the learning).

However, how to allow an AI chatbot to utilize the ontologies domain to learn” remains a question. In this case, the answer relies on the operation of the

Natural Language Processing (NLP) algorithms demonstrated in Colace et al. s (2018) reference framework shown in Figure 2.4.

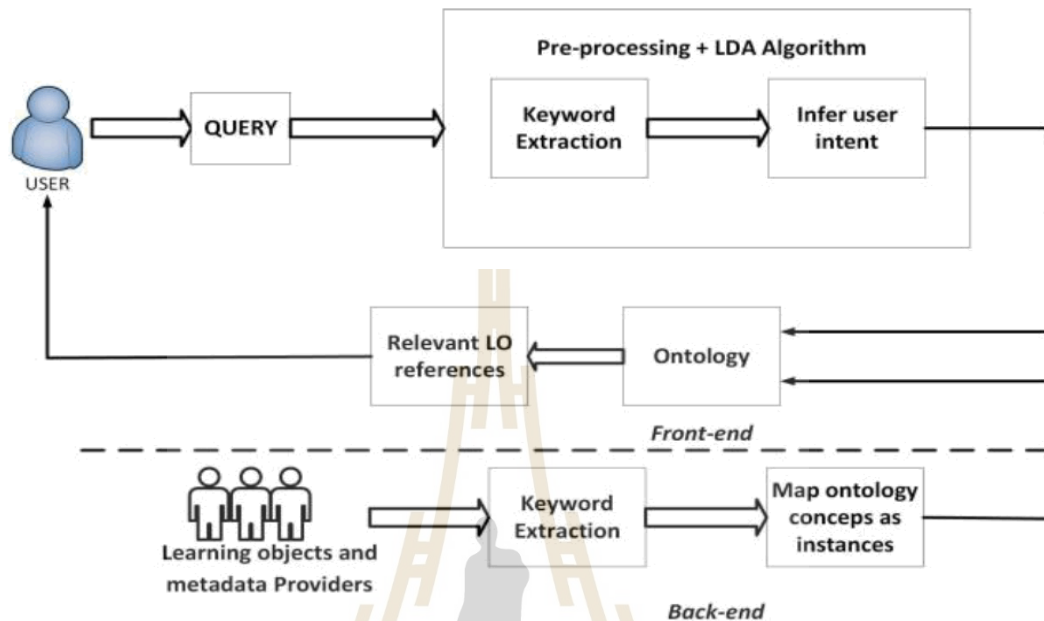


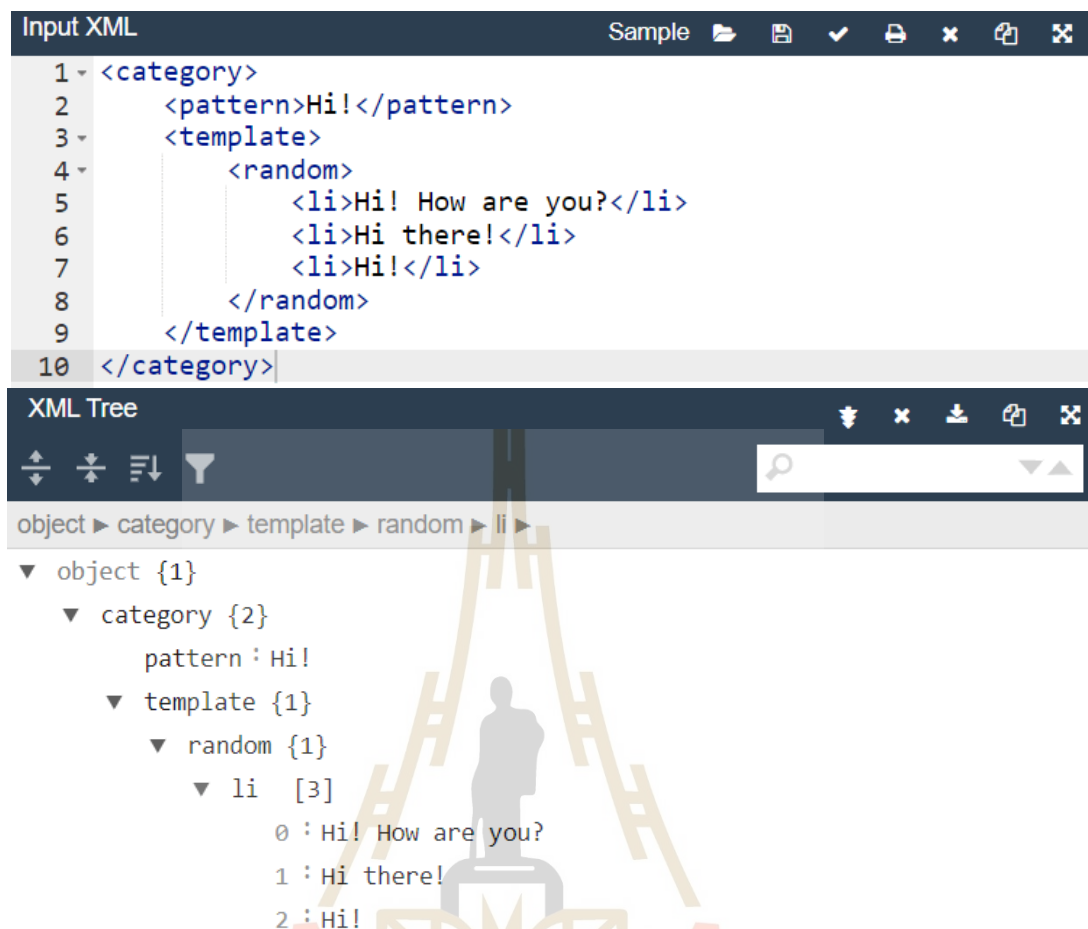
Figure 2.4 The Reference Framework (Colace et al., 2018, p. 531)

Within this framework, the NLP algorithms fulfill three tasks. The first task is to build an e-learning ontology which has the capacity of addressing the relationship between learning objects, evaluating the semantic distance between them, and then inferring from a user's intention to clarify the user's question. The second task is to manage the questions asked by users to serve the purpose of extracting keywords to thereby infer user intention. This task is done by a semantic inference engine administered by the Latent Dirichlet Allocation (LDA) algorithm which helps an AI chatbot connect the questions made by users with the existing learning object data to perform successful semantic inference. The final task is to map the instances of the existing learning object into an ontology so that an AI chatbot can have concepts or examples to infer the relations between learning objects. One thing to be aware of is that AI chatbots at that time were still unable to perform oral communication with the user. However, the text-to-speech technology was considered to be a promising factor which might be integrated into an AI chatbot so that it could respond to the user in both text and voice forms.

Adamopoulou and Moussiades (2020) conducted a study about the history, technology, and applications of an AI chatbot to perform an in-depth investigation on its development, technological design, and applications in various

industries. The researchers took a wide range of AI chatbot applications into the investigation such as MSN ELIZA, Apple Siri, IBM Watson, Google Now, Google Assistant, Microsoft Cortana, Microsoft Xiaolce, and Amazon Alexa. They further classified AI chatbots into specific categories: knowledge, service, goals, response generation, human-aid, permissions, and communication. Adamopoulou and Moussiades (2020) clarified that a communicative AI chatbot is used as a communication channel which involves text, voice, and image into the communication and interaction processes with their user. The findings of the research shed light on the AI chatbot s techniques and architecture. If compared with the work of Ahmad et al. (2018) and Colace et al. (2018), it can be assumed that Adamopoulou and Moussiades (2020) work on AI chatbots has been a crucial development because it includes nearly all aspects of the other contributors while adding more innovative approaches in terms of pattern matching and machine learning techniques.

In Adamopoulou and Moussiades s (2020) model, the pattern matching technique involves three programming approaches which include Artificial Intelligence Markup Language (AIML), RiveScript, and ChatScript. The AIML is an open-source language which was developed in 1995, using XML (eXtensible Markup Language) for implementing codes for the purpose of creating natural language within dialogues to facilitate the communication between human and AI chatbots (Adamopoulou & Moussiades, 2020; Marietto et al., 2013). XML is easy for designing and implementing the codes thanks to its tag structure. For the example shown in Figure 2.5, if we wish the system to start text processing when a user types "Hi!", then the tag structure will be `<pattern>Hi!</pattern>`. After that, with the attribute named `<template>...</template>` as the leading tag, developers can list all possible random discourses such as "Hi! How are you?", "Hi there!", or just "Hi!" into subordinate tag `<random>...</random>` from which the bot can randomly select the listed discourses to respond to the user. Within this structure, `<template>` or `<random>` is an opening tag, `</template>` or `</random>` is a closing tag, and the "..." stands for the information or discourse to be used by an AI chatbot. The reason why tag structure is preferred by developers is because of the simple structure of opening tag `< >` and closing tag `</ >`. Developers only need to remember the correct name of the commands to put into the tags and then the system will carry out the rest process. ALICE and ELIZA were the predecessors of AI chatbots which were integrated with the AIML language.



The image shows a screenshot of an XML editor. The top pane, titled "Input XML", displays the following code:

```

1 <category>
2   <pattern>Hi!</pattern>
3   <template>
4     <random>
5       <li>Hi! How are you?</li>
6       <li>Hi there!</li>
7       <li>Hi!</li>
8     </random>
9   </template>
10 </category>

```

The bottom pane, titled "XML Tree", shows the hierarchical structure of the XML document:

```

object {1}
├── category {2}
│   ├── pattern : Hi!
│   └── template {1}
│       ├── random {1}
│       │   └── li [3]
│       │       ├── 0 : Hi! How are you?
│       │       ├── 1 : Hi there!
│       │       └── 2 : Hi!

```

Figure 2.5 The eXtensible Markup Language (XML)

If compared the AIML approach with the other two, we may find that the RiveScript and ChatScript are more innovative but complex. This should not be surprising as the AIML language was created more than twenty years ago while the RiveScript and ChatScript namely were invented in 2009 and 2011. Moreover, the AIML is an open source language working under XML environment while the RiveScript operates under Java and Python environments, and ChatScript utilizes a completely new programming language. What is similar between the three approaches is that all of them are open-source languages which can be modified or coded by any developer or coder. Adamopoulou and Moussiades (2020) described the RiveScript as *a line-based scripting language implementing the Knowledge Base in rule-based chatbots*. In the Java and Python coding environments, the RiveScript appears in the form of `+`, `-`, and `*`. The symbol `+` means the user's input and the `-` refers to an AI chatbot's responses. Finally, the `*` is to mark important conversations where the input of the interpreter unit operated by the Natural Language Processing algorithms in an AI chatbot's Knowledge Base is required. To

provide a further explanation of RiveScript, this study will adopt the sample script of Adamopoulou and Moussiades (2020) shown in Figure 2.6 and give a detailed explanation about the script.

```

+Hello
- Hi! What's your name?
- Hello, How are you?
-Hello!

+ my name is *
-<set name=<formal>>Nice to meet you, <get name>!

+ Goodnight
- Goodnight <getname>

```

Figure 2.6 Example of RiveScript code (Adamopoulou & Moussiades, 2020, p. 6)

We may separate the conversation into three phases. The first phase starts from “+Hello” and ends with “-Hello!” The second phase comprises two next lines, and the final phase is the balance. In the first phase, the code structure “+Hello” means that, when a user types “Hello”, an AI chatbot may respond with either “Hi! What’s your name?” or “Hello, How are you?” or just “Hello!”. In the second phase, the “*” infers that whenever a user types their name in that space, another process will be carried out by the interpreter unit in the Knowledge Base. This process will be implemented by Java or Python coding language, read by the formula “<set name=<formal>>...<get name>!” in which the “...” stands for the discourse to be uttered by an AI chatbot. Within this aspect, the conversation in the second phase will start when a user introduces his name (i.e. My name is Peter). At that time, the formula mentioned above will inform an AI chatbot that the name of the user is ‘Peter,’ and that this type of introduction is a formal type. Therefore, the suitable response should be “Nice to meet you, Peter!”. In the final phase, when a user types “Goodnight”, then the same process will take place as in the first and the second phases so that the AI chatbot can respond appropriately by saying “Goodnight Peter.”

The ChatScript technique as mentioned previously is the latest programming language for operating an AI chatbot. It is also an open-source

programming language as RiveScript but it is much more complicated as it further involves the embedding of tagger and parser analyses (Adamopoulou & Moussiades, 2020) into the process. This allows an AI chatbot to further analyze the linguistic features and semantic meanings from the input to enable a high-proficient response. Dramatically, ChatScript was reported to be able to handle both short-term memory processing and long-term memory processing to help the AI chatbot memorize the user's information from previous conversations for later use. Unfortunately, this programming language was reported to be only suitable for the developers who already possessed a high level of expertise in programming languages due to its complexity.

Regarding learning machine techniques, the AI chatbot models by Adamopoulou and Moussiades (2020) support dramatic development because they operate with a wide range of AI approaches including the Natural Language Processing (NLP), the Natural Language Understanding (NLU), the Artificial Neural Networks, the Recurrent Neural Networks, the Sequence-to-Sequence model, and the Deep Seq2seq Models . In contrast, Ahmad et al. (2018) and Colace et al. (2018) models only utilized the Natural Language Processing system. The Natural Language Understanding algorithms work on an AI chatbot's response generation process which serves to support the Frequently Ask Question functions. Accordingly, the NLP and the NLU are significant approaches of machine learning. In fact, if the NLP is to help the system to translate and monitor natural language, then the NLU helps the system to understand the natural language, and therefore, the NLU can be seen as an inseparable companion of the NLP. The Natural Language Understanding (NLU) is the algorithm which helps with the process of classifying the intents and extracting the entities in accordance within a context. At this point, there are two things we need to consider: the intents and the entities. The intents are seen as specific topics which may be derived from the prediction of the user's messages. Figure 2.7 provides an example for an intent which has been developed for this study, powered by Google Dialogflow, which is a platform for creating an AI chatbot.

Default Greeting Intent SAVE

” Add user expression

” Hi!

” What's up?

” How is it going?

” How are you?

” Hello!

Default Greeting Intent SAVE

Responses ? ^

DEFAULT +

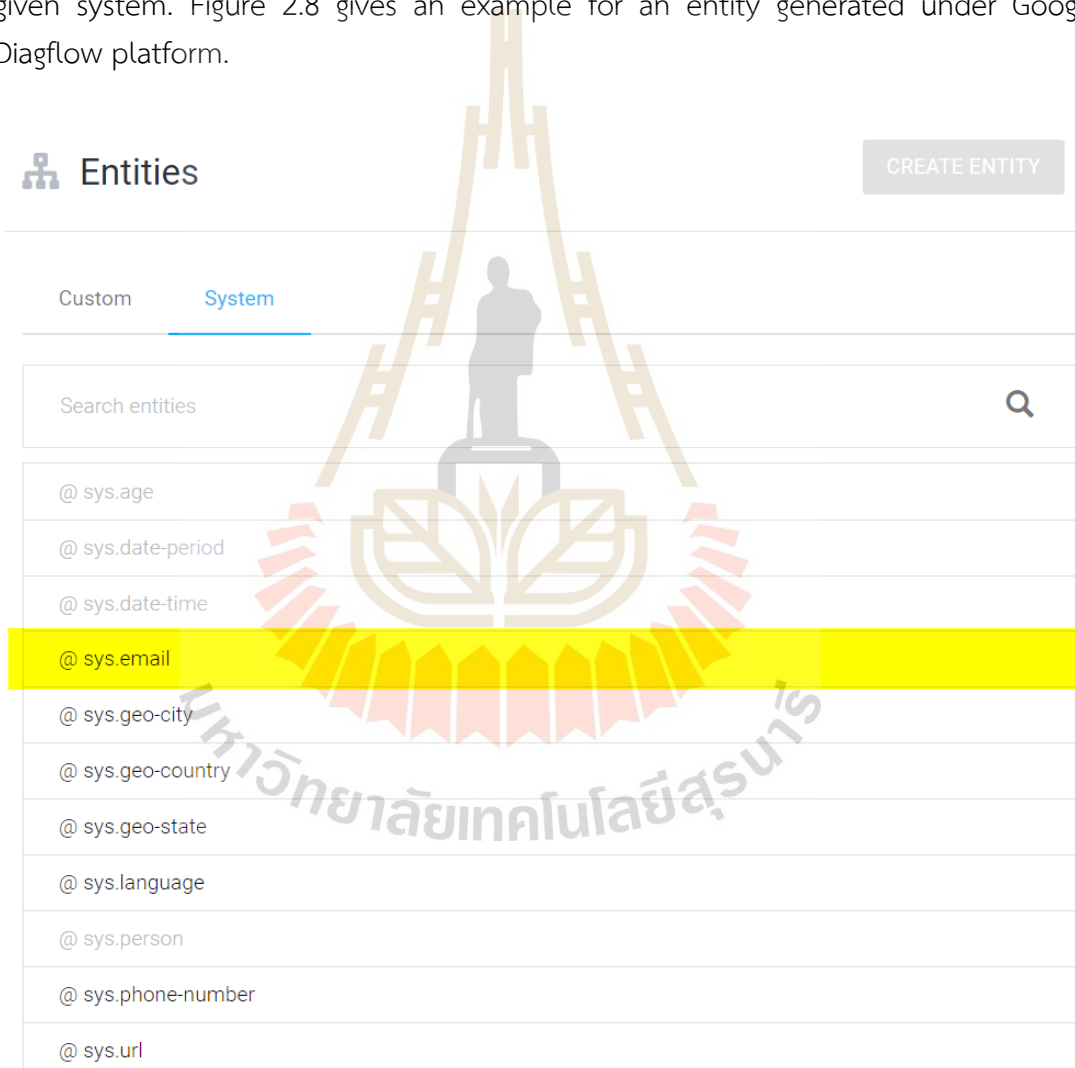
Text or SSML Response 🗑️

- 1 I'm fine. Thank you.
- 2 I'm fine. Thanks.
- 3 Everything is great.
- 4 Fine, thanks.
- 5 I'm alright.
- 6 Not so well.
- 7 I'm good.
- 8 Everyone is fine. Thank you.
- 9 Enter a text or SSML response variant ⌵

Figure 2.7 Example of An Intent

Within this example, we can predict that when performing greetings, one can use a variety of discourses to say “hello!” such as “hello!”, “hi!”, “how are you?”, “how are you doing?”, or the like. Within this situation, an intent can be created and labelled as “Greeting” in which the possible discourses related to greeting will be stored in the intent together with possible responses that an AI chatbot can utilize to respond to the user.

The entities are seen as the word blocks which contain the definition within a context and can be defined manually by developers or automatically by a given system. Figure 2.8 gives an example for an entity generated under Google Dialogflow platform.



The screenshot shows the 'Entities' page in Google Dialogflow. The 'System' tab is selected, and a list of system entities is displayed. The '@ sys.email' entity is highlighted in yellow. A watermark for 'มหาวิทยาลัยเทคโนโลยีสุรนารี' is visible in the background.

Entity Name
@ sys.age
@ sys.date-period
@ sys.date-time
@ sys.email
@ sys.geo-city
@ sys.geo-country
@ sys.geo-state
@ sys.language
@ sys.person
@ sys.phone-number
@ sys.url

sys.email SAVE ⋮

Define synonyms ? Regexp entity ?

i Separate synonyms by pressing the enter, tab or ; key. ×

@gmail.com	@gmail.com
@yahoo.com	@yahoo.com, @yahoo.com.vn
@outlook.com	@outlook.com
Enter reference value	Enter synonym

Figure 2.8 Example of Entities

This example is used to explain that when a user provides his email address during the conversation with an AI chatbot, the email addresses may come with the formats of @gmail.com, @yahoo.com, @outlook.com, or the like, and therefore each email format may have millions of characters put before the @” letter because there are millions email users who use these various email services provided those providers. The users must use a different email address from each other. The entities, in this case, will help an AI chatbot to recognize and understand that email addresses which contain the formats mentioned above are email addresses, and produce an appropriate response rather than become absorbed in detailed processing for meaning.

The last but important aspect to be considered in Adamopoulou and Moussiades s (2020) AI chatbot model is the system structure which contains five elements: user interface, user message analysis, dialogue management, the backend, and response generation. Figure 2.9 illustrates the system structure of an AI chatbot in Adamopoulou and Moussiades s (2020) model.

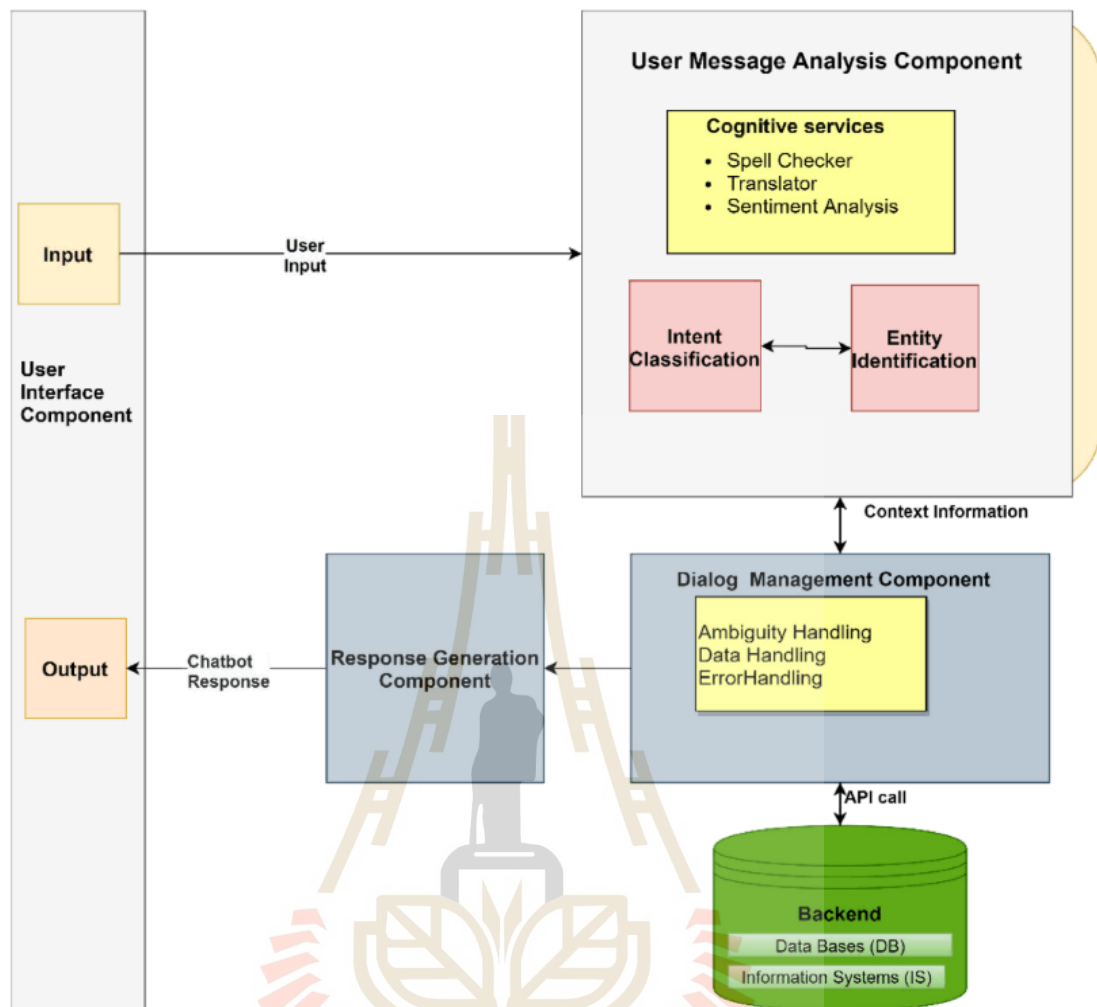


Figure 2.9 General Chatbot Structure (Adamopoulou & Moussiades, 2020, p. 11)

Basically, this AI chatbot structure shares some similar features as Colace et al. s (2018) model in terms of its Front-End design. The Back-End design where many tasks are processed in the background seems to be somehow different. For the Front-End, both models have the same structure of User Input and Response Output which is responsible for receiving user s messages and producing responses to those messages in a visible way. For the Back-End, Adamopoulou and Moussiades s (2020) model has four main components. They are the User Message Analysis, the Dialog Management, the Backend, and Response Generation. First, the User Message Analysis is similar to the E-Learning Bot component found in Colace et al. s (2018) AI chatbot structure because this component does the task of message analysis based on the Natural Language Processing (NLP) algorithms. In detail, when a user types a message through the user interface, that message will be sent to the User Message Analysis Component in which the NLP algorithms will analyze the message by

classifying it with the intents and then extracting the entities in accordance with the built-in pattern. This allows the AI chatbot to make inferences from the message before responding to it. A new item found in Adamopoulou and Moussiades s (2020) model is that apart from NLP, the User Message Analysis component also utilizes the Natural Language Understanding to provide an additional function for an AI chatbot, called the Cognitive Services. With the NLU, the Cognitive Services can perform further tasks such as checking spelling, translating, and semantic analyzing. This additional function has not been found in either Ahmad et al. (2018) or Colace et al. s (2018) models. Second, the Dialog Management is responsible for managing and updating the context of a conversation through three modules: Ambiguity Handling, Data Handling, and Error Handling. These modules work by retaining previous conversational information, updating new conversational intents, and retrieving identified entities to suggest necessary discourses for an AI chatbot to inquire for a user s clarification whenever it fails to understand the context of the conversation.

This component is quite similar to Colace et al. s (2018) E-Learning BOT component which is also responsible for handling the issues of conversational contexts through three units: Context-Aware Manager, H/C Interaction Supervisor, and Interaction Quality Checker. Third, the Backend is where the knowledge or ontologies domain is stored in a database or in information systems so that an AI chatbot can select suitable information to answer the user s queries. Colace et al. s (2018) model also has this component with the same function, called the Knowledge Base. Finally, the Response Generation generates the appropriate responses for an AI chatbot, retrieved from the Backend and analyzed by the Dialog Management. This component is quite similar to Colace et al. s (2018) Inference Engine which also works on creating suitable responses to user s messages. One difference between the two components is that Colace et al. s (2018) model used the Latent Dirichlet Allocation (LDA) algorithm to perform the task while Adamopoulou and Moussiades s (2020) model utilized the Natural Language Generation (NLG) algorithm.

Having carefully investigated three AI chatbot structures, it is reasonable to conclude that AI chatbots have witnessed significant development in recent years as they feature innovative techniques such as pattern matching and machine learning. In addition, the AI chatbot has been developed with more functionalities, driven by both conventional and modern programming languages such as AIML, RivesScript, or ChatScript, operated under innovative coding environments such as XML, Java, or Python, and processed by the AI algorithms such as the Natural Language Processing (NLP), the Natural Language Understanding (NLU), the Latent Dirichlet Allocation

(LDA), or the Natural Language Generation (NLG) which serve both pattern matching and machine learning techniques. Moreover, among the three structures analyzed, we find that the Adamopoulou and Moussiades s (2020) AI chatbot model is preferred as it refers to most of the innovative technologies we have just mentioned even though it has some complex programming languages which are only suitable for senior developers. Moreover, programming languages may not be crucial for the researchers on AI chatbots because there are ready-made platforms which can handle all the coding and programming tasks. Such may include Google Dialogflow, Dasa, or NativeChat. Within these platforms, researchers need not compose the scripts anymore. Instead, they provide the discourses (also known as parameters) into the given sections, which include intents and identified entities. Then those platforms will automatically conduct the rest of the process. The next section will discuss the difference between an AI text chatbot and an AI voice chatbot.

2.3.3 The Difference Between an AI Text Chatbots and an AI Voice Chatbots

Basically, AI voice chatbots have similar operational structures to those discussed previously in this chapter. The only difference between the two types is that an AI voice chatbot is integrated with voice recognition and text-to-speech algorithms which help the system to recognize words and sentences from spoken speech instead of text entry. Some AI chatbots are classified as AI text chatbots because they only support conversations with users in form of texts. The well-known examples include ELIZA, ALICE, and NativeChat. Apart from those, a variety of AI chatbots are considered to be AI voice chatbots due to their capability of communicating with users by both voice and text. The popular names of these are Cleverbot, Andy English Bot, Replika, Apple Siri, Google Assistant, and Amazon Alexa. An interesting point which differentiates an AI text chatbot from an AI voice chatbot is that the latter has the function of speaking, illustrated by the symbol of a speaker or a microphone next to the chat line in the user interface, as illustrated in Figure 2.10.

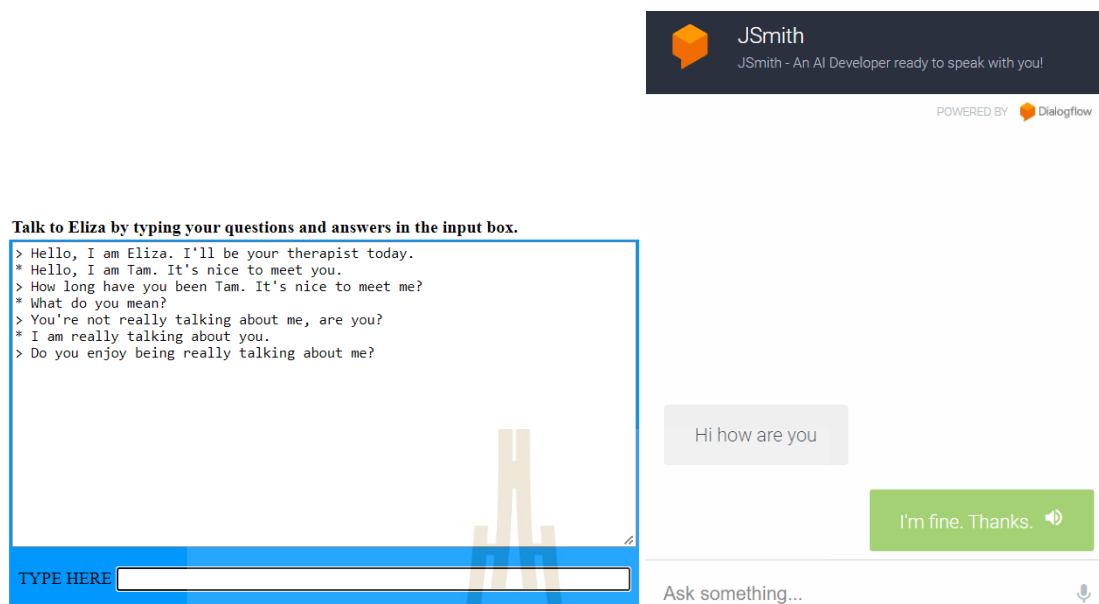


Figure 2.10 An AI Text Chatbot (left) and An AI Voice Chatbot (right)

Based on the functions of the two types of AI chatbots, it can be assumed that an AI voice chatbot is more innovative for virtual communications as it can communicate with users through both text and voice messages. Therefore, it should be considered for integration into online learning. The next section will discuss possible applications of various AI voice chatbots in the context of EFL, and students' perceptions, to investigate the most suitable AI voice chatbot to employ in this study.

2.4 Previous Studies about AI Voice Chatbots in EFL Contexts

AI voice chatbots have become common for customer assistant services. For example, Google Assistant and Apple Siri are two customer service assistants which can assist Android or iOS users. They are able to perform simple tasks like opening an app, finding a song or a piece of online news through voice commands. Amazon Alexa with voice assistance is also integrated into Amazon's products so that it can support customers to perform actions such as turning on a product or switching it off, playing a song or moving forward to another song, as examples. However, less is known about the role of an AI voice chatbot in the educational context, especially in EFL. This section will explore the research studies related to EFL.

Kim et al. (2019) conducted a study reviewing different types of AI chatbots used for language learning. A variety of AI chatbots were accounted in the study such as ELIZA, ALICE, Cleverbot, Elbot, Talk to Eve, Replika, Lyra, Andy English Bot, Poket

Friend, Mondly, and Duolingo. Within the study, the researchers employed nine criteria to evaluate the AI chatbots. Those criteria were (1) the ability of chatbots to understand complex user input, (2) the turn-taking scheme, (3) the ability to recall user's name, (4) the capability of multilanguage support, (5) the feature of voice input and output, (6) the feature of text input and output, (7) the ability to access historical conversations, (8) the ability to deal with strange questions, and (9) the ability to overcome typographical errors. The findings revealed that most of the AI chatbots could perform turn-taking and feature text I/O while some of them could support voice I/O, multilanguage, and recalling historical conversations. Few of the AI chatbots were able to perform other tasks like understanding complex user's input, recalling user's name, answering complicated questions, or overcoming typographical errors. Kim et al.'s (2019) AI chatbots data when related to the identified functionalities, found that ALICE, Replika, and Andy English Bot contained many more features than the other programs, especially the feature of voice input and output. However, it is curious that Kim et al.'s (2019) finding identified ALICE as the AI chatbot which supported voice input and output because no researchers have mentioned this feature of the program. To test this statement, we attempted to verify the claim by having a chat with ALICE and found that it could not perform any voice input or output. In Figure 2.11, even though it has a button labeled "Say", the button had no functional capacity. The only thing happened after pressing "Say" was that the text we input would appear in the chat session. Therefore, Replika and Andy English Bot remain the primary AI voice chatbots.

Nghi et al. (2019) implemented an experimental study on applying an AI chatbot for helping students to learn prepositions. The study involved 200 undergraduate students divided into two groups: 100 students in the control group and 100 in the experimental group. The researchers of the study utilized the Facebook chatbot developed by Chatfuel as a research instrument to conduct the experiment. The English Pronouns and Prepositions book was used as an instructional instrument during the experiment period. The control group took 15 periods to learn and the experimental group took 10 periods. The results of the study provided three useful pieces of information. First, the AI chatbot brought a new learning experience to the students. The new learning experience was the combination of both the class meeting and the use of the AI chatbot. This combination was observed to accelerate the students' performance in learning. Second, the students' perception of using AI chatbot tools for the learning process was increased because they felt comfortable to self-practice and self-access in learning with the AI chatbot. Third, the AI chatbot

created a fun and exciting environment for the students to learn, which motivated them to share their understandings and performance with classmates.

El Shazly (2020) conducted a research study investigating the role of AI chatbots in English speaking practice for foreign language anxiety (FLA) management of the Egyptian EFL students. The research invited 48 undergraduate Korean EFL students to take part in the eight-week quasi-experimental research. Pre-speaking and post-speaking tests were developed to evaluate the students' speaking proficiency level before and after the intervention of AI chatbots. The IELTS speaking test rubric was used and developed for the speaking evaluation. The FLA questionnaire containing 33 items was used to explore the students' anxiety levels at pre-intervention and post-intervention. Mondly and some web chatbots such as Audrey, Charles, Cristal, and Mike were used as AI chatbot tools for the intervention. The results of the research revealed that the students' speaking proficiency level improved after the intervention. However, it was found that the students experienced a degree of anxiety and that their anxiety increased slightly even after the pre-intervention of the AI chatbots. Therefore, this finding was reported to contrast with other research studies as those always found positive effects of AI chatbots toward EFL students' anxiety level. Nevertheless, El Shazly (2020) also explained for this contradictory finding could possibly be explained. First, eight-week intervention was a short duration. Therefore, the students might not have enough time to get familiar with the AI chatbots. Second, the emotional deficit of the AI chatbots might fall short of learners' emotional needs, while a real teacher could do so. Third, unfamiliar and irrelevant topics along with repeated conversations with the virtual partners might also demotivate the students, even though the students did improve their language proficiency. Disregarding those negative factors, El Shazly (2020) reported that the AI chatbots could help increasing the students' effort. The chatbots were also reported to be able to enhance the students' oral performance by improving their cognitive faculties and linguistic abilities.

Kim et al. (2021) conducted research aiming at exploring the effects of voice chat on the communication skills of the EFL learners. The research involved 49 undergraduate EFL students in two different English proficiency levels who took an experiment with fourteen weeks of sessions speaking in voice chat sessions with the AI voice chatbots named respectively Replika, Andy, and Google Assistant. The participants could choose to talk with any of the given AI voice chatbots they preferred. The researcher used questionnaires, pre-speaking and post-speaking, tests, and interviews as the main instruments for data collection procedure. The results of

the research revealed that all students improved their pronunciation, intonation, and stress after the intervention. However, insignificant differences were found on the low-level students' fluency improvement while the students at intermediate level could significantly improve their speaking fluency. In exploring the students' perceptions on communicating with the AI voice chatbots, Kim et al. (2021) found both positive and negative perceptions from the students. In light of the positive perceptions, the study reported that they could find more opportunities to practice speaking without the fear of losing face or worrying about using the wrong sentences. They could also improve pronunciation and become more confident, active, and interested in speaking due to the unlimited time available when speaking with the AI voice chatbots. Regarding negative perceptions, the study found that the students were not quite comfortable with having voice chats with the AI voice chatbots. Moreover, the study also reported the AI voice chatbots' limitations in voice recognition when the students' voices were not recognized correctly by the AI voice chatbots. The study also mentioned that the AI voice chatbots could not help the students with English correction.

In investigating the effect of an AI voice chatbot interaction on EFL students' L2 speaking performance and anxiety, Çakmak (2022) selected 90 EFL students from a Turkish university and let them use Replika for practicing English speaking during a twelve-week period. A questionnaire was utilized to explore the students' perceptions of using the AI voice chatbot for practicing English speaking. The results of the study revealed both positive and negative findings. For the positive findings, Çakmak (2022) found that using Replika to practice English speaking could help enhance the students' English speaking performance more effectively than having them practice speaking face-to-face with peers. However, the research reported that there was not a significant difference with the students' anxiety reduction because of challenges in getting the AI voice chatbot to understand them correctly, which then increased their anxiety more when speaking. Therefore, Çakmak (2022) suggested that using chatbot interaction proved not to be an effective way to help the EFL students reduce anxiety with speaking English.

2.5 How Vietnamese Undergraduate EFL Students Learn English Speaking in Traditional Classrooms

Learning English speaking skills in Vietnamese universities is performed in traditional classrooms in which students rely on teachers and textbooks to learn English speaking. However, this traditional learning style creates some difficulties for

Vietnamese undergraduate EFL students to achieve successful English speaking skills. To seek solutions, this study will explore the context of learning English speaking skills in traditional English speaking classrooms and the issues faced by Vietnamese Undergraduate EFL Students.

2.5.1 Traditional English Speaking Classrooms in Vietnam

Duyen and Huan (2017) determined the elements of traditional English speaking classrooms by identifying four elements which included context, teacher, student, and tasks. Within this framework, English speaking classrooms were often conducted in terms of group work containing three to five students to allow them to develop communication proficiency in English. However, in such classrooms, the teacher was seen as the one who had authority over students throughout the students learning process, causing the students to passively learn. Tasks and detailed instructions were given and guided by the teacher in which students were instructed to do the group work. Sufficient time was provided for the students to complete the tasks. Although in the group work, communication between students was crucial, they tended not to communicate with each other in terms of asking questions, giving comments, responding with gesture like nodding heads, or by using sounds to show their agreement or disagreement to other members points of view. In these classrooms, the students rarely used English to communicate or shared ideas with their classmates.

Dao (2017) described the traditional English speaking classrooms in terms of material, method, environment, teacher, and student. The study reported that in traditional classrooms the students learned English speaking with the textbook as the only material and very few communicative activities were organized in such classrooms. Furthermore, CD or audio files spoken by native speakers were often used by teachers as the speaking models for students to imitate and the speaking exercises taken from textbooks were given to the students to practice English speaking in front of the class. In such classrooms, the students studied in large-size classrooms accommodating around thirty to forty seats, and teachers used both English and Vietnamese to teach English speaking skills or give instructions to ensure the students understanding of the lesson and tasks. The students used Vietnamese in traditional English speaking classrooms when discussing or debating in group. Regarding the use of Vietnamese in English speaking classrooms, Anh (2010) confirmed that using the Vietnamese language in English speaking classrooms was part of the teaching method. This could be positive in some specific contexts when teachers could decide when to use Vietnamese.

Nghi (2019) reported that in traditional classrooms, language teachers normally taught a foreign language lesson with four stages including warm-up, pre-practice, control practice and post-practice. During classroom time, the students were provided with handouts, teachers gave lectures throughout the lesson, and the classroom time was spent strictly on the course procedure. Thao and Nguyet (2019) reported that at the end of each English speaking course, the students were required to take an individual verbal test with lecturers by having a face-to-face conversation with their lecturers on a chosen topic.

In observation of the traditional English speaking classrooms, Nghi (2019) reported that the normal and repeated procedure from one lesson to another caused boredom for the students. Tuyen and Loan (2019) added that in such classrooms, most of the students never or rarely volunteered to speak or answer the teacher's questions, and often avoided the teacher asking additional questions. As a result, the students often only sat silently in the class and listened to the teacher. Trinh and Pham (2021) reported that in traditional classrooms, the students tended to translate the information in the textbook into Vietnamese before they did the tasks assigned by teacher.

In consideration of the types of traditional English speaking classrooms, a clearer picture of how Vietnamese undergraduate EFL students learn English speaking has been revealed. The traditional classrooms were mainly driven by teachers. Textbooks, groupwork tasks, handouts, and media files were used as the main tools for the teaching and learning process. Moreover, the students still did not feel comfortable speaking in the class and the teacher still had to provide all information during class time. Moreover, the use of the Vietnamese language in those English speaking classrooms was found to happen regularly. Consequently, these limitations might challenge EFL students who planned to achieve successful English speaking skills.

2.5.2 Difficulties of Vietnamese Undergraduate EFL Students in Achieving English Speaking Skill in the Traditional Classrooms

In investigating the difficulties of Vietnamese undergraduate EFL students in achieving speaking English successfully, previous researchers found that the students' low proficiency in English, their unwillingness to speak, and the lack of English speaking environment were the primary problems in the traditional classroom model.

Dao (2017) conducted a study to investigate the key factors which had negative effects on Vietnamese undergraduate EFL students. The study had 108 non-

English majored students at a university in Vietnam. A questionnaire and classroom observation were employed as the research instruments for data collection. The findings revealed that the lack of English speaking environment, the students insufficient English language competence, and their unwillingness to speak English caused difficulties for students who wished to achieve English speaking skill.

In more detail, the research found that the frequent use of Vietnamese in English classrooms created the habit of using Vietnamese in verbal interactions among the students, which then caused a lack of exposure to English even though they were studying in an English classroom. Regarding the students themselves, the study found that the students were reluctant to speak English due to their insufficiency of vocabulary and grammar to express their thoughts. Furthermore, due to lack of practice in speaking English, the students were then further unwilling to speak English in the class. In addition, factors such as inadequate materials and activities for practicing English speaking, the insufficient time the students were allowed to practice English, and the limitation of space for students to practice speaking English when communicating with teachers and with each other were all found to limit the students' exposure and success in speaking English.

As a result, those students failed to realize if they had spoken English successfully or if there were areas needed for improvement. As a solution for the problems found, the study suggested engaging both teachers and students in overcoming the difficulties. Specifically, recommendations for teachers were to speak English in the class more frequently and encourage the students to exclusively use English rather than their mother tongue language in classroom discussion. In addition, the study also mentioned the need for the students to actively speak English in class to form the habit of speaking English among themselves. Moreover, the students were encouraged to actively join English clubs and societies outside the classrooms to increase chances to use English for speaking in the real-life. In line with this study, Nguyen and Nguyen (2016) suggested raising the students' awareness of using communication strategies by integrating those strategies into the teaching curriculum to help the students successfully overcome the difficulties of speaking English.

Similarly, Quyen et al. (2018) carried out a study aiming to investigate the factors contributing to the difficulties Vietnamese undergraduate EFL students have been facing when they spoke English. There were 131 students participating in the research: a questionnaire, individual interviews, and class observation were used as the research instruments. The findings revealed that the lack of vocabulary,

confidence, strategies, chances, and motivation contributed to the students' difficulties in English speaking.

Furthermore, the factors such as teachers' methods, teaching curriculum, in-class environment, and limited extra-curricular activities were found to have created a challenge for the students' English speaking skills. Among those factors, the lack of in-class environment and extra-curricular activities outweighed the others because most of the students agreed that they lacked native lecturers to practice English speaking with. They also mentioned that English use was limited after the class. In short, they lacked environments to practice English speaking skill and became fluent.

Some solutions to overcome these difficulties included suggestions that teachers should use various teaching strategies in a flexible manner to get the students to use less mother tongue based language in the class so that they would have more opportunity to speak English. In addition, the self-awareness of the students regarding independent learning was also mentioned as one of the possible solutions for the problem.

Duc (2017) further suggested the use of a computer-based model to help the students self-practice English speaking skills. The research used a computer-based model for the pre-speaking and post-speaking tests as the research instruments to investigate if the students were able to self-practice and self-assess English speaking skill after the intervention. There were 40 undergraduate students joining the research which lasted five months. Among the students, 21 were in the experimental group and got familiarized with computer-based speaking tests while the remaining 19 students were in the control group and were not trained to be familiarized with computer-based speaking tests.

The results of the study revealed that using the computer-based model could effectively help the students with self-practice and self-assessment of their speaking and thereby become responsible for learning English speaking skills.

An inference can be drawn that using technology can help with improving Vietnamese undergraduate EFL students' English speaking skill by giving them more chances to be more self-aware in learning English speaking skill.

Tuyen and Loan (2019) conducted a study in a university in Vietnam to investigate the EFL students' willingness to communicate in English and the factors contributing to this. The research chose 200 EFL students and used questionnaire surveys, semi-structured interviews, and class observation as the instruments for the research. The results revealed that most of the students were not willing to

communicate in English speaking classes and their unwillingness was resulted from their insufficiency in English language and shortage of confidence. Furthermore, the research provided an explanation that the students' low confidence in English speaking was because the students felt that their English speaking was not proficient enough, and mistakes might lead to the teachers' misunderstanding or might make them feel embarrassed.

In addition, difficult tasks and unfamiliar topics were found to impede the students from speaking due to their low proficiency in grammar and vocabulary. To improve the students' willingness to communicate in English classrooms, the research suggested improving syllabus design, correct placement of students at the correct level of English skills to help them feel more comfortable.

Also, changing teaching strategies were recommended to help improving the students' willingness to communicate in English classes. For this aspect, Anh and Nhu (2021) also stated that visual aids can help speakers release their stress when speaking by shifting the audience's attention from the speaker to the video aids. Therefore, it is understandable that visual aids may also be a useful application of technology to help the students become more willing to communicate in English classrooms as they feel more relaxed in this situation.

Thao and Nguyet (2019) implemented research on examining the four aspects of speaking difficulties faced by the EFL students in a university in Vietnam. The research involved 150 participants who were undergraduates. It employed a questionnaire as the instrument for investigation. Throughout the data collection procedure, the research found that the students had difficulties speaking English although they had already spent considerable time in practice. Importantly, the difficulties were found to be triggered because the students did not have sufficient chances to practice English speaking when they left the classrooms, and they found it difficult to cooperate with their classmates. In addition, the duration of English learning was reported to impair improvement if the time for practice was too short. Long English learning times might allow the students to explore helpful strategies to overcome their difficulties in English speaking. The researchers recommended the students' use of recording software on a computer to help them self-create opportunities for practicing English speaking and then having a native speaker evaluate the voice.

The research studies mentioned above suggest that it can be observed that although the students' lack of English speaking exposure, their insufficient English language proficiency, and their unwillingness to communicate prevented

them from achieving English speaking skills successfully, there are, nevertheless, possible interventions to help them overcome these problems.

Apart from applying traditional methods such as the application of intensive speaking in English classrooms and the encouragement of frequent speaking in English after class (Dao, 2017), the improvement of students' self-awareness (Nguyen & Nguyen, 2016), or the use of speaking strategies (Quyen et al., 2018), recent researchers have gradually shifted their attention to the use of technology to help solving the students' difficulties in English speaking.

Some of the technological applications can be found in previous research such as the use of videos to help the students prepare for speaking in English classrooms (Quyen & Loi, 2018), the application of computer-based models to improve the students' autonomy in practicing English speaking skill (Duc, 2017), the use of visual aids to reduce the students' stress when speaking in English classrooms (Anh & Nhu, 2021), or the use of computer software to evaluate the students' English speaking skill and provide feedback to them for improvement in English speaking skills (Thao & Nguyet, 2019).

However, none of the studies have ever considered the use of an AI voice chatbot as a possible solution. Therefore, it is a reasonable consideration to conduct research on an AI voice chatbot to investigate its effects on the English speaking skills of Vietnamese undergraduate EFL students. Technology-enhanced language learning (TELL) provides the theoretical guideline for using AI voice chatbots in learning English speaking process, which is the next area of discussion.

2.6 Technology-Enhanced Language Learning (TELL)

Technology-enhanced language learning (TELL) fosters the use of technology in the English learning and teaching process which has been applied widely in institutions to facilitate the learning objectives of students (Devlin, Feldhaus, & Bentrem, 2013). Theoretically, TELL is a wider range of computer-assisted language learning (CALL) engaging not only computers into the language learning process but also various types of technological devices such as phones, game consoles, and tablets (Walker & White, 2013).

2.6.1 The TELL Model

In exploration of the framework applied for the research, the TELL model proposed by Walker and White (2013) will be discussed in detail, illustrated in Table 2.1.

Table 2.1 TELL Model (Walker & White, 2013)

Approach	Structural/ restricted CALL	Communicative CALL Open CALL	Integrative CALL	TELL
Technology	From mainframe to mobile	PCs	Multimedia, internet	Mobile devices, tablets, multiplayer games, virtual worlds
English-teaching paradigm	Grammar-translation and audio-lingual	Communicative language teaching	Content-based ESP/EAP	Communication, interaction
View of language	Structural (a formal structural system)	Cognitive (a mentally constructed system)	Socio-cognitive (developed in social interaction)	Structural, cognitive, socio-cognitive, adaptable
Principal use of technology	Drill and practice	Communicative exercises	Authentic discourse	Normalized
Principal objective	Accuracy	Fluency	Agency	Autonomy within community
View of learning	Behaviourism	Constructivism	Social constructivism/ situated learning	Connectivism
Role of technology	Tutor	Tutee	Mediational tool	Environment, resource

In this model, there are seven approaches throughout the development stages from restricted CALL to TELL. In Restricted CALL approach, CALL is determined under the behaviorist view in which the teaching method focuses on grammar translation and audio-lingual data. The language learning comprises both a formal and structural system. The model by Walker and White (2013) seems to further engage accuracy to be the objective of language use and technology plays the role of a tutor who teaches learners. This can be implied that within the Restricted CALL approach, technology has already played a significant role during the complete teaching process.

In the Open CALL approach, the main technological device is the personal computer (PC) and communicative language learning is the teaching method. Within this stage, the language is supposed to deal with not only structures but also cognitive factors. There is a significant development within this approach because learning is determined with constructivist perspective focusing on communicative exercises facilitated by technology. Moreover, accuracy is not considered to be the only factor, as focus is shifted to language fluency. Consequently, technology plays a role of a computer learner which adapts from human learners through communicative exercises, which helps construct knowledge via communication during the learning process.

In the integrated CALL approach, the model centers upon multimedia and the internet as the means of technology. Within this approach, English teaching methods shift concern toward content-based learning, which features English for Specific Purposes (ESP) and English for Academic Purposes (EAP). This learning theory stands with a social constructivist view. As a result, language learning is moved from a cognitive to socio-cognitive perspective which focuses on the development of language through social interaction. Therefore, the principal use of technology does not only anchor communication exercises but it also does so with authentic discourse, the principal objective in which agency is seen as the focal point of language learning. Within this approach, technology plays the role of a mediating tool which helps learners achieve learning goals.

In the final and latest approach, CALL is transformed into TELL as technology can be featured by a variety of devices and tools such as mobile phones, tablets, games, and virtual worlds. This stage also witnesses the vast development among the approaches. The first and foremost transformation is from social constructivist view into connectivism, in which the central point of teaching method is about communication and interaction. This also leads to a change of the language learning perspective which sees language learning as both the socio-cognitive processes and the structural, cognitive, and adaptable processes. Interestingly, technology is normalized in this stage, and autonomy within community is the main principal objective of TELL. In the next step, we will further explore the learning theories behind TELL to gain theoretical support for our research.

2.6.2 Learning Theories behind TELL

According to Walker and White (2013), TELL was developed under the a connectivist framework inherited from the theories of behaviorism, constructivism, and social constructivism.

2.6.2.1 Behaviorism

The behaviorist view claimed that the development of language learning was through stimulus and achieved by both positive and negative reinforcement (Skinner, 1957 as cited by Chomsky, 1959). Within this framework, learning a foreign language refers to habit formation and the main objective of language learning is to use language as a native speaker. Therefore, behaviorism only centers the training of listening comprehension, accurate pronunciation, and letter recognition and reproduction. Within the behaviorist paradigm, the principle is that machine is seen to contain the knowledge and it delivers the knowledge to learners in terms of drill and practice applications or small chunks which need to be

reinforced frequently. Based on this learning theory, the machine helps learners learn by using drill and practice programs focused on the exercises like multiple-choice, short answer questions, gap-filling, re-arranging words. Learners can learn and revise their understanding independently as well as feel more secure by those mentioned activities. By all appearance, the knowledge is right there for learners, safe and sound, and what learners need to do is practice these exercises for learning to occur. The machine plays the role of an instructor to the students (Walker & White, 2013) in behaviorist paradigm.

2.6.2.2 Constructivism

In the constructivist framework, teaching follows a communicative language teaching (CLT) approach (Richards & Rodgers, 2014) which identifies communication and interaction as the main goals of learning language. This model focuses on learners construction of knowledge through communication and interaction. This approach is based on the theory of constructivism as suggested by Piaget (1964), that learners can best develop their learning when they can construct their knowledge by themselves from the existing knowledge which they have gained. Consequently, language learning involves the development of linguistic skills, communication, and interaction rather than simply constructing texts through drills and practice as guided by the behaviorism theory. Within constructivist paradigm, the principle is that machine no longer helps learners to learn by its ready-made knowledge but by allowing learners to teach, learn, and share so that learners can construct knowledge by themselves. Based on this learning theory, the software used for teaching is not only the software used by teachers but also by learners so that both teachers and learners can create activities for each other to construct knowledge together. Learning within this paradigm focuses on the learners ability to develop knowledge. Therefore, Walker and White (2013) simulated the role of the machine in constructivist perspective to be the learner in learning process.

2.6.2.3 Social Constructivism

The social constructivism originated from Vygotsky's (1978) learning theory claims that learning is shaped through social interaction before being formed by the individual aspect. With social interaction, the goal of language learning is to engage language learners in communication activities in which all factors such as text construction, linguistic skills, and interaction are combined to build up a successful language learning process. Vygotsky (1978) stated that learning is triggered by the zone of proximal development (ZPD) which refers to the connection between the ability and knowledge of learners with those of others who have more knowledge

and ability. This statement can be understood as learners learn and improve their knowledge through interacting and communicating with those who have more expertise than the learners themselves. The social constructivist approach was also reflected in Krashen's (2013) input theory in which comprehensible input plays an important role in language acquisition as learners need to acquire language through receiving the input from hearing and reading. However, this input must be at a higher level than learners' knowledge and ability. Under the social constructivist framework, machines were acknowledged as "tools" which facilitate communication and interaction between people. Based on the social constructivist paradigm, the machine can be used to teach, be taught, and facilitate learning process. Word-processing programs or editing-software are some examples of the "tool" for learning.

2.6.2.4 Connectivism

The theory of connectivism suggested by Siemens (2005) explains that learning is the process which involves teachers, learners, community, and technology together in a process. Within connectivist theory, learners learn with various ideas and opinions stored in machines, and with that stored knowledge, machines play the role of "tutor". Therefore, it can be implied that the behaviorist paradigm is reflected in connectivism. The constructivist paradigm is also reflected in connectivism. This statement can be supported by the assumption of Discourse Community mentioned by Walker and White (2013) relating learning to the glue which connects people in a community together. Within this perspective, an overlooked component which should be involved is the environment. The environment includes both physical and virtual entities and it is the place in which people can learn through developing the relationship with each other when they share similar interests. Since the discourse community relates the learners' shared knowledge in the connected community and involves the transformation of learners' knowledge into their construction of knowledge through the action of building relationship and communicating during the learning process, it can be implied that this assumption is connected with the constructivist perspective discussed by Piaget (1964). Finally, if the knowledge was stored in machines, then it can be inferred that people may use the machine to store their knowledge as well. If so, the machine can also be seen as a technological resource facilitating the learning process. Therefore, it can also play the role of a "tool" as assumed in the social constructivist paradigm as mentioned above.

Walker and White (2013) utilized the Communities of Practice to explain the phenomenon in which learning is connected with practicing communities. Within this aspect, the learners who belong to a community learn by engaging in the community. This can be implied that the more learners who engage in the community, the more their expertise is developed. Learners may need to be supported by the members whose expertise is far beyond what learners currently know. However, previous researchers argued that conventional classrooms could not apply this concept for learning because there were no such expert members in the classrooms to help learners to go beyond what they already know. Fortunately, with the development of technology, learners can easily find an authentic community for learning and practicing with expert members in the community through the internet. They can use Facebook, Youtube, and Google to find support from the community for their learning. Therefore, it can be seen that technology plays a role of providing resources within the learning environment.

Based on the discussion mentioned above, it can be seen that this assumption is in line with the view of social-constructivism as it likely refers to the zone of proximal development by Vygotsky (1978) and the comprehensible input by Krashen (2013). Siemens (2005) suggested some learning principles guided by the connectivist paradigm, including learning determination, learning objectives, learning method, and learning process. In determining learning, connectivism suggests that learning is to understand connections in fields, ideas, and concepts. In determining learning objectives, the learner's ability to know more is very important. It's more crucial than learner's ability to know what. Therefore, the activities of learning must be accurate and the knowledge base needs to be updated continuously.

Additionally, learning should be guided by various opinions and connection which should be focused, nurtured, and maintained. Within the connectivist paradigm, learning process is conducted and associated with non-human appliances. The process should be continual and connect sources of information and specific nodes. Moreover, during the learning process, learners should be allowed to learn from reality and to decide what to learn. There are not only computers which can serve the language learning process, but other technological devices such as mobile phones, smartphones, and tablets can also play the same role, which highlighted the need of TELL in providing principles and guidelines for the learning process through technology.

Based on the above discussion, it can be concluded that TELL is constructed under the connectivist paradigm in which foundational theories and

principles of behaviorism, constructivism, and social constructivism are reflected. Therefore, technology can play many roles in the learning process. It can be the instructor to teach learners, the learner to learn from learners, or the tool to facilitate learners in the learning community with both physical and virtual entities.

2.7 Conceptual Framework for Developing the Lessons

The conceptual framework of this study was constructed based on the connectivist paradigm guided by TELL (Walker and White, 2013). The framework provided a clear concept of teaching English speaking skill in three stages: (1) pre-speaking, (2) while-speaking, and (3) post-speaking. Figure 2.14 is the conceptual framework for developing the lessons.

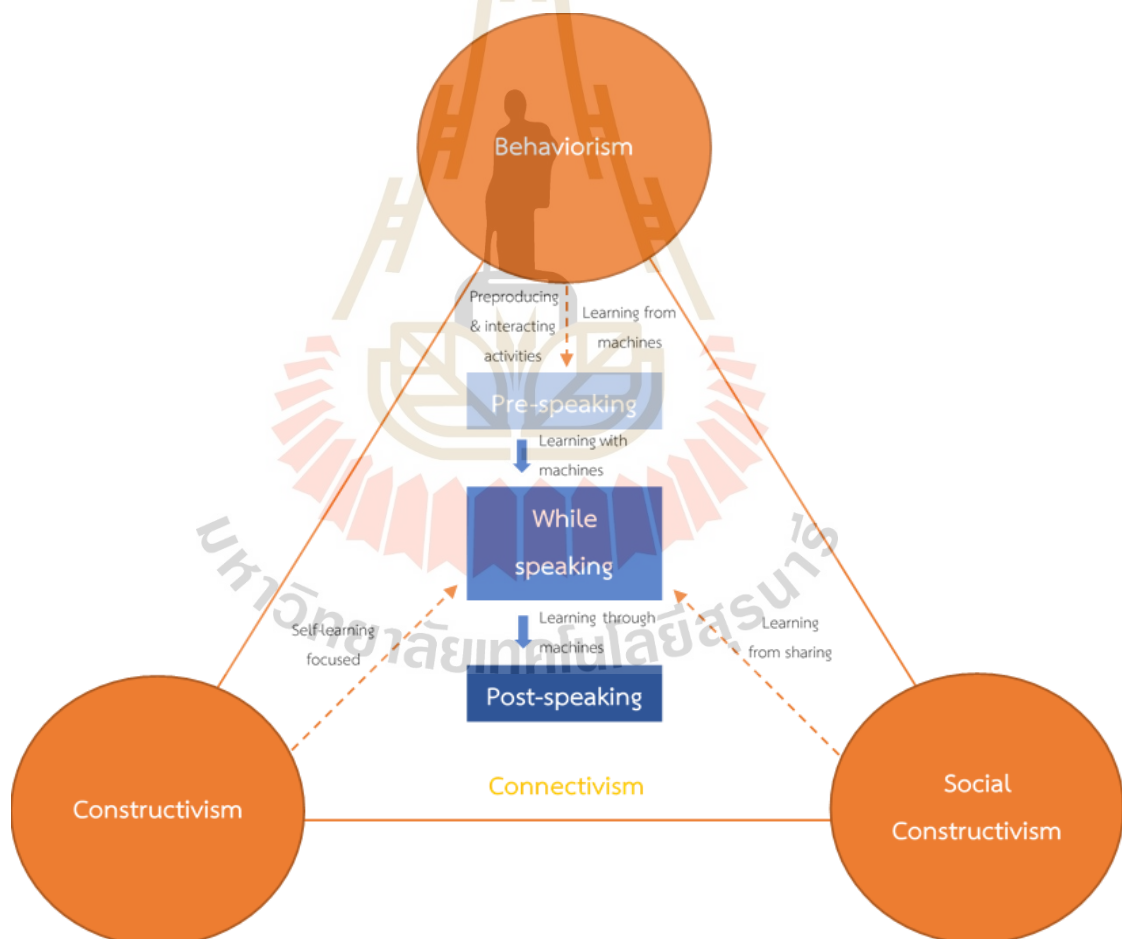


Figure 2.14 Conceptual Framework for Developing Lessons

2.7.1 Pre-Speaking

In pre-speaking stage, the behaviorism (Skinner, 1957) and connectivism (Siemens, 2005) theories should be applied in terms of warm-up and action-based

activities. Within this stage, the students learn to form habits through reproducing and interactive activities assigned in group work in which machines play the role of instructors. Accordingly, warm-up activities should be useful for students to recall and find ideas and knowledge related to the given tasks to share and support each other in their group in achieving the learning objectives. Apart from warm-up activities, the students can be taught with linguistic knowledge by a machine to have a bottom-up understanding before practicing.

2.7.2 While-Speaking

The while-speaking stage should be associated with constructivism (Piaget, 1964), social constructivism (Vygotsky, 1978; Krashen, 2013) and connectivism (Siemens, 2005) based theories reflected in practicing activities. Within the constructivist, social constructivist, and connectivist frameworks, the students learn and construct knowledge by themselves, learn from developing connection with people in the community, learn from those who possess expertise knowledge, and learn with non-human appliances. Therefore, the students should be provided with opportunities to learn with the machine in individual work to be autonomous in learning and be assigned with paired work to be able to learn from each other and share various ideas during the learning process. Students should also learn with the machine, and then in paired work activities to share their opinions with a partner. Sharing each other's experiences in speaking allowed each student to learn from others to improve English speaking skills.

2.7.3 Post-Speaking

In the post-speaking stage, the theories of behaviorism (Skinner, 1957), constructivism (Piaget, 1964), social constructivism (Vygotsky, 1978; Krashen, 2013) and connectivism (Siemens, 2005) are integrated into individual tasks to help the students develop their self-study ability so that they can learn and construct their knowledge of vocabulary, pronunciation, and grammar by themselves. This supports their developing existing knowledge through communication and connection with the community through a non-human appliance whose knowledge is more proficient than the students' own knowledge.

2.8 Summary

In conclusion, this chapter has reviewed the information on English speaking skills and how Vietnamese undergraduate EFL students learn English speaking in traditional classrooms, where many problems were discovered. The chapter has also reviewed the approaches of TELL which are considered to be the theoretical

guideline for conducting the research. Significantly, the in-depth review of AI chatbot characteristics and structures together with information about how AI voice chatbots contribute to English language learning has been discussed. In the next chapter, this research continues with the description of the research methodology.



CHAPTER 3

RESEARCH METHODOLOGY

This chapter is focused upon the research methodology for this study. This chapter is organized with nine sections including research context, research design, participants, instruments, data collection procedure, data analysis, trustworthiness and validity, and pilot study.

3.1 Research Context

The study was conducted at Can Tho University in Vietnam. Data collection was limited to the Faculty of Foreign Languages because this faculty is in charge of the education and training of English to all students, regardless of major. The Vietnamese undergraduate EFL students who participated in this research came from any non-English degree based academic programs, with any major. These students were enrolled in the English 2 level course. Each week, the students were required to be in the classroom for nine hours to learn English basics. The Faculty was responsible for screening the students by using TOEIC Bridge as the placement test to enter the course. The main type of training is regular teaching, and the English classes occur onsite, which requires the students to be in the classroom to study. The English 2 course is taught in traditional classrooms in which the teacher gives lectures while learners listen and take notes. More recently, learners have been also allowed to use smartphones to facilitate their learning process. This was an advantage for this study because we could employ Andy English Bot which supports Android and iOS to integrate with in-class teaching. Therefore, with available smartphones, the target students could easily access the chatbot to speak English.

3.2 Research Design

This study applied an experimental design with one independent variable and two dependent variables. The independent variable was the use of an AI voice chatbot in teaching English speaking skill. The dependent variables included (1) the students' speaking test scores from the pre-speaking and post-speaking tests and (2) their opinions about using an AI voice chatbot. A mixed method design was conducted because this study contained two factors: (1) the between-subjects factor

which was the students speaking in the traditional classrooms and those speaking with an AI voice chatbot; and (2) the within-subjects factor which was the measurement before and after the intervention. According to Bordens and Abbott (2011), a mixed method design is one that includes a between-subjects factor and a within-subjects factor, used to assess the effects of the variables which cannot be controlled within subjects so as to maintain statistical advantage for studying the remaining variables. For quantitative design, the researcher employed the pre-speaking and post-speaking tests and a survey questionnaire to collect numerical data of the students participating in the experiment. With the pre-speaking and post-speaking tests, the researcher would like to compare whether the interventions led to significant differences in learning. The survey questionnaire would help to explore the students' opinions of the intervention. For the qualitative design, the researcher used a list of interview questions to collect descriptive data from the participants. Based on the participants' responses to the interview questions, the researcher provided an in-depth investigation on the students' opinions about the intervention. The research was designed within six steps starting from instrumental design, then the pilot study, the pre-speaking test, the in-class teaching, the data collection, and finally the data analysis. In the first step, the researcher designed research instruments and instructional instruments to facilitate the experiment. After that, a pilot study was conducted to assess the initial plan of the experiment and evaluate the accuracy and validity of the instruments. In the third step, the researcher employed the pre-speaking tests to evaluate the students' English speaking ability before the intervention. The in-class teaching was conducted in the fourth step so that the researcher could have an intervention in the learning process of the students. In the fifth step, the data from the speaking tests together with the students' responses to the questionnaire and interview were collected for data analysis.

3.3 Participants

In selection of the participants, the purposive sampling technique introduced by Schreiber and Asner-Self (2011) was employed. The participants were assigned to two different groups which were the control group and the experimental group. Sample size was an initial step to discover the sufficient number of the necessary participants. Schreiber and Asner-Self (2011) stated that selecting sample size can be conducted by analyzing certain statistical measurements or by following the an agreed upon rule set. Accordingly, the current study consisted of 60 students who

were the non-English majored undergraduate students and were studying English 2 in Can Tho University in Vietnam. As the courses of English 2 are open every trimester in Can Tho University, it was convenient to select the target students who enrolled in this course. Moreover, the homogeneity of the students was ensured because it depended on their English placement test score which decided by testing to the L2 level of fluency. Regarding the placement test score, the Rector of Can Tho University issued the Decision No. 3003/QD-DHCT on July 22nd, 2021 regulating the course exemption and certification of course completion in academic programs for regular undergraduate education and training. Within this Decision, the TOEIC bridge was taken as a placement test to consider if a non-English majored undergraduate student was permitted to be exempted from a General English course including English 1, 2, and 3 or he/she was required to take some of the General English courses. Within this regulation, if the students TOEIC Bridge score is from 90 to 100, they will be exempted from taking the General English courses including English 1, 2, and 3. If their score is under 89, then the Faculty of Foreign Languages will decide if the students will be placed in either English 1, 2, or 3 based on their overall TOEIC Bridge score. In consideration of the overall scores, English 1 is required for those who scored between 30 and 42. English 2 is mandatory for the students who scored from 43 to 75. English 3 is mandatory for those who scored from 76 to 89. Table 3.1 illustrates the standard scale for evaluating English proficiency. Accordingly, the students English level was equivalent to A2 when they participated in the research.

Table 3.1. The Standard Scale for Evaluating English Proficiency

Standard scale for 6-level of English proficiency used in Vietnam	Equivalent with CEFR standard	English levels A, B, C	IELTS	TOEIC (L&R)	TOEIC Bridge	TOEFL IPT	TOEFL CBT	TOEFL iBT	Cambridge Tests
Level 1	A1	A		120-220	30-42		60		100-120 KET
Level 2	A2	B	4.0	225-445	43-75	360-449	96	30	120-140 KET 120-140 PET
Level 3	B1	C	4.5-5.0	450-595	76-89	450-499	133	31-45	140-150 KET 140-160 PET 140-160 FCE
Level 4	B2		5.5-6.5	600-845	90-100	500-589	173	46-93	160-170 PET 160-180 FCE 160-180 CAE
Level 5	C1		7.0-7.5	850-940		590-649	213	94-109	180-190 FCE 180-200 CAE 180-200 CPE
Level 6	C2		8.0-9.0	945-900		650-677	250	110-120	200-210 CAE 200-230 CPE
			Top Score 9	Top Score 990	Top Score 100	Top Score 677	Top Score 300	Top Score 120	

(Decision No. 3003/QD-DHCT on July 22nd, 2021 by the Rector of Can Tho University).

3.4 Instruments

This study utilized a variety of instruments to conduct the experiment. The instruments were divided into two types: research instruments and instructional instruments. The research instruments included the pre-speaking and post-speaking tests, the questionnaire, and the semi-structured interview while the instructional instruments included an AI voice chatbot, a learning website, a speaking rubric for assessment, and finally, lesson plans for the control group and the experimental group.

3.4.1 Research Instruments

Pre-speaking and post-speaking tests, the questionnaire and the semi-structured interview were the primary research instruments. Accordingly, the pre-speaking and post-speaking tests were used for evaluating the participants' English speaking skill before and after the experiment. The questionnaire and the semi-structured interview were utilized to explore the participants' opinions about using an AI voice chatbot for improving English speaking skill.

3.4.1.1 The Pre-speaking and post-speaking Tests

The pre-speaking and post-speaking tests (See Appendix A) were developed based on the CEFR's principles for assessment to evaluate the students' speaking skill. Those principles suggest that the assessment should be based upon five aspects, including context, purpose, linkability, production, and standards. For the context, the speaking tests were designed for the non-English majored undergraduate EFL students who learn English 2. The purpose of the tests was to evaluate the students' speaking skill based on two criteria, fluency and accuracy. The test results were linkable because they might help examiners to predict the students' English speaking skill. Additionally, the test questions and topics could be used and re-used for many times to serve for English speaking assessment, which met the principle of production. Finally, the tests' structure was standardized as it was designed based on the IELTS speaking test structure developed by the British Council, consisting of three parts. Part 1 included ten questions in which every two questions referred to one familiar topic such as food, study, hometown, daily activities, or family and friends. Part 2 included five topics for individual discussion in which the students talked about one selected topic among the five such as books, people, country, exercise, or festival. Part 3 featured two-way discussion and contained ten questions. Every two questions referred to one topic in part 2.

The IELTS speaking test structure was adopted for this research because it was investigated and confirmed as practical for scoring and interpreting

the results (Karim & Haq, 2014). Besides, the reliability and validity of the IELTS speaking test were investigated and confirmed by Li (2019) who reported that the test could assess candidates' speaking ability in terms of fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation.

In addition, due to the observation that test takers' scores may not necessarily demonstrate their true competence, multiple other measures may help to provide more reliable and valid assessment than a single measure. As a result, three speaking parts were utilized to assure test reliability and validity. It is worth mentioning that in speaking part 1, the students were asked to answer only one topic, and in part 3, only two questions related to the topic the students selected in part 2 were asked.

The speaking test pack containing 25 questions was provided to three experts for validity evaluation and all questions were evaluated with the score of +1 by the experts (see Appendix N). This suggested that the questions clearly measured the students' speaking skills in terms of accuracy and fluency.

3.4.1.2 The Questionnaire

A questionnaire was utilized in this study as it allowed anonymous responses, to improve honesty of the feedback. The questionnaire was developed based on Bordens and Abbott's (2011) framework (See Appendix B). Within this framework, three features were considered before designing the questionnaire. First, keep the survey short to not overburden subjects. Second, provide a broad enough range of questions to keep track of the focused phenomenon. Third, responses are focused to avoid extraneous information. Accordingly, the questionnaire designing process started with defining the topic of this study to have clear and concise definitions to avoid ambiguity when interpreting the data. After that the questionnaire was developed, consisting of three parts including demographic information, closed-ended items and open-ended items. There were three reasons for this. First, the demographic information should be used to investigate the participants' background before the intervention. Second, if only closed-ended items were provided in the questionnaire, misunderstanding might occur because different items might be understood differently between subjects. Finally, with open-ended questions, the participants could communicate more accurate understanding. However, open-ended questions needed to be limited in number to prevent fatigue of the subjects.

The demographic section included six questions which related to the participants' general information: gender, age, major, year of study, English

entrance score, and technological experience to explore the participants demographic information. The demographic information set was used as a predictor variable.

The second part contained five close-ended questions to explore the participants' opinions on their improvement in English speaking skill after the experiment. The close-ended items were employed for ease in summarizing and analyzing responses. The items in the second part were classified into two aspects – fluency and accuracy. Accordingly, questions 7-8 were to explore the participants' opinions on improving English speaking skill in terms of fluency, reflected by the factors such as hesitations or pauses, and their frequency of using interjections such as ah..., uhm..., oh... Questions 9-11 explored the aspect of accuracy in terms of pronunciation, grammar, and vocabulary.

The third part allowed free writing responses for the participants to share opinions or suggestions on using an AI voice chatbot to improve English speaking skill. The open-ended items were used to allow the participants to elaborate their further opinions. The questionnaire was evaluated by experts who were the lecturers in the field of English speaking and familiar with using technology in language learning to ensure the reliability and validity.

A five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) following the framework of Cohen et al. (2007) was utilized to measure the students' statements on the questionnaires. The five-point Likert scale is a way to measure participants' opinions in the research and it was widely used for quantitative data collection in previous studies by (Each & Suppasetserree, 2021; Mukhallafi, 2020; Thao et al., 2019). Moreover, the five-point scale is simple to understand and rapid to answer. Actually, five-point scale was reported to be user friendly (Bordens & Abbott, 2011). Notably, Bordens and Abbott (2011) also suggested the importance of organizing the questionnaire items into a coherent and easy-to-follow format. The questionnaire items were translated into Vietnamese so that the participants easily respond.

In validating the questionnaire items, the researcher delivered all eleven questionnaire items to three experts for their evaluation. As a result, all questionnaire items were rated +1 by all three experts (see Appendix N), which meant that the questionnaire item clearly measured the objectives of the research.

The students' responses to Part 2 of the questionnaire referred to their opinions on using the AI voice chatbot and were interpreted with ordinal scale ranked by the rules of "less than" or "greater than" instructed by Cohen et al.

(2007). Since the questionnaire items in Part 2 were ranked in the order of (1) strongly disagree, (2) disagree, 3 (neutral), 4 (agree), and 5 (strongly agree), the responses to such items could also be interpreted with an ordinal scale in association with the rules of "less than" or "greater than" as long as the interpretation complied with the rule: "crude data can yield only crude interpretation" (Cohen et al., 2007). Accordingly, the responses to each questionnaire item were interpreted by adapting the interpretation framework by Banditvilai (2016) as "strongly disagree" if the mean score was lesser than 1.51 but greater than 0.99, "disagree" if lesser than 2.51 but greater than 1.50, "neutral" if lesser than 3.51 but greater than 2.50, "agree" if lesser than 4.51 but greater than 3.50 and "strongly agree" if lesser than 5.10 but greater than 4.50. Table 3.2 shows the interpretation of the questionnaire responses.

Table 3.2. The Interpretation of Points in the Questionnaire Part 2

Mean scores (M) scale	1.00-1.50	1.51-2.50	2.51-3.50	3.51-4.50	4.51-5.00
Interpretation	Strongly disagree	Disagree	Neutral	Agree	Strongly agree

3.4.1.3 The Semi-Structured Interview

The interviews were conducted to investigate more opinions from twelve participants among the thirty who were in the experimental group because an individual interview might allow the participants to share their opinions more completely. A semi-structured interview technique was applied as this was more open to respondents and they could fully share their opinions on specific issues (Cohen, Manion, & Morrison, 2007). For the semi-structured interview, the informant sampling method suggested by Schreiber and Asner-Self (2011) was utilized as the researcher would like to purposively select and talk with twelve participants, six of whom scored the highest and six of whom scored the lowest in the post-test. The participants' highest and lowest post-speaking scores suggested that these participants appeared to be distinct after taking the intervention in the experimental group. Therefore, it was predicted that they could share significant opinions after practicing English speaking with an AI voice chatbot.

The semi-structured interview (See Appendix C) was constructed following the non-directive framework suggested by Cohen et al. (2007). The semi-structured interview contained a smaller size of the sample and more open questions. Within the non-directive framework, the respondents felt free to share their opinions based on a few set questions, and the researcher could also prompt

for clarification on items. Accordingly, the semi-interview was conducted with seven guided questions to explore the students' process and opinions about using an AI voice chatbot to improve English speaking skill so that sufficient information could be collected for interpretation and exploration. These guided questions aimed at exploring the participants' practicing process, their opinions on speaking improvement in terms of accuracy (grammar, pronunciation, vocabulary) and fluency, and exploring the positive and negative aspects of an AI voice chatbot.

Zoom was employed to conduct the interviews to improve flexibility in terms of time and place for both the interviewees and the researcher. The interviewees were informed that their voice would be recorded during the interview to serve for transcription and interpretation.

The interview questions were also checked and validated by the experts to ensure their reliability and validity. Accordingly, the interview questions including seven guided questions were sent to three experts for evaluation, and all seven questions were rated +1 by the three experts, which indicated that those seven questions could clearly measure the research's objectives (see Appendix N).

3.4.2 Instructional Instruments

The instructional instruments used in this study included an AI voice chatbot named Andy English Bot, a learning website, a speaking assessment rubric, and the lesson plans for the control group and the experimental group. Andy English Bot served for the students' practice of English speaking, the learning website facilitated the students to learn English speaking skill on their own, the speaking assessment rubric provided grading criteria for speaking assessment after the pretest and posttest, and the lesson plans served for instructing the students in the control group and those in the experimental group.

3.4.2.1 The AI Voice Chatbot

This study used an AI voice chatbot named Andy - English Speaking Bot (Andy English Bot) as a tool for the students to practice English speaking. There were several purposes for selecting Andy English Bot as the practicing tool in the experiment.

First, Andy English Bot could be found and downloaded easily on Google Play or App Store. The students only needed to visit Google Play or App Store and type the keyword "Andy". The app would appear immediately on the first search result on Google Play or App store. Thus, most of the participants who joined the research could have Andy English Bot installed on their smartphone with ease.

Second, Andy English Bot was powered by the ZTO Labs which is a trusted technological company. With the company's policies on privacy and security, it was safe for users to install Andy English Bot on their smartphone as this AI voice chatbot did not collect or share any user data with third parties and it was also free for use unless the users preferred premium features such as complicated speaking topics, more exercises allotted per day, more grammar lessons, and unlimited vocabulary learning. As this study focused on less complex oral tasks, the basic and free features provided by Andy English Bot were considered adequate for the students to use.

Lastly, Andy English Bot received many positive reviews from previous research, one of which was that this AI voice chatbot could act as a virtual English native speaker and an English tutor (Kim et al., 2019) which could facilitate users in learning English speaking, vocabulary, and grammar. Therefore, this study expected that Andy English Bot could provide authentic inputs for the target students when they practiced English speaking.

In addition, Andy English Bot could communicate with users by both text and voice chats, which was appropriate in this study because its focus was the intervention of an AI voice chatbot into teaching English speaking skill.

One drawback of Andy English Bot reported in a previous study was that it failed to remember or save users' conversations (Kim et al., 2019). However, the mentioned drawback might be seen as an advantage because the students would not get bored with repeated answers made up as a result of the chatbot's remembering previous conversations.

3.4.2.2 The Learning Website for The Experimental Group

A learning website was created to facilitate the students in the experimental group to study English speaking. The lessons in the website were developed based on the lesson plans for the experimental group including five lessons featured with five topics: Journeys, Interests, Entertainment, Learning, and Tourism. Each lesson had specific instructions for the students to learn by themselves. Accordingly, the students could use their smartphone to access the website with the link <https://aispeaks.mobirisesite.com/> to learn these English speaking lessons during class time. The teacher could stay in the class to facilitate the students in technical training and solving technical issues. Moreover, this website also included the links for the students to download Andy English Bot and supportive apps such as Voice Recorder and Voice Access to facilitate the students' learning process. During the instrumental evaluation process, the website was sent to

three experts in English Language Teaching field for their evaluating together with the lesson plans to make sure it was validated to apply in teaching.

3.4.2.3 The Speaking Rubric

The analytic speaking rubric (See Appendix E) was designed as an analytic rubric whose columns stood for levels of achievement and rows for criteria of assessment. The analytic rubric was used because it allowed examiners to assess the participants' level of achievement based on multiple criteria. This type of rubric also allowed the designer to assign various values to different criteria to determine an overall achievement by totaling the subscores. The designed rubric was adapted from the Tohoku Fukushi University (TFU) Foreign Language Assessment Rubrics designed by TFU Language Educators Group (TFU-LEG). It was used in an examination context for the one-on-one interview test.

The rubric was adapted because it mainly served for grading learners' English speaking skill in terms of accuracy and fluency. This was in line with our objectives which focused on evaluating the students' English speaking skill based on their accuracy and fluency. However, some changes were made so that the speaking rubric could better facilitate the speaking assessment process. For the first update, examples for each criterium were added so that examiners would have a clearer description for assessment. For the second update, the test takers' English speaking skill was classified into the three levels recommended by the Common European Framework of Reference for Languages (CEFR) standards which included: insufficient user, basic user (A1), and independent user (A2).

Actually, the TFU's speaking assessment rubric classified the test takers' proficiency into five scoring levels which are "0-not able to perform", "1-inadequate", "2-needs improvement", "3-meets expectation", and "4-exceeds expectations". However, when compared their descriptions with the CEFR standards (Cambridge, 2011), this study found that those five levels appear to be equivalent to A1 and A2 levels. According to Cambridge (2011), level A1 was described in terms of learners' understanding and using language satisfied the level of using everyday expressions, introducing, giving and asking about personal details, speaking slowly and clearly. The TFU-LEG's speaking rubric also used similar measures for the speaking scores achieved by test takers in terms of their ability to understand and use basic, simple, and appropriate language and expressions. To simplify the scoring process, a range from 0 to 5 points was used for evaluating the test takers' levels of achievement based on two criteria, fluency and accuracy.

The speaking rubric was delivered to the experts for validation and the evaluation; results revealed that most of the criteria were evaluated to clearly help the examiners to evaluate the participants' English speaking skill because the average scores of those criteria were 1.00 collected from the three experts (see Appendix N). However, criterion number 1, 13, and 20 got the average score of 0.67, which meant that two experts agreed on the possibility of those criteria in facilitating the examiners in the scoring process for the speaking tests but the remaining expert was unconvinced. Based on the pre-determined convention mentioned previously in section 3.3.1, it could be concluded that the speaking assessment rubric could help the examiners with the scoring process because no criteria was below a 0.67. However, one expert suggested simplification of the assessment rubric so that the prospective examiners would find it easy to respond when giving score to the participants' speaking skill. As a result, the researchers of this study decided to make the speaking assessment rubric simpler by reducing the criteria to two main criteria which were speaking fluency and speaking accuracy. Based on the literature review mentioned in Chapter 2, it was reasonable that speaking fluency could be observed based on the speaker's pauses and hesitations while speaking accuracy could be observed based on the speaker's accurate use of vocabulary, grammar, and pronunciation. Each criteria also had five levels but the levels were not mentioned in the assessment rubric. Instead, those five levels were replaced with the band score 1-5 so that the examiners would find it simple to refer to when they gave scores for the participants' speaking skill. Within this aspect, the score of 1 means "not able to perform", the score of 2 means "inadequate", the score of 3 means "needs improvement", the score of 4 means "meets expectation", and finally the score of 5 means "exceeds expectations".

3.4.2.4 The Lesson Plans

In instructing the English speaking lessons, there was a lesson plan for teaching the students in the control group and a different one for teaching the students in the experimental group. The former lesson plan was used by a teacher in a traditional classroom of English speaking while the latter one was used to instruct the students to learn English speaking by themselves with teacher's facilitation in technical issues..

3.4.2.4.1 The Lesson Plan for the Control Group

The lesson plan for the control group included five units featured with five topics: Journeys, Interests, Entertainment, Learning, and Tourism taught in a traditional class. Each unit took three periods in which 45 minutes were

be spent for each. In each period, there were specific activities of teaching and learning. Normally, the teacher introduced the lessons and let students learn vocabulary and grammar from the textbook. Then some paired work or group work activities were given to students to do. After that, the teacher asked students to listen to sample dialogues from the textbook to practice listening. Next, the teacher asked students to practice English speaking with their partners while the teacher observed students speaking performance. During practice time, the teacher instructed students to take turns speaking with each other based on the textbook's suggestion to practice English speaking. Finally, the teacher gave feedback to students' English speaking performance at the end of each lesson so that students could become aware of their strengths and weaknesses in English speaking for future improvement.

3.4.2.4.2 The Lesson Plans for The Experimental Group

In the experimental group, the lesson plan was designed with the purpose of getting the students practice English speaking with Andy English Bot to improve their English speaking skill in terms of fluency and accuracy as Kim (2016) reported that an AI voice chatbot could enhance the students' English speaking skills. Accordingly, the structure and contents of the lesson plans were designed based on the CEFR's Can-Do Statement suggested by Cambridge (2011) with the application of Task-Based Language Learning method.

Regarding the Can-Do Statements, it is explained as the description of what language learners can do at different stages of their learning across five language skills: spoken interaction, spoken production, listening, reading, and writing. The overall focus of Can-Do Statements is on the learners' production and their ability to engage in conversations and discussion (Cambridge, 2011). For example, an A2 learner at pre-intermediate level of English proficiency can apply simple techniques to begin, continue, and leave the conversation while a B2 learner at intermediate English proficiency level can appropriately use complex phrases or sentences to develop their conversation.

There are six target levels concerned in the CEFR: levels A1, A2, B1, B2, C1, and C2. Level A1 evaluates the learners in terms of their ability to understand and use basic expressions and phrases when communicating while level A2 additionally evaluates their ability to make sentences to communicate. Level B1 evaluates the learners in their ability to understand key points of standard input, dealing with common situations, producing simple connected texts, and using language to describe and share familiar topics. Level B2 shifts the evaluation to the

learners' ability to understand complex text, fluent interaction, and production of detailed text for discussion. Level C1 and C2 drive the evaluation on learners' speaking proficiency in terms of understanding contextual meaning and well-organized discussion using language.

As the subjects in this study were the non-English majored undergraduate students, the lesson plan focused on getting the students to achieve level A1 and A2 only, which meant that the students were expected to be able to understand basic topics and use appropriate phrases or sentences to talk about familiar topics..

In addition, the lesson plan was designed with five lessons but those lessons were taught from a learning website. Each lesson was separated into three periods which took 45 minutes each. There were three stages of teaching and learning which were before speaking, while speaking, and after speaking. In the first stage, students were given a task to do on the website based on the current topics of the lesson to learn vocabulary, pronunciation, and grammar. In the second stage, students practiced English speaking with Andy English Bot and focused on the correct use of vocabulary, grammar, and pronunciation when speaking. In the final stage, students took a speaking exercise with Andy English Bot and then self-evaluated their English speaking skill by looking at the chat screenshots and listening to the voice chat recordings to see how they improved throughout the learning stages before getting feedback from the teacher.

The lesson plan focused on five main topics which were journeys, interests, entertainment, learning, and tourism. The reason for choosing these topics was that they might be familiar to most of undergraduate EFL students at A2 level. These topics were chosen based on the finding of Ellis (2009) that less complex tasks might enhance oral fluency.

The lesson plans contained eight survey statements and each statement was rated by the five-point likert scales from 1 to 5, respectively equivalent to strongly disagree, disagree, neutral, agree, and strongly agree. The evaluation results revealed that there were three statements which received a 100% agreement from three experts: statement numbers 3, 7, and 8. Statements 1, 5, and 6 received a 93.33% agreement from the experts. Finally, statements 2 and 4 received an 86.67% agreement from the experts (see Appendix N). The results implied that the lesson plans met the teaching objectives.

3.5 Data Collection Procedure

Overall, this study collected both quantitative and qualitative data to achieve the research objectives and answer the research questions. The instruments used thorough the study were evaluated by experienced experts to ensure the validity and reliability of each instrument before being applied into the experiment.

For general process, sixty participants participated in the eight-week experimental period starting from the tenth of April, 2023 with introduction sessions, orientation sessions, and pre-speaking tests after the researcher had got an approval from the Ethnics Committee for conducting the experiment. The number of participants and the term of experiment were determined based on the research design by Han (2020); Kim (2017); and Podesva and Sharma (2013). In addition, an eight-week experimental period was taken for the research to attempt to avoid Hawthorne effect which could possibly cause participants' over achievement due to their knowledge of being in an experiment (Kim, 2018).

Before participating in the research, the subjects were given an information sheet and a consent form for their review and signature to ensure that they agreed to participate in the experiment by their own free will. This process was to satisfy the requirement for ethics in human research.

The study collected the quantitative data from the students pre-speaking and post-speaking test results to analyze and find answers to research question 1. The students responses to the questionnaire were collected and analyzed quantitatively while the results from the interview were collected and analyzed qualitatively to find answers to research question 2.

3.5.1 Quantitative Data Collection

In collecting the quantitative data, 60 participants took a pre-speaking test at the beginning so that the researcher could have baseline data on the students English speaking proficiency before the intervention. They were then randomly assigned in two different groups: a control group and an experimental group. Each group contained 30 students.

The students in the control group learned with a textbook and spoke English in a traditional English speaking classroom while the experimental group learned on the website and practiced English speaking with Andy English Bot. Kim et al. (2019) reported that this AI voice chatbot could serve as communicative and learning tools for the students to learn English speaking.

The researcher was in charge of teaching English speaking to the students from the lesson plans in both groups. This ensured the research framework was

strictly followed. It eliminated extraneous variables which could have occurred if other teachers provided instruction.

Instructional sessions on how to use Andy English Bot had been conducted before the students in the experimental group began the intervention.

During class time, all students learnt English speaking about some familiar topics such as journeys, interests, entertainment, learning, tourism, with a focus on speaking with accuracy and fluency.

The participants in the control group learned and practiced English speaking face-to-face with their partners as usual in traditional classrooms while those in the experimental group were asked to practice speaking with Andy English Bot for around 30 minutes in class for each unit following Cambridge's (2017) suggestion that about 25% of the classroom time is ideal for communicative speaking tasks. While speaking, the students were encouraged to record their voice and take screenshots of the chat sessions to provide tracking data.

After five weeks of learning English speaking, all students took a post-speaking test. For conducting the speaking tests, the researcher provided the questions to the students and recorded the responses of the students in the experimental group while the students in the control group took the speaking test in a pair following the traditional manner in which one student asked and another one answered.

The topics for the pre-speaking and post-speaking tests were similar to the topics the students learned during class time. The study used familiar topics for the students to practice English speaking so that they would not find it too difficult to speak. This decision was supported by Ellis's (2009) research finding which revealed that with less complex tasks, the students could achieve greater oral fluency and accuracy. Finally, the participants' scores were graded by the three examiners from the pretest and posttest were collected for analysis.

The process of grading was conducted by three examiners, one English native speaker and two non-native speakers of English. The examiners were selected under the criteria for English speaking examiners suggested by the researcher in Appendix F. This study had adapted and adjusted the criteria from the IELTS to test examiners' knowledge and skill requirements. Since this research used the English speaking tests for the students at A1 and A2 levels, the criteria for the speaking examiners in this research was more relaxed than the IELTS's criteria for speaking examiners. The speaking tests were structured following the actual IELTS speaking test structure. It is also worth mentioning that the English speaking examiners were

trained by the researcher before grading the students in the pre-speaking and post-speaking tests.

The scores of the students in the pre-speaking and post-speaking tests were graded based on the speaking assessment rubric. For grading procedure, in each speaking element, the students were scored based on two criteria – fluency and accuracy. After that, the average scores for each criterium graded by the three examiners were averaged to ensure reliability for each criterium. The total score was then summed from the average scores for each criterium to ensure consistency. The scores of the examiners were often similar from one examiner to another. No reconsideration of individual scores was requested.

The test for the students in the experimental group was online and took around 5-10 minutes while the test for the students in the control group was conducted in class traditionally. The double-blind rating method was applied to avoid raters' bias when they graded the students' speaking skill in both pre-speaking and post-speaking tests. The term "double-blind rating" was implemented in this study based on the definition of "double-blind peer review" defined by Casserly (2017) in which the status of both the author and the reviewer were kept unknown. Double blind style assessment ensures no bias from the evaluators since they do not personally know who is being evaluated.

There were some concerns on whether the process of reviewing was actually blinding based on this approach or whether the reviewers could potentially guess about the subject's work and style. This concern was reduced by preventing the reviewers from referencing other materials during the tests.

The double-blind model was applied in this study in two ways. First, the pre-speaking and post-speaking test recordings from the students were kept secret until the end of the experimental period and then mixed up together before they were delivered to the raters. This ensured there was no information available about the subjects for the pretests and posttests provided. Second, the information whether the recordings came from the control group or the experimental group was not known by the examiners. This prevented experimenter bias.

After taking the post-speaking tests, the students in the experimental group further responded to the questionnaire so that the research could collect quantitative data for analysis. Within this process, quantitative data were collected from the scores of the students in the pre-speaking and post-speaking tests and from their responses to the questionnaire after the post-speaking test.

3.5.2 Qualitative Data Collection

In collecting the qualitative data, a semi-structured interview was conducted under the informant sampling framework in which the researcher selected six students in the experimental group who got the highest score and six who got the lowest scores in the post-speaking test to take the interview.

Some guided questions were asked to explore in-depth information of the students' opinions on using an AI voice chatbot to improve English speaking skill. Throughout this process, qualitative data was collected from the participants' responses in the interview.

In summary, this study collected the quantitative data from the participants' pre-speaking and post-speaking tests to answer research question 1. The quantitative data from the students' responses to the questionnaire and the qualitative data from their responses to the semi-structured interview were collected and analyzed to answer research question 2. Table 3.3 summarises the research instruments and data analysis methods used in answering two research questions.

Table 3.3. Summary of Research Instruments and Data Analysis

Research Question (RQ)	Instrument	Data Analysis
RQ1: What are the effects of an AI voice chatbot on English speaking skill of Vietnamese undergraduate EFL students?	<ul style="list-style-type: none"> ▪ Pretest and posttest speaking (quantitative data) 	<ul style="list-style-type: none"> ▪ Paired sample t-test
RQ2: What are the opinions of Vietnamese undergraduate EFL students on using an AI voice chatbot?	<ul style="list-style-type: none"> ▪ Questionnaire (quantitative data) ▪ Semi-structured interview (qualitative data) 	<ul style="list-style-type: none"> ▪ Descriptive analysis ▪ Thematic analysis

3.6 Data Analysis

The data analysis process was conducted with the quantitative data and the qualitative data. Accordingly, the quantitative data included the students' scores from the pre-speaking and post-speaking tests. The students' scores were analyzed with the paired sample t-test provided by SPSS to answer for research question 1

What are the effects of an AI voice chatbot on English speaking skill of Vietnamese undergraduate EFL students?". After that, the data collected from the students' responses to the questionnaire were quantitatively analysed with frequencies and descriptive statistic methods available in SPSS. Finally, the students' responses to the semi-structured interview were qualitatively analyzed with a thematic method to provide answers for research question 2 "What are the opinions of Vietnamese undergraduate EFL students on using an AI voice chatbot?".

3.6.1 Quantitative Data Analysis

In analyzing the quantitative data, the paired sample t-tests following Podesva & Sharma (2013) s guideline was utilized to analyze the students pretest and posttest scores. The paired sample t-tests were used in this study because there was only one independent variable which was the AI voice chatbot and two samples to be compared which were the control sample and the intervention sample. With the paired sample t-test, the p values derived from the comparisons were assessed in finding statistically significant differences between the two samples.

The analysis was conducting by using SPSS version 20. Furthermore, to compare the control group and the experimental group from various values, five measures were implemented. The first measure was the p value from comparing the mean scores of students in both control and experimental groups in the pre-speaking test to see if the students in different groups were significantly different from each other. The second measure was the p value from comparing the mean scores of the control group in the post-speaking test with those in the pre-speaking test to see if the students got significantly improved after the experiment. The third measure was the p value from comparing the mean scores of the experimental group in the post-speaking test with those in the pre-speaking test to see if the students got significantly improved after the experiment. The fourth measure was the p value from comparing the mean scores of two groups in the post-speaking test to see if the intervention was significantly superior from the control group after the experimental period. The last measure was the p value from comparing the mean scores growth of the control group from pre- to post-speaking test and those of the experimental group from the same process to see if the experimental period had significant effects on both control and experimental groups.

3.6.2 Qualitative Data Analysis

In analyzing the qualitative data, the students responses to the semi-structured interview were analyzed deductively starting from interpreting the general opinions of the students and then extracting their specific opinions on each theme. The thematic analysis method discussed by Heigham and Croker (2009) was utilized. Within the thematic analysis method, five themes were suggested by the study itself in consideration of the relevance of each theme to the research objectives. Accordingly, the themes included (1) the students practice process, (2) speaking accuracy improvement (grammar, pronunciation, vocabulary), (3) speaking fluency improvement, (4) positive aspects of an AI voice chatbot, and (5) negative aspects of an AI voice chatbot.

Within the analysis process, the first theme contained the information of the students' process of practicing English speaking with Andy English Bot, the second and third themes referred to the students' points of view on their improvement in English speaking fluency and accuracy (grammar, pronunciation, vocabulary) after the intervention, and the final theme focused on the positive and negative aspects of the AI voice chatbot. The thematic analysis method was reported to be appropriate for grouping data within specific themes and this method would allow for data interpretation.

3.7 Trustworthiness and Validity

3.7.1 Trustworthiness

This current study employed the indexes of items-objective congruence (IOC) model introduced by Turner and Carlson (2003) to evaluate item validity of the instruments, including the speaking test, the questionnaire, the semi-structured interview, the lesson plans, and the speaking assessment rubric. The index of item-objective congruence (IOC) was defined by Turner and Carlson (2003) as *a procedure used in test development for evaluating content validity at the item development stage*. Within this definition, the evaluation of an independent expert panel to rate the suitability of the items used for conducting measurement of one or more objectives can be checked to obtain evidence of item validity. The IOC was conducted by experts evaluating each item by ranking the item with a rating point with the values of -1, 0, or 1. The rating of 1 meant "clearly measuring" while -1 meant "clearly not measuring", and 0 meant "unclear in measurement."

Along with the rating point system, the independent status of experts was maintained when conducting the IOC. Turner and Carlson (2003) suggested that the experts should not be informed of the target constructs which the individual items were intended to measure.

In this process, a set of items was created to measure the mentioned constructs. After the creation, each item was listed in the rows of a table and the possible measuring objectives were put in the columns. Next, the created list was distributed to the experts for rating each item by utilizing the rating point system as described above. Finally, the ratings were combined to give IOC measurement results for each item on each objective. The combined score of each item ranged from -1 to 1. The score of 1 meant all experts agreed on the item validity and concluded that the item was not measuring any irrelevant items, while -1 meant the item validity was not valid as it was not measuring the hypothesized objective.

The IOC was also utilised in Each & Suppasetseree s (2021) s research for evaluating item trustworthiness and validity of their instruments. They had the forms evaluated by three experts in the field of English language teaching. Those experts were reported to have at least five-years of experience in teaching at the tertiary level. The results from the forms were confirmed to be reliable. This study also used the IOC model with the same number of evaluating experts who had equivalent qualification in the field of English language teaching to evaluate the instrument for trustworthiness.

3.7.2 Validity

For the validity of the instruments, a team of three experts who had more than five years of experience in the fields of technology-enhanced language learning (TELL) and English speaking took part in the evaluation process. The researcher provided the experts with the IOC evaluation forms for four instruments: (1) the questionnaire (Appendix G1), (2) the semi-structured interview (Appendix G2), (3) the speaking test pack (Appendix G3), and (4) the speaking assessment rubric (Appendix G4) so that they could provide a score of -1 (clearly not measuring), 0 (unclear), or 1 (clearly measuring) for each item based on their opinions about their validity.

The experts were trained on how to perform inter-rating assessments based on the index of item-objective congruence value proposed by Turner & Carlson (2003). Specifically, they were trained to do the ratings for each item which was then calculated into an average, which might result in -1, -0.67, -0.33, 0, 0.33, 0.67, or 1. In the case in which the average of the items was less than 0, the item would be eliminated. If the average score was between 0 and 0.33, that item was revised to make it clearer in meaning. The average score between 0.67 and 1 provided the best result because all experts agreed on the effectiveness of the item.

For the lesson plans validity, the researcher sent a survey on training needs to the three experts for evaluation (Appendix H). The survey was designed based on Laksana s Training Need Survey for Teachers in which the five-point rating scale was provided to the experts so that they could give their opinions about each item response by rating 1 (strongly disagree), 2 (disagree), 3 (neutral), 4 (agree), and 5 (strongly agree), regarding the lesson plan meeting teaching objectives.

3.8 Pilot Study

A pilot study was conducted so that this research could evaluate the reliability of the instruments used in our research and predict possible issues of concern for

the experimental design. The pilot study generated data to assess general process, specific process, and problems discovered.

3.8.1 General Process

For the general process, the pilot study was conducted during four weeks with 30 voluntary Vietnamese undergraduate EFL students. Two units of the eight were taught during the pilot study via the website.

The students took a pre-speaking test and were trained on how to have conversations with an AI voice chatbot before they learned the English speaking lessons. During class time, the students practiced speaking with Andy English Bot and Google Dialogflow Bots on the website.

After finishing two units, the students took the post-speaking test to assess improvement in English speaking. This allowed the researcher to investigate the reliability of the test pack and the lesson plans.

After the post-speaking test, the students further answered the questionnaire and took part in an interview so that the reliability of the questionnaire and the semi-structured interview could be evaluated.

Based on the results of the pilot study, the researcher made some adjustments in the instruments to improve study design. The questionnaire which was evaluated by the experts as valid, was found to be easy for students to complete in all three sections. The semi-structured interview was assessed by the students as reasonable, and they were able to give direct answers and elaborate appropriately when asked.

3.8.2 Specific Process

Regarding the specific process, the pilot study was conducted within three various stages. At first, all instruments were submitted to three experts to review and evaluate before beginning the pilot study. Three experts selected had at least five years of teaching English speaking and held at least an MA degree in English studies and English education. The detailed results of the experts' evaluation were reported in Appendix I.

For the second stage, there were 29 students who studied English 2 in a university participating in the experiment during the pilot study time. However, only 24 out of 29 participated in the pilot study until the end. The other six students dropped out due to their overall course load. Those students participated in a pre-speaking test before they attended the four-week pilot. After that, they took a post-speaking test, answered the questionnaire, and participated in an interview with the researcher.

For the final stage, the pilot study invited three examiners to grade the participants' English speaking skills to evaluate the data process. One examiner was a native speaker of English who came from America and other two examiners were non-native lecturers of English who were teaching English at a university in Vietnam. All examiners had at least five years of teaching experience and held at least an MA degree. The researchers created a Google Drive folder, uploaded all recorded audio files of the students from the pre-speaking and post-speaking tests including the grading sheets and the speaking assessment rubric in to the folder. A link was shared with the three examiners for their assessment. All audio files were put into specific subfolders labelled "PARxx" in which "xx" stood for the number of the subject. Each subfolder contained two audio files recorded from the pre-speaking and post-speaking tests of the subjects. The examiners noted scores directly to the grading sheet without the need for downloading. After the grading process, the scores given by the three examiners were averaged. The average scores were selected as the final scores to be analyzed. The detailed results of each stage throughout the pilot study are located in Appendix J, containing three main parts: the research instrument evaluation, the pilot study, and the pilot results.

3.8.3 Pilot Results

The results of the pilot study are discussed below, including quantitative data and qualitative.

3.8.3.1 Quantitative Data Analysis Results

For the results of quantitative data analysis, the students' scores from the pre-speaking and post-speaking tests were collected and analyzed using the paired sample t-test in SPSS to investigate if there were any significant differences between the Vietnamese undergraduate EFL students who learnt English speaking skill in the traditional classroom and those who used an AI voice chatbot to learn English speaking skill. The results of the speaking tests showed that there were significant differences between the pre-speaking and post-speaking tests of the students in the experimental group ($p = 0.00$).

The questionnaire responses in Q7 and Q8 revealed that more than 70% of the students believed that speaking with the AI voice chatbots helped them to improve their fluency in speaking because they could speak with less hesitation and pauses and use hedging words more appropriately.

The responses in Q9, Q10, and Q11 revealed that around 60% of the students believed in their improvement in speaking accuracy after speaking with

the AI voice chatbots as they could pronounce correctly, use appropriate sentence structures, and choose appropriate words and vocabulary.

3.8.3.2 Qualitative Data Analysis Results

For the results of qualitative data analysis, the interview responses revealed that most of the students believed their improvement in grammar and pronunciation was supported by the interaction with the AI voice chatbots. Some of the students reported that when they uttered a sentence, and their sentence appeared on the screen, this was useful feedback because they could see whether the grammatical structures were correct. For example, PAR03 said: *“When I spoke with the AI voice chatbot, I found that my grammar was improved very much because when I produced a sentence, that sentence appeared on the chat screen between me and the AI voice chatbot. Therefore, I knew what grammar mistakes I was making and how to overcome.”*

The participants also reported that they could learn new grammatical structures after talking with the AI voice chatbots because it also used new grammar structures and new vocabulary. For instance, PAR16 reported: *“I found that my grammar had been improved because every time the AI voice chatbot talked to me, they used new grammar structures which encouraged me to learn more.”*

Regarding vocabulary improvement, the participants reported that their improvement in vocabulary required both practice with the AI voice chatbots and their own knowledge of vocabulary. Actually, PAR03 mentioned: *“I think that my vocabulary has not improved much because my available vocabulary is still limited.”* Even when asked if PAR09 found that his vocabulary was improved after speaking with an AI voice chatbot, he responded: *“I think no because it did not explain in Vietnamese. Therefore, I did not know the words, too.”*

The students' responses in the interview also revealed that after speaking with an AI voice chatbot, they found themselves speaking with less pauses and hesitations because they could practice speaking repeatedly until the AI voice chatbots could understand them. This repetition helped them become familiar with the ideas they wished to convey which allowed them to speak more fluently. PAR11 reported: *“When I practiced with the AI voice chatbot, if I spoke with too many hesitations, then it would not understand at all. Therefore, I had to speak slowly and repeat many times so that it could understand.”*

When the students were asked about their feeling about using the AI voice chatbots, they reported a range of responses. Some students reported that

they were interested in using the AI voice chatbots for improving English speaking skill because they believed that the AI voice chatbots provided a friendly interaction in English which was comfortable. The AI voice chatbots could also speak clearly and loudly enough to make the students feel like they were talking with a native speaker. PAR01 said: *“When I spoke with the AI voice chatbot, this made me feel friendly. The AI voice chatbot pronounced clearly and loudly enough, and it could also repeat many times.”* The students also reported that they thought the AI voice chatbots could be their tireless friends because those chatbots could talk with them tirelessly at any time in any places. PAR05 reported: *“When I practiced with the AI voice chatbot, the thing I like best is that the restless time we could take for asking and answering.”*

However, there were two things that the students did not like about using the AI voice chatbots for practicing English speaking. The first was the limitation of speaking topics, and the second was technological issues regarding the AI voice chatbots’ ability to recognize Vietnamese names. Most of the students reported that they felt quite bored because they could not talk about many topics when speaking with the AI voice chatbots. For instance, PAR05 shared: *“The thing I dislike is that because the AI voice chatbot is a robot, it can only reply to me with fixed sentences. Its responses are not plentiful. I hope that there will be more topics and a variety of responses so that people can have more choices. Sometimes, the AI voice chatbot only asked for my repetition because it failed to answer me.”* Other students reported that the AI voice chatbots could not recognize Vietnamese names even though they repeated it several times. For example, PAR03 said: *“The thing I like best is that I can improve my English and speak better. However, there are some sections which I speak again and again but the AI voice chatbot still could not get it.”*

Based on the quantitative and qualitative results, there were three findings which could relate the research questions. First, there were significant differences between Vietnamese undergraduate EFL students who studied in the traditional classroom and those who studied using an AI voice chatbot in terms of speaking fluency and accuracy, demonstrated by the p values of 0.00 in the paired sample t-test when comparing the English speaking scores of the students in the pretest with the posttest. Second, the AI voice chatbots were found to help the students improve their English speaking skill because most of them reported that after the experiment period, they could speak English with less hesitation and pauses and use interjections appropriately. Moreover, they also reported that speaking with

an AI voice chatbot helped them speak more accurately in terms of pronunciation, grammar, and vocabulary. Finally, most of the students expressed positive opinions on using the AI voice chatbots while some shared constructive opinions.

3.8.4 Problems Found

During the process of the pilot study, there were some problems found in terms of the speaking tests, the practice time requirement, and technical issues.

The first problem related to the questions for the speaking tests in Part 1 and Part 3. In Part 1, the questions in the topics of food, study, and daily activities were unfamiliar to the students because there were some words used at a higher level than the students' current level of fluency. Therefore, these questions were adjusted appropriately. In Part 3, the structures of some questions related to the topics of people, country, exercise, and festivals appeared too complicated for the students because these questions were designed with an advanced grammatical structure. These questions were downgraded to a simple structure so that these questions would be more comfortable for students at this level.

The second problem related to the time the students were assigned to practice and record audio files of their speaking with an AI voice chatbot. At first, the students were expected to record their speaking session with an AI voice chatbot after class for about ten minutes per week. They later believed this request was too burdensome because they had duties from other coursework. To solve this problem, the researcher decided to reduce practice time in English speaking with an AI voice chatbot to class time only. They could then voluntarily provide screenshots of speaking sessions with an AI voice chatbot if they have time after class, without it being a course requirement.

The final problem related to the Google Dialogflow bots utilized in the pilot. At the beginning, we decided to use both Andy English Bot (Andy) and Google Dialogflow bots (namely Peter and Elsa) to speak with the students. Therefore, the students had the choice to practice and speak English with either Andy, Peter, or Elsa. However, some students did not have laptop accessibility, which was required for two of the programs. As a result, we decided to utilize Andy as the primary AI voice chatbot for the students to practice English speaking because all of the students possessed a smartphone compatible with Andy. Moreover, Peter and Elsa were reported to have limited topics for students to practice English when compared to Andy. This was a second reason the researcher decided to use Andy in the study. The final problem was that both Andy English Bot and Google Dialogflow bots failed to recognize the students' Vietnamese names. To solve this problem, the students

should consider using an English nickname when speaking with Andy rather than their Vietnamese name.

3.9 Summary

In summary, this section has discussed the research methodology and design. The next section will delve into results of the study and discussion of the findings.



CHAPTER 4

RESULTS AND DISCUSSION

This section will review the results of the study, and discuss the meaning thereof to address the two research questions: (1) What are the effects of an AI voice chatbot on English speaking skill of Vietnamese undergraduate EFL students? and (2) What are the opinions of Vietnamese undergraduate EFL students on using an AI voice chatbot?

4.1 Results

The experiment was conducted within an eight weeks timeframe with thirty participants in the control group and thirty in the experimental group. The pre-speaking and post-speaking tests and the questionnaire items were used as the quantitative data while the semi-structured interview was employed as the qualitative data. Overall, the students scores from the pre-speaking and post-speaking tests in the experimental group showed significant difference with the p value of 0.00 ($p < 0.05$) and all students agreed that using the AI voice chatbot could improve their English speaking skill, demonstrated by the total mean score of 3.90 ($M = 3.90$) after interpretation. The focus of the results of the study will be described in terms of the following two questions:

(1) What are the effects of an AI voice chatbot on English speaking skill of Vietnamese undergraduate EFL students?

(2) What are the opinions of Vietnamese undergraduate EFL students on using an AI voice chatbot?

Both quantitative and qualitative data from the pretest and posttest, the questionnaire, and the interview will be reported. The pretest and posttest results will be compared to discover the effects of an AI voice chatbot on English speaking skills of Vietnamese undergraduate EFL students. The questionnaire and the semi-structured interview responses will be used to explore the students opinions on using the AI voice chatbot.

4.1.1 The Effects of an AI Voice Chatbot on English Speaking Skill of Vietnamese Undergraduate EFL Students

In investigating the effects of an AI voice chatbot on English speaking skill of Vietnamese undergraduate EFL students, the quantitative results were analyzed.

Accordingly, the scores of the students in the pretest and posttest were analyzed by using paired sample t-tests.

The detailed results are found in Table 4.1 and revealed the students improvement in English speaking skill generally and their speaking accuracy and fluency particularly. Some significant differences between the student s English speaking improvement in the control group and in the experimental group were also found in the table.

Table 4.1. Results of the Paired sample T-tests

		Paired Samples Test							t	df	Sig. (2-tailed)
		Paired Differences					Lower	Upper			
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference						
Pair 1	PretestCG - PretestEG	-1.0667	3.0954	.5651	-2.2225	.0892	-1.887	29	.069 <i>p>0.05</i>		
Pair 2	PosttestCG - PosttestEG	-1.5667	3.2129	.5866	-2.7664	-.3669	-2.671	29	.012 <i>p<0.05</i>		
Pair 3	PretestEG - PosttestEG	-.8167	.4676	.0854	-.9913	-.6421	-9.565	29	.000 <i>p<0.05</i>		
Pair 4	PretestCG - PosttestCG	-.3167	.5676	.1036	-.5286	-.1047	-3.056	29	.005 <i>p<0.05</i>		
Pair 5	GrowthCG - GrowthEG	-.5000	.6565	.1199	-.7452	-.2548	-4.171	29	.000 <i>p<0.05</i>		
Pair 6	AccuracyPretestEG - AccuracyPosttestEG	-.2833	.2780	.0508	-.3872	-.1795	-5.582	29	.000 <i>p<0.05</i>		
Pair 7	FluencyPretestEG - FluencyPosttestEG	-.4333	.3198	.0584	-.5528	-.3139	-7.421	29	.000 <i>p<0.05</i>		

In pair 1, comparison of the students scores in the pretest for both groups showed no significant difference between the control group and the experimental group, demonstrated by a $p=0.69$ ($p>0.05$). This measure demonstrated that students were equivalent in English speaking before the intervention.

In pair 2, the students post-test scores were compared which revealed that the scores of students in the experimental group significantly different from the control group, demonstrated by the p value of 0.012 ($p<0.05$).

In pair 3, students in the experimental group showed significant improvement in their English speaking scores in the posttest compared with the pretest with the p value of 0.000 ($p<0.05$).

In Pair 4, those in the control group also improved their English speaking scores in the posttest with $p=0.005$ ($p<0.05$). These p values suggest that after eight

weeks of learning English speaking skills, the students in both groups could improve their English speaking skill.

However, as demonstrated by pair 5, the students who used an AI voice chatbot to practice English speaking in an AI-aided classroom were had significantly more improvement than the control group, as demonstrated by the p value of 0.000 ($p < 0.05$).

In pair 6, the scores of speaking accuracy compared pre-test and post-test of students in the experimental group. The results demonstrated that speaking accuracy was significantly improved after use of an AI voice chatbot to learn English speaking, suggested by the p value of 0.000 ($p < 0.05$).

In pair 7, the scores of the students in the experimental group were compared pre-test and post-test for fluency. The results suggested that students fluency was significantly improved after they used an AI voice chatbot to learn English speaking, indicated by the p value of 0.000 ($p < 0.05$).

4.1.2 The Students' Opinions on Using an AI Voice Chatbot

The questionnaire responses were analyzed using descriptive and frequencies statistical method. The results from the semi-structured interview were qualitatively analyzed with the thematic analysis method.

4.1.2.1 Questionnaire Responses

The results of the pre-test and post-test responses were analyzed using the frequencies and descriptive statistical method to provide quantitative results. The questionnaire was used to collect two sorts of information. The first was the students' demographic information and the second was the students' opinions on using an AI voice chatbot to learn English speaking skill based on the questionnaire.

A. The Students' Demographic Information

The demographic information of the students included their gender, age, major, year of study, placement test score, and technological experience. In line with the demographic information, the results of the students' gender, major, and technological experience will be illustrated in percentages while the results of age, year of study, and English placement test score will be presented using mean scores.

Table 4.2. The Students' Personal Information

Students' Personal Information						
Gender (%)		Age (Mean)	Non-English Major (%)	Year of study (Mean)	English placement test score (Mean)	Technological experience (%)
Male	Female					
13 (43.3%)	17 (56.7%)	20.67	100%	2.73	53.8	100%

In Table 4.2, there were 30 students participating in the study. Among the students, there were 13 males and 17 females which were 43.3% and 56.7% of the sample, respectively. The students' average age was around 21 and most were in the 2nd or the 3rd year of study. Their average English placement test score in the TOEIC Bridge test was of 53.8, which was at the A2 level. All students came from non-English majors and had technological experience prior to the intervention.

B. The Students' Opinions on Using an AI Voice Chatbot

The five items in Part 2 of the questionnaire were analyzed by the use of a descriptive statistic method to find out the students' opinions on using an AI voice chatbot to learn English speaking skills. The analyzed results were interpreted based on Table 3.2 (The Interpretation of Points in the Questionnaire Part 2) explained in Section 3.4.1.2 previously. The interpreted results are provided in Table 4.3.

Table 4.3. Students' Opinions on Using the AI Voice Chatbot

	Descriptive Statistics			Interpretation
	N	Mean	Std. Deviation	
Q7 - After practicing with the AI voice chatbot, I think I could speak English without making too many pauses and hesitations.	30	3.73	0.74	Agree
Q8 - After practicing with the AI voice chatbot, I think I could speak English with appropriate hedging words such as uhm... ah... oh....	30	4.07	0.74	Agree
Q9 - After practicing with the AI voice chatbot, I think I could produce correct pronunciation when I spoke English.	30	3.87	0.90	Agree

Table 4.3. Students' Opinions on Using the AI Voice Chatbot (Cont.)

Descriptive Statistics				Interpretation
	N	Mean	Std. Deviation	
Q10 - After practicing with the AI voice chatbot, I think I could use appropriate sentence structures when I spoke English.	30	3.90	0.71	Agree
Q11 - After practicing with the AI voice chatbot, I think I could use appropriate words and vocabulary when I spoke English.	30	3.93	0.78	Agree
Valid N (listwise)	30			
	Total:	3.90	0.77	Agree

In Table 4.3, the results showed that all students agreed with all five items in Part 2 of the questionnaire with the total mean score of $M=3.90$ and $SD=0.77$. The results indicated that the students agreed that they had improved English speaking skill in terms of fluency and accuracy after practicing English speaking with the AI voice chatbot. Among the five items, the highest mean score fell into item 8 - “*After practicing with the AI voice chatbot, I think I could speak English with appropriate hedging words such as uhm... ah... oh....*” which received a score of 4.07 out of 5.00. Item 7 “*After practicing with the AI voice chatbot, I think I could speak English without making too many pauses and hesitations.*” received the lowest mean score of 3.73 while the items 9, 10, and 11 received a medium mean score of 3.87, 3.90, and 3.93 respectively. However, the score of 3.73 did not affect the interpretation. All mean scores obtained 3.50, or greater, which was interpreted as “agree” based on the definitions provided in Table 3.1 mentioned in Section 3.4.1.2 in Chapter 3 of this document. The results suggested that all students agreed that using the AI voice chatbot could help them improve English speaking skill in terms of fluency and accuracy.

4.1.2.2 Semi-Structured Interview Responses

The semi-structured interviews were employed to provide greater insight of the students' opinions about using an AI voice chatbot to learn English speaking skills. There were 12 students selected based on their post-test scores, six of whom were in highest scoring group and six of whom were in the lowest scoring group. This selection relied on the premise that these students might have significantly different opinions on the use of an AI voice chatbot to learn English speaking skills due to their post-test results. The responses of the students in the

interview were collected and interpreted by using the thematic method. The themes were organized into (A) the students practice process, (B) the improvement in speaking fluency, (C) the improvement in speaking accuracy (grammar, pronunciation, vocabulary), (D) the positive aspects of the AI voice chatbot, and (E) the negative aspects of the AI voice chatbot.

A. The Students' Practice Time with AI Voice Chatbot

All 15 participants were asked how often they practiced English speaking with the AI voice chatbot every week. Their answers revealed that most spent time speaking with the AI voice chatbot during free time, but the length of time between groups was different. Some spent 10 to 15 minutes to speak every week while others spent 30 to 60 minutes or more to practice and record their speaking sessions.

PAR05: *“Normally I spoke around 3-4 times a week when I was free. It helped me to kill time and learn more. I usually spoke for around 10-15 minutes each time.”*

PAR06: *“I practiced English speaking with the AI voice chatbot everyday for around 30-60 minutes depending on each day availability.”*

B. The Improvement in Speaking Accuracy

When the participants were asked if they thought their grammar had improved after speaking with the AI voice chatbot, most of them found that they could improve grammar more effectively after speaking with the AI voice chatbot. This was in part because they could look at the chat screen to evaluate their use of grammar within an uttered sentence. Moreover, the AI voice chatbot used correct grammar structures which could be used as a model for students to learn from.

In terms of accurate pronunciation in speaking, the students reported that most found that they could speak with better pronunciation because they could listen to the native voice of the AI voice chatbot to practice pronunciation multiple times. The visual display on the chat screen also provided feedback about pronunciation.

For accuracy in using vocabulary when speaking, the students' answers revealed improvement because the AI voice chatbot often suggested new vocabulary with an explanation while talking with the students. As a result, they were exposed to additional vocabulary.

The following responses give examples of the students' opinions regarding improving speaking accuracy in terms of grammar, pronunciation, and vocabulary.

PAR19: *"I found that I could speak with more correct grammar and learn more structures from the AI voice chatbot because it used grammatical-corrected sentences when chatting with me. In addition, I could look at the chat contents displayed on the chat screen to find out how correctly I used grammar so that I could correct it by myself."*

PAR04: *"When I pronounced wrongly, the chat screen displayed the wrong word. When I pronounced correctly, the chat screen displayed the right word. Therefore, I could look at that to self-evaluate my pronunciation. One more, the AI voice chatbot spoke with native voice, which supported me in learning pronunciation."*

PAR03: *"The AI voice chatbot suggested more vocabulary related to our conversation whenever speaking, which inspired me to learn the words to be able to chat with it."*

C. The Improvement in Speaking Fluency

Most of the participants said that they could speak with less hesitations compared practicing speaking with the AI voice chatbot. They explained that the first time, they spoke with hesitations and pauses. This prevented the AI voice chatbot from recognizing what they said. As a result, they had to practice speaking again and again until they could speak more fluently so that the AI voice chatbot properly recognize their speech. Students also reported that they felt free to speak with the AI voice chatbot and liked having unlimited time to speak repeatedly to help find speaking ideas. This reduced embarrassment when speaking with a real person. The reduced stress reduced their hesitations and pauses.

PAR03: *"I can speak with less pauses compared with I did before. I could also speak more smoothly because I had time to find ideas. The AI voice chatbot could wait for me for a long time without making me embarrassed. I could also speak again and again until I got familiar with the speech flow gradually."*

PAR05: *“The AI voice chatbot did not limit the speaking time, so I could speak again and again many times, which helped me speak with less pauses and hesitations.”*

D. Positive Aspects of the AI Voice Chatbot

When the participants were asked what they liked most about speaking with the AI voice chatbot, they reported that they liked to speak with the AI voice chatbot for a number of reasons. First, they said that the AI voice chatbot made them feel comfortable and gave them the sense of chatting with a human, which was inspiring. Second, they liked to speak with the AI voice chatbot because it pronounced accurately, like a native speaker which helped them improve their speaking skill in terms of pronunciation, grammar, and vocabulary without the embarrassment of speaking to a human. Participants said that the AI voice chatbot could speak for long time periods, which gave them more experience to practice English speaking.

PAR03: *“I found that the AI voice chatbot spoke very friendly and it could wait for me to repeat again and again. I could also speak with it at any time and any places. It could also speak like a native speaker and give me time to find ideas. Speaking with a real foreigner usually made me embarrassed with my incorrect pronunciation and long time for finding ideas.”*

PAR06: *“Whenever we spoke, the AI voice chatbot frequently spoke, asked, and changed topics to keep the communication unstoppable.”*

E. Negative Aspects of the AI Voice Chatbot

When asked about what they disliked, the students reported several reasons. First, the AI voice chatbot could not accurately understand the names of some places and names in Vietnamese. Second, the topics were still limited for students wanted more range in the content of practice conversations. Some did not like the app’s frequent suggestions for buying premium features and its lack of in-app microphone.

PAR03: *“The AI voice chatbot was still limited in terms of speaking topics. I wanted to speak about more and more topics but it could not support. One more thing is that it could*

not recognize the name of Vietnamese places. Therefore, it usually displayed wrong names of such places.”

PAR06: *“What I dislike is that the app did not have integrated micro. I had to use the micro icon on my phone keyboard to speak. It would be great if a micro had been already integrated into the app directly.”*

4.2 Discussion

The results of the study suggested that the students in both groups improved in English speaking skills. However, those in the experimental group obtained higher scores than those in the control group. Furthermore, the students in the experimental group improved in English speaking skills in terms of fluency and accuracy after they learned English speaking with an AI voice chatbot. Finally, they also agreed that using the AI voice chatbot could help them improve English speaking skills in terms of fluency and accuracy.

4.2.1 The Effects of an AI Voice Chatbot on English Speaking Skill of Vietnamese Undergraduate EFL Students

In discussing the effects of the AI voice chatbot on the students' English speaking skill, two areas will be discussed, which include improvement in speaking accuracy and fluency, and differences in English Speaking Skills of the Students in the control group and in the experimental group.

4.2.1.1 The Students' Improvement in Speaking Accuracy and Fluency

The students' speaking accuracy and fluency were improved, as demonstrated by improvement in scores of speaking accuracy and speaking fluency in the post-test, evaluated by the examiners using the speaking assessment rubric mentioned in Section 3.4.2.3. The reasons for the students' improvement relate to the functions of the AI voice chatbot. First, the AI voice chatbot could correct students' grammar when they spoke. The AI voice chatbot could inform the students of their mistakes in using quantifiers for uncountable and countable nouns and in differentiating between different verb tenses. This finding is quite a new discovery because previous studies (Fryer and Carpenter, 2006; Kim, 2016; Kim, 2017; Kim, 2018; Ahmad et al, 2018; Colace et al., 2018; Adamopoulou and Moussiades, 2020) did not clearly mention this function of the AI voice chatbots. Second, the AI voice chatbot could suggest vocabulary to the students to increase vocabulary when they practiced speaking. Within this function, the AI voice chatbot suggested words

randomly throughout the chat session along with definition in English so that the students could learn more vocabulary while speaking. In this way, the students knowledge of vocabulary was improved after they learned new words through conversation with AI voice chatbot. This finding provides a tool to help the students overcome their difficulties in achieving English speaking skill caused by insufficient vocabulary, mentioned by Dao (2017), Quyen et al, (2018), Tuyen and Loan (2019), Thao and Nguyet (2019). Finally, the AI voice chatbot had the ability to maintain the conversation with the students as it could continuously ask questions, respond to the students statements, and change speaking topics so that the conversation flow was always available. This finding is in line with Ahmad et al. (2018) that AI voice chatbots could speak with humans in answering and asking questions for extended periods of time.

4.2.1.2 The Differences in English Speaking Skill of the Students in the Control Group and in the Experimental Group

In investigating the students speaking scores in the pre-test, the scores were homogeneous between the control group and the experimental group. This finding agreed with Tuyen and Loan s (2019) suggestion that the students needed to be assigned to classrooms where the English level was at a similar level to provide ease and comfort in speaking. Moreover, the students scores were low in both groups, which was in line with Dao (2017) who noted that speaking skills were very difficult for students, and thus, they often received low scores in speaking tests.

The scores of the students in the experimental group, significantly improved on the post-test after they had practiced English speaking with the Andy English Bot app on their smartphone in the class. This finding suggests that using an AI voice chatbot can help the students improve their English speaking skills. This improvement came from the students increased practice time in actual English speaking during and after class thorough five weeks of experiment. This practice time met with Cambridge s (2017) suggestion that 25% of class time should be spent for practicing speaking to help students improve. In addition, the AI voice chatbot possessed a native voice which made the students feel as if they were speaking with a native speaker, and it brought about a stress-free speaking environment as they could speak freely and have few time limitations, both of which encouraged the students to speak more and thereby become more confident in speaking. Normally, students were nervous speaking with native speakers because they were afraid of spending too much time to find ideas to discuss, which might cause the discomfort

in the native speaker. However, when speaking with the AI voice chatbot, those difficulties were overcome.

This finding is in line with Nghi et al. (2019), that an AI voice chatbot could provide the students with new learning experiences which could improve the students' learning performance when combined with class meetings. This was the reason why the students in the experimental group tended to demonstrate improvement in their speaking scores significantly after the intervention. Moreover, the finding is also in line with Kim's (2017) that using an AI voice chatbot could help the students improve their English speaking skills. In addition, when the students learned and practiced English speaking with the Andy English Bot app on the smartphone, they were instructed carefully by a learning website where videos, images, and sounds were integrated. The students could also record their voice for self-assessment regularly. Therefore, this study also confirms the suggestions of Thao and Nguyet (2019) and Anh and Nhu (2021) that using visual aids could help the students speak with less stress, and that recording the students' voices for later assessment could help them to be more self-aware in practicing English speaking.

Finally, the results of the improvement in speaking scores revealed that the students in the experimental group had higher scores than those in the control group in the post-test after the ten-week intervention. This finding suggests that using an AI voice chatbot or learning English speaking skills in a traditional classroom can improve the students' speaking skills. However, using an AI voice chatbot can improve the students' speaking skill more than learning English speaking skill in a traditional classroom because the AI voice chatbot could talk with the students at anytime and anywhere. This reduced barriers to practice such that students found it both easier and more motivating to speak when compared with speaking in traditional classrooms. This finding also agrees with Fulton's (2019) recommendation that using AI technologies like AI voice chatbots could provide the students with more chances to have English exposure to improve the students' proficiency in the English language and help them to learn English more successfully.

4.2.2 The Students' Opinions on Using an AI Voice Chatbot

Having investigated the students' practice process with the students' speaking scores in the post-test, the study found that those who spent their time speaking with the Andy English Bot app on their smartphone for 10 minutes to 15 minutes a week got higher speaking scores than those who spent less than 10 minutes speaking per week. For example, PAR05 who spent 15 minutes speaking with the AI voice chatbot obtained 17.7 in the post-test while PAR10 who obtained 4.7 in

the post-test spent 5 minutes speaking with the AI voice chatbot every week. This can be explained by referring Cambridge's (2017) recommendation that 25% of the classroom time spent for practice is sufficient for the students to improve English speaking skills. Since the students learned English speaking for 135 minutes (three periods) a week, which took 25% of the total time of 540 minutes (twelve periods) for learning other English skills, this amount of time was sufficient for the students to improve their English speaking skills in general. However, if the students spent more time to practice English speaking with the AI voice chatbot after class, their improvement was more significant. This is why the students who spent more time practicing English speaking after class tended to get higher scores in the post-speaking test.

For the students' opinions on improving English speaking accuracy, the study found that most of the students could speak with correct grammar, vocabulary, and pronunciation because they could self-evaluate their speaking performance based on the chat screen and audio recordings. Students reported that when they could speak and their utterances on the chat screen, it helped them to know if they had spoken with correct grammar and vocabulary. Moreover, they could also look at their utterances on the chat screen to see if they had just pronounced a word correctly because it would provide immediate feedback. In addition, the students' self-study and self-evaluation abilities were reflected in their responses to the interview questions as they reported that when they listened to the AI voice chatbot's responses, they found that the AI voice chatbot used a variety of grammar and vocabulary from which they could learn more grammar and vocabulary simply by chatting with the AI voice chatbot. The students' self-practice ability also enhanced pronunciation with practice. The AI voice chatbot could not understand what they said with poor pronunciation, which led them to improve. This finding is in line with the observations of Dao (2017), Quyen et al. (2018), Duc (2017), Tuyen and Loan (2019), and Thao and Nguyet (2019) who found that fostering the students' self-awareness in learning enhanced their abilities of self-learning, self-practicing, self-evaluating, and self-creating.

In investigating the students' opinions about their improvement in English speaking fluency, the study found that all students could speak better with less pauses and hesitations after they spoke with the AI voice chatbot during the 8-week intervention. This finding was demonstrated by students' reports that they could not speak too fast if they wanted the AI voice chatbot to understand them because the faster they spoke the more wrong pronunciations they made, reflected by the wrong

words appeared on the chat screen. Therefore, they had to speak carefully and repeatedly until the AI voice chatbot could recognize their words and understand what they had just said. The students said that they could speak again and again but the AI voice chatbot did not feel bored. It responded to the students at any time. This triggered a positive finding that the AI voice chatbot can be an attentive listener and speaker for the students to practice English speaking. In terms of theory, this finding agrees with Fryer and Carpenter (2006) and Ahmad et al. (2018) that the AI voice chatbot could tirelessly and attentively talk with students and be at their service all day. Moreover, the students reported that they felt like they were talking with a native speaker during conversations with the AI voice chatbot as it could speak smoothly with excellent pronunciation. This highlighted the characteristic of the AI voice chatbot of its ability to perform human-like English conversation as a native speaker, which is in line with Kim et al. (2021) and Çakmak (2022) that the AI voice chatbot had the ability to provide students with authentic inputs by speaking like a native speaker.

In exploring the students' feelings of using the AI voice chatbot to learn English speaking, the study found that almost all students agreed that they could improve English speaking skill after practicing speaking with the chatbot because they felt relaxed when speaking with the AI voice chatbot as it was very friendly and patient. This finding agreed with the findings of Nghi et al. (2019), Kim et al. (2021), and Çakmak (2022) that an AI voice chatbot could bring relaxing and interesting learning environments to students to help them learn speaking more successfully and motivate them to learn and share their knowledge. However, there were some negative aspects of the AI voice chatbot reported by students such as the AI voice chatbot's failure in recognizing names of some places in Vietnam, its frequent popup advertisement, and its limited topic choice, although these drawbacks did not affect the students' English speaking skill.

4.3 Summary

In summary, this section has reviewed the results and of experiment and found crucial answers for the three research questions. The pedagogical implications provide useful clues for teachers and students to apply the AI voice chatbot to teach and learn English speaking skill successfully.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

This research provided findings that the AI voice chatbot is effective in improving English speaking skills of Vietnamese undergraduate EFL students. There was a significant difference in the English speaking skill between Vietnamese undergraduate EFL students who study in the traditional classroom and those who study by using an AI voice chatbot. The positive opinions of the students toward using the AI voice chatbot for learning English speaking were discussed in Chapter 4. This chapter will summarize the major findings, and suggest implications of the study, in addition to considering the study limitations and provide recommendations for future study.

5.1 Summary of the Major Findings

The first finding of the study refers to the significant differences between the students' English speaking skill in traditional classroom in comparison with the classroom learning English speaking skill with an AI voice chatbot. The study found that the students in both the control group and in the experimental group improved English speaking skills at the post-test. However, the students who were in the experimental group had a tendency to speak English better than those in the control group, as demonstrated by their English speaking post-test scores.

Second, when investigating the effects of an AI voice chatbot on English speaking skills of the students, this study found that the AI voice chatbot had useful functions for students to learn English speaking. It could help the students correct grammar mistakes, suggest vocabulary to students when speaking, and maintain the conversation flow with the students by asking follow-up questions or changing to a new topic to speak. Consequently, the longer time the student practiced English speaking with the AI voice chatbot, the higher speaking score they obtained in speaking fluency and accuracy, as demonstrated scores in their post-test. Their improvement was related to self-learning regarding grammar mistakes, updating their vocabulary knowledge, and speaking with the AI voice chatbot continuously.

The students' opinions on using the AI voice chatbot to learn English speaking skill were generally positive because they found that they could improve their grammar, pronunciation, vocabulary, and fluency after speaking with the AI voice chatbot. Furthermore, the AI voice chatbot was also liked by the students because it

created a friendly and stress-free atmosphere while speaking with them like a native speaker.

5.2 Implications of the Study

The study's findings suggest that the AI voice chatbot has improved students' speaking skills in terms of accuracy and fluency. It has also given the students more English speaking opportunities. More importantly, the results and the findings of this study suggest educational implications for students, pedagogical implications, and research implications.

5.2.1 Educational Implications for Students

For learning purposes, it is recommended that the students learn English speaking with others who are at an equivalent English level so that they can practice speaking easily with each other. This can also improve their confidence in speaking because it will lower the risk of feeling inferior to students.

When learning English speaking skill, students should use the AI voice chatbot to practice English speaking because the AI voice chatbot can help them speak and to self-assess their grammar, vocabulary, pronunciation, and pauses and hesitations. This self-awareness enables and encourages them to practice, evaluate, and learn more.

In addition, the students should also make use of the user interface or the chat screen to learn English speaking because the AI voice chatbot not only speaks with the students with a native voice, it can also display the chat contents as text messages which visually supports the students in both self-learning and self-practicing.

As the study found that the more time students spent practicing English speaking, the higher scores they obtained in the post-test, this suggests that when learning English speaking, students need to spend as much time as possible in practicing English speaking.

5.2.2 Pedagogical Implications

For pedagogical purposes, the first finding that teachers should be aware of is that they should assure equivalence of students' skills when they join an English class so that teachers can support students, and students can support other students.

As this study found that the students' English speaking skill improved more when they practiced English speaking with an AI voice chatbot, teachers are

encouraged to consider allowing students to use the AI voice chatbots in English speaking classrooms to practice English in addition to practicing with partners.

Sufficient time for practice should also be considered by teachers so that the practice time can be reasonably balanced. This means that time for practicing English speaking with the AI voice chatbot should be balanced with time for practicing with a partner. Teachers should not use only one method of practicing in such classrooms, because this would limit students' chance to speak English.

Although the AI voice chatbot was found to improve the students' accuracy and fluency in English speaking, it is important that teachers understand its use and its basic technology before introducing it into a classroom setting. Students need to be instructed and trained carefully before using the AI voice chatbot so that any technological problems can be avoided.

Regarding the AI voice chatbot's features, this study recommends that teachers take advantages of the AI voice chatbot's features such as text and voice support, unlimited conversation time, native voice, and attentive listener to give students more chances to speak and practice grammar, vocabulary, pronunciation. This allows them to improve their speaking accuracy first, after which they can speak with fluency. For example, when teachers teach a grammar point to students, teachers can ask students to practice a speaking exercise with the AI voice chatbot. After students finish speaking, they should then look at the chat screen to confirm if they have used correct grammar or not. For vocabulary and pronunciation, the process is similar.

5.2.3 Research Implications

For future research opportunities, this study first notes that AI voice chatbot is still a new field of TELL research, and therefore, the number of publications on AI voice chatbots remain limited. There will need to be many more future studies to provide additional data in this growing field.

Future research may delve into different aspects such as investigating how to measure students' speaking fluency and accuracy when practicing English speaking with an AI voice chatbot, exploring how an AI voice chatbot helps students to practice language skills, exploring students' attitudes toward using an AI voice chatbot to learn English speaking skills or other skills, or investigating how effectively an AI voice chatbot can improve students' English proficiency in speaking, listening, writing, and reading. However, there are three things future researcher should consider.

The first is about technology. As AI voice chatbot is new in the educational area, technological problems will challenge both researchers and participants. Therefore, future researchers actively review and test various AI voice chatbots before selecting the one for their research. Within the scope of this study, Andy English Speaking Bot is a good suggestion. However, there are still many more AI voice chatbots which may work better such as ChatGPT, Nova AI voice chatbots, or GPT Chatvoice – AI Open Wisdom, although these chatbots require users to purchase access for use in research.

The second consideration future researchers should be related to the timeframe of the study. Ideally, a longer time period should be used for future experiments because 8 weeks may be inadequate to collect data on a newer technology like the AI voice chatbot.

Finally, future researchers should investigate the effects of the AI voice chatbot on other English language skills of students, especially in receptive skills like reading or listening. As Brown (2000) suggested, listening and speaking are two intertwined skills learners acquire when learning a language. It is strongly suggested that future researchers consider investigating the effects of the AI voice chatbot on English listening skills of students. This will bring about new vistas of exploration to this field of research.

5.3 Limitations and Recommendations for Future Studies

During the experimental period, there were some limitations which this study could not avoid. First, the experiment was conducted in eight weeks, which may have been too short a time for observing long-term development of the students. Further research should consider a longer timeframe to measure results. Second, this study had a small sample size, with 30 subjects in each arm. This may have limited some of the statistical analysis. Future researchers should consider studying a larger sample size to improve research data. Finally, technological issues were a challenge during the study because using an AI voice chatbot for learning English speaking skill is a very new field in Vietnam. Most of the English classrooms in Vietnam have been conducted traditionally. Therefore, both teachers and learners may lack familiarity with this new technology. As a result, many students had difficulties when using the AI voice chatbot for learning English speaking such as installing the app, activating the keyboard microphone to have a voice chat with the AI voice chatbot, and failing to activate the English voice so that the AI voice chatbot could recognize English words.

Fortunately, those particular problems were found during the orientation period. Hence, the researcher could help the students to solve each problem before they started using the AI voice chatbot for learning English speaking during the experiment.

Another technological issue was found but this issue could not be solved. It was about the inability of the AI voice chatbot to recognize Vietnamese names. Most of the time the AI voice chatbot could recognize Vietnamese names of locations in Vietnam such as Hue, Cham Temple, Hau Giang, Cai Rang. Regarding this issue, this author would like to suggest that future researchers pay careful attention to the technical aspects of the AI voice chatbot when conducting a study on using AI voice chatbots for learning other English skills.

5.4 Summary

This study investigated the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL students. This study compared the significant differences in the English speaking skills between Vietnamese undergraduate EFL students who studied in the traditional classroom and those who studied by using an AI voice chatbot. Students' opinions on using AI voice chatbots to improve English speaking skills have been also explored.

The results revealed that students improved their English speaking after learning English speaking and practicing English speaking with the AI voice chatbots in comparison to English speaking skills in a traditional classroom.

From the questionnaire and semi-structured interviews, students also reported positive opinions on using the AI voice chatbots in English speaking to improve speaking fluency and accuracy in terms of grammar, vocabulary, pronunciation, pauses and hesitations.

For pedagogical applications, the AI voice chatbots are considered a good learning tool for the Vietnamese undergraduate EFL students as it is one of the technology-enhanced language learning methods. It also has support of the government and institutions. If teachers are ready and familiar with an AI voice chatbot, they can enhance interest and fun in classroom environments and engage students in learning English speaking skills. For students, the use of this technology provides a comfortable learning environment to practice English speaking and enhance English speaking skills.

REFERENCES

- Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, 2. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Ahmad, N. A., Che, M. H., & Zainal, A. (2018). Review of chatbots design techniques. *International Journal of Computer Applications*, 18(8). <https://doi.org/10.5120/ijca2018917606>
- Anh, K. H. K. (2010). Use of Vietnamese in English language teaching in Vietnam: Attitudes of Vietnamese university teachers. *English Language Teaching*, 3(2). <https://doi.org/10.5539/elt.v3n2p119>
- Anh, V. T. Q., & Nhu, P. T. T. (2021). Using technology to teach speaking skill online during time of Covid-19 outbreak in Vietnam- facts and recommendation. *PalArch's Journal of Archaeology of Egypt / Egyptology (PJAEE)*, 18(7).
- Baek, J., & Lee, C. H. (2018). University students' perceptions and engagement in mobile assisted blended learning in English speaking classes. *Multimedia-Assisted Language Learning*, 21(4). <https://doi.org/10.15702/mall.2018.21.4.11>
- Banditvilai, C. (2016). Enhancing students' language skills through blended learning. *The Electronic Journal of e-Learning*, 14(3), 220-229. <https://files.eric.ed.gov/fulltext/EJ1107134.pdf>
- Bordens, K. S., & Abbott, B. B. (2011). Research design and methods: A process approach. *New York: McGraw-Hill*.
- Brown, H. D. (2000). Teaching by principles: An interactive approach to language pedagogy. *Longman*.
- Brown, H. D., & Lee, H. (2015). Teaching by principles an interactive approach to language pedagogy (4th ed.). *Pearson Education, Inc*.
- Çakmak, F. (2022). Chatbot-human interaction and its effects on EFL students' L2 speaking performance and anxiety. *Novitas-ROYAL (Research on Youth and Language)*, 16(2), 113-131. https://files.eric.ed.gov/fulltext/EJ1365002.pdf?fbclid=IwAR3dyW7lgCAJjktzcDIL_DuCo7Ot7_ldl4a_-Uj7izAO_bKKh4xKat6lGvM
- Cambridge ESOL (2011). Using the CEFR: Principles of good practice. *ESOL Examinations*. <https://www.cambridgeenglish.org/Images/126011-using-cefr-principles-of-good-practice.pdf>
- Cambridge University Press (2017). How much time should we give to speaking practice? *Cambridge: Cambridge University Press*.

- https://www.cambridge.org/us/files/8415/7488/4554/CambridgePapersinELT_TimeForSpeaking_2017_ONLINE.pdf
- Casserly, P. (2017). Benefits and drawbacks of double-blind peer review. *Ex Ordo Conference Software*. <https://www.exordo.com/blog/double-blind-peer-review/#:~:text=Double%20blind%20peer%20review%20means,reviewer%20from%20forming%20any%20bias>
- Cheng, X., Sun, J., & Zarifis, A. (2020). Artificial intelligence and deep learning in educational technology research and practice. *British Journal of Educational Technology*, 51(5), 1653-1656. <https://doi.org/10.1111/bjet.13018>
- Chomsky, N. (1959). Reviews: Verbal behavior by B.F. Skinner. *Language*, 35(1), 26-58. <https://doi.org/10.2307/411334>
- Cohen, L., Manion, L., & Morrison, K. (2007). Research methods in education (6th ed.). London & New York: Routledge. <https://doi.org/10.4324/9780203029053>
- Colace, F., Santo, M. D., Lombardi, M., Pascale, F., Pietrosanto, A., & Colace et al., S. (2018). Chatbot for e-learning: A case of study. *International Journal of Mechanical Engineering and Robotics Research*, 7(5), 528-533. <https://doi.org/10.18178/ijmerr.7.5.528-533>
- Dao, H. (2017). Identify factors that negatively influence: Non-English major students speaking skills. *Higher Education Research*, 2(2), 35-43. <https://doi.org/10.11648/j.her.20170202.12>
- Devlin, T. J., Feldhaus, C. R., & Bentrem, K. M. (2013). The evolving classroom: A study of traditional and technology-based instruction in a STEM classroom. *Journal of Technology Education*, 25(1), 34-54. <https://doi.org/10.21061/jte.v25i1.a.3>
- Duc, P. H. (2017). Towards a computer-based model to enhance Vietnamese EFL learners' autonomy in learning oral skills. *International Journal of Language and Linguistics*, 4(4).
- Duyen, L. T. N., & Huan, N. B. (2017). Factors influencing group work of students in learning English as a foreign language (EFL): A case study at a Vietnamese university. *Can Tho University Journal of Science*, 6, 9-16. <https://doi.org/10.22144/ctu.jen.2017.021>
- Each, N., & Suppasetsee, S. (2021). The effects of mobile-blended cooperative learning on EFL students' listening comprehension in Cambodian context. *Language Institute, Thammasat University*, 14, No. 2.
- El Shazly, R. (2020). Effects of artificial intelligence on English speaking anxiety and speaking skill: A case study. *Expert Systems*, 38(3). <https://doi.org/10.1111/exsy.12667>

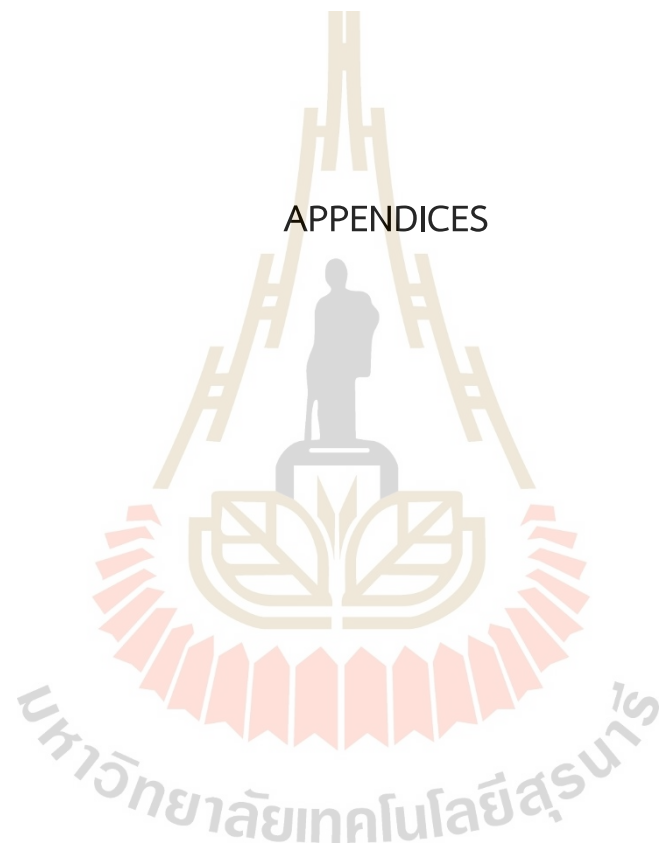
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 30(4), 474-509. <https://doi.org/10.1093/applin/amp042>
- Fryer, L., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology*, 10(3), 8-14. <http://dx.doi.org/10125/44068>
- Fulton, J. (2019). 7 examples of innovative educational technology. *Classcraft Studios Inc.* <https://www.classcraft.com/blog/examples-of-educational-technology/>
- Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2006). Educating English language learners. *Cambridge University Press.* <https://doi.org/10.1017/CBO9780511499913>
- Han, D.-E. (2020). The effects of voice-based AI chatbots on Korean EFL middle school students speaking competence and affective domains. *Asia-pacific Journal of Convergent Research Interchange*, 6(7), 71-80. <https://doi.org/10.47116/apjcri.2020.07.07>
- Haristiani, N. (2019). Artificial intelligence (AI) chatbot as language learning medium: An inquiry. *Journal of Physics: Conference Series.* <https://doi.org/10.1088/1742-6596/1387/1/012020>
- Hedge, T. (2000). Teaching and learning in the language classroom. *Oxford: Oxford University Press.*
- Heigham, J., & Croker, R. A. (2009). Qualitative research in applied linguistics: A practical introduction. *London: Palgrave Macmillan.* <https://doi.org/10.1057/9780230239517>
- Kaplan, J. (2016). Artificial intelligence what everyone needs to know *Oxford: Oxford University Press.* <https://doi.org/10.1093/wentk/9780190602383.001.0001>
- Karim, S., & Haq, N. (2014). An Assessment of IELTS Speaking Test. *International Journal of Evaluation and Research in Education (IJERE)*, 3(3), 152-157. <https://doi.org/10.11591/ijere.v3i3.6047>
- Kim, H.-S., Cha, Y., & Kim, N. Y. (2021). Effects of AI chatbots on EFL students communication skills. *Korean Journal of English Language and Linguistics*, 21, 712-734. <https://doi.org/10.15738/kjell.21..202108.712>
- Kim, N.-Y. (2016). Effects of voice chat on EFL learners speaking ability according to proficiency levels. *Multimedia-Assisted Language Learning*, 19(4), 63-88. <https://doi.org/10.15702/mall.2016.19.4.63>
- Kim, N.-Y. (2017). Effects of types of voice-based chat on EFL students negotiation of meaning according to proficiency levels. *English Teaching*, 72(1), 159-181. <https://doi.org/10.15858/engtea.72.1.201703.159>

- Kim, N.-Y. (2018). A study on chatbots for developing Korean college students' English listening and reading skills. *Journal of Digital Convergence*, 16(8), 19-26. <https://doi.org/10.14400/JDC.2018.16.8.019>
- Kim, N.-Y., Cha, Y., & Kim, H.-S. (2019). Future English learning - chatbots and artificial intelligence. *Multimedia-Assisted Language Learning*, 22(3), 32-53. <https://doi.org/10.15702/mall.2019.22.3.32>
- Koudenburg, N., Postmes, T., & Gordijn, E. H. (2014). Conversational flow and entitativity: the role of status. *Br J Soc Psychol*, 53(2), 350-366. <https://doi.org/10.1111/bjso.12027>
- Krashen, S. (2013). *Second language acquisition: theory, applications, and some conjectures*. Cambridge University Press.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. New York: Pergamon Press.
- Laksana, S. Instrumentation. Assumption University of Thailand. http://www.education.au.edu/wp-content/uploads/2016/09/instrumentation.pdf?fbclid=IwAR1M7SqDOiYlo2O4Hl1KaAfW21X3iJ7SN81231q1fg_R2G8OlafHnGRHPSc
- Li, J. (2019). An evaluation of IELTS speaking test. *OALib*, 06(12), 1-17. <https://doi.org/10.4236/oalib.1105935>
- Liddy, E. D. (2001). Natural language processing. In *Encyclopedia of Library and Information Science*. Ed. New York: Marcel Decker, Inc.
- Manurung, K. (2015). Improving the speaking skill using reading contextual internet-based instructional materials in an EFL class in Indonesia. *Procedia - Social and Behavioral Sciences*, 176, 44-51. <https://doi.org/10.1016/j.sbspro.2015.01.442>
- Marietto, M. d. G. B., Aguiar, R. V., Barbosa, G. d. O., Botelho, W. T., Pimentel, E., França, R. d. S., & Silva, V. L. d. (2013). Artificial intelligence markup language: A brief tutorial. *International Journal of Computer Science & Engineering Survey*, 4(3), 1-20. <https://doi.org/10.5121/ijcses.2013.4301>
- Najeeb, S. S. R. (2013). Learner autonomy in language learning. *Procedia - Social and Behavioral Sciences*, 70, 1238-1242. <https://doi.org/10.1016/j.sbspro.2013.01.183>
- Nazara, S. (2011). Students' perception of EFL speaking skill development. *Journal of English Teaching*, 1(1), 28-43. <https://doi.org/10.33541/jet.v1i1.50>
- Nergis, A. (2020). Second language speaking skill: The role of L2 working memory capacity. *Konin Language Studies*, 8(4), 349-368. <https://doi.org/10.30438/ksj.2020.8.4.1>

- Nga, P. T. (2022). Artificial intelligence (A.I) application in foreign language teaching and learning. *Advances in Image and Video Processing*, 10(5). <https://doi.org/10.14738/aivp.105.13061>
- Nghi, T. T., Phuc, T. H., & Thang, N. T. (2019). Applying AI chatbot for teaching a foreign language: An empirical research. *International Journal of Scientific & Technology Research*, 8(12).
- Nguyen, T., & Nguyen, T. (2016). Oral English communication strategies among Vietnamese non-majors of English at intermediate level. *American Journal of Educational Research*, 4(3), 283-287. <http://pubs.sciepub.com/education/4/3/9/>
- Nguyen, T., & Tran, M. (2015). Factors affecting students' speaking skill at Le Thanh Hien high school. *Asian Journal of Educational Research*, 3(2).
- Piaget, J. (1964). Cognitive development in children: Piaget development and learning. *Journal of Research in Science Teaching*, 2, 176-186. <https://doi.org/10.1002/tea.3660020306>
- Podesva, R. J., & Sharma, D. (2013). Research methods in linguistics. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139013734>
- Quyen, T. T. T., & Loi, N. V. (2018). Flipped model for improving students' English speaking performance. *Can Tho University Journal of Science*, 54(2), 90-97. <https://doi.org/10.22144/ctu.jen.2018.012>
- Quyen, V. P., Nga, P. T. M., & Nguyen, H. T. (2018). Challenges to speaking skills encountered by English-majored students: A story of one Vietnamese university in the Mekong Delta. *Can Tho University Journal of Science*, 54(5), 38-44. <https://doi.org/10.22144/ctu.jen.2018.022>
- Richards, J. C., & Rodgers, T. S. (2014). Approaches and methods in language teaching (3rd ed.). New York: Cambridge University Press. <https://doi.org/10.1017/9781009024532>
- Schreiber, J., & Asner-Self, K. (2011). Educational Research. USA: John Wiley & Sons, Inc.
- Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2(1). http://www.itdl.org/Journal/Jan_05/article01.htm
- Skinner, B. F. (1957). Verbal behavior. Appleton Century Crofts. <https://doi.org/10.1037/11256-000>
- Socheath, M. (2018). Simple ways to improve English speaking in teaching and learning. *Jurnal Online Universitas Muhammadiyah Purwokerto*.
- Tohoku Fukushi University. TFU foreign language assessment rubrics. *Tohoku Fukushi University*. https://www.tfu.ac.jp/students/arnp890000001rch-att/rubric_languages_english.pdf

- Thao, T. Q., & Nguyet, D. T. N. (2019). Four aspects of English speaking difficulties encountered by tertiary English-majored students. *Ho Chi Minh City Open University Journal of Science*, 9(2), 53-64. <https://doi.org/10.46223/HCMCOUJS.soci.en.9.2.261.2019>
- Tram, H. T. A. (2020). Problems of learning speaking skills encountered by English major students at Ba Ria-Vung Tau university, Vietnam. *European Journal of English Language Teaching*, 5(4). <https://doi.org/10.46827/ejel.v5i4.3144>
- Thao, T. Q., Tham, D. M., & Ngan, H. T. T. (2019). Attitudes toward the use of TELL tools in English language learning among Vietnamese tertiary English majors. *VNU Journal of Social Sciences and Humanities*, 5(5), 581-596. <https://doi.org/10.33100/jossh5.5.TranQuocThao.etal>
- Trinh, N. B., & Pham, D. T. T. (2021). Challenges in speaking classrooms among non-English majors. *Vietnam Journal of Education*, 5(2), 37-42. <https://doi.org/10.52296/vje.2021.52>
- Turner, R. C., & Carlson, L. (2003). Indexes of item-objective congruence for multidimensional items. *International Journal of Testing*, 3(2), 163-171. https://doi.org/10.1207/s15327574ijt0302_5
- Tuyen, L. V., & Loan, T. T. T. (2019). Factors affecting EFL students willingness to communicate in speaking classes at the Vietnamese tertiary level. *International Journal of English Literature and Social Sciences*, 4(2), 252-262. <https://doi.org/10.22161/ijels.4.2.10>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press. <https://doi.org/10.2307/j.ctvjf9vz4>
- Walker, A., & White, G. (2013). *Technology enhanced language learning: Connecting theory and practice*. Oxford University Press.
- Wang, Z. (2014). Developing accuracy and fluency in spoken English of Chinese EFL learners. *English Language Teaching*, 7(2). <https://doi.org/10.5539/elt.v7n2p110>

APPENDICES



APPENDIX A
THE SPEAKING TEST PACK
 FOR THE SPEAKING TESTS

1. Speaking Part 1

In this part, the test takers will answer two oral questions to a topic randomly picked by the examiner. At first, the examiner will introduce himself/herself, then he/she will explain briefly about the test. After that, the examiner will ask for the test taker's name before starting part 1 of the speaking test.

Topic	Question	Specification
Food	Do you like street food or homecooked food? Why?	The test takers are expected to be able to talk about <i>whether they like street food or homecooked food and why they like it.</i>
	Where do you usually go to buy food?	The test takers are expected to be able to talk about <i>where they usually go to buy food, and if possible, share special things about food in that place.</i>
Study	What major do you study?	The test takers are expected to be able to talk about <i>what major, subject, or program they learn or have learnt, what the major, subject, or program is about, and if possible, share their thoughts about the major, subject, or program.</i>
	Where do you study?	The test takers are expected to be able to talk about <i>where they study, which school, university, or college they study at, and if possible, share special things about their school, university, or college.</i>
Hometown	Where is your hometown?	The test takers are expected to be able to talk about <i>where they were born and grew up, where it is located, and if possible, sharing special things about their hometown.</i>
	How often do you visit your hometown?	The test takers are expected to be able to talk about <i>how frequently they visit their hometown, when they visit, and if possible, share some reasons for their visit.</i>
Daily activities	What do you usually do every day?	The test takers are expected to be able to list the activities they do every day, and if possible, explain <i>why they have that habit.</i>
	Do you like outdoor or indoor activities? Why?	The test takers are expected to be able to tell <i>whether like to do outdoor or indoor activities, and if possible, explain why they think so.</i>

Topic	Question	Specification
Family and friends	Do you like to spend time with your family or friends? Why?	The test takers are expected to be able to tell <i>whether they like to be with their family or friends</i> and if possible, explain <i>why they believe so</i> .
	What do you usually do with your friends?	The test takers are expected to be able to list the activities they do with friends, and if possible, <i>share their thoughts</i> about those activities.

2. Speaking Part 2

In this part, the test takers will select one random topic among the six and are given one minute for preparation and two minutes for talking about the selected topic.

2.1 Book

Describe a book you have recently read.

You should say:

- What kind of book it is.
- What it is about.
- Where you found it.
- What sort of people would enjoy it.
- And explain why you like it.

Specification: The test takers can speak in two minutes and respond to all given clues in the topic. They are encouraged to describe clearly the book they have just read (what it is, what it is special for, etc.). They are also expected to tell a little about the book's overview and where or from whom they got it. In addition, they should be able to discuss briefly the target audience of the book. Importantly, it is required that they can give a clear explanation for the reason why they like the book.

2.2 People

Describe a person you like best.

You should say:

- Who he/she is.
- When you met that person.
- What he/she looks like.
- How long you have known him/her.
- And explain why you like that person.

Specification: the test takers can speak in two minutes and respond to all given clues in the topic. They are encouraged to describe clearly about a person they like best (who he/she is, what he/she is special for, etc.). They are also expected to tell a short story about the occasion they met that person and how long they have known that one. In addition, they should be able to describe briefly the person's appearance and personality. Importantly, it is required that they can give a clear explanation for the reason why they like that person.

2.3 Country

Describe a country you like best.

You should say:

- What that country is.
- Where it is located.
- What it is special for.
- When you knew about that country.
- And explain why you like it.

Specification: the test takers can speak in two minutes and respond to all given clues in the topic. They are encouraged to describe clearly the country they live in (what its name, where it is located, etc.). They are also expected to talk about the specialties of the country. In addition, they should be able to mention briefly the occasion they happened to know the country. Importantly, it is required that they can give a clear explanation for why they like the country.

2.4 Exercise

Describe an exercise you are doing for fitness.

You should say:

- What it is.
- How often do you do it.
- When you first tried it.
- What kind of people it is suitable for.
- And explain why you think it is good.

Specification: the test takers can speak in two minutes and respond to all given clues in the topic. They are encouraged to describe clearly a sports game they are doing for fitness (what its name, what it is about, etc.) and how frequently they

do the game. They are also expected to tell a short story about the occasion they came to it. In addition, they should be able to describe briefly the target players of the game. Importantly, it is required that they can give a clear explanation for the reason why they like the game.

2.5 Festival

Describe a festival you have ever joined in your country.

You should say:

- What it is.
- When it is celebrated.
- What it is special for.
- How people celebrate the festival.
- And explain why you like it.

Specification: the test takers can speak in two minutes and respond to all given clues in the topic. They are encouraged to describe clearly a festival they have joined (its name, what it is about, etc.). They are expected to tell when the festival takes place and what is special. In addition, they should be able to describe briefly how people enjoy the festival. Importantly, it is required that they can give a clear explanation for the reason why they like it.

3. Speaking Part 3

In this part, the test takers will answer two oral questions asked by the examiner about the selected topic in Part 2.

Topic	Question	Specification
Book	What kind of books do you think are good for reading?	The test takers are expected to be able to talk about <i>what types of books are meaningful for reading, what is good or bad side, and if possible, sharing their thoughts</i> about the criteria for selecting a good book.
	Is reading a book more interesting than watching a movie?	The test takers are expected to be able to tell <i>whether reading a book outweighs watching a movie, why it is interesting or boring, what example can be drawn, and if possible, share some of their experiences</i> about reading a book or watching a movie.
People	What are the factors that make a person famous?	The test takers are expected to be able to talk about <i>what contributes to one's reputation, how it happens, and if possible, give an example</i> for their beliefs.
	Are famous people always happy?	The test takers are expected to be able to tell <i>whether famous people always feel happy with their life, why they are happy or</i>

Topic	Question	Specification
		<i>unhappy, what example can be drawn, and if possible, share their thoughts about being a famous person.</i>
Country	Where is the best place for living?	The test takers are expected to be able to talk about <i>what place is the best for living, what contributes to the decision, and if possible, share their imagination about an ideal place for living.</i>
	Should people study and work inside or outside their country?	The test takers are expected to be able to tell <i>whether people should study and work around their country or go abroad, why they should stay in their country or go abroad, what example can be drawn, and if possible, share their thoughts about themselves studying and working inside their country or going abroad.</i>
Exercise	Should people do exercise at home or in a gym?	The test takers are expected to be able to tell <i>whether people should do exercises at home or in a gym, why they should do there and if possible, share their thoughts about doing exercises at home or in a gym.</i>
	What do people usually do to be healthy?	The test takers are expected to be able to talk about <i>what young people often do to stay healthy, what benefits should be considered, and if possible, share their thoughts about keeping fit at a young age.</i>
Festival	Do people have enough time for festivals in this modern life?	The test takers are expected to be able to tell <i>whether people still have time to enjoy festivals nowadays, why it is so, what example can be drawn, and if possible, share their thoughts about enjoying festivals.</i>
	How can festivals bring money to your country?	The test takers are expected to be able to talk about <i>how festivals create income for their country, and if possible, give an example for the explanation.</i>

APPENDIX B
THE QUESTIONNAIRE (ENGLISH VERSION)
 REFLECTION AND SELF-EVALUATION OF PARTICIPANTS

You have just completed all the speaking lessons and practices with the AI voice chatbot of this study, from which you have gained lots of experience in English speaking performance. Therefore, this questionnaire aims to explore your reflections and self-evaluations about the speaking practices with the AI voice chatbot.

It should take about 05 minutes to complete the questions. There is no wrong or right answer, so feel free to respond to the questions by typing your answers or clicking on the right options.

All your responses to these questions will be strictly kept confidential and anonymous. When all the answers are completed, click "SUBMIT".

Thank you very much for taking your time to participate in this survey. If you have any questions about the questions, contact me by email dvttam@outlook.com or call me at 0819759388.

Part 1: General information	
Please provide your detail information below.	
Description	Detail
1. Gender:	<input type="checkbox"/> Male <input type="checkbox"/> Female
2. Age:	<input type="checkbox"/> 18 <input type="checkbox"/> 19 <input type="checkbox"/> 20 <input type="checkbox"/> 21 <input type="checkbox"/> 22 <input type="checkbox"/> 23
3. Major:	<input type="checkbox"/> Medicinal Chemistry <input type="checkbox"/> Primary Education <input type="checkbox"/> Other (please specify):
4. Year of study:	<input type="checkbox"/> 1 st year <input type="checkbox"/> 2 nd year <input type="checkbox"/> 3 rd year <input type="checkbox"/> 4 th year
5. English placement test score:	
6. What technology have you ever used for learning English speaking? <i>You can select more than one option listed in this section.</i>	<input type="checkbox"/> Podcast <input type="checkbox"/> Voice chat apps (i.e. Zoom, Zalo, LINE, etc.) <input type="checkbox"/> Chatbots (i.e. Andy English Bot, ELIZA, Siri, Google Assistant, etc.) <input type="checkbox"/> Learning websites (i.e. SpeechAce.com, IELTS Liz.com, etc.) <input type="checkbox"/> Youtube <input type="checkbox"/> Social network (i.e. Facebook, Twitter, Reddit, etc.) <input type="checkbox"/> Others (Please specify):

Part 2: Opinions on English speaking performance					
Please choose ONE response that best matches with your opinion.					
Statements	Responses				
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	(1)	(2)	(3)	(4)	(5)
7. After practicing with the AI voice chatbot, I think I could speak English without making too many pauses and hesitations.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. After practicing with the AI voice chatbot, I think I could speak English with appropriate hedging words such as uhm... ah... oh....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. After practicing with the AI voice chatbot, I think I could produce correct pronunciation when I spoke English.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. After practicing with the AI voice chatbot, I think I could use appropriate sentence structures when I spoke English.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. After practicing with the AI voice chatbot, I think I could use appropriate words and vocabulary when I spoke English.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Part 3: Additional suggestions or comments on learning English speaking skill by using an artificial intelligence voice chatbot (if any):					

BẢNG CÂU HỎI (Vietnamese Version)
NHẬN XÉT VÀ TỰ ĐÁNH GIÁ CỦA NGƯỜI THAM GIA

Bạn đã hoàn thành tất cả bài học và thực hành nói với AI trong nghiên cứu này, từ hoạt động này bạn đã có nhiều trải nghiệm trong việc thể hiện nói tiếng Anh. Do đó, bảng câu hỏi này hướng tới việc tham khảo nhận xét và tự đánh giá của bạn về việc thực hành nói tiếng Anh với AI.

Bạn sẽ mất khoảng 05 phút để hoàn thành các câu hỏi. Không có câu trả lời nào là sai hay đúng, vì vậy bạn có thể tự do trả lời các câu hỏi bằng cách đánh máy hoặc nhấn vào lựa chọn phù hợp.

Mọi câu trả lời cho các câu hỏi đều sẽ được đảm bảo tuyệt đối bảo mật và ẩn danh. Hãy nhấn nút "GỬI" khi bạn đã hoàn thành các câu trả lời.

Chân thành cảm ơn các bạn đã dành thời gian tham gia vào khảo sát này. Hãy liên hệ tôi qua địa chỉ email dvttam@outlook.com hoặc gọi số điện thoại 0819759388 nếu bạn có thêm thắc mắc nào về các câu hỏi.

Phần 1: Thông tin chung	
Vui lòng cung cấp thông tin chi tiết của bạn dưới đây.	
Mô tả	Chi tiết
1. Giới tính:	<input type="checkbox"/> Nam <input type="checkbox"/> Nữ
2. Tuổi:	<input type="checkbox"/> 18 <input type="checkbox"/> 19 <input type="checkbox"/> 20 <input type="checkbox"/> 21 <input type="checkbox"/> 22 <input type="checkbox"/> 23
3. Ngành học:	<input type="checkbox"/> Hóa Dược <input type="checkbox"/> Giáo dục Tiểu học <input type="checkbox"/> Khác (ghi rõ):
4. Sinh viên năm:	<input type="checkbox"/> Năm nhất <input type="checkbox"/> năm hai <input type="checkbox"/> năm ba <input type="checkbox"/> năm bốn
5. Điểm Anh Văn thi xếp lớp:	
6. Bạn hiện đã từng sử dụng công nghệ nào để học tiếng Anh? Bạn có thể chọn nhiều hơn một tùy chọn liệt kê trong phần này:	<input type="checkbox"/> Podcast <input type="checkbox"/> Các ứng dụng gọi thoại (Zoom, Zalo, LINE, v.v.) <input type="checkbox"/> Các Chatbots (Andy English Bot, ELIZA, Siri, Google Assistant, v.v...) <input type="checkbox"/> Các trang web học tập (SpeechAce.com, IELTS Liz.com, v.v...) <input type="checkbox"/> Youtube <input type="checkbox"/> Mạng xã hội (Facebook, Twitter, Reddit, v.v...) <input type="checkbox"/> Khác (ghi rõ):

Phần 2: Các nhận định về việc thể hiện nói tiếng Anh					
Vui lòng chọn MỘT câu trả lời đúng nhất với nhận định của bạn.					
Nhận định	Câu trả lời				
	Hoàn toàn không đồng ý	Không đồng ý	Trung lập	Đồng ý	Hoàn toàn đồng ý
	(1)	(2)	(3)	(4)	(5)
7. Sau khi thực hành với AI, tôi nghĩ tôi đã có thể nói tiếng Anh mà không vấp và ngập ngừng quá nhiều.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Sau khi thực hành với AI, tôi nghĩ tôi đã có thể nói tiếng Anh và dùng các từ phù hợp như um... ah... oh... để thể hiện sự không chắc chắn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Sau khi thực hành với AI, tôi nghĩ tôi đã có thể phát âm đúng khi nói tiếng Anh.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Sau khi thực hành với AI, tôi nghĩ tôi đã có thể dùng đúng cấu trúc câu khi nói tiếng Anh.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. Sau khi thực hành với AI, tôi nghĩ tôi đã có thể dùng đúng từ và từ vựng khi nói tiếng Anh.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Phần 3: Đề xuất hoặc nhận xét bổ sung về việc sử dụng AI trong việc học nói tiếng Anh (nếu có):					

APPENDIX C

THE SEMI-STRUCTURED INTERVIEW (ENGLISH VERSION)

The interview would take about 15 minutes. Please be noted that there is no right or wrong answer. Your responses to the interview questions will be recorded to an audio file and kept strictly confidential, only for research purposes and only the researcher of this study can get access to the recorded files. Thank you for your participation. Below is a list of semi-structured interview questions.

Interview questions
<ol style="list-style-type: none">1. How often and how long did you practice speaking with the AI voice chatbot every week?2. Do you think that your grammar has been improved after speaking with the AI voice chatbot? Why do you think so?3. Do you think that your pronunciation has been improved after speaking with the AI voice chatbot? Why do you think so?4. Do you think that your vocabulary has been improved after speaking with the AI voice chatbot? Why do you think so?5. Do you think that you can speak with fewer hesitations or pauses after speaking with the AI voice chatbot? Why do you think so?6. Do you think that you can speak better after speaking with the AI voice chatbot? Why do you think so?7. What do you like most about speaking with the AI voice chatbot? And what do you dislike?

PHẦN PHỎNG VẤN (Vietnamese Version)

Lưu ý: Buổi phỏng vấn sẽ mất khoảng 15 phút. Bạn cần biết rằng không có câu trả lời nào là đúng hay sai. Câu trả lời của bạn cho các câu hỏi phỏng vấn sẽ được ghi âm vào tệp âm thanh và bảo mật tuyệt đối. Các tệp này chỉ sử dụng cho mục đích nghiên cứu và duy nhất nghiên cứu viên trong nghiên cứu này mới có thể tiếp cận các tệp âm thanh đã ghi âm. Cảm ơn bạn đã tham gia vào buổi phỏng vấn này. Dưới đây là danh mục các câu hỏi phỏng vấn.

Các câu hỏi phỏng vấn
<ol style="list-style-type: none"> 1. Bạn thực hành nói với AI bao lâu và thường xuyên ra sao ở mỗi tuần? 2. Bạn có nghĩ rằng ngữ pháp của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy? 3. Bạn có nghĩ rằng phát âm của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy? 4. Bạn có nghĩ rằng từ vựng của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy? 5. Bạn có nghĩ rằng bạn có thể nói ít vấp hay ít ngập ngừng hơn sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy? 6. Bạn có nghĩ rằng bạn có thể nói tốt hơn sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy? 7. Bạn thích điều gì nhất khi nói chuyện với AI và điều gì bạn không thích?

APPENDIX D
 EXAMPLES OF LESSON PLANS FOR CONTROL GROUP
 Lesson Plan
 English Speaking in The Traditional Classroom

Unit 1

1. **Topic:** Journeys
2. **Level of learners:** Undergraduate students at level A2 based on CEFR standard
3. **Class accommodation:** 30 students
4. **Number of teaching hours:** 3 periods (135 minutes)
5. **Objectives:** By the end of the unit, students are able to:
 - 5.1. describe their hometown orally.
 - 5.2. compare two or more things orally.
6. **Target tasks:**
 - 6.1. Students understand how to use “am, is, are” and “there is, there are” to describe places around their hometown.
 - 6.2. Students understand how to use comparative and superlative comparison to compare two or more things.
7. **Teaching procedure:**

Period	Learning and Teaching Activities	Assessment
1 (45 minutes)	<p>Warm-up: Groupwork</p> <p>Students play a game of following-up words in two groups. Each group starts with an adjective shouted by the first member, then the second member shouted another adjective whose starting letter matches with the ending letter of the previous adjective. The third member through the final member do the same as the first and the second members.</p> <p>Ex: P1: long ⇨ P2: gentle ⇨ P3: enormous ⇨ P4: safe ⇨ Pn: ...</p>	<p>Students are given 1 minute 30 seconds to finish the round. A group wins if all members finish the round within the given time or which group has more adjectives shouted within the given time. The teacher records the scores of each group to reward by the end of the course.</p>
	<p>Step 1: Presentation: Individual work</p> <p>Learning how to make descriptive sentences.</p> <p>The teacher asks students to look at the board to see the guideline on how to make descriptive sentences. The teacher explains in class that we usually use am, is, are with the adjectives to make descriptive sentences.</p> <ul style="list-style-type: none"> ● am/is/are + adjectives <p>The teacher further explains in class that a complete sentence should include subject(s), a verb, and object(s). For description, the verb am/is/are should be used. Therefore, we have the structure:</p>	

Period	Learning and Teaching Activities	Assessment
	<ul style="list-style-type: none"> Subject + Verb (am/is/are) + Object Ex: This is a fun party. <p>The teacher notes in class that “am” goes with “I”, “is” goes with singular subjects, and “are” goes with plural subjects.</p> <p>Step 2: Practice: Individual & pair work The teacher asks students to do Exercise 1 on p. 92 to practice describing the given things. After that, the teacher asks students to take turn talking with each other in describing those things.</p> <p>Step 3: Production: Individual & Pair work The teacher asks students to look at the things given in the textbook in Exercise 7 on p. 83 and talk with a partner next to them in describing those things. The given things include: Australia, Antarctica, a car, a bicycle, rock-climbing, surfing, travel by air, travel by sea, an elephant, a lion, a holiday in the city, camping in the countryside, Paris, New York, train journeys, plane journeys.</p>	
2 (45 minutes)	<p>Warm-up: Group work Students play a game of following-up sentences in two groups. Each group starts with a description of a place said by the first member, then the second member said another sentence. The third member through the final member do the same as the first and the second members. Ex: P1: Mount Everest is dangerous ⇨ P2: Swan Lake is beautiful ⇨ Pn: ...</p> <p>Step 1: Presentation: Individual & Pair work Learning how to use comparative to compare two things. The teacher asks students to look at the textbook p. 83 to learn how to make comparative sentences. After that, the teacher asks students to share with a partner next to them of what they understand about using comparative adjectives. Finally, the teacher asks students to look at the board and explain how to use comparative and superlative adjectives to describe and compare things.</p> <ul style="list-style-type: none"> Adjectives with more than 2 syllables: [more + adjective + (than)] Ex: My mother is more careful than me. Adjectives with 2 syllables or less: [adjective-er + (than)] Ex: My mother is taller than me. <p>Step 2: Practice: Individual & Pair work The teacher asks students to do Exercise 6 on p. 83 to practice comparing the given things.</p>	<p>Students are given 1 minute 30 seconds to finish the round. A group wins if all members finish the round within the given time or which group has more sentences said within the given time. The teacher records the scores of each group to reward by the end of the course.</p> <p>Students are given 5 minutes to prepare and write down their sentences on a notebook before speaking. The teacher goes around to observe and assess students’ speaking skill with respect of using correct grammar, vocabulary, pronunciation and their dealing with pauses or hesitations.</p>

Period	Learning and Teaching Activities	Assessment
	<p>After that, the teacher asks students to take turn talking with each other in describing those things.</p> <p>Step 3: Production: Individual & Pair work</p> <p>The teacher asks students to look at the things given in the textbook in Exercise 11 on p. 83 and talk with a partner next to them in comparing those things. The given things include:</p> <ul style="list-style-type: none"> • places or cities • sports or free-time activities • types of travel • types of holiday • places in the city • types of transport • famous people 	<p>errors in pronunciation and grammar.</p> <p>Students are given 5 minutes to prepare and write down their sentences on a notebook before speaking. The teacher goes around to observe and assess students' speaking skill with respect of using correct grammar, vocabulary, pronunciation and their dealing with pauses or hesitations.</p>
<p>3 (45 minutes)</p>	<p>Warm-up: Group work</p> <p>Students play a game of fast description in two groups. The teacher shows 5 pictures in which each picture includes two items for comparison. Each group tries to raise hand the fastest to take the right to orally compare the items in the picture.</p> <p>Step 1: Presentation: Individual & Pair work</p> <p>Learning how to use comparative and superlative to compare two or more things.</p> <p>The teacher asks students to look at the textbook p. 85 to learn how to make superlative sentences.</p> <p>After that, the teacher asks students to share with a partner next to them of what they understand about using superlative adjectives.</p> <p>Finally, the teacher asks students to look at the board and explain how to use superlative adjectives to describe and compare things.</p> <ol style="list-style-type: none"> 1. Adjectives with more than 2 syllables: [the + most + adjective] Ex: She is the most beautiful student in my class. 2. Adjectives with 2 syllables or less: [the + adjective-est] Ex: I am the shortest one in my family. <p>Step 2: Practice: Individual & Pair work</p> <p>The teacher asks students to do Exercise 7 on p. 85 to practice comparing the given things.</p> <p>After that, the teacher asks students to take turn talking with each other in describing those things.</p> <p>Step 3: Production: Individual & Pair work</p> <p>The teacher asks students to look at the things given in the textbook in Exercise 9 on p. 85 and talk with a partner next to them in describing those things. The given things include:</p>	<p>A picture scores 10 marks. A group wins if getting higher scores than the opponent. The teacher records the scores of each group to reward by the end of the course.</p> <p>The teacher goes around and listen to students' practice to fix and give feedback on their errors in pronunciation and grammar.</p> <p>Students are given 5 minutes to prepare and write down their sentences on a notebook before speaking. The teacher goes around to observe and assess students' speaking skill with respect of using correct grammar, vocabulary, pronunciation and their dealing</p>

Period	Learning and Teaching Activities	Assessment
	<ul style="list-style-type: none">• the most famous city• the oldest city• the most beautiful place• the most popular place for tourists• the best place to visit• the hottest month• the coldest month• the cheapest way to travel	with pauses or hesitations.



Unit 2

1. **Topic:** Appearance
2. **Level of learners:** Undergraduate students at level A2 based on CEFR standard
3. **Class accommodation:** 30 students
4. **Number of teaching hours:** 3 periods (135 minutes)
5. **Objectives:** By the end of the unit, students are able to:
 - 5.1. ask and describe a person orally.
 - 5.2. ask and answer about daily routine or ongoing activities orally.
6. **Target tasks:**
 - 6.1. Students understand how to use simple tense and “have/has got” to ask and describe a person.
 - 6.2. Students understand how to use continuous tense to ask and answer about ongoing activities.
7. **Teaching procedure:**

Period	Learning and Teaching Activities	Assessment
1 (45 minutes)	<p>Warm-up: Groupwork</p> <p>Students play a game of adjective teller in two groups. Each group starts with shouting out adjectives describing appearance or personality. However, the next adjectives must be different from previous adjectives of both groups. whose starting letter matches with the ending letter of the previous adjective. The third member through the final member do the same as the first and the second members.</p> <p>Ex: G1: long, short, intelligent...; G2: fast, slow, beautiful...</p>	<p>Students are given 1 minute 30 seconds to finish the round. Each group has a secretary to the board to write down the adjectives he/she heard from group members. A group wins if giving more different adjectives than the opponent. The wrong adjectives written on the board by the group secretary will not be counted. The teacher records the scores of each group to reward by the end of the course.</p>
	<p>Step 1: Presentation: Individual & Pair work</p> <p>Learning how to use present continuous to describe ongoing activities.</p> <p>The teacher asks students to look at the textbook p. 95 to learn how to use present continuous to describe ongoing activities.</p> <p>After that, the teacher asks students to share with a partner next to them of what they understand about using present continuous to describe ongoing activities.</p> <p>Finally, the teacher asks students to look at the board and explain how to use present continuous to ask and answer about ongoing things.</p> <ul style="list-style-type: none"> ● Affirmative: [S + am/is/are + V_{ing}] 	

Period	Learning and Teaching Activities	Assessment
	<p>Ex: I am going to the cinema.</p> <ul style="list-style-type: none"> Negative: [S + am/is/are + NOT + V_{ing}] Ex: I am NOT going to the cinema. I am at home. Question: [(Wh-Question) + am/is/are + S + V_{ing} ... ?] Ex1: Where are you going? Ex2: Are you going to the cinema? 	
	<p>Step 2: Practice: Individual & pair work</p> <p>The teacher asks students to do Exercise 8, 12 on p. 95 to practice using present continuous tense. After that, the teacher asks students to do Exercise 9 about “Describe what clothes...” on p.95 and take turn talking with each other in describing the clothes people in the exercise are wearing.</p>	<p>The teacher goes around and listen to students’ practice to fix and give feedback on their errors in pronunciation and grammar.</p>
	<p>Step 3: Production: Individual & Pair work</p> <p>The teacher asks students to do Exercise 13 on p. 95 and talk with a partner next to them by asking the given questions in the exercise and answering with the students’ true situation.</p>	<p>Students are given 5 minutes to prepare and write down their sentences on a notebook before speaking. The teacher goes around to observe and assess students’ speaking skill with respect of using correct grammar, vocabulary, pronunciation and their dealing with pauses or hesitations.</p>
<p>2 (45 minutes)</p>	<p>Warm-up: Group work</p> <p>Students play a game of fast description in two groups. The teacher shows 5 pictures of people. Each group tries to raise hand the fastest to take the right to orally describe the people in the picture.</p>	<p>A picture scores 10 marks. A group wins if getting higher scores than the opponent. The teacher records the scores of each group to reward by the end of the course.</p>
	<p>Step 1: Presentation: Individual & Pair work</p> <p>Learning how to use present simple to describe present activities.</p> <p>The teacher asks students to look at the textbook p. 95 to learn how to use present simple to describe present activities. After that, the teacher asks students to share with a partner next to them of what they understand about using present simple to describe present activities. Finally, the teacher asks students to look at the board and explain how to use present simple to ask and answer about present activities.</p> <ul style="list-style-type: none"> Affirmative: [S + V_s/V_{es}/V_{inf} + ...] Ex: He normally designs clothes for the rich. Negative: [S + do/does + NOT + V_{inf}] Ex: She does NOT drink beer. 	

Period	Learning and Teaching Activities	Assessment
	<ul style="list-style-type: none"> Question: [(Wh-Question) + do/does + S + V_{inf} ... ?] Ex: What does she do for living? <p>Step 2: Practice: Individual & Pair work The teacher asks students to do Exercise 6 on p. 104 to practice using present continuous and present simple tenses to describe ongoing and present activities. After that, the teacher asks students to take turn talking with each other in describing the picture in the exercise.</p> <p>Step 3: Production: Individual work The teacher asks students to look at the photo on p. 100 and do the activity named “TALKING ABOUT PICTURES AND PHOTOS” on the same page with a partner next to them in describing the photo following the instruction of the activity.</p>	<p>The teacher goes around and listen to students’ practice to fix and give feedback on their errors in pronunciation and grammar.</p> <p>Students are given 5 minutes to prepare and write down their sentences on a notebook before speaking. The teacher goes around to observe and assess students’ speaking skill with respect of using correct grammar, vocabulary, pronunciation and their dealing with pauses or hesitations.</p>
<p>3 (45 minutes)</p>	<p>Warm-up: Group work Students play a game of following-up sentences in two groups. Each group starts with a description of an ongoing activity said by the first member, then the second member, the third member, and the rest. Ex: P1: I am learning English. ⇨ P2: She is playing a guitar ⇨ Pn: ...</p>	<p>Each group is given 1 minute 30 seconds to finish the round. A group wins if all members finish the round within the given time or which group has more sentences said within the given time. Wrong sentences in terms of grammar will not be counted. The teacher records the scores of each group to reward by the end of the course.</p>
	<p>Step 1: Presentation: Individual & Pair work Learning how to use “has/have got” to describe people. The teacher asks students to look at the textbook p. 97 to learn how to use “has/have got” to describe people. After that, the teacher asks students to share with a partner next to them of what they understand about using “has/have got” to describe people. Finally, the teacher asks students to look at the board and explain how to use “has/have got” to ask and answer about ongoing things.</p> <ul style="list-style-type: none"> Affirmative: [S + has/have + got ...] Ex: He has got brown short hair. Negative: [S + has/have + NOT + got ...] Ex: I have NOT got thick eyebrows. Question: [Has/Have + S + got ... ?] Ex: Has your father got black beard? 	

Period	Learning and Teaching Activities	Assessment
	<p>Step 2: Practice: Individual & Pair work</p> <p>The teacher asks students to do Exercise 2 on p. 96 and Exercise 8, 9 on p. 97 to practice using “has/have got”.</p> <p>After that, the teacher asks students to do Exercise 10 about “Description of your appearance” on p.97 and take turn talking with each other in describing each other appearance.</p>	<p>The teacher goes around and listen to students’ practice to fix and give feedback on their errors in pronunciation and grammar.</p>
	<p>Step 3: Production: Individual & Pair work</p> <p>The teacher asks students to look at Exercise 12 on p. 97 and talk with a partner next to them in describing some people in the class.</p>	<p>Students are given 5 minutes to prepare and write down their sentences on a notebook before speaking. The teacher goes around to observe and assess students’ speaking skill with respect of using correct grammar, vocabulary, pronunciation and their dealing with pauses or hesitations.</p>

APPENDIX E
EXAMPLES OF LESSON PLANS FOR EXPERIMENTAL GROUP

Lesson Plan
English Speaking with AI Voice Chatbots

Unit 1

1. **Topic:** Journeys
2. **Level of learners:** Undergraduate students at level A2 based on CEFR standard
3. **Class accommodation:** 30 students
4. **Number of teaching hours:** 3 periods (135 minutes)
5. **Objectives:** By the end of the unit, students are able to:
 - 5.1. describe their hometown orally.
 - 5.2. compare two or more things orally.
6. **Target tasks:**
 - 6.1. Students understand how to use “am, is, are” and “there is, there are” to describe places around their hometown.
 - 6.2. Students understand how to use comparative and superlative comparison to compare two or more things.
7. **Teaching procedure:**

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
1 (45 minutes)	Warm-up: Groupwork Students play a game of following-up words in two groups. Each group starts with an adjective shouted by the first member, then the second member shouted another adjective whose starting letter matches with the ending letter of the previous adjective. The third member through the final member do the same as the first and the second members. Ex: P1: long ⇨ P2: gentle ⇨ P3: enormous ⇨ P4: safe ⇨ Pn: ...	<i>Connectivism by Siemens, 2005:</i> Students learn by connecting and interacting with each other.	Students are given 1 minute 30 seconds to finish the round. A group wins if all members finish the round within the given time or which group has more adjectives shouted within the given time. The teacher records the scores of each group to reward by the end of the course.
	Step 1: Presentation: Individual work & pair work Learning how to make descriptive sentences. Students watch a video clip in the section titled “Describing Your Hometown” on the website https://aispeaks.mobirisesite.com/unit-1.html to learn how to use am, is, are and there is, there are to describe places around their hometown.	<i>Behaviorism by Skinner, 1957:</i> Students learn to use accurate pronunciation through listening, recognize and reproduce words. Machine contains knowledge and deliver the knowledge to learners. <i>Connectivism by Siemens, 2005:</i> Students learn from a variety of opinions <i>Behaviorism by Walker &</i>	

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
	<p>After that, students share with a partner next to them of what they understand about the ways they think could be used to describe places based on the video clip they have watched.</p> <p>Finally, students look at the section titled “How to Make Descriptive Sentences” on the website https://aispeaks.mobirisesite.com/unit-1.html to learn that we can use a sentence with “to be” (am, is, are) or “there is...”, “there are...” to make descriptive sentences.</p> <ul style="list-style-type: none"> ● I + am + ... ● He/She/It/Singular subjects + is + ... ● You/We/They/Plural subjects + are + ... ● There is + singular noun + ... ● There are + plural noun + ... <p>Ex: There are three city parks in my hometown. They are huge.</p>	<p><i>White, 2013:</i> Machine plays the role of an instructor to students.</p>	
	<p>Step 2: Practice: Individual & pair work <i>* Note: The teacher asks students to repeat the words below after the teacher to practice pronunciation before talking to Andy.</i></p> <p>Students open the app Andy English Bot (Andy) on their smartphone to talk with Andy in describing the following places in their hometown:</p> <ul style="list-style-type: none"> ● Supermarkets ● A city park ● Movie theatres ● Local markets ● Schools and universities. <p><i>Ex: Can Tho University is very large. There are many schools in the university.</i></p> <p>Students show their chat screen with Andy to a partner next to them or to the teacher to share opinions about each other’s correct or incorrect use of descriptive sentences.</p>	<p><i>Constructivism by Piaget, 1964:</i> Students learn by constructing knowledge by themselves from the existing knowledge.</p> <p><i>Social constructivism by Vygotsky, 1978; and Krashen, 2013:</i> Students learn by developing communication with those who have more expertise.</p> <p><i>Connectivism by Siemens, 2005:</i> Students learn from a variety of opinions. Students learn in a physical or virtual environment built on non-human appliances.</p>	<p>The teacher listens to students’ recorded voices and investigate their chat screenshots after class to evaluate their speaking accuracy and fluency. Students receive teachers’ grades and oral feedback by the end of the unit.</p>
	<p>Step 3: Production: Individual work Students select ONE of the given topics in the production section on the website https://aispeaks.mobirisesite.com/unit-1.html and open the app Andy English Bot (Andy) on their smartphone to talk with Andy in describing the places and then submitting the recordings and screenshots of the chat sessions to teacher for assessment. The given topics include:</p> <ul style="list-style-type: none"> ● Tourist attractions in Vietnam 	<p><i>Constructivism by Piaget, 1964:</i> Students learn by constructing knowledge by themselves from the existing knowledge.</p> <p><i>Social constructivism by Piaget, 1964; Vygotsky, 1978; and Krashen, 2013:</i> Students learn by developing communication with those who have more expertise.</p> <p><i>Connectivism by Siemens,</i></p>	<p>Students are given 5 minutes to prepare and write down their sentences on a notebook or a phone note before speaking. The teacher guides students on how to record their voice and take screenshots of the chat sessions with Andy</p>

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
	<ul style="list-style-type: none"> Shopping malls in hometown Eating places around Can Tho City Entertaining places around Can Tho City 	<p>2005: Students learn from a variety of opinions and decide what to learn. Students learn in a physical or virtual environment built on non-human appliances.</p>	to submit via Google form for assessment before making production.
2 (45 minutes)	<p>Warm-up: Group work Students play a game of following-up sentences in two groups. Each group starts with a description of a place said by the first member, then the second member said another sentence. The third member through the final member do the same as the first and the second members. Ex: P1: Mount Everest is dangerous ⇨ P2: Swan Lake is beautiful ⇨ Pn: ...</p>	<p>Connectivism by Siemens, 2005: Students learn by connecting and interacting with each other.</p>	Students are given 1 minute 30 seconds to finish the round. A group wins if all members finish the round within the given time or which group has more sentences said within the given time. The teacher records the scores of each group to reward by the end of the course.
	<p>Step 1: Presentation: Individual work Learning how to use comparative and superlative to compare two or more things. Students watch a video clip in the section titled “Comparative and Superlative Adjectives” on the website https://aispeaks.mobirisesite.com/unit-1.html to learn how to use comparative and superlative to compare two or more things.</p>	<p>Behaviorism by Walker & White, 2013: Machine plays the role of an instructor to students. Connectivism by Siemens, 2005: Students learn from a variety of opinions. Students learn in a physical or virtual environment built on non-human appliances.</p>	
	<p>Step 2: Practice: Individual & pair work <i>* Note: The teacher asks students to repeat the words below after the teacher to practice pronunciation before talking to Andy.</i> Students open the app Andy English Bot (Andy) on their smartphone to talk with Andy in comparing the following things: 1 Australia / Antarctica 2 a car / a bicycle 3 rock-climbing / surfing 4 travel by air / by sea 5 an elephant / a lion 6 a holiday in the city / camping in the countryside 7 Paris / New York 8 train journeys / plane journeys Ex: Australia is smaller than Antarctica ⇨ Antarctica is the biggest. Students show their chat screen with Andy to a partner next to them or to the teacher to share opinions about each other's correct or incorrect use of comparative adjectives.</p>	<p>Constructivism by Piaget, 1964: Students learn by constructing knowledge by themselves from the existing knowledge. Social constructivism by Vygotsky, 1978; and Krashen, 2013: Students learn by developing communication with those who have more expertise. Connectivism by Siemens, 2005: Students learn from a variety of opinions. Students learn in a physical or virtual environment built on non-human appliances.</p>	Students are given 5 minutes to prepare and write down their sentences on a notebook or a phone note before speaking. The teacher guides students on how to record their voice and take screenshots of the chat sessions with Andy to submit via Google form for assessment before making production.
	<p>Step 3: Production: Individual work Students select ONE of the given topics in the production section on the website</p>	<p>Constructivism by Piaget, 1964: Students learn by constructing knowledge by themselves from</p>	Students are given 5 minutes to prepare and write down their

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
	<p>https://aispeaks.mobirisesite.com/unit-1.html and open the app Andy English Bot (Andy) on their smartphone to talk with Andy in comparing the things and then submitting the recordings and screenshots of the chat sessions to teacher for assessment. The given topics include:</p> <ul style="list-style-type: none"> • places or cities • sports or free-time activities • types of travel • types of holiday • places in the city • types of transport • famous people • the most famous city • the oldest city • the most beautiful place • the most popular place for tourists • the best place to visit • the hottest month • the coldest month • the cheapest way to travel 	<p>the existing knowledge.</p> <p><i>Social constructivism by Vygotsky, 1978; and Krashen, 2013:</i> Students learn by developing communication with those who have more expertise.</p> <p><i>Connectivism by Siemens, 2005:</i> Students learn from a variety of opinions and decide what to learn. Students learn in a physical or virtual environment built on non-human appliances.</p>	<p>sentences on a notebook or a phone note before speaking. The teacher guides students on how to record their voice and take screenshots of the chat sessions with Andy to submit via Google form for assessment before making production.</p>
3 (45 minutes)	<p>Warm-up: Group work Students play a game of fast description in two groups. The teacher shows 5 pictures in which each picture includes two items for comparison. Each group tries to raise hand the fastest to take the right to orally compare the items in the picture.</p>	<p><i>Connectivism by Siemens, 2005:</i> Students learn by connecting and interacting with each other.</p>	<p>A picture scores 10 marks. A group wins if getting higher scores than the opponent. The teacher records the scores of each group to reward by the end of the course.</p>
	<p>Step 1: Presentation: Teacher's Feedback The teacher gives oral feedback in class on students' speaking skill based on the recordings and chat screenshots they submitted. The feedback is about:</p> <ol style="list-style-type: none"> 3. How fluently they spoke in terms of pauses and hesitations. 4. How accurately they spoke in terms of grammar, vocabulary, and pronunciation. 5. How to improve English speaking skill in terms of fluency and accuracy – speak slowly but clearly enough, focus on meaning first, then grammar. 	<p><i>Constructivism by Piaget, 1964:</i> Students learn by constructing knowledge by themselves from the existing knowledge.</p> <p><i>Social constructivism by Vygotsky, 1978; and Krashen, 2013:</i> Students learn by developing communication with those who have more expertise.</p>	
	<p>Step 2: Practice: Students' Self-Study Students are given 10 minutes to discuss with a partner about the words they found difficult to pronounce based on the productions they made in Period 1 & 2 and then write those words down in a piece of paper or type in the chat group.</p> <p>After that, students click on the link provided in the section titled "Practice Pronunciation" on the website https://aispeaks.mobirisesite.com/unit-1.html to be directed to Cambridge Online Dictionary.</p>	<p><i>Constructivism by Piaget, 1964:</i> Students learn by constructing knowledge by themselves from the existing knowledge.</p> <p><i>Social constructivism by Vygotsky, 1978; and Krashen, 2013:</i> Students learn by developing communication with those who have more expertise.</p> <p><i>Behaviorism by Walker & White, 2013:</i> Machine plays the role of an instructor to students.</p>	

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
	<p>Finally, students type in the words they found difficult on Cambridge Online Dictionary to listen to the pronunciation of the words and orally repeat after listening to practice pronunciation.</p>		
	<p>Step 3: Production: Students select ONE of the given topics in the production section on the website https://aispeaks.mobirisesite.com/unit-1.html and open the app Andy English Bot (Andy) on their smartphone to talk with Andy in describing and comparing the things and then submitting the recordings and screenshots of the chat sessions to teacher for assessment. The given topics include:</p> <ul style="list-style-type: none"> ● Describe and compare students' hometown with a friend's hometown. ● Describe and compare a supermarket with a local market in students' hometown. ● Describe and compare Vietnam with a country in the world. 	<p><i>Constructivism by Piaget, 1964:</i> Students learn by constructing knowledge by themselves from the existing knowledge.</p> <p><i>Social constructivism by Vygotsky, 1978; and Krashen, 2013:</i> Students learn by developing communication with those who have more expertise.</p> <p><i>Connectivism by Siemens, 2005:</i> Students learn from a variety of opinions and decide what to learn. Students learn in a physical or virtual environment built on non-human appliances.</p>	<p>Students are given 5 minutes to prepare and write down their sentences on a notebook or a phone note before speaking. The teacher guides students on how to record their voice and take screenshots of the chat sessions with Andy to submit via Google form for assessment before making production.</p>


Unit 2

1. **Topic:** Appearance
2. **Level of learners:** Undergraduate students at level A2 based on CEFR standard
3. **Class accommodation:** 30 students
4. **Number of teaching hours:** 3 periods (135 minutes)
5. **Objectives:** By the end of the unit, students are able to:
 - 5.1. ask and describe a person orally.
 - 5.2. ask and answer about daily routine or ongoing activities orally.
6. **Target tasks:**
 - 6.1. Students understand how to use simple tense and “have/has got” to ask and describe a person.
 - 6.2. Students understand how to use continuous tense to ask and answer about ongoing activities.
7. **Teaching procedure:**

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
1 (45 minutes)	<p>Warm-up: Groupwork Students play a game of adjective teller in two groups. Each group starts with shouting out adjectives describing appearance or personality. However, the next adjectives must be different from previous adjectives of both groups. whose starting letter matches with the ending letter of the previous adjective. The third member through the final member do the same as the first and the second members. Ex: G1: long, short, intelligent...; G2: fast, slow, beautiful...</p>	<p><i>Connectivism</i> by <i>Siemens, 2005:</i> Students learn by connecting and interacting with each other.</p>	<p>Students are given 1 minute 30 seconds to finish the round. Each group has a secretary to the board to write down the adjectives he/she heard from group members. A group wins if giving more different adjectives than the opponent. The wrong adjectives written on the board by the group secretary will not be counted. The teacher records the scores of each group to reward by the end of the course.</p>
	<p>Step 1: Presentation: Individual work & pair work Learning how to use present continuous to describe ongoing activities. Students watch a video clip in the section titled “Describing Ongoing Activities” on the website https://aispeaks.mobirisesite.com/unit-2.html to learn how to use present continuous tense to describe ongoing activities. After that, students share with a partner next to them of what they understand about the ways they think could be used to ask and describe ongoing activities.</p>	<p><i>Behaviorism</i> by <i>Skinner, 1957:</i> Students learn to use accurate pronunciation through listening, recognize and reproduce words. Machine contains knowledge and deliver the knowledge to learners. <i>Connectivism</i> by <i>Siemens, 2005:</i></p>	

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
	<p>Finally, students look at the section titled “How to Use Present Continuous Tense” on the website https://aispeaks.mobirisesite.com/unit-2.html to learn that we can use present continuous tense to describe ongoing activities.</p> <ul style="list-style-type: none"> ● Affirmative: [S + am/is/are + V_{ing}] Ex: I am going to the cinema. ● Negative: [S + am/is/are + NOT + V_{ing}] Ex: I am NOT going to the cinema. I am at home. ● Question: [(Wh-Question) + am/is/are + S + V_{ing} ... ?] Ex1: Where are you going? Ex2: Are you going to the cinema? 	<p>Students learn from a variety of opinions Behaviorism by Walker & White, 2013: Machine plays the role of an instructor to students.</p>	
	<p>Step 2: Practice: Individual & pair work <i>* Note: The teacher asks students to repeat the words below after the teacher to practice pronunciation before talking to Andy.</i></p> <p>Students open the app Andy English Bot (Andy) on their smartphone to talk with Andy in describing the following ongoing activities:</p> <ul style="list-style-type: none"> ● The clothes you are wearing. ● The lessons you are learning. ● The clothes your friend is wearing. ● The clothes your teaching is wearing. <p>Ex: I am wearing a blue T-shirt. I am learning English in the class...</p> <p>Students show their chat screen with Andy to a partner next to them or to the teacher to share opinions about each other's correct or incorrect use of present continuous tense.</p>	<p>Constructivism by Piaget, 1964: Students learn by constructing knowledge by themselves from the existing knowledge. Social constructivism by Vygotsky, 1978; and Krashen, 2013: Students learn by developing communication with those who have more expertise. Connectivism by Siemens, 2005: Students learn from a variety of opinions. Students learn in a physical or virtual environment built on non-human appliances.</p>	<p>The teacher listens to students' recorded voices and investigate their chat screenshots after class to evaluate their speaking accuracy and fluency. Students receive teachers' grades and oral feedback by the end of the unit.</p>
	<p>Step 3: Production: Individual work Students select ONE of the given topics in the production section on the website https://aispeaks.mobirisesite.com/unit-2.html and open the app Andy English Bot (Andy) on their smartphone to talk with Andy in describing ongoing activities and then submitting the recordings and screenshots of the chat sessions to teacher for assessment. The given topics include:</p> <ul style="list-style-type: none"> ● A festival taking place in your hometown. ● An activity you are doing. ● A movie you are watching. ● A genre of music you are listening. 	<p>Constructivism by Piaget, 1964: Students learn by constructing knowledge by themselves from the existing knowledge. Social constructivism by Piaget, 1964; Vygotsky, 1978; and Krashen, 2013: Students learn by developing communication with those who have more expertise. Connectivism by</p>	<p>Students are given 5 minutes to prepare and write down their sentences on a notebook or a phone note before speaking. The teacher guides students on how to record their voice and take screenshots of the chat sessions with Andy to submit via Google form for assessment before making production.</p>

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
		<i>Siemens, 2005:</i> Students learn from a variety of opinions and decide what to learn. Students learn in a physical or virtual environment built on non-human appliances.	
2 (45 minutes)	<p>Warm-up: Group work Students play a game of fast description in two groups. The teacher shows 5 pictures of people. Each group tries to raise hand the fastest to take the right to orally describe the people in the picture.</p>	<i>Connectivism by Siemens, 2005:</i> Students learn by connecting and interacting with each other.	A picture scores 10 marks. A group wins if getting higher scores than the opponent. The teacher records the scores of each group to reward by the end of the course.
	<p>Step 1: Presentation: Individual & Pair work Learning how to use present simple to describe present activities. Students watch a video clip in the section titled “Describing Present Activities” on the website https://aispeaks.mobirisesite.com/unit-2.html to learn how to use present simple tense to describe present activities.</p> <p>After that, students share with a partner next to them of what they understand about the ways they think could be used to ask and describe present activities.</p> <p>Finally, students look at the section titled “How to Use Present Simple Tense and Has/Have Got” on the website https://aispeaks.mobirisesite.com/unit-2.html to learn that we can use present simple tense and “has/have got” to describe present activities.</p> <ul style="list-style-type: none"> ● Affirmative: [S + V_s/V_{es}/V_{inf} + ...] [S + has/have + got ...] Ex1: He normally designs clothes for the rich. Ex2: He has got three years of experience. ● Negative: [S + do/does + NOT + V_{inf}] [S + has/have + NOT + got ...] Ex1: She does NOT drink beer. Ex2: She has not got comfortable. ● Question: [(Wh-Question) + do/does + S + V_{inf} ... ?] [Has/Have + S + got ... ?] Ex1: What does she do for living? Ex2: What has she got from the job? 	<p><i>Behaviorism by Walker & White, 2013:</i> Machine plays the role of an instructor to students.</p> <p><i>Connectivism by Siemens, 2005:</i> Students learn from a variety of opinions. Students learn in a physical or virtual environment built on non-human appliances.</p>	
	<p>Step 2: Practice: Individual & pair work <i>* Note: The teacher asks students to repeat the words below after the teacher to practice pronunciation before talking to Andy.</i></p>	<i>Constructivism by Piaget, 1964:</i> Students learn by constructing knowledge by themselves from the	Students are given 5 minutes to prepare and write down their sentences on a notebook or a phone

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
	<p>Students open the app Andy English Bot (Andy) on their smartphone to talk with Andy in describing the following present activities:</p> <ul style="list-style-type: none"> ● What you do in the morning. ● What you do in the afternoon. ● What you do in the evening. ● What you do on weekends. <p>Ex: I normally get up early in the morning. I have lunch in the afternoon. I have got lunch with rice and stewed pork...</p> <p>Students show their chat screen with Andy to a partner next to them or to the teacher to share opinions about each other's correct or incorrect use of present simple tense.</p>	<p>existing knowledge.</p> <p>Social constructivism by Vygotsky, 1978; and Krashen, 2013: Students learn by developing communication with those who have more expertise.</p> <p>Connectivism by Siemens, 2005: Students learn from a variety of opinions. Students learn in a physical or virtual environment built on non-human appliances.</p>	<p>note before speaking. The teacher guides students on how to record their voice and take screenshots of the chat sessions with Andy to submit via Google form for assessment before making production.</p>
	<p>Step 3: Production: Individual work</p> <p>Students look at the photo in the production section on the website https://aispeaks.mobirisesite.com/unit-2.html and open the app Andy English Bot (Andy) on their smartphone to talk with Andy in describing activities and then submitting the recordings and screenshots of the chat sessions to teacher for assessment. The photo provided:</p>  <p>Ex: I see many people in the picture. A man is standing behind a woman. He has got a hat. Another woman normally looks at somewhere...</p>	<p>Constructivism by Piaget, 1964: Students learn by constructing knowledge by themselves from the existing knowledge.</p> <p>Social constructivism by Vygotsky, 1978; and Krashen, 2013: Students learn by developing communication with those who have more expertise.</p> <p>Connectivism by Siemens, 2005: Students learn from a variety of opinions and decide what to learn. Students learn in a physical or virtual environment built on non-human appliances.</p>	<p>Students are given 5 minutes to prepare and write down their sentences on a notebook or a phone note before speaking. The teacher guides students on how to record their voice and take screenshots of the chat sessions with Andy to submit via Google form for assessment before making production.</p>
3 (45 minutes)	<p>Warm-up: Group work</p> <p>Students play a game of following-up sentences in two groups. Each group starts with a description of an ongoing activity said by the first member, then the second member, the third member, and the rest.</p> <p>Ex: P1: I am learning English. ⇨ P2: She is playing a guitar ⇨ Pn: ...</p>	<p>Connectivism by Siemens, 2005: Students learn by connecting and interacting with each other.</p>	<p>Each group is given 1 minute 30 seconds to finish the round. A group wins if all members finish the round within the given time or which group has more sentences said within the given time. Wrong sentences in terms of grammar will not be counted. The teacher records the scores of each group to</p>

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
			reward by the end of the course.
	<p>Step 1: Presentation: Teacher's Feedback The teacher gives oral feedback in class on students' speaking skill based on the recordings and chat screenshots they submitted. The feedback is about:</p> <ol style="list-style-type: none"> 6. How fluently they spoke in terms of pauses and hesitations. 7. How accurately they spoke in terms of grammar, vocabulary, and pronunciation. 8. How to improve English speaking skill in terms of fluency and accuracy – speak slowly but clearly enough, focus on meaning first, then grammar. 	<p><i>Constructivism</i> by <i>Piaget, 1964</i>: Students learn by constructing knowledge by themselves from the existing knowledge. <i>Social constructivism</i> by <i>Vygotsky, 1978; and Krashen, 2013</i>: Students learn by developing communication with those who have more expertise.</p>	
	<p>Step 2: Practice: Students' Self-Study Students are given 10 minutes to discuss with a partner about the words they found difficult to pronounce based on the productions they made in Period 1 & 2 and then write those words down in a piece of paper or type in the chat group.</p> <p>After that, students click on the link provided in the section titled "Practice Pronunciation" on the website https://aispeaks.mobirisesite.com/unit-2.html to be directed to Cambridge Online Dictionary.</p> <p>Finally, students type in the words they found difficult on Cambridge Online Dictionary to listen to the pronunciation of the words and orally repeat after listening to practice pronunciation.</p>	<p><i>Constructivism</i> by <i>Piaget, 1964</i>: Students learn by constructing knowledge by themselves from the existing knowledge. <i>Social constructivism</i> by <i>Vygotsky, 1978; and Krashen, 2013</i>: Students learn by developing communication with those who have more expertise. <i>Behaviorism</i> by <i>Walker & White, 2013</i>: Machine plays the role of an instructor to students.</p>	

Period	Learning and Teaching Activities	Learning Theory Applied	Assessment
	<p>Step 3: Production: Students select ONE of the given pictures in the production section on the website https://aispeaks.mobirisesite.com/unit-2.html and open the app Andy English Bot (Andy) on their smartphone to talk with Andy in describing a picture and then submitting the recordings and screenshots of the chat sessions to teacher for assessment. The given pictures include:</p> <ul style="list-style-type: none"> ● Picture 1:  ● Picture 2:  ● Picture 3:  	<p>Constructivism by Piaget, 1964: Students learn by constructing knowledge by themselves from the existing knowledge.</p> <p>Social constructivism by Vygotsky, 1978; and Krashen, 2013: Students learn by developing communication with those who have more expertise.</p> <p>Connectivism by Siemens, 2005: Students learn from a variety of opinions and decide what to learn. Students learn in a physical or virtual environment built on non-human appliances.</p>	<p>Students are given 5 minutes to prepare and write down their sentences on a notebook or a phone note before speaking. The teacher guides students on how to record their voice and take screenshots of the chat sessions with Andy to submit via Google form for assessment before making production.</p>

APPENDIX F
THE SPEAKING ASSESSMENT RUBRIC
THE ASSESSMENT RUBRIC FOR THE SPEAKING TESTS

Speaking Performance (Test domain)					
Categories (Concepts & Constructs)	Score and interpretation				
	1	2	3	4	5
Speaking accuracy	<p>The test taker can only use few words, insufficient grammar structure, and inaccurate pronunciation, or the test taker fails to answer the question properly (out of the topic or wrong content).</p> <p><i>Ex:</i> Q: How many students are there in your class? A: Their twenty.</p>	<p>The test taker can only use simple words and grammar structures. Sometimes he/she fails to pronounce accurately or fails to answer the question properly (out of the topic or wrong content).</p> <p><i>Ex:</i> Q: Do you like indoor or outdoor activities? A: I like outdoor, ... activity.</p>	<p>The test taker can use a small range of words and makes errors in grammar structures and pronunciation. Sometimes he/she fails to answer the question properly (out of the topic or wrong content).</p> <p><i>Ex:</i> Q: What do you study? A: I'm study English.</p>	<p>The test taker can use a wide range of words and makes only a few grammar errors. However, the test taker often makes clear and accurate pronunciation.</p> <p><i>Ex:</i> Q: Do you prefer PCs or laptops? A: I have both PCs and laptops at home but I prefer laptops more than PCs.</p>	<p>The test taker can use appropriate expressions and a wide range of vocabulary with correct grammar and pronunciation.</p> <p><i>Ex:</i> Q: What do you study? A: I study English. It is one of my favorite subjects I have learnt in my course.</p>
Speaking fluency	<p>The test taker speaks very little or speaks nothing.</p> <p><i>Ex:</i> Q: How do you go to school? A: Uh... uhm...</p>	<p>The test taker can speak with much hesitation interfering with communication.</p> <p><i>Ex:</i> Q: How do you go to school? A: Uh... I... uh... go to school... uhm... by bus.</p>	<p>The test taker can speak with some hesitation, which sometimes interferes with communication.</p> <p><i>Ex:</i> Q: How do you go to school? A: Uh... I go to school... uhm... by bus.</p>	<p>The test taker can speak with some hesitation without usually interrupting the conversation flow.</p> <p><i>Ex:</i> Q: How do you go to school? A: Uh... I go to school by bus, yeah.</p>	<p>The test taker can speak smoothly with confidence but little hesitation, without interrupting the conversation flow.</p> <p><i>Ex:</i> Q: How do you go to school? A: I go to school by bus.</p>

*Adapted from the Tohoku Fukushi University (TFU) Foreign Language Assessment
 Rubrics by TFU Language Educators 'Group (TFU-LEG)*

APPENDIX G
THE CRITERIA FOR ENGLISH SPEAKING EXAMINERS

The examiners for the pre-speaking and post-speaking tests in this study will be selected under the criteria for English speaking examiner suggested by the researchers in consideration of the IELTS speaking test examiner knowledge and skill. However, as this research uses the IELTS-emulated speaking tests, the criteria for speaking examiners in this research is proposed to be more relaxing than the IELTS speaking examiner criteria. Therefore, the examiners for the pretest and posttest in this research will be selected based on the following criteria:

No.	Criteria for IELTS English speaking examiners proposed by the IELTS Committee	Criteria for English speaking examiners proposed in this research
1	Be native speakers or non-native speakers whose IELTS scores for speaking and writing components are 9.0.	Be native speakers or non-native speakers. The non-native speakers should have got an IELTS speaking minimum score of 6.5 or got at least 02 years of experience in teaching English speaking.
2	Hold tertiary qualifications or equivalent.	Hold tertiary qualifications or equivalent.
3	Hold relevant qualifications in Teaching English to Speakers of Other Languages (TESOL) or equivalent.	Hold relevant qualifications in the following fields: English language studies, English studies, or English language education.
4	Have substantial relevant teaching experience.	Have at least 03 years of teaching experience.

APPENDIX H

THE IOC VALUE FOR THE QUESTIONNAIRE ITEMS

To investigate the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL students.

This is aimed at measuring validity of the questionnaire items to investigate the target students' opinions learning English speaking by using an AI voice chatbot to see how they think about their improvement in English speaking in terms of fluency and accuracy after they spoke English with an AI voice chatbot. The definitions of the key terms related to this research are also listed in this evaluation form.

The artificial intelligence voice chatbot or the AI voice chatbot in this study is the program integrated with artificial intelligence algorithms and can talk with users through the text-to-speech and voice recognition processes operating under internet connection. According to Kaplan (2016) and Lemma (2018), an AI voice chatbot is a kind of program or app which works under artificial intelligence algorithms like machine learning and natural language processing operated with internet connection and can have a chat with users via a mobile. In this study, Andy English Bot, called Andy, is an AI voice chatbot used through the experiment as it was reported by Kim et al. (2019) that it could chat with users in form of text chat and voice chat and can help users learn and practice English conversation, vocabulary, and grammar.

The Vietnamese undergraduate EFL students in the study means the non-English majored undergraduate students whose English level is equivalent to A2 level in CEFR standard, classified by a TOEIC Bridge test. Those students are in the first year or second year and learn English 2 which is one of the three modules of General English taught by Faculty of Foreign Languages at Can Tho University, Vietnam. They come from various schools and colleges in Can Tho University such as College of Natural Sciences, School of Education, College of Technology, College of Agriculture, and more. Accordingly, Vietnamese undergraduate EFL students are referred by those who major in non-English academic programs such as Chemistry Education, Agriculture, Primary Education, Computer Network, and so on.

Traditional classrooms mentioned in this study refers to the English classrooms in Can Tho University where students study on textbooks and follow teachers guiding in the classrooms. Traditional classrooms are conducted in a face-to-face manner and English language skills are taught by a teacher using oral instruction, handouts, or electronic media such as CDs, MP3 files, or presentation slides throughout the lesson.

English speaking skill in this study refers to the students' competence of using accurately and fluently spoken language to communicate during a conversation in English as a foreign language, from which the hearer can understand the speaker's message. This key term is based on Dao (2017) mentioning that English speaking skill is achieved when the speaker can successfully convey his meaning to the listener during a conversation by speaking the language fluently and accurately.

Fluency is defined in this study as one of the distinctive criteria demonstrating learners' proficient speaking skill. Fluency is the flowing or the natural language produced by the EFL students when they speak English, demonstrated through speaker's hesitations and pauses. Brown and Lee (2015) explained that fluency in speaking is best achieved when the speech flow is performed smoothly. In addition, Wang (2014) determined hesitations and pauses in speaking as the main obstructions to speaking fluency.

Accuracy in this study is defined as the EFL students' correct use of pronunciation, vocabulary, and grammar when they speak English. According to Brown and Lee (2015), accuracy focuses on the accurate use of phonology, grammar, morphosyntax, and discourse in the spoken output. Wang (2014) also stated that speaking accuracy involves the correct use of pronunciation, vocabulary, and grammar in English speaking. Together with fluency accuracy is seen in this study as another distinctive criterion to demonstrate learners' proficient speaking skill.

Directions: Please refer to each statement on the evaluation form and then make a tick (✓) in the rating box that well describes your opinions on each statement. The rating criteria are described as follows:

Description:

+1: The statement is appropriate and relevant to the researchers' objectives.

0: The statement sounds irrelevant to the researchers' objectives.

-1: The statement is inappropriate and irrelevant to the researchers' objectives.

Part 1: General information							
<i>The general information is to investigate the demographic of the participants regarding their age, major, year of study, English placement test score, and technological background in the undergraduate program so that the researchers can have an overview of the participants' background.</i>							
No.	Description	Experts' Opinions			Remarks		
		-1	0	+1			
1	Gender <input type="checkbox"/> Male <input type="checkbox"/> Female	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			
2	Age <input type="checkbox"/> 18 <input type="checkbox"/> 19 <input type="checkbox"/> 20 <input type="checkbox"/> 21 <input type="checkbox"/> 22 <input type="checkbox"/> 23	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			
3	Major <input type="checkbox"/> Medicinal Chemistry <input type="checkbox"/> Primary Education <input type="checkbox"/> Other (please specify):	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			
4	Year of study <input type="checkbox"/> 1 st year <input type="checkbox"/> 2 nd year <input type="checkbox"/> 3 rd year <input type="checkbox"/> 4 th year	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			
5	English placement test score	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			
6	What technology have you ever used for learning English speaking? You can select more than one option listed in this section. <input type="checkbox"/> Podcast <input type="checkbox"/> Voice chat apps (i.e. Zoom, Zalo, LINE, etc.) <input type="checkbox"/> Chatbots (i.e. Andy English Bot, ELIZA, Siri, Google Assistant, etc.) <input type="checkbox"/> Learning websites (i.e. SpeechAce.com, IELTS Liz.com, etc.) <input type="checkbox"/> Youtube <input type="checkbox"/> Social network (i.e. Facebook, Twitter, Reddit, etc.) <input type="checkbox"/> Others (Please specify):	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			
Part 2: Opinions on English speaking performance							
<i>The questionnaire items are to investigate the participants' opinions on English speaking skill in terms of fluency and accuracy AFTER they take the experiment speaking with an AI voice chatbot.</i>							
No.	Statement	Use the same Likert's scale	Students' opinions on English speaking	Experts' Opinions			Remarks
				-1	0	+1	
1	After practicing with the AI voice chatbot, I think I could speak English without making too many pauses and hesitations.	1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; (5) Strongly agree	Fluency	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2	After practicing with the AI voice chatbot, I think I could speak English with appropriate hedging words such as uhm... ah... oh....	1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; (5) Strongly agree	Fluency	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3	After practicing with the	1) Strongly disagree; (2)	Accuracy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

	AI voice chatbot, I think I could produce correct pronunciation when I spoke English.	Disagree; (3) Neutral; (4) Agree; (5) Strongly agree				
4	After practicing with the AI voice chatbot, I think I could use appropriate sentence structures when I spoke English.	1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; (5) Strongly agree	Accuracy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	After practicing with the AI voice chatbot, I think I could use appropriate words and vocabulary when I spoke English.	1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; (5) Strongly agree	Accuracy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
No.	Description	Experts' Opinions			Remarks	
		-1	0	+1		
	Part 3: Additional suggestions or comments on learning English speaking skill by using an artificial intelligence voice chatbot (<i>if any</i>):	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		



APPENDIX I

THE IOC VALUE FOR THE INTERVIEW QUESTIONS

To investigate the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL students.

This is aimed at measuring validity of the interview questions to investigate the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL students through exploring their opinions on the improvement of their English speaking skill in terms of fluency and accuracy AFTER learning English speaking by using an AI voice chatbot. The definitions of the key terms related to this research are also ed in this evaluation form.

The artificial intelligence voice chatbot or **the AI voice chatbot** in this study is the program integrated with artificial intelligence algorithms and can talk with users through the text-to-speech and voice recognition processes operating under internet connection. According to Kaplan (2016) and Lemma (2018), an AI voice chatbot is a kind of program or app which works under artificial intelligence algorithms like machine learning and natural language processing operated with internet connection and can have a chat with users via a mobile. In this study, Andy English Bot, called Andy, is an AI voice chatbot used thorough the experiment as it was reported by Kim et al. (2019) that it could chat with users in form of text chat and voice chat and can help users learn and practice English conversation, vocabulary, and grammar.

The Vietnamese undergraduate EFL students in the study means the non-English majored undergraduate students whose English level is equivalent to A2 level in CEFR standard, classified by a TOEIC Bridge test. Those students are in the first year or second year and learn English 2 which is one of the three modules of General English taught by Faculty of Foreign Languages at Can Tho University, Vietnam. They come from various schools and colleges in Can Tho University such as College of Natural Sciences, School of Education, College of Technology, College of Agriculture, and more. Accordingly, Vietnamese undergraduate EFL students are re ed by those who major in non-English academic programs such as Chemistry Education, Agriculture, Primary Education, Computer Network, and so on.

Traditional classrooms mentioned in this study refers to the English classrooms in Can Tho University where students study on textbooks and follow teachers guiding in the classrooms. Traditional classrooms are conducted in a face-to-face manner and English language skills are taught by a teacher using oral instruction,

handouts, or electronic media such as CDs, MP3 files, or presentation slides throughout the lesson.

English speaking skill in this study refers to the students' competence of using accurately and fluently spoken language to communicate during a conversation in English as a foreign language, from which the hearer can understand the speaker's message. This key term is based on Dao (2017) mentioning that English speaking skill is achieved when the speaker can successfully convey his meaning to the listener during a conversation by speaking the language fluently and accurately.

Fluency is defined in this study as one of the distinctive criteria demonstrating learners' proficient speaking skill. Fluency is the flowing or the natural language produced by the EFL students when they speak English, demonstrated through speaker's hesitations and pauses. Brown and Lee (2015) explained that fluency in speaking is best achieved when the speech flow is performed smoothly. In addition, Wang (2014) determined hesitations and pauses in speaking as the main obstructions to speaking fluency.

Accuracy in this study is defined as the EFL students' correct use of pronunciation, vocabulary, and grammar when they speak English. According to Brown and Lee (2015), accuracy focuses on the accurate use of phonology, grammar, morphosyntax, and discourse in the spoken output. Wang (2014) also stated that speaking accuracy involves the correct use of pronunciation, vocabulary, and grammar in English speaking. Together with fluency, accuracy is seen in this study as another distinctive criterion to demonstrate learners' proficient speaking skill.

Directions: Please refer to each statement on the evaluation form and then make a tick (✓) in the rating box that well describes your opinions on each statement. The rating criteria are described as follows:

Description:

+1: The statement is appropriate and relevant to the researchers' objectives.

0: The statement sounds irrelevant to the researchers' objectives.

-1: The statement is inappropriate and irrelevant to the researchers' objectives.

Interview questions						
<p>The section is to investigate the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL students through exploring the participants' opinions on their improvement in English speaking performance in terms of fluency, accuracy, and interaction AFTER they spoke with the AI voice chatbot.</p>						
No.	Statement	Students' opinions on English speaking	Experts' Opinions			Remarks
			-1	0	+1	
1	How often and how long did you practice speaking with the AI voice chatbot every week?	General process	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2	Do you think that your grammar has been improved after speaking with the AI voice chatbot? Why do you think so?	Accuracy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3	Do you think that your pronunciation has been improved after speaking with the AI voice chatbot? Why do you think so?	Accuracy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4	Do you think that your vocabulary has been improved after speaking with the AI voice chatbot? Why do you think so?	Accuracy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
5	Do you think that you can speak with fewer hesitations or pauses after speaking with the AI voice chatbot? Why do you think so?	Fluency	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6	Do you think that you can speak better after speaking with the AI voice chatbot? Why do you think so?	Fluency	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7	What do you like most about speaking with the AI voice chatbot? And what do you dislike?	Feeling	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

APPENDIX J

THE IOC VALUE FOR THE SPEAKING TEST PACK

To evaluate speaking test questions over assessing the test takers English proficiency based on their performance in answering the speaking questions.

This is aimed at measuring validity of the test pack to provide questions and topics for evaluating the proficiency level of the test takers when they take the pre-speaking and post-speaking tests. It is worth to note that the test pack follows the IELTS speaking test structure.

Directions: Please refer to each statement on the evaluation form and then make a tick (✓) in the rating box that well describes your opinions on each statement. The rating criteria are described as follows:

Description:

+1: The statement is appropriate and relevant to the researchers objectives.

0: The statement sounds irrelevant to the researchers objectives.

-1: The statement is inappropriate and irrelevant to the researchers objectives.

Speaking test pack								
<i>The speaking test pack is to provide questions and topics for evaluating the proficiency level of the test takers when they take the pre-speaking and post-speaking tests. It is worth to note that the test pack follows the IELTS speaking test structure.</i>								
No.	Questions & Topics	Specification	Topic	Speaking part	Experts' Opinions			Remarks
					-1	0	+1	
1	Do you prefer street food or homecooked food? Why	The test takers are expected to be able to talk about <i>whether they like street food or homecooked food</i> and <i>why they like it</i> .	Food	1	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
2	Where do you usually go to buy food?	The test takers are expected to be able to talk about <i>where they usually go to buy food</i> , and if possible, <i>share special things</i> about food in that place.	Food	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3	What do you study?	The test takers are expected to be able to talk about <i>what major, subject, or program they learn or have learnt</i> , <i>what the major, subject, or program is about</i> , and if possible, <i>share their thoughts</i> about the major, subject, or program.	Study	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4	Where do you study?	The test takers are expected to be able to talk about <i>where they study, which school, university, or college they study at</i> , and if possible, <i>share special things</i> about their school, university, or college.	Study	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Speaking test pack								
The speaking test pack is to provide questions and topics for evaluating the proficiency level of the test takers when they take the pre-speaking and post-speaking tests. It is worth to note that the test pack follows the IELTS speaking test structure.								
No.	Questions & Topics	Specification	Topic	Speaking part	Experts' Opinions			Remarks
					-1	0	+1	
5	Where is your hometown?	The test takers are expected to be able to talk about <i>where they were born and grew up, where it is located</i> , and if possible, <i>sharing special things about their hometown</i> .	Hometown	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6	How often do you visit your hometown?	The test takers are expected to be able to talk about <i>how frequently they visit their hometown, when they visit</i> , and if possible, <i>share some reasons for their visit</i> .	Hometown	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7	What do you usually do every day?	The test takers are expected to be able to list the activities they do every day, and if possible, explain <i>why they have that habit</i> .	Daily activities	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
8	Do you prefer outdoor or indoor activities? Why?	The test takers are expected to be able to tell <i>whether like to do outdoor or indoor activities</i> , and if possible, explain <i>why they think so</i> .	Daily activities	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
9	Do you like to spend time with your family or friends? Why?	The test takers are expected to be able to tell <i>whether they like to be with their family or friends</i> and if possible, explain <i>why they believe so</i> .	Family and friends	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
10	What do you usually do with your friends?	The test takers are expected to be able to list the activities they do with friends, and if possible, <i>share their thoughts</i> about those activities.	Family and friends	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
11	Describe a book you have recently read. You should say: • What kind of book it is. • What it is about. • Where you found it. • What sort of people would enjoy it. • And explain why you like it.	The test takers can speak in two minutes and respond to all given clues on the topic. They are encouraged to describe clearly the book they have just read (what it is, what it is special for, etc.). They are also expected to tell a little about the book's overview and where or from whom they got it. In addition, they should be able to discuss briefly the target audience of the book. Importantly, it is required that they can give a clear explanation for the reason why they like the book.	Book	2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
12	Describe a person you like best. You should say: • Who he/she is. • When you met	The test takers can speak in two minutes and respond to all given clues on the topic. They are encouraged to describe clearly about a person they like best (who he/she is, what he/she is special for, etc.). They are also expected to tell a short story about the occasion they met that person and how long they have known that one.	People	2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Speaking test pack								
<i>The speaking test pack is to provide questions and topics for evaluating the proficiency level of the test takers when they take the pre-speaking and post-speaking tests. It is worth to note that the test pack follows the IELTS speaking test structure.</i>								
No.	Questions & Topics	Specification	Topic	Speaking part	Experts' Opinions			Remarks
					-1	0	+1	
	that person. • What he/she looks like. • How long you have known him/her. • And explain why you like that person.	In addition, they should be able to describe briefly the person's appearance and personality. Importantly, it is required that they can give a clear explanation for the reason why they like that person.						
13	Describe a country you like best. You should say: • What that country is. • Where it is located. • What it is special for. • When you knew about that country. • And explain why you like it.	The test takers can speak in two minutes and respond to all given clues on the topic. They are encouraged to describe clearly the country they live in (what its name, where it is located, etc.). They are also expected to talk about the specialties of the country. In addition, they should be able to mention briefly the occasion they happened to know the country. Importantly, it is required that they can give a clear explanation for why they like the country.	Country	2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
14	Describe an exercise you are doing for fitness. You should say: • What it is. • How often do you do it. • When you first tried it. • What kind of people it is suitable for. • And explain why you think it is good.	The test takers can speak in two minutes and respond to all given clues on the topic. They are encouraged to describe clearly a sports game they are doing for fitness (what its name, what it is about, etc.) and how frequently they do the game. They are also expected to tell a short story about the occasion they came to it. In addition, they should be able to describe briefly the target players of the game. Importantly, it is required that they can give a clear explanation for the reason why they like the game.	Exercise	2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
15	Describe a festival you have ever joined in your country.	The test takers can speak in two minutes and respond to all given clues on the topic. They are encouraged to describe clearly a festival they have joined (its name, what it is about, etc.). They	Festival	2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Speaking test pack								
The speaking test pack is to provide questions and topics for evaluating the proficiency level of the test takers when they take the pre-speaking and post-speaking tests. It is worth to note that the test pack follows the IELTS speaking test structure.								
No.	Questions & Topics	Specification	Topic	Speaking part	Experts' Opinions			Remarks
					-1	0	+1	
	You should say: • What it is. • When it is celebrated. • What it is special for. • How people celebrate the festival. • And explain why you like it.	are expected to tell when the festival takes place and what is special. In addition, they should be able to describe briefly how people enjoy the festival. Importantly, it is required that they can give a clear explanation for the reason why they like it.						
16	What kind of books do you think are good for reading?	The test takers are expected to be able to talk about <i>what types of books are meaningful for reading, what is good or bad side</i> , and if possible, <i>sharing their thoughts</i> about the criteria for selecting a good book.	Book	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
17	Is reading a book more interesting than watching a movie?	The test takers are expected to be able to tell <i>whether reading a book outweighs watching a movie, why it is interesting or boring, what example can be drawn</i> , and if possible, <i>share some of their experiences</i> about reading a book or watching a movie.	Book	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
18	What are the key factors that make a person famous?	The test takers are expected to be able to talk about <i>what contributes to one's reputation, how it happens</i> , and if possible, <i>give an example</i> for their beliefs.	People	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
19	Are famous people always happy?	The test takers are expected to be able to tell <i>whether famous people always feel happy with their life, why they are happy or unhappy, what example can be drawn</i> , and if possible, <i>share their thoughts</i> about being a famous person.	People	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
20	Where is the best place for living?	The test takers are expected to be able to talk about <i>what place is the best for living, what contributes to the decision</i> , and if possible, <i>share their imagination</i> about an ideal place for living.	Country	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
21	Should people study and work inside their country or abroad?	The test takers are expected to be able to tell <i>whether people should study and work around their country or go abroad, why they should stay in their country or go abroad, what example can be drawn</i> , and if possible, <i>share their thoughts</i> about themselves studying and working inside their country or going abroad.	Country	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
22	Should people do exercise at	The test takers are expected to be able to tell <i>whether people should do exercises at home or</i>	Exercise	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Speaking test pack								
<i>The speaking test pack is to provide questions and topics for evaluating the proficiency level of the test takers when they take the pre-speaking and post-speaking tests. It is worth to note that the test pack follows the IELTS speaking test structure.</i>								
No.	Questions & Topics	Specification	Topic	Speaking part	Experts' Opinions			Remarks
					-1	0	+1	
	home or in a gym?	<i>in a gym, why they should do there and if possible, share their thoughts about doing exercises at home or in a gym.</i>						
23	What do people usually do to keep fit?	The test takers are expected to be able to talk about <i>what young people often do to stay healthy, what benefits should be considered, and if possible, share their thoughts about keeping fit at a young age.</i>	Exercise	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
24	Do people have enough time for festivals in this modern life?	The test takers are expected to be able to tell <i>whether people still have time to enjoy festivals nowadays, why it is so, what example can be drawn, and if possible, share their thoughts about enjoying festivals.</i>	Festival	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
25	How can festivals create income for your country?	The test takers are expected to be able to talk about <i>how festivals create income for their country, and if possible, giving an example for the explanation.</i>	Festival	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	



APPENDIX K

THE IOC VALUE FOR THE SPEAKING ASSESSMENT RUBRIC

To evaluate test takers English speaking proficiency based on their performance in English speaking in terms of accuracy and fluency

This is aimed at measuring validity of the speaking rubric to evaluate the proficiency level of the test takers when they take the pre-speaking and post-speaking tests.

Directions: Please refer to each statement on the evaluation form and then make a tick (✓) in the rating box that well describes your opinions on each statement. The rating criteria are described as follows:

Description:

+1: The statement is appropriate and relevant to the researchers objectives.

0: The statement sounds irrelevant to the researchers objectives.

-1: The statement is inappropriate and irrelevant to the researchers objectives.

Speaking rubric							
<i>The speaking rubric is to evaluate the proficiency level of the test takers when they take the pre-speaking and post-speaking tests.</i>							
No.	Statement	Aspects of English speaking performance	Test taker's proficiency level	Experts' Opinions			Remarks
				-1	0	+1	
1	The test taker can only use few words and expressions or the test takers' vocabulary knowledge is inadequate. Ex: Q: What do you study? A: English.	Accuracy in terms of vocabulary	Insufficient performance <i>(Insufficient user)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2	The test taker can only use basic or simple vocabulary and expressions. Sometimes, the test taker fails to respond properly due to inadequate vocabulary knowledge. Ex:	Accuracy in terms of vocabulary	Inadequate performance <i>(Basic user)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

No.	Statement	Aspects of English speaking performance	Test taker's proficiency level	Experts' Opinions			Remarks
				-1	0	+1	
	Q: What do you study? A: I'm study English.						
3	The test taker can use limited vocabulary and expressions. The test taker makes frequent errors in choosing words. The test taker does not try to use a wide range of words or expand vocabulary and expressions. Ex: Q: What do you study? A: I study English, my favorite English.	Accuracy in terms of vocabulary	Needs improvement (Basic user)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4	The test taker can use a wide range of vocabulary and expressions, and makes only a few errors in choosing words. Ex: Q: What do you study? A: I study English which is my favorite.	Accuracy in terms of vocabulary	Meets expectations (Independent user)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
5	The test taker can use appropriate expressions and a wide range of vocabulary. Ex: Q: What do you study? A: I study English. It is one of my favorite subjects I have learnt in my course.	Accuracy in terms of vocabulary	Exceeds expectations (Independent user)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6	The test taker fails to use appropriate sentence structures and cannot make proper word orders. Ex: Q: Do you prefer PCs or laptops? A: Yes/No/Laptops my prefer.	Accuracy in terms of grammar	Insufficient performance (Insufficient user)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7	The test taker can only use basic structures and makes frequent errors.	Accuracy in terms of grammar	Inadequate performance (Basic user)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

No.	Statement	Aspects of English speaking performance	Test taker's proficiency level	Experts' Opinions			Remarks
				-1	0	+1	
	Ex: Q: Do you prefer PCs or laptops? A: I prefer PCs./PCs is my prefer.						
8	The test taker can use varied structures with frequent errors, or basic structures with only a few errors. Ex: Q: Do you prefer PCs or laptops? A: I prefer PCs, not laptops.	Accuracy in terms of grammar	Needs improvement <i>(Basic user)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
9	The test taker can use varied sentence structures but still makes some errors. Ex: Q: Do you prefer PCs or laptops? A: I have both PCs and laptops at home but I prefer laptops more than PCs.	Accuracy in terms of grammar	Meets expectations <i>(Independent user)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
10	The test taker can use a wide range of structures based on contexts with only a few grammatical errors. Ex: Q: Do you prefer PCs or laptops? A: I think I prefer laptops to PCs because I can take the laptops to everywhere but PCs, no.	Accuracy in terms of grammar	Exceeds expectations <i>(Independent user)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
11	The test taker cannot pronounce words correctly for the examiner to understand what the test taker says. Ex: Q: How many students are there in your class? A: Their... trenty... tudent.	Accuracy in terms of pronunciation	Insufficient performance <i>(Insufficient user)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
12	The test taker faces frequent problems with pronunciation and intonation. The test taker speaks too quietly to hear and too hard to understand.	Accuracy in terms of pronunciation	Inadequate performance <i>(Basic user)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

No.	Statement	Aspects of English speaking performance	Test taker's proficiency level	Experts' Opinions			Remarks
				-1	0	+1	
	Ex: Q: How many students are there in your class? A: Their... there'e... there trenty... twenty student.						
13	The test taker makes some errors in pronunciation, rhythm, and intonation which make it difficult to understand. Ex: Q: How many students are there in your class? A: There trenty student.	Accuracy in terms of pronunciation	Needs improvement (Basic user)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
14	The test taker can almost make clear and accurate pronunciation, rhythm, and intonation but only occasionally difficult to understand. Ex: Q: How many students are there in your class? A: There're trenty students.	Accuracy in terms of pronunciation	Meets expectations (Independent user)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
15	The test taker can almost always make clear and accurate pronunciation, rhythm, and intonation. Ex: Q: How many students are there in your class? A: There're twenty students in my class.	Accuracy in terms of pronunciation	Exceeds expectations (Independent user)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
16	The test taker speaks very little or speaks nothing. Ex: Q: How do you go to school? A: Uh... uhm...	Fluency in terms of hesitation, pauses, and self-corrections	Insufficient performance (Insufficient user)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
17	The test taker can speak with much hesitation interfering with communication.	Fluency in terms of hesitation, pauses, and self-corrections	Inadequate performance (Basic user)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

No.	Statement	Aspects of English speaking performance	Test taker's proficiency level	Experts' Opinions			Remarks
				-1	0	+1	
	Ex: Q: How do you go to school? A: Uh... I... uh... go to school... uhm... by bus.						
18	The test taker can speak with some hesitation, which sometimes interferes with communication. Q: How do you go to school? A: Uh... I go to school... uhm... by bus.	Fluency in terms of hesitation, pauses, and self-corrections	Needs improvement <i>(Basic user)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
19	The test taker can speak with some hesitation without usually interrupting the conversation flow. Q: How do you go to school? A: Uh... I go to school by bus, yeah.	Fluency in terms of hesitation, pauses, and self-corrections	Meets expectations <i>(Independent user)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
20	The test taker can speak smoothly with confidence but little hesitation, without interrupting the conversation flow. Q: How do you go to school? A: I go to school by bus.	Fluency in terms of hesitation, pauses, and self-corrections	Exceeds expectations <i>(Independent user)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

APPENDIX L

THE RATING SCALE FOR THE LESSON PLANS

To evaluate the adequacy of the English speaking lessons used to teach Vietnamese undergraduate EFL students.

This is aimed at rating the adequacy of the lesson plans for teaching English speaking to Vietnamese undergraduate EFL students to see if the lesson plans can meet with the teaching and learning objectives. The definitions of the key terms related to this research are also used in this evaluation form.

The artificial intelligence voice chatbot or the AI voice chatbot in this study is the program integrated with artificial intelligence algorithms and can talk with users through the text-to-speech and voice recognition processes operating under internet connection. According to Kaplan (2016) and Lemma (2018), an AI voice chatbot is a kind of program or app which works under artificial intelligence algorithms like machine learning and natural language processing operated with internet connection and can have a chat with users via a mobile. In this study, Andy English Bot, called Andy, is an AI voice chatbot used thorough the experiment as it was reported by Kim et al. (2019) that it could chat with users in form of text chat and voice chat and can help users learn and practice English conversation, vocabulary, and grammar.

The Vietnamese undergraduate EFL students in the study means the non-English majored undergraduate students whose English level is equivalent to A2 level in CEFR standard, classified by a TOEIC Bridge test. Those students are in the first year or second year and learn English 2 which is one of the three modules of General English taught by Faculty of Foreign Languages at Can Tho University, Vietnam. They come from various schools and colleges in Can Tho University such as College of Natural Sciences, School of Education, College of Technology, College of Agriculture, and more. Accordingly, Vietnamese undergraduate EFL students are referred by those who major in non-English academic programs such as Chemistry Education, Agriculture, Primary Education, Computer Network, and so on.

Traditional classrooms mentioned in this study refers to the English classrooms in Can Tho University where students study on textbooks and follow teachers guiding in the classrooms. Traditional classrooms are conducted in a face-to-face manner and English language skills are taught by a teacher using oral instruction, handouts, or electronic media such as CDs, MP3 files, or presentation slides throughout the lesson.

English speaking skill in this study refers to the students' competence of using accurately and fluently spoken language to communicate during a conversation in English as a foreign language, from which the hearer can understand the speaker's message. This key term is based on Dao (2017) mentioning that English speaking skill is achieved when the speaker can successfully convey his meaning to the listener during a conversation by speaking the language fluently and accurately.

Fluency is defined in this study as one of the distinctive criteria demonstrating learners' proficient speaking skill. Fluency is the flowing or the natural language produced by the EFL students when they speak English, demonstrated through speaker's hesitations and pauses. Brown and Lee (2015) explained that fluency in speaking is best achieved when the speech flow is performed smoothly. In addition, Wang (2014) determined hesitations and pauses in speaking as the main obstructions to speaking fluency.

Accuracy in this study is defined as the EFL students' correct use of pronunciation, vocabulary, and grammar when they speak English. According to Brown and Lee (2015), accuracy focuses on the accurate use of phonology, grammar, morphosyntax, and discourse in the spoken output. Wang (2014) also stated that speaking accuracy involves the correct use of pronunciation, vocabulary, and grammar in English speaking. Together with fluency accuracy is seen in this study as another distinctive criterion to demonstrate learners' proficient speaking skill.

Directions: Please refer to each statement on the evaluation form and then make a tick (✓) in either 1, 2, 3, 4, or 5 in the rating box that well describes your opinions on each statement. The rating criteria are described as follows:

Description:

- 1: Strongly disagree.
- 2: Disagree.
- 3: Neutral.
- 4: Agree.
- 5: Strongly agree.

Items responses							
<i>The lessons plans are to teach English speaking to the EFL students in English classrooms in Vietnam by providing the students with necessary knowledge of English speaking through teaching, learning, and practicing activities.</i>							
No.	Statement	Experts' Opinions					Remarks
		1	2	3	4	5	
1	The lesson plans contain knowledge of what students should learn	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2	The lesson plans provide knowledge of subject matter.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3	The lesson plans use instructional design principles.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4	The lesson plans use appropriate assessment to plan and improve teaching.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
5	The activities are designed to support academic growth	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6	The activities are designed to use modern media/technology	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7	The lesson plans enable students to engage more in the activities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
8	The activities are designed to develop students' communication skills.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

APPENDIX M
THE EXPERTS' EVALUATION ON THE INSTRUMENTS

M1. The Experts' Evaluation on The Questionnaire

Part 1: General information							
<i>The general information is to investigate the demographic of the participants regarding their age, major, year of study, English placement test score, and technological background in the undergraduate program so that the researchers can have an overview of the participants' background.</i>							
No.	Description	Expert's Evaluation			Remarks		
		1	2	3			
1	Gender <input type="checkbox"/> Male <input type="checkbox"/> Female	+1	+1	+1			
2	Age <input type="checkbox"/> 18 <input type="checkbox"/> 19 <input type="checkbox"/> 20 <input type="checkbox"/> 21 <input type="checkbox"/> 22 <input type="checkbox"/> 23	+1	+1	+1			
3	Major <input type="checkbox"/> Medicinal Chemistry <input type="checkbox"/> Primary Education <input type="checkbox"/> Other (please specify):	+1	+1	+1			
4	Year of study <input type="checkbox"/> 1 st year <input type="checkbox"/> 2 nd year <input type="checkbox"/> 3 rd year <input type="checkbox"/> 4 th year	+1	+1	+1			
5	English placement test score	+1	+1	+1			
6	What technology have you ever used for learning English speaking? <i>You can select more than one option listed in this section.</i> <input type="checkbox"/> Podcast <input type="checkbox"/> Voice chat apps (i.e. Zoom, Zalo, LINE, etc.) <input type="checkbox"/> Chatbots (i.e. Andy English Bot, ELIZA, Siri, Google Assistant, etc.) <input type="checkbox"/> Learning websites (i.e. SpeechAce.com, IELTS Liz.com, etc.) <input type="checkbox"/> Youtube <input type="checkbox"/> Social network (i.e. Facebook, Twitter, Reddit, etc.) <input type="checkbox"/> Others (Please specify):	+1	+1	+1			
Part 2: Opinions on English speaking performance							
<i>The questionnaire items are to investigate the participants' opinions on English speaking skill in terms of fluency and accuracy AFTER they take the experiment speaking with an AI voice chatbot.</i>							
No.	Statement	Use the same Likert's scale	Students' opinions on English speaking	Expert's Evaluation			Remarks
				1	2	3	
7	After practicing with the AI voice chatbot, I think I could speak English without making too many pauses and hesitations.	1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; (5) Strongly agree	Fluency	+1	+1	+1	

8	After practicing with the AI voice chatbot, I think I could speak English with appropriate hedging words such as uhm... ah... oh....	1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; (5) Strongly agree	Fluency	+1	+1	+1	
9	After practicing with the AI voice chatbot, I think I could produce correct pronunciation when I spoke English.	1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; (5) Strongly agree	Accuracy	+1	+1	+1	
10	After practicing with the AI voice chatbot, I think I could use appropriate sentence structures when I spoke English.	1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; (5) Strongly agree	Accuracy	+1	+1	+1	
11	After practicing with the AI voice chatbot, I think I could use appropriate words and vocabulary when I spoke English.	1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; (5) Strongly agree	Accuracy	+1	+1	+1	
No.	Description	Expert's Evaluation			Remarks		
		1	2	3			
	Part 3: Additional suggestions or comments on learning English speaking skill by using an artificial intelligence voice chatbot (if any):		+1	+1	+1		

M2. The Experts Evaluation on The Interview Questions

Interview questions						
The section is to investigate the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL students through exploring the participants' opinions on their improvement in English speaking performance in terms of fluency, accuracy, and interaction AFTER they spoke with the AI voice chatbot.						
No.	Statement	Students' opinions on English speaking	Expert's Evaluation			Remarks
			1	2	3	
1	How often and how long did you practice speaking with the AI voice chatbot every week?	General process	+1	+1	+1	
2	Do you think that your grammar has been improved after speaking with the AI voice chatbot? Why do you think so?	Accuracy	+1	+1	+1	
3	Do you think that your pronunciation has been improved after speaking with the AI voice chatbot? Why do you think so?	Accuracy	+1	+1	+1	
4	Do you think that your vocabulary has been improved after speaking with the AI voice chatbot? Why do you think so?	Accuracy	+1	+1	+1	

5	Do you think that you can speak with fewer hesitations or pauses after speaking with the AI voice chatbot? Why do you think so?	Fluency	+1	+1	+1	
6	Do you think that you can speak better after speaking with the AI voice chatbot? Why do you think so?	Fluency	+1	+1	+1	
7	What do you like most about speaking with the AI voice chatbot? And what do you dislike?	Feeling	+1	+1	+1	

M3. The Experts Evaluation on The Speaking Test Pack

Speaking test pack								
The speaking test pack is to provide questions and topics for evaluating the proficiency level of the test takers when they take the pre-speaking and post-speaking tests. It is worth to note that the test pack follows the IELTS speaking test structure.								
No.	Questions & Topics	Specification	Topic	Speaking part	Expert's Evaluation			Remarks
					1	2	3	
1	Do you prefer street food or homecooked food? Why	The test takers are expected to be able to talk about <i>whether they like street food or homecooked food</i> and <i>why they like it</i> .	Food	1	+1	+1	+1	
2	Where do you usually go to buy food?	The test takers are expected to be able to talk about <i>where they usually go to buy food</i> , and if possible, <i>share special things</i> about food in that place.	Food	1	+1	+1	+1	
3	What do you study?	The test takers are expected to be able to talk about <i>what major, subject, or program they learn or have learnt</i> , <i>what the major, subject, or program is about</i> , and if possible, <i>share their thoughts</i> about the major, subject, or program.	Study	1	+1	+1	+1	
4	Where do you study?	The test takers are expected to be able to talk about <i>where they study</i> , <i>which school, university, or college they study at</i> , and if possible, <i>share special things</i> about their school, university, or college.	Study	1	+1	+1	+1	
5	Where is your hometown?	The test takers are expected to be able to talk about <i>where they were born and grew up</i> , <i>where it is located</i> , and if possible, <i>sharing special things</i> about their hometown.	Hometown	1	+1	+1	+1	
6	How often do you visit your hometown?	The test takers are expected to be able to talk about <i>how frequently they visit their hometown</i> , <i>when they visit</i> , and if possible, <i>share some reasons for their visit</i> .	Hometown	1	+1	+1	+1	
7	What do you	The test takers are expected to be able to list	Daily	1	+1	+1	+1	

	usually do every day?	the activities they do every day, and if possible, explain <i>why they have that habit</i> .	activities					
8	Do you prefer outdoor or indoor activities? Why?	The test takers are expected to be able to tell <i>whether like to do outdoor or indoor activities</i> , and if possible, explain <i>why they think so</i> .	Daily activities	1	+1	+1	+1	
9	Do you like to spend time with your family or friends? Why?	The test takers are expected to be able to tell <i>whether they like to be with their family or friends</i> and if possible, explain <i>why they believe so</i> .	Family and friends	1	+1	+1	+1	
10	What do you usually do with your friends?	The test takers are expected to be able to list the activities they do with friends, and if possible, <i>share their thoughts</i> about those activities.	Family and friends	1	+1	+1	+1	
11	Describe a book you have recently read. You should say: • What kind of book it is. • What it is about. • Where you found it. • What sort of people would enjoy it. • And explain why you like it.	The test takers can speak in two minutes and respond to all given clues on the topic. They are encouraged to describe clearly the book they have just read (what it is, what it is special for, etc.). They are also expected to tell a little about the book's overview and where or from whom they got it. In addition, they should be able to discuss briefly the target audience of the book. Importantly, it is required that they can give a clear explanation for the reason why they like the book.	Book	2	+1	+1	+1	
12	Describe a person you like best. You should say: • Who he/she is. • When you met that person. • What he/she looks like. • How long you have known him/her. • And explain why you like that person.	The test takers can speak in two minutes and respond to all given clues on the topic. They are encouraged to describe clearly about a person they like best (who he/she is, what he/she is special for, etc.). They are also expected to tell a short story about the occasion they met that person and how long they have known that one. In addition, they should be able to describe briefly the person's appearance and personality. Importantly, it is required that they can give a clear explanation for the reason why they like that person.	People	2	+1	+1	+1	
13	Describe a country you like	The test takers can speak in two minutes and respond to all given clues on the topic. They	Country	2	+1	+1	+1	

	<p>best.</p> <p>You should say:</p> <ul style="list-style-type: none"> • What that country is. • Where it is located. • What it is special for. • When you knew about that country. • And explain why you like it. 	<p>are encouraged to describe clearly the country they live in (what its name, where it is located, etc.). They are also expected to talk about the specialties of the country. In addition, they should be able to mention briefly the occasion they happened to know the country. Importantly, it is required that they can give a clear explanation for why they like the country.</p>					
14	<p>Describe an exercise you are doing for fitness.</p> <p>You should say:</p> <ul style="list-style-type: none"> • What it is. • How often do you do it. • When you first tried it. • What kind of people it is suitable for. • And explain why you think it is good. 	<p>The test takers can speak in two minutes and respond to all given clues on the topic. They are encouraged to describe clearly a sports game they are doing for fitness (what its name, what it is about, etc.) and how frequently they do the game. They are also expected to tell a short story about the occasion they came to it. In addition, they should be able to describe briefly the target players of the game. Importantly, it is required that they can give a clear explanation for the reason why they like the game.</p>	Exercise	2	+1	+1	+1
15	<p>Describe a festival you have ever joined in your country.</p> <p>You should say:</p> <ul style="list-style-type: none"> • What it is. • When it is celebrated. • What it is special for. • How people celebrate the festival. • And explain why you like it. 	<p>The test takers can speak in two minutes and respond to all given clues on the topic. They are encouraged to describe clearly a festival they have joined (its name, what it is about, etc.). They are expected to tell when the festival takes place and what is special. In addition, they should be able to describe briefly how people enjoy the festival. Importantly, it is required that they can give a clear explanation for the reason why they like it.</p>	Festival	2	+1	+1	+1
16	<p>What kind of books do you</p>	<p>The test takers are expected to be able to talk about <i>what types of books are</i></p>	Book	3	+1	+1	+1

	think are good for reading?	<i>meaningful for reading, what is good or bad side, and if possible, sharing their thoughts about the criteria for selecting a good book.</i>						
17	Is reading a book more interesting than watching a movie?	The test takers are expected to be able to tell <i>whether reading a book outweighs watching a movie, why it is interesting or boring, what example can be drawn, and if possible, share some of their experiences about reading a book or watching a movie.</i>	Book	3	+1	+1	+1	
18	What are the key factors that make a person famous?	The test takers are expected to be able to talk about <i>what contributes to one's reputation, how it happens, and if possible, give an example for their beliefs.</i>	People	3	+1	+1	+1	
19	Are famous people always happy?	The test takers are expected to be able to tell <i>whether famous people always feel happy with their life, why they are happy or unhappy, what example can be drawn, and if possible, share their thoughts about being a famous person.</i>	People	3	+1	+1	+1	
20	Where is the best place for living?	The test takers are expected to be able to talk about <i>what place is the best for living, what contributes to the decision, and if possible, share their imagination about an ideal place for living.</i>	Country	3	+1	+1	+1	
21	Should people study and work inside their country or abroad?	The test takers are expected to be able to tell <i>whether people should study and work around their country or go abroad, why they should stay in their country or go abroad, what example can be drawn, and if possible, share their thoughts about themselves studying and working inside their country or going abroad.</i>	Country	3	+1	+1	+1	
22	Should people do exercise at home or in a gym?	The test takers are expected to be able to tell <i>whether people should do exercises at home or in a gym, why they should do there and if possible, share their thoughts about doing exercises at home or in a gym.</i>	Exercise	3	+1	+1	+1	
23	What do people usually do to keep fit?	The test takers are expected to be able to talk about <i>what young people often do to stay healthy, what benefits should be considered, and if possible, share their thoughts about keeping fit at a young age.</i>	Exercise	3	+1	+1	+1	
24	Do people have enough time for festivals in this modern life?	The test takers are expected to be able to tell <i>whether people still have time to enjoy festivals nowadays, why it is so, what example can be drawn, and if possible, share their thoughts about enjoying festivals.</i>	Festival	3	+1	+1	+1	
25	How can	The test takers are expected to be able to	Festival	3	+1	+1	+1	

festivals create income for your country?	talk about <i>how festivals create income for their country</i> , and if possible, <i>giving an example</i> for the explanation.							
---	--	--	--	--	--	--	--	--

M4. The Expert s Evaluation on The Speaking Assessment Rubric

Speaking rubric							
<i>The speaking rubric is to evaluate the proficiency level of the test takers when they take the pre-speaking and post-speaking tests.</i>							
No.	Statement	Aspects of English speaking performance	Test taker's proficiency level	Expert's Evaluation			Remarks
				1	2	3	
1	The test taker can only use few words and expressions or the test takers' vocabulary knowledge is inadequate. Ex: Q: What do you study? A: English.	Accuracy in terms of vocabulary	Insufficient performance <i>(Insufficient user)</i>	+1	+1	0	Expert 3: "A few words" not "few words"
2	The test taker can only use basic or simple vocabulary and expressions. Sometimes, the test taker fails to respond properly due to inadequate vocabulary knowledge. Ex: Q: What do you study? A: I'm study English.	Accuracy in terms of vocabulary	Inadequate performance <i>(Basic user)</i>	+1	+1	+1	
3	The test taker can use limited vocabulary and expressions. The test taker makes frequent errors in choosing words. The test taker does not try to use a wide range of words or expand vocabulary and expressions. Ex: Q: What do you study? A: I study English, my favorite English.	Accuracy in terms of vocabulary	Needs improvement <i>(Basic user)</i>	+1	+1	+1	
4	The test taker can use a wide range of vocabulary and expressions, and makes only a few errors in choosing words. Ex:	Accuracy in terms of vocabulary	Meets expectations <i>(Independent user)</i>	+1	+1	+1	

	Q: What do you study? A: I study English which is my favorite.					
5	The test taker can use appropriate expressions and a wide range of vocabulary. Ex: Q: What do you study? A: I study English. It is one of my favorite subjects I have learnt in my course.	Accuracy in terms of vocabulary	Exceeds expectations (<i>Independent user</i>)	+1	+1	+1
6	The test taker fails to use appropriate sentence structures and cannot make proper word orders. Ex: Q: Do you prefer PCs or laptops? A: Yes/No/Laptops my prefer.	Accuracy in terms of grammar	Insufficient performance (<i>Insufficient user</i>)	+1	+1	+1
7	The test taker can only use basic structures and makes frequent errors. Ex: Q: Do you prefer PCs or laptops? A: I prefer PCs./PCs is my prefer.	Accuracy in terms of grammar	Inadequate performance (<i>Basic user</i>)	+1	+1	+1
8	The test taker can use varied structures with frequent errors, or basic structures with only a few errors. Ex: Q: Do you prefer PCs or laptops? A: I prefer PCs, not laptops.	Accuracy in terms of grammar	Needs improvement (<i>Basic user</i>)	+1	+1	+1
9	The test taker can use varied sentence structures but still makes some errors. Ex: Q: Do you prefer PCs or laptops? A: I have both PCs and laptops at home but I prefer laptops more than PCs.	Accuracy in terms of grammar	Meets expectations (<i>Independent user</i>)	+1	+1	+1
10	The test taker can use a wide range of structures based on contexts with only a few grammatical errors. Ex: Q: Do you prefer PCs or laptops? A: I think I prefer laptops to PCs because I can take the laptops to everywhere but PCs, no.	Accuracy in terms of grammar	Exceeds expectations (<i>Independent user</i>)	+1	+1	+1

11	<p>The test taker cannot pronounce words correctly for the examiner to understand what the test taker says.</p> <p>Ex: Q: How many students are there in your class? A: Their... trenty... tudent.</p>	Accuracy in terms of pronunciation	Insufficient performance (<i>Insufficient user</i>)	+1	+1	+1	
12	<p>The test taker faces frequent problems with pronunciation and intonation.</p> <p>The test taker speaks too quietly to hear and too hard to understand.</p> <p>Ex: Q: How many students are there in your class? A: Their... there'e... there trenty... twenty student.</p>	Accuracy in terms of pronunciation	Inadequate performance (<i>Basic user</i>)	+1	+1	+1	
13	<p>The test taker makes some errors in pronunciation, rhythm, and intonation which make it difficult to understand.</p> <p>Ex: Q: How many students are there in your class? A: There trenty student.</p>	Accuracy in terms of pronunciation	Needs improvement (<i>Basic user</i>)	+1	+1	0	Expert 3: "The test taker makes some pronunciation, rhythm, and intonation errors, making it difficult to understand".
14	<p>The test taker can almost make clear and accurate pronunciation, rhythm, and intonation but only occasionally difficult to understand.</p> <p>Ex: Q: How many students are there in your class? A: There're trenty students.</p>	Accuracy in terms of pronunciation	Meets expectations (<i>Independent user</i>)	+1	+1	+1	
15	<p>The test taker can almost always make clear and accurate pronunciation, rhythm, and intonation.</p> <p>Ex: Q: How many students are there in your class? A: There're twenty students in my class.</p>	Accuracy in terms of pronunciation	Exceeds expectations (<i>Independent user</i>)	+1	+1	+1	
16	<p>The test taker speaks very little or speaks nothing.</p>	Fluency in terms of hesitation, pauses,	Insufficient performance	+1	+1	+1	

	Ex: Q: How do you go to school? A: Uh... uhm...	and self-corrections	<i>(Insufficient user)</i>				
17	The test taker can speak with much hesitation interfering with communication. Ex: Q: How do you go to school? A: Uh... I... uh... go to school... uhm... by bus.	Fluency in terms of hesitation, pauses, and self-corrections	Inadequate performance <i>(Basic user)</i>	+1	+1	+1	
18	The test taker can speak with some hesitation, which sometimes interferes with communication. Q: How do you go to school? A: Uh... I go to school... uhm... by bus.	Fluency in terms of hesitation, pauses, and self-corrections	Needs improvement <i>(Basic user)</i>	+1	+1	+1	
19	The test taker can speak with some hesitation without usually interrupting the conversation flow. Q: How do you go to school? A: Uh... I go to school by bus, yeah.	Fluency in terms of hesitation, pauses, and self-corrections	Meets expectations <i>(Independent user)</i>	+1	+1	+1	
20	The test taker can speak smoothly with confidence but little hesitation, without interrupting the conversation flow. Q: How do you go to school? A: I go to school by bus.	Fluency in terms of hesitation, pauses, and self-corrections	Exceeds expectations <i>(Independent user)</i>	+1	+1	0	Expert 3: "But with little..." Expert 3: <i>To me evaluating 1-20 may be difficult since some are quite close.</i> <i>Why don't you have bands 1-5 and X4 to get 20 marks?</i>

M5. The Expert s Evaluation on The Lesson Plans

Items responses					
<i>The lessons plans are to teach English speaking to the EFL students in English classrooms in Vietnam by providing the students with necessary knowledge of English speaking through teaching, learning, and practicing activities.</i>					
No.	Statement	Expert's Opinions			Remarks
		1	2	3	
1	The lesson plans contain knowledge of what students should learn	4	5	5	
2	The lesson plans provide knowledge of subject matter.	4	4	5	
3	The lesson plans use instructional design principles.	5	5	5	
4	The lesson plans use appropriate assessment to plan and improve teaching.	4	4	5	
5	The activities are designed to support academic growth	5	4	5	
6	The activities are designed to use modern media/technology	4	5	5	
7	The lesson plans enable students to engage more in the activities.	5	5	5	
8	The activities are designed to develop students' communication skills.	5	5	5	



APPENDIX N
INSTRUMENTAL EVALUATION RESULTS

After the evaluation process, the validity of each item was reported and ed in detail with the following order:

a) Validity of The Questionnaire Items

Table N1. Validity of The Questionnaire Items

Evaluation report on the questionnaire items				
Item No.	Exp01	Exp02	Exp03	AVR
1	1	1	1	1.00
2	1	1	1	1.00
3	1	1	1	1.00
4	1	1	1	1.00
5	1	1	1	1.00
6	1	1	1	1.00
7	1	1	1	1.00
8	1	1	1	1.00
9	1	1	1	1.00
10	1	1	1	1.00
11	1	1	1	1.00

b) Validity of The Semi-Structured Interview Questions

Table N2. Validity of The Semi-Structured Interview Questions

Evaluation report on the interview questions				
Item No.	Exp01	Exp02	Exp03	AVR
1	1	1	1	1.00
2	1	1	1	1.00
3	1	1	1	1.00
4	1	1	1	1.00
5	1	1	1	1.00
6	1	1	1	1.00
7	1	1	1	1.00

c) Validity of The Speaking Test Pack

Table N3. Validity of The English Speaking Test Pack

Evaluation report on the English speaking test pack				
Item No.	Exp01	Exp02	Exp03	AVR
1	1	1	1	1.00
2	1	1	1	1.00
3	1	1	1	1.00
4	1	1	1	1.00
5	1	1	1	1.00
6	1	1	1	1.00
7	1	1	1	1.00
8	1	1	1	1.00
9	1	1	1	1.00
10	1	1	1	1.00
11	1	1	1	1.00
12	1	1	1	1.00
13	1	1	1	1.00
14	1	1	1	1.00
15	1	1	1	1.00
16	1	1	1	1.00
17	1	1	1	1.00
18	1	1	1	1.00
19	1	1	1	1.00
20	1	1	1	1.00
21	1	1	1	1.00
22	1	1	1	1.00
23	1	1	1	1.00
24	1	1	1	1.00
25	1	1	1	1.00

d) Validity of The Speaking Assessment Rubric

Table N4. Validity of The Speaking Assessment Rubric

Evaluation report on the speaking assessment rubric				
Item No.	Exp01	Exp02	Exp03	AVR
1	1	1	0	0.67
2	1	1	1	1.00
3	1	1	1	1.00
4	1	1	1	1.00
5	1	1	1	1.00
6	1	1	1	1.00
7	1	1	1	1.00
8	1	1	1	1.00
9	1	1	1	1.00
10	1	1	1	1.00
11	1	1	1	1.00
12	1	1	1	1.00
13	1	1	0	0.67
14	1	1	1	1.00
15	1	1	1	1.00
16	1	1	1	1.00
17	1	1	1	1.00
18	1	1	1	1.00
19	1	1	1	1.00
20	1	1	0	0.67
21	1	1	1	1.00
22	1	1	1	1.00
23	1	1	1	1.00
24	1	1	1	1.00
25	1	1	1	1.00

e) *Validity of the Lesson Plans*

Table N5. Validity of The Lesson Plans

Evaluation report on the lesson plans					
Statement No.	Exp01	Exp02	Exp03	AVR	%
1. The lesson plans contain knowledge of what students should learn	4	5	5	4.67	93.33
2. The lesson plans provide knowledge of subject matter	4	4	5	4.33	86.67
3. The lesson plans use instructional design principles	5	5	5	5.00	100.00
4. The lesson plans use appropriate assessment to plan and improve teaching	4	4	5	4.33	86.67
5. The activities are designed to support academic growth	5	4	5	4.67	93.33
6. The activities are designed to use modern media/technology	4	5	5	4.67	93.33
7. The lesson plans enable students to engage more in the activities	5	5	5	5.00	100.00
8. The activities are designed to develop students' communication skills	5	5	5	5.00	100.00

In summary, the questionnaire got the total score of +1 for each item, which meant that these items could clearly measure the objectives. Regarding the guided questions for the interview, all the experts evaluated +1 for each question to confirm that those questions could guide to the research objectives. For the speaking assessment rubric, the three experts agreed that it could meet the research objectives (+1). In addition, one out of the three experts suggested making the criteria and the scoring paradigm of the rubric simpler to make ease of assessment. Regarding this process, Appendix I led the detail of the experts' evaluation results. Finally, all the experts agreed (+4) and strongly agreed (+5) that the lesson plan met with teaching goals.

APPENDIX O
THE PILOT STUDY RESULTS

O1 Changes in The Speaking Questions

In Table O1, each item got the average score of 1.00, suggesting that those questions could be used to clearly measure the participants' English speaking skill. However, during the pre-speaking test, the researcher found that some questions caused difficulties to the participants when they tried to understand the meaning of the questions. Therefore, some questions in Part 1 and Part 3 of the speaking tests were revised a little bit to make them more friendly to the participants. Table O1 ed the changes made from the pre-speaking test.

Table O1. Changes in The Speaking Test Questions

PART 1		
Topic	Before	After
Food	Do you prefer street food or homecooked food? Why	Do you <u>like</u> street food or homecooked food? Why?
Study	What do you study?	What <u>major</u> do you study?
Daily activities	Do you prefer outdoor or indoor activities? Why?	Do you <u>like</u> outdoor or indoor activities? Why?
PART 3		
Topic	Before	After
People	What are the key factors that make a person famous?	What are <u>the factors</u> that make a person famous?
Country	Should people study and work inside their country or abroad?	Should people study and work inside or <u>outside</u> their country?
Exercise	What do people usually do to keep fit?	What do people usually do to <u>be healthy</u> ?
Festival	How can festivals create income for your country?	How can festivals <u>bring money to</u> your country?

O2 The Pilot Study

The pilot study took place within four weeks. Before the pilot started, the researchers had sent out the Information Sheet for Participants and Informed consent Form to the participants and ed about the information of the research and participant s rights so that the participants could know the necessary information of the research as well as their rights when joining the research. In the pre-speaking test, 29 students agreed to take the pre-speaking test.

However, after the pretest, five students requested to quit the pilot because they were too busy with reviewing lessons and preparing final projects of other subjects in the semester. As a result, only 24 students followed the pilot study to the end.

In the pre-speaking test, the participants were requested to speak in three parts. In Part 1, the researcher randomly selected a topic among the five and asked two questions belonging to the topic. The test takers listened to the questions and gave suitable answers. These topics were about food, study, hometown, daily activities, family and friends. In the Part 2, the test takers were requested to talk about a random topic selected by the researcher. Speaking clues were suggested in each topic so that the test takers could think about and get the ideas to speak. Five topics provided in Part 2 included book, people, country, exercise, and festival. In Part 3, the researcher asked two more questions related to the topic the test takers had been given in Part 2.

After the pre-speaking test, the participants were trained on how to use the AI voice chatbots to practice English speaking. For classroom teaching, the researcher selected two first units among the eight to teach English speaking to the participants in class. The first unit was about new friends (Unit 1) and the second unit was about interests (Unit 2). During the pilot time, the participants were taught English speaking in class but some days they had to study via Zoom because the region got flooded.

During the class time, the participants learnt English speaking from the website and practiced English speaking with Andy English Bot because this AI voice chatbot supported Android and iOS, which was easy for the students to access through their smartphones. To provide additional support for the participants to do warm-up activities and practice tasks, the researcher used Diagflow platform to develop two more AI voice chatbots which were named Elsa and Peter. Elsa was designed to talk with the students thorough the warm-up activities in each lesson while Peter was designed to talk with the students when they would like to further practice outside the classroom. While Andy English Bot supported Android and iOS, Elsa and Peter

supported Windows operation. Therefore, the students were encouraged to prioritize Andy English Bot as the first choice when they practiced English speaking to because most of the participants possessed smartphones rather than a laptop.

After class, the participants were encouraged to further practice English speaking with Andy English Bot, Elsa, and Peter to improve their speaking skill. Additionally, the researcher also recommended the participants to record audio files while practicing and take some screenshots from the chatting screen to submit to the researcher for following-up. However, not many of them agreed to do that because most of them had to spend time preparing for many term projects by the end of the semester and did not have enough time to submit the recording files and the screenshots. Even five participants requested to stop joining the pilot study so that they could have time to prepare for their homework in other subjects. Fortunately, apart from those busy participants, the researchers could get some audio recordings and screenshots from a few students to serve for data analysing purposes. Figure O1 ed a chat session between a participant and Andy English Bot when he practiced English speaking after class.



Figure O1. Participant s Chat Sessions with Andy English Bot in The Pilot Study

Generally, Andy English Bot could answer most of the questions asked by the student. In some cases, this AI voice chatbot failed to answer but it was intelligent enough to drive the student's question to another topic so that both Andy and that student could keep talking. For example, when the student asked "Why do you want to learn English?", Andy failed to directly answer for the reason why but it could drive the student to a new speaking topic which was about learning some new words. In this case, the talk would be maintained smoothly between the student and Andy. Figure O2 showed other chat session between another participant and Andy English Bot.



Figure O2. Participant's Chat Sessions with Andy English Bot in the Pilot Study

The conversation between the student and Andy English Bot in this figure looked more natural and plausible because they could ask and answer naturally and smoothly even though Andy English Bot seemed to make long answers and questions rather than the student. The interesting point was that Andy could suggest the student to learn new vocabulary by driving the talk to the vocabulary learning

session very easily. For this function, Andy may be a useful tool to help the students practice speaking while learning vocabulary. Figure O3 described another chat session conducted by one of the participants when he practiced speaking with Andy.

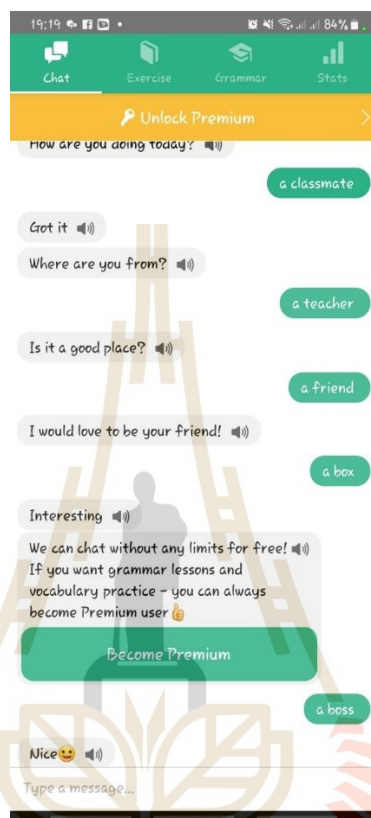


Figure O3. Participant's Chat Sessions with Andy English Bot in the Pilot Study

Based on the illustrative figure, it seemed that when the student only pronounced a word, Andy might check if the student pronounced correctly or incorrectly by responding “got it” or “interesting”. This function might help the student know whether he pronounced the word with accurate pronunciation or not. However, there were some words that Andy could not respond. For example, when the student pronounced the word “a teacher”, Andy responded by asking back “is it a good place?”. This ask back question revealed that Andy was not fully ready for its vocabulary bank. Moreover, because Andy English Bot was a commercial app, sometimes it popped up advertising messages throughout the chat sessions, which might annoy its user. For example, in the descriptive figure, the chat session was interrupted with the message “We can chat without any limits for free! If you want grammar lessons... become Premium user” which encouraged its user to upgrade the

app to Premium version. Certainly, the upgrading cost some money. Figure O4 illustrated another chat session between a student and Andy.



Figure O4. Participant's Chat Sessions with Andy English Bot in the Pilot Study

Similar to the issues discussed previously, sometimes Andy did not respond sufficiently to the student's questions although most of the speaking time, the communication took place quite naturally. For example, when asked "How long have you been learning English?", the student actually would like to mention the time Andy had been learning English. However, it responded "Yeah, especially when people are nice to me", which was not plausible enough.

Apart from Andy English Bot, Elsa and Peter were also used by the participants to practice English speaking. Figure O5 showed the chat sessions between a participant with Peter, an AI voice chatbot developed from Diagflow.

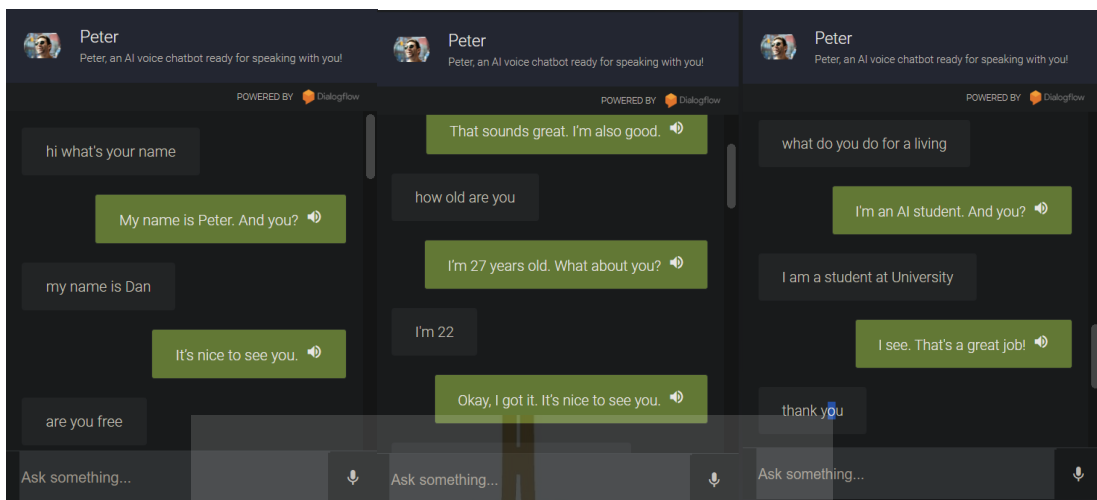


Figure O5. Participant's Chat Sessions with Peter in the Pilot Study

Peter seemed to understand the participant's saying pretty well in familiar topics like greetings and job. Furthermore, Peter could ask follow-up questions such as “and you?” or “what about you?” so that it could trigger a two-way discussion with the participant. The limitation of this AI voice chatbot was that it failed to display the question marks (?) or the full stop (.) after the participant finished a sentence or a question. This limitation was also found at Andy English Bot. However, Peter revealed its weakness when talking about other topics such as sports, movies, and time. In Figure O6, Peter failed to answer most of the questions and statements from another participant.

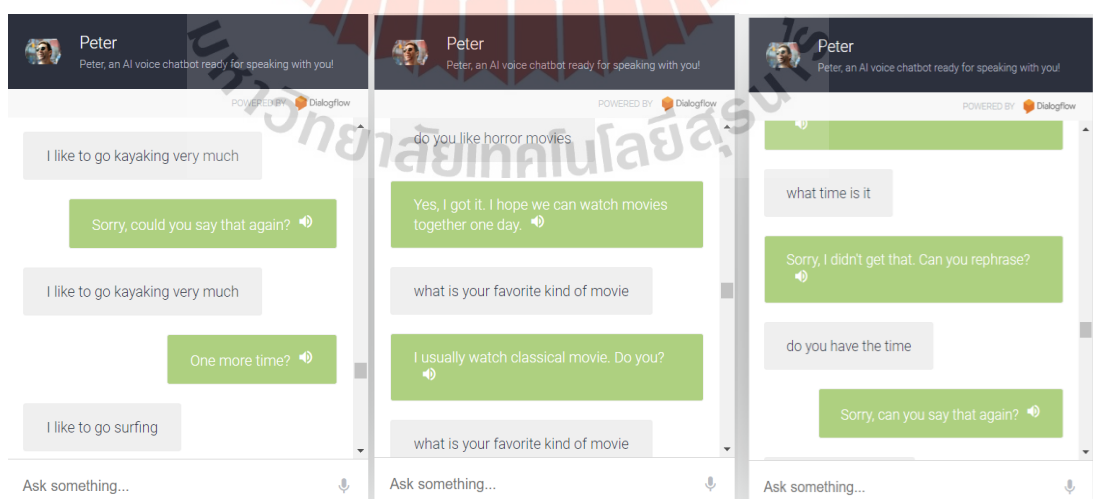


Figure O6. Participant's Chat Sessions with Peter in the Pilot Study

For example, when the participant said “I like to go kayaking very much.”, Peter failed to get the idea and made a fallback question: “Sorry, could you say that again?”. Even the participant repeated her previous statement, Peter still did not get it and asked again “One more time?”. This limitation was due to the limitation of the topics on which Peter had been trained. However, it could do better if it was trained with more topics and more discourses. Luckily, as discussed previously in the literature review, Google Dialogflow provided a platform which allowed a developer to create and keep training and improving the AI voice chatbots to make them understand more topics and discourses. Therefore, the limitations of the AI voice chatbots can be overcome in the future.

Another worth-mentioned AI voice chatbot was Elsa. This AI voice chatbot was additionally developed by the researcher to specifically work with the students during the warm-up activities throughout the lessons. As a result, instead of doing the warm-up activities with a partner, students could do it with Elsa. Figure O7 illustrated one of the warm-up sessions between a participant and Elsa.



Figure O7. Participant's Warm-Up Sessions with Elsa in the Pilot Study

This activity belonged to unit 1, period 1 on the website <https://aispeaks.mobirisesite.com/> in which the learner was requested to try to pronounce some sentences with correct pronunciation. During this warm-up activity, the participant had already pronounced some of the sentences such as “Hello, how are you?”, “How are you doing?”, and “How is it going?” with correct pronunciation.

Therefore, Elsa could recognize all sentences and respond appropriately. Within this aspect, Elsa seemed to be able to play a role of a partner who could do English speaking activities with learners along teaching and learning activities. Another demonstration for Elsa's ability was reported in Figure O8. The warm-up session belonged to unit 1, period 4 which requested the learner to practice talking with Elsa about how frequently he/she did the mentioned activities (e.g. listen to music, watch movies, get together with family, earn money, and play sports).

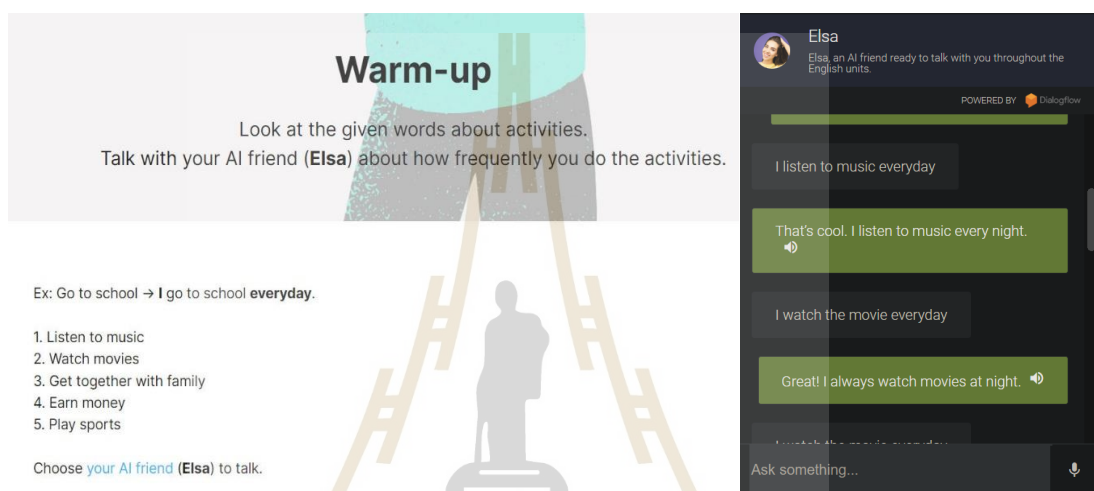


Figure O8. Participant s Warm-Up Sessions with Elsa in the pilot study

The screenshot was provided by one of the participants in which he did the warm-up activity with Elsa. Although he did not try to develop the adverbs of frequency to talk about how frequently he listened to music and how frequently he watched the movie, Elsa could help its user to find more adverbs of frequency by using “every night” or “always” in every of its responses. This might trigger the participant's use of a various adverbs of frequency in future talks. There was an interesting situation in which one of the participants tried to do warm-up activity in unit 7, period 2 which was not required during the pilot study. However, he just tried it and provided the research with the screenshot taken from the warm-up activity (Figure O9).

Let's try!

Try saying the following sentences to your AI friend (**Elsa**).

1. Butter.
2. Tomatoes.
3. Pepper.
4. Salt.
5. Flour.
6. Eggs.
7. Oil.
8. Onions.
9. Olives.
10. Beef.

*Please select an AI friend (**Elsa**) to check if you have spoken accurately as described in the picture on your right.*

Select an AI friend

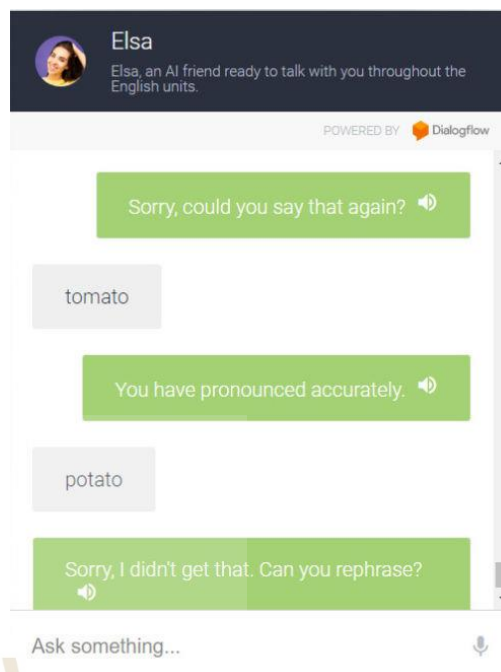


Figure O9. Participant's Warm-Up Sessions with Elsa in the pilot study

In this activity the learner was asked to practice pronounce the given words with correct pronunciation. It could be seen that whenever the participant pronounced a word among the given words correctly, Elsa could inform him that he had pronounced correctly. In case he failed to pronounced the word, Elsa might ask for his repetition. For example, when he pronounced the word “tomato” into “potato”, Elsa asked him to repeat. For this function, Elsa was able to help the participant with pronunciation when he spoke English. However, Elsa still revealed some weaknesses when speaking with the participants about the topics which were not contained in warm-up activities. In such cases, Elsa misrecognized participants' statements and questions. For example, in Figure O10, a participant talked with Elsa about her family and daily routine such as “My sister is funny”, “my sister is as funny as yours”, “I usually make up at the beauty salon”, or “I will buy it with \$130”, but Elsa provided irrelevant responses to those statement or frequently asked the participant to repeat and repeat.

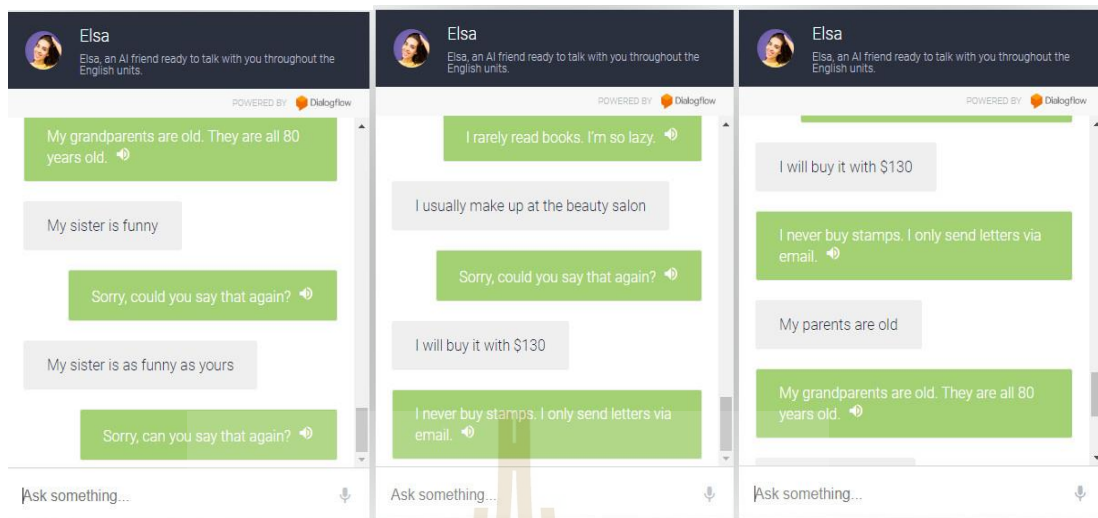


Figure O10. Participant's Warm-Up Sessions with Elsa in the pilot study

This was understandable because Elsa was not designed for purely chatting purposes. It was only designed to help learners to practice the contents which were available in the warm-up activities. In the final week of the pilot time, all participants took a post-speaking test. As mentioned previously, five out of twenty-nine students requested to leave the pilot study. Therefore, only twenty-four students took the post-speaking test. During the speaking test, the test takers were asked similar questions and given the same topic as in the pre-speaking test. All speaking test sessions including the pretest and posttest were recorded into audio files. However, only when the post-speaking test was done, all audio files from the pre-speaking and post-speaking tests were sent to three examiners for scoring purposes.

The researcher applied the double-blind rating method when sending the audio files to the examiners. During this process, the examiners were not told about the information of the participants, they were not informed of what was the pre-speaking test and what was the post-speaking test. All that the examiners were informed was the speaking assessment rubric and the topics which the participants were asked and given along with the deadline at which the scores were supposed to be provided.

Within the final week, the participants were also requested to answer the questionnaire items and finally took an interview with the researcher after they finished the post-speaking test.

O3 The Pilot Results

O3.1 The Pre-speaking and post-speaking Test Results

The pre-speaking and post-speaking tests were used as a measure to cross check with the results received from the questionnaire and from the interview. A paired samples t-test was applied to compare between the participants' scores from the pre-speaking and post-speaking tests to find out if the speaking results revealed statistically significant between the pre-speaking and post-speaking tests to confirm the hypothesis of this study. The researchers of this study used SPSS as a tool to analyse the research data collected from the pre-speaking and post-speaking test scores of the participants. Moreover, as there were three examiners taking part in the scoring process, their scores for each test taker in the pretest and posttest were collected and calculated in average to have the final score of each test taker in the pretest and posttest. Table O2 showed the mean scores of the participants in the pre-speaking and post-speaking tests after measured with the paired samples t-test. Particularly, the speaking tests evaluated the English speaking skill of 24 test takers based on two criteria which were fluency and accuracy. For each criterium, the mean scores of the participants were namely 2.333 (M=2.333) and 2.996 (M=2.996) in the pretest and posttest in terms of fluency. In terms of accuracy, the mean scores were respectively 2.621 (M=2.621) and 3.125 (M=3.125) in the pretest and posttest. As a result, the total mean scores of the participants in the pretest and posttest were namely 4.958 (M=4.958) and 6.121 (M=6.121).

Table O2. The Mean Scores of The Participants in the Speaking Tests in the Pilot Study

		Paired Samples Statistics			
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	FluencyPretest	2.333	24	.7081	.1445
	FluencyPosttest	2.996	24	.6307	.1287
Pair 2	AccuracyPretest	2.621	24	.6782	.1384
	AccuracyPosttest	3.125	24	.6088	.1243
Pair 3	Pretest	4.958	24	1.3548	.2766
	Posttest	6.121	24	1.1710	.2390

Based on the total mean scores, it can be predicted that there is an increase in English speaking skill among the participants from the pretest to the posttest. Hence, it is worth conducting a statistical investigation to see if the participants' increase is statistically significant by looking at the results of p value (or

the *Sig.* value) from the paired samples t-test (Table O3). Based on statistic theory, if the *p* value is ≤ 0.05 , then it means that the result is statistically significant. If the *p* value is < 0.05 , then the result is seen not statistically significant. The result of the paired sample t-test revealed that the overall *p* value between the pretest and the posttest was 0.000 (*Sig.* = 0.000). Therefore, it can be concluded that the result of the pretest and the posttest was statistically significant ($p < 0.050$), which means that the participants significantly improved their English speaking skill after the experiment. In addition, the *p* value of Pair 1 (the result of fluency in the pretest and posttest) and Pair 2 (the result of accuracy in the pretest and posttest) was namely 0.000 (*Sig.* = 0.000) and 0.000 (*Sig.* = 0.000), which was also < 0.05 . Therefore, it can also be confirmed that the speaking skill of the participants was significantly improved in terms of fluency and accuracy. Having triangulated with the results of the questionnaire and the interview, the researchers of this study found that using AI voice chatbots could help the students improve their English speaking skill.

Table O3. The Result of The Paired Samples T-Test in The Pilot Study

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	FluencyPretest - FluencyPosttest	-.6625	.3998	.0816	-.8313	-.4937	-8.118	23	.000
Pair 2	AccuracyPretest - AccuracyPosttest	-.5042	.4563	.0931	-.6969	-.3115	-5.412	23	.000
Pair 3	Pretest - Posttest	-1.1625	.6736	.1375	-1.4469	-.8781	-8.455	23	.000

J3.2 The Questionnaire Responses

The questionnaire responses were collected using Google Form and analysed using SPSS version 20. Having investigated the frequencies of the responses, the researchers of this study found that there were nine males and fifteen females participating in the pilot study (Table O4) and their ages ranged from 19 to 23 (Table O5).

Table O4. The Number of Males and Females in the Pilot Study

		Q1 - Gender			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	9	37.5	37.5	37.5
	Female	15	62.5	62.5	100.0
	Total	24	100.0	100.0	

Table O5. The Range of Ages of the Participants in the Pilot Study

Q2 - Age

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 19	17	70.8	70.8	70.8
20	3	12.5	12.5	83.3
21	3	12.5	12.5	95.8
23	1	4.2	4.2	100.0
Total	24	100.0	100.0	

In addition, almost all participants came from non-English majors and others (Table O6). There were nineteen participants who were at the 2nd year of study and three at the 3rd year while only one participant was at the 1st year and another one did not report (Table O7). Regarding the English placement score of the participants, the data revealed that there were two participants who got the level of under average (under 5.0) and sixteen participants who got the level of above average (from 5.0 to 8.0). However, there were six participants who did not report their scores (Table O8). For the participants' technological experience, almost all of them used to try technology (Table O9).

Table O6. The Majors of the Participants in the Pilot Study

Q3 - Major

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Medicinal Chemistry	2	8.3	8.3	8.3
Primary Education	4	16.7	16.7	25.0
Chemistry Education	1	4.2	4.2	29.2
Agriculture	1	4.2	4.2	33.3
Civics	6	25.0	25.0	58.3
Software Engineering	1	4.2	4.2	62.5
Aquatic Product Processing Technology	1	4.2	4.2	66.7
Vietnamese Studies	1	4.2	4.2	70.8
Computer Network & Information Communication	1	4.2	4.2	75.0
Applied Biology	1	4.2	4.2	79.2
Japanese Language Studies	1	4.2	4.2	83.3
Information Technology	1	4.2	4.2	87.5
Other	3	12.5	12.5	100.0
Total	24	100.0	100.0	

Table O7. The Participants' Year of Study in The Pilot Study

Q4 - Year of Study

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	4.2	4.3	4.3
	2	19	79.2	82.6	87.0
	3	3	12.5	13.0	100.0
	Total	23	95.8	100.0	
Missing	System	1	4.2		
Total		24	100.0		

Table O8. The Participants' English Placement Score in the Pilot Study

Q5 - English Placement Test Score

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4.0	1	4.2	5.6	5.6
	4.7	1	4.2	5.6	11.1
	5.0	3	12.5	16.7	27.8
	5.3	2	8.3	11.1	38.9
	5.6	1	4.2	5.6	44.4
	6.0	1	4.2	5.6	50.0
	6.1	1	4.2	5.6	55.6
	6.5	1	4.2	5.6	61.1
	6.8	1	4.2	5.6	66.7
	7.0	2	8.3	11.1	77.8
	7.5	1	4.2	5.6	83.3
	8.0	3	12.5	16.7	100.0
	Total	18	75.0	100.0	
	Missing	System	6	25.0	
Total		24	100.0		

Table O9. The Participants' Technological Experience in The Pilot Study

Q6 - Technological Experience

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Podcast	4	16.7	16.7	16.7
	Voice chat apps	13	54.2	54.2	70.8
	Chatbots	2	8.3	8.3	79.2
	Learning websites	3	12.5	12.5	91.7
	Other	2	8.3	8.3	100.0
	Total	24	100.0	100.0	

Regarding the participants' opinions on their English speaking skill after they practiced English speaking with the AI voice chatbots, this study reported in detail by providing the results of the participants' responses from question 7 to question 15. In question 7, there were 50% of the participants agreed that they could speak English without making too many pauses and hesitations after they practiced with the AI voice chatbots while 33.3% of them kept neutral opinion about that. Only 8.3% of them gave negative opinions (Table O10).

Table O10. The participants' Responses to Question 7 in the Pilot Study

Q7 - After practicing with the AI voice chatbot, I think I could speak English without making too many pauses and hesitations.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Strongly disagree	2	8.3	8.3	8.3
Disagree	2	8.3	8.3	16.7
Neutral	8	33.3	33.3	50.0
Agree	12	50.0	50.0	100.0
Total	24	100.0	100.0	

In question 8, there were namely 45.8% and 29.2% of the participants agreed and strongly agreed that they could speak English with appropriate hedging words after they practiced with the AI voice chatbots while 16.7% of them kept neutral opinion about that. Only 8.3% of them gave negative opinions (Table O11).

Table O11. The Participants' Responses to Question 8 in The Pilot Study

Q8 - After practicing with the AI voice chatbot, I think I could speak English with appropriate hedging words such as uhm... ah... oh....

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Strongly disagree	2	8.3	8.3	8.3
Neutral	4	16.7	16.7	25.0
Agree	11	45.8	45.8	70.8
Strongly agree	7	29.2	29.2	100.0
Total	24	100.0	100.0	

Since questions 7 and 8 were to explore if the AI voice chatbots could improve the students' English speaking skill in terms of fluency, the results revealed that most of the participants agreed that they could improve their fluency in English speaking because they could speak English without making too many pauses or

hesitations and could use appropriate hedging words such as uhm... ah... oh... when speaking. In question 9, there were namely 41.7% and 8.3% of the participants agreed and strongly agreed that they could produce correct pronunciation after they practiced with the AI voice chatbots while 37.5% of them kept neutral opinion about that. However, there were respectively 4.2% and 8.3% of the participants disagreed or even strongly disagreed with the statement (Table O12).

Table O12. The Participants' Responses to Question 9 in the Pilot Study

Q9 - After practicing with the AI voice chatbot, I think I could produce correct pronunciation when I spoke English.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Strongly disagree	2	8.3	8.3	8.3
Disagree	1	4.2	4.2	12.5
Neutral	9	37.5	37.5	50.0
Agree	10	41.7	41.7	91.7
Strongly agree	2	8.3	8.3	100.0
Total	24	100.0	100.0	

In question 10, there were namely 58.3% and 4.2% of the participants agreed and strongly agreed that they could use appropriate sentence structures when they spoke English after practicing with the AI voice chatbots while 25% of them kept neutral opinion about that. However, there were respectively 4.2% and 8.3% of the participants disagreed or even strongly disagreed with the statement (Table O13).

Table O13. The Participants' Responses to Question 10 in the Pilot Study

Q10 - After practicing with the AI voice chatbot, I think I could use appropriate sentence structures when I spoke English.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Strongly disagree	1	4.2	4.2	4.2
Disagree	2	8.3	8.3	12.5
Neutral	6	25.0	25.0	37.5
Agree	14	58.3	58.3	95.8
Strongly agree	1	4.2	4.2	100.0
Total	24	100.0	100.0	

In question 11, there were namely 54.2% and 12.5% of the participants agreed and strongly agreed that they could use appropriate words and vocabulary when they spoke English after practicing with the AI voice chatbots while 25% of them kept neutral opinion about that. However, there were only 8.3% of the participants strongly disagreed with the statement (Table O14).

Table O14. The Participants' Responses to Question 11 in the Pilot Study

Q11 - After practicing with the AI voice chatbot, I think I could use appropriate words and vocabulary when I spoke English.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly disagree	2	8.3	8.3	8.3
	Neutral	6	25.0	25.0	33.3
	Agree	13	54.2	54.2	87.5
	Strongly agree	3	12.5	12.5	100.0
	Total	24	100.0	100.0	

Since questions 9 to 11 were to explore if the AI voice chatbots could improve the students' English speaking skill in terms of accuracy, the results revealed that most of the participants agreed that they could improve their accuracy in English speaking because they could speak with correct pronunciation and use appropriate sentence structures as well as words and vocabulary when speaking.

J3.3 The Interview Responses

After answering the questionnaire, the participants were further asked to take an interview with the researcher. All interview sessions were recorded to serve for transcribing and interpreting purposes. At first, the researchers of this study preferred to use NVivo to interpret the interview responses, but Microsoft Word was used to replace for NVivo because it was quite expensive for us to afford for NVivo. The responses of the participants were transcribed into text using Microsoft Word 2019 and then translated from Vietnamese to English to serve for interpreting purposes. The results of the interview were reported in detail. In question 1, the participants were asked if the frequently practiced speaking with the AI voice chatbots every week. Their answers revealed that almost all of them spent time speaking with the AI voice chatbots but the length of time of each other was quite different. Some spent around 5 to 15 minutes to speak every week but some even spent an hour or more to practice and record their speaking sessions (Table O15).

Table O15. The Participants' Responses to Interview Question 1 in The Pilot Study

Question 1: How often and how long did you practice speaking with the AI voice chatbots every week? <i>Bạn thực hành nói với AI bao lâu và thường xuyên ra sao ở mỗi tuần?</i>		
Participant ID	VN	EN
PAR01	Có, 1 tuần khoảng 2-3 lần. 1 lần khoảng 5 phút.	Yes, about twice or three times a week. Each time took around 5 minutes.
PAR02	Chắc em làm được 1-2 buổi/tuần khoảng từ 15-30 phút.	I did it around once or twice a week and it took around 15-30 minutes.
PAR03	Em thường sử dụng AI mỗi tối khi em rảnh và ko có bài trên lớp hoặc khi em có thời gian rảnh và cảm thấy chán. Thường em nói khoảng 5-10 phút.	I usually spoke with the AI voice chatbots every night when I had free time or did not have any homework or when I felt bored. It took around 5-10 minutes.
PAR04	Cũng hok thường xuyên lắm nhưng em thấy học trên AI rất là tốt.	Not very often but I believed that learning with the AI voice chatbots must be very good.
PAR05	Em nói với AI khoảng 3 lần/tuần.	I spoke with the AI voice chatbots around three times a week.
PAR06	Thường thì em làm khoảng 1 tiếng đến 1 tiếng rưỡi vào cuối tuần. Do mỗi lần em làm thì em ghi âm và chụp màn hình rồi đưa lên Word thành file để gửi thầy luôn nên hơi lâu.	Normally did I do it around an hour or an hour and a half on weekends. It was quite long because I recorded my voice and took the chatting screenshots everytime I spoke and then I designed those screenshots on Word to submit to you.
PAR07	Có, dạ thì khi nào em rảnh em nhớ thì em thực hành. Mỗi lần thực hành em thích nói lắm.	Yes, I practiced speaking whenever I was free and remembered. I really love it when I spoke.
PAR08	Dạ ko thường xuyên lắm.	Not very often.
PAR09	Em thực hành mỗi tuần khoảng 3 tiếng.	I practiced around three hours every week.
PAR10	Ko thường xuyên ạ.	Not very often.
PAR11	Trong 1 tuần em cũng ko nhớ lắm. Cũng ko nhiều lắm vì em còn phải đi học các môn.	I did not remember clearly how long I practiced in a week. I know it was not much because I had to learn other subjects.
PAR12	Ko thường xuyên ạ.	Not very often.
PAR13	Em luyện cỡ 5 phút mỗi ngày.	I practiced around 5 minutes everyday.
PAR14	Em thường dành ra 2 buổi trong tuần để nói và mỗi lần vậy là em nói khoảng 10-15 phút.	I usually spoke twice a week and it took me around 10-15 minutes for each time.
PAR15	1 tuần em luyện tầm 3-4 buổi, khoảng 20-30 phút/buổi.	I practiced around three times to four times a week, and it took around 20-30 minutes/time.
PAR16	Em luyện 1 tuần khoảng 1-2 ngày, 2 lần/ngày và mỗi lần luyện khoảng 5-10 phút.	I practiced around once to twice a day. I did it in one or two days a week. It took about 5-10 minutes for practicing.
PAR17	Khoảng 15-30 phút và khoảng 2-3 buổi/tuần.	About 15-30 minutes with twices or three times a week.
PAR18	Dạ thực hành ngày nào cũng nhắn tin với AI. AI cũng thông minh vì em hỏi thì trả lời được hết và AI cũng có chỉ em về ngữ	I chatted with the AI voice chatbots everyday. It was also intelligent because when I asked questions, it could answer all, and it also instructed me some

Question 1: How often and how long did you practice speaking with the AI voice chatbots every week? <i>Bạn thực hành nói với AI bao lâu và thường xuyên ra sao ở mỗi tuần?</i>		
Participant ID	VN	EN
	pháp nữa.	grammar structures.
PAR19	Mỗi lần thực hành em làm khoảng 5 phút khi nào rảnh em thực hành.	I practiced whenever I have free time, and it took me around 5 minutes each time.
PAR20	Mỗi tuần em sắp xếp từ 30 đến 1 tiếng để luyện nghe và nói với AI.	I spent around 30 minutes to an hour to practice listening and speaking with the AI voice chatbots.
PAR21	Thường thì chắc phải cách ngày em mới có luyện, mỗi lần khoảng hơn 30 phút xiu.	I practiced once after days, and it took approximately 30 minutes each time.
PAR22	Em luyện cũng thường, lúc nào rảnh là em luyện. 1 ngày chắc tầm 30 phút đến 1 tiếng.	I usually practiced. I practiced whenever I was free. It was about 30 minutes to an hour a day.
PAR23	Thường em luyện nói với AI khoảng 4 ngày trong tuần khi em có thời gian rảnh vào buổi tối. Em luyện khoảng 5 đến 15 phút/lần.	I usually practiced speaking with the AI voice chatbots four days a week when I had spare time at night. I practiced around 5 to 15 minutes each time.
PAR24	Sau 7 giờ tối thì em dành từ 5-10 phút để luyện nói với AI vì em muốn cải thiện kỹ năng nói tiếng Anh của em để sau này em sẽ thi 1 chứng chỉ quốc tế nào đó.	After 7:00 p.m., I spent about 5-10 minutes to practice speaking with the AI voice chatbots because I wanted to improve my English speaking skill so that I could take an international exam in the future.

In question 2, the participants were asked if they thought their grammar had been improved after speaking with the AI voice chatbots. Their answers revealed that most of them found that they could use grammar more accurately after speaking with the AI voice chatbots because those chatbots used correct grammar structures from which the participants could learn and follow. Moreover, they reported that when they uttered a sentence, their sentence appeared on the screen so that they could see grammar structures they had made and know if they had spoken with accurate or inaccurate grammar structures. One of the participants also reported that they could learn new grammar structures after talking with the AI voice chatbots because those chatbots also used new grammar structures and new vocabulary. In contrast, there were two participants (PAR05 and PAR12) saying that they only found their vocabulary improved rather than grammar because the AI voice chatbots used simple grammar structures which they had already known (Table O16).

Table O16. The Participants' Responses to Interview Question 2 in the Pilot Study

Question 2: Do you think that your grammar has been improved after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng ngữ pháp của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
PAR01	Có, cải thiện được xiu. Như trong bài test em thấy nói được nhiều hơn.	Yes, a little bit. I found that I could speak more in the speaking test.
PAR02	Em nói câu nó chuẩn hơn và hiểu câu hơn.	I found that I could speak more accurately and understand the sentences much more.
PAR03	Khi nói chuyện với AI thì em thấy ngữ pháp của em rất cải thiện do khi nói ra thì phần nội dung của em và AI sẽ hiện lên màn hình và nhờ đó em biết được ngữ pháp mình sai chỗ nào và cần khắc phục ra sao.	When I spoke with the AI voice chatbots, I found that my grammar was improved very much because when I produced a sentence, that sentence appeared on the chat screen between me and the AI voice chatbots. Therefore, I knew what grammar mistakes I was making and how to overcome.
PAR04	Em thấy cũng cải thiện chút chút.	I found a little improved.
PAR05	Ngữ pháp khi nói với AI thường dùng cấu trúc đơn giản thôi. Em nghĩ em được cải thiện phần từ vựng nhiều hơn và nói tự nhiên hơn thôi.	The AI voice chatbots only used basic grammar structures. I thought vocabulary was the part I could improve much more, and I could also speak more naturally.
PAR06	Về ngữ pháp em thấy được cải thiện ở mấy câu trả lời về sở thích hay đại loại vậy với lại phần hỏi giờ.	I found that I could improve my grammar in the cases of asking and answering about likes and dislikes or the like, and also the part of asking time.
PAR07	Có cải thiện vì em nói sai nhưng AI trả lời đúng ngữ pháp thì em nhìn em có thể học theo.	It was improved because when I spoke with wrong grammar, the AI voice chatbots responded to me with correct grammar. Therefore, I could look at that and learn.
PAR08	Có, em thấy em dùng đúng ngữ pháp hơn.	Yes, I found that I used much more accurate grammar.
PAR09	Có, em nói với AI rồi AI sửa lỗi ngữ pháp cho em.	Yes, I spoke with the AI voice chatbots and then they corrected grammatical mistakes for me.
PAR10	Có, từ vựng em cũng biết nhiều hơn.	Yes, I also knew more vocabulary.
PAR11	Tại vì khi luyện với AI nó có hiện phần chữ nên khi nói ra chữ sẽ hiện ra giúp em dễ nhớ ngữ pháp hơn.	Yes, because when I practiced speaking with the AI voice chatbots, my sentences appeared on the screen, so it helped me remember grammar more easily.
PAR12	Em chỉ thấy có thêm vốn từ vựng thôi.	I only found that I knew more vocabulary.
PAR13	Có, em nói thì AI cũng hiểu em nói gì. Khi em nói nó hiện ra đúng ngữ pháp và AI cũng hiểu.	Yes, because when I spoke, the AI voice chatbots could understand what I was saying, and my sentences appeared on the screen with correct grammar, so the AI voice chatbots could understand, too.
PAR14	Em nghĩ là có vì khi em đọc xong em tra thì thấy cú pháp cũng đúng.	I think yes because after I had spoken, I found that my grammar on the screen was correct.
PAR15	Em nghĩ là có vì qua các bài học và	I think yes because throughout the lessons and speaking

Question 2: Do you think that your grammar has been improved after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng ngữ pháp của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
	nói với AI thì AI cũng nói lại đúng ngữ pháp và từ vựng để em có thể ghi nhớ.	sessions with the AI voice chatbots, I found that they spoke with correct grammar and vocabulary so that I could remember.
PAR16	Em thấy có cải thiện ngữ pháp vì mỗi lần AI chat với em thì sẽ có những điểm ngữ pháp mới làm em phải tìm hiểu thêm.	I found that my grammar had been improved because every time the AI voice chatbots talked with me, they used new grammar structures which encouraged me to learn more.
PAR17	Có nhiều ạ vì khi em nói thì AI có trả lời lại và nhìn câu trả lời em biết được ngữ pháp đúng để em tự sửa lại.	Yes, very much because when I spoke, the AI voice chatbots replied. When I looked at their replies, I knew the correct grammar structures so that I could self-correct for my speaking.
PAR18	Có cái voice lúc em đọc AI nhận được phát âm lúc thì AI ko nhận được. Nhất là tên tiếng Việt và phát âm chữ "it's", "hi", "hello". Chắc do mic ko nhận. Luyện với AI thì em tra thêm trên Google và Youtube nên được cải thiện ngữ pháp luôn.	Yes, but about the voices. When I read, the AI voice chatbots could recognize my voice, but sometimes they could not. Particularly, when I read my Vietnamese name and pronounced "it's", "hi", or "hello", they could not hear. It might be due to the problem of the microphone. When I practiced speaking with the AI voice chatbots, I looked up vocabulary on Google and Youtube, so I could improve both grammar and vocabulary.
PAR19	Em thấy có cải thiện vì trên AI dùng đúng ngữ pháp và sau khi em học ngữ pháp em có thể thực hành ngay nên cải thiện hơn.	I found improved because the AI voice chatbots used correct grammar and after I had learnt about grammar, I could practice immediately, which helped me improved.
PAR20	Có tại vì nói với AI em thấy được một số từ vựng và điểm ngữ pháp mới để em học thêm. Hơn nữa nói với AI như nói với người nước ngoài nên em nhận biết được mặt chữ rõ hơn và biết từ đó là gì. Việc này làm em nghe dễ hơn so với nói với người thật.	Yes, because when I spoke with the AI voice chatbots, I could see and learn new vocabulary and grammar structures. Moreover, speaking with the AI voice chatbots gave the feeling like speaking with foreigners so I could recognize the metaphors more easily and knew the words. This will help me to listen easily rather than speaking with human.
PAR21	Cũng có đôi chút. Thường thì em học theo cấu trúc mẫu gợi ý từ AI và trên web học rồi từ từ em mở rộng ra thêm.	A little bit. I usually imitated the structures used by the AI voice chatbots and suggested on the learning website, and then I developed my knowledge wider and wider.
PAR22	Có vì lúc trước khi thường xuyên luyện tập thì câu nói của em lủng củng và nói toàn câu ngắn. Sau khi luyện nói với AI thì phần ngữ pháp em chuẩn hơn và mạch văn em có thể nói dài hơn.	Yes, because in the past, I did not practice speaking frequently, so my sentences were mostly run-on sentences and I could only speak short sentences. But after I practiced with the AI voice chatbots, my grammar was more accurate and my sentences were longer.
PAR23	Em thấy có 1 số cấu trúc ngữ pháp nhờ AI mà em biết thêm. Tuy nhiên	I found that I could learn some grammar structures from the AI voice chatbots. However, there were some problems

Question 2: Do you think that your grammar has been improved after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng ngữ pháp của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
	cũng có 1 số vấn đề là khi em nhắn 1 đấng AI trả lời 1 nẻo.	that when I spoke with them, I talked about one thing, but they answered to me with another thing.
PAR24	Em thấy cải thiện nhiều vì lúc trước em hơi rụt rè vì phần nói của em nhưng sau khi dành thời gian luyện tập với AI thì em thấy cải thiện nhiều về khả năng nói và khả năng đặt câu của em.	I found that I could get improved much because I used to be embarrassed with my speaking. However, when I spent time practicing speaking with the AI voice chatbots, I could improve much in my speaking ability and the ability of making sentences.

In question 3, the participants were asked if they thought their pronunciation had been improved after speaking with the AI voice chatbots. Their answers revealed that most of them found that they could speak with better pronunciation because of four reasons. First, they could listen to the AI voice chatbots' speaking voice again and again and repeat the voice to improve pronunciation because those chatbots could speak with perfect pronunciation. For example, PAR06 reported that after he had listened to the AI voice chatbots, he realized that the article "a" should be pronounced /ei/ not /oh/. Second, they had to try to speak slowly for many times so that the AI voice chatbots could recognize their sentences and respond. Third, before they spoke with the AI voice chatbots, they had to look up some words on online dictionaries and practice pronunciation so that the AI voice chatbots would not fail to understand. Finally, when they uttered a sentence with wrong pronunciation, they could see the wrong sentence appear on the screen, from which they could try to speak more carefully with correct pronunciation (Table O17).

Table O17. The Participants' Responses to Interview Question 3 in the Pilot Study

Question 3: Do you think that your pronunciation has been improved after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng phát âm của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
PAR01	Có, cải thiện xiu. Khi nói với AI có thể bấm nghe lại AI nói nhiều lần sau đó em đọc theo nên cải thiện được.	Yes, a little bit. When I spoke with the AI voice chatbots, I could press the button to listen again for many times. Therefore, I could repeat and improve my pronunciation.
PAR02	Em thấy mình đọc nó chuẩn hơn, phát âm chuẩn hơn.	I found that I could speak and pronounce more accurately.

Question 3: Do you think that your pronunciation has been improved after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng phát âm của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
PAR03	Em thấy có vì khi nói với AI nếu nói đúng thì AI sẽ trả lời lại, còn khi phát âm sai thì AI ko phản hồi lại em được và khi đó em biết em phát âm sai.	I think yes because when I spoke with the AI voice chatbots, if I spoke accurately, then they would reply. If I pronounced wrongly, then they did not reply to my saying. At that time, I knew I pronounced incorrectly.
PAR04	Sau khi học với AI thì phát âm của em cũng chuẩn hơn tí. Em nghe AI đọc rồi em đọc lại theo.	After practicing with the AI voice chatbots, I found that my pronunciation was improved a bit because I could listen to the AI voice chatbots and repeat after them.
PAR05	Có, vì khi em phát âm ko đúng sẽ ko hiện ra đúng câu em nói nên lúc nói em phải để ý phát âm hơn.	Yes, because when I pronounced incorrectly, my saying appeared wrongly. This made me become more careful when speaking.
PAR06	Em thấy 1 phần là phát âm được cải thiện. Đặc biệt là chữ “a” đọc là “ei” chứ ko phải “ờ”.	I found that my pronunciation was improved more or less. Interestingly, I have just known that “a” was pronounced /ei/ not /oh/.
PAR07	Có, khi nói với AI phản hồi lại thì em nghe được giọng của AI.	Yes, when speaking with the AI voice chatbots, they responded and I could hear their voice.
PAR08	Có, em thấy em đọc phát âm đúng hơn nhờ AI có thể lặp lại nhiều lần 1 câu.	Yes, I found that I could pronounce more accurately thanks to many times the AI voice chatbots repeated the sentences.
PAR09	Có tại vì em nói ko được em sẽ nói hoài đến khi nào phát âm đúng thì mới thôi.	Yes, because when I failed to pronounce correctly, I had to speak again and again until I could pronounce correctly to get the AI voice chatbots understand me.
PAR10	Có cải thiện hơn.	Much more improved.
PAR11	Có, vì khi em phát âm sai nó sẽ ko hiện ra chữ đúng và AI cũng ko hiểu nên em phải cố gắng luyện và nói nhiều lần đến khi AI hiểu thì em đã nói được tốt hơn.	Yes, because when I pronounced inaccurately, my saying did not appear correctly on the screen and the AI voice chatbots did not understand what I was saying, too. Therefore, I had to try to practice and speak for many times until the AI voice chatbots understand me. At that time, my pronunciation had been improved.
PAR12	Có 1 chút thôi.	Yes, but only a bit.
PAR13	Có luôn, em nói xong nó hiện ra đúng câu em muốn nói.	Yes, sure. After I spoke, I found my saying appeared correctly on the screen.
PAR14	Chắc chắn có vì em đọc ra AI nghe hiểu.	Definitely yes because when I pronounced, the AI voice chatbots could hear and understand.
PAR15	Có cải thiện vì em có tập luyện và nói tự tin hơn cũng như dùng ah uhm để tìm ý nói.	Yes, because I spent time practicing and could speak with more confidence. I could also say ah... uhm... to find ideas to speak.
PAR16	Khi em nói với AI thì em phải tập phát âm cho chuẩn để nói AI hiểu.	When I talked with the AI voice chatbots, I had to practice pronunciation many times so that I could pronounce correctly and get the AI voice chatbots understand me.
PAR17	Có vì khi phát âm thì AI cũng	Yes, because when I listen to the AI voice chatbots, I could

Question 3: Do you think that your pronunciation has been improved after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng phát âm của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
	phát âm chuẩn nên em có thể nghe và luyện theo.	imitate their pronunciation thanks to their perfect pronunciation voice.
PAR18	Phát âm là ok nhất. Có nhiều từ lúc trước em đọc nó ko nhận, sau này khi em đọc nó nhận được do em luyện phát âm được từ từ.	The best improvement is my pronunciation. There are many words I pronounced but the AI voice chatbots did not recognize before. But now when I pronounced those words, my words were recognized. I could gradually practice pronunciation.
PAR19	Cũng có, vì nếu em nói đúng thì AI mới nhận diện được từ.	Somehow yes, because I need to pronounce correctly to get the AI voice chatbots understand me.
PAR20	Phát âm em có cải thiện nhưng mà ít hơn nói trực tiếp với người nước ngoài do ngôn ngữ AI ko hoàn thiện lắm. Khi nói với người nước ngoài thì ngôn ngữ khác so với nói với AI.	My pronunciation had been improved but not as much as when I spoke directly with foreigners because the language of the AI voice chatbots were not completely developed. When I spoke with foreigners, their languages were quite different compared to those of the AI voice chatbots.
PAR21	Phát âm thì em cũng hok rõ vì em ko nghe lại nên cũng ko biết có cải thiện ko.	I am not clear if my pronunciation had been improved because I did not listen to my voice again.
PAR22	Có luôn vì em có sử dụng mic để ghi âm lại giọng em để AI nghe và thấy AI nhận diện được câu em nói sau vài lần.	Yes, because I used a microphone to speak with the AI voice chatbots and recorded my voice in the microphone. I found that they could recognize my sentences after some tries.
PAR23	Có vì em thường chat thì em bấm vào biểu tượng loa để nghe lại AI nói và em lặp lại chừng nào giống mới thôi.	Yes, because when I talked with the AI voice chatbots, I usually clicked on the speaker icon to hear their voice again and repeat until I could pronounce similarly to them.
PAR24	Có ví dụ khi gặp 1 từ em ko biết cách phát âm thì em có thể nghe AI đọc hoặc tra từ điển từ đó phát âm của em được cải thiện nhiều hơn.	Yes. For example, when I found a word which I did not know how to pronounce, I could listen to the AI voice chatbots' voice or look up the word in the dictionary. Therefore, my pronunciation had been improved much more.

In question 4, the participants were asked if they thought their vocabulary had been improved after speaking with the AI voice chatbots. Their answers revealed three important points. First, most of them reported that their vocabulary had been improved because the AI voice chatbots used quite a lot of new words to them. Therefore, when they talked with the AI voice chatbots, they had to look up the new words using dictionaries. As a result, their vocabulary was improved. Second, for those who did not have enough vocabulary to speak, they had learnt new words from the learning websites before speaking with the AI voice chatbots so that they

could have enough words to speak. Finally, some of the participants reported that their vocabulary had not been improved yet because their available vocabulary knowledge was too limited. As a result, they still could not understand the meanings even though they asked the chatbots for explanation because the explanation was still in English, not in Vietnamese (Table O18). Based on the participants' reports, it can be seen that the AI voice chatbots can indirectly improve vocabulary of learners by encouraging the learners to use the dictionary. The AI voice chatbots themselves cannot directly improve vocabulary of learners if those learners are not supported by Vietnamese dictionaries.

Table O18. The Participants' Responses to Interview Question 4 in the Pilot Study

Question 4: Do you think that your vocabulary has been improved after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng từ vựng của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
PAR01	Có, vì em có thể nói với AI nhiều chủ đề khác nhau. Khi nói em học được nhiều từ vựng hơn.	Yes, because I could speak with the AI voice chatbots about a variety of topics. When I spoke, I could learn more vocabulary.
PAR02	Có, em thấy từ vựng có nâng lên đôi chút.	Yes, I found that my vocabulary has increased a little.
PAR03	Em nghĩ là từ vựng em chưa được cải thiện lắm do vốn từ vựng em vẫn chưa nhiều.	I think that my vocabulary has not improved much because my available vocabulary is still limited.
PAR04	Có.	Yes.
PAR05	//	//
PAR06	Có, từ vựng em biết được thêm từ vựng như về âm nhạc, thời gian. Trong các đoạn hội thoại nhỏ em cũng học được 1 số từ vựng.	Yes, I have known more vocabulary about music and time. In small talks, I could also learn some vocabulary.
PAR07	Cũng có khi AI phản hồi có những từ em ko biết thì em lên Google Translate dịch ra thì em sẽ học được từ đó.	Somehow yes, because when the AI voice chatbots responded, there were some words which I did not know. Therefore, I looked them up using Google Translate, then I could learn those words.
PAR08	Có, em biết nhiều từ vựng hơn.	Yes, I knew more vocabulary.
PAR09	Em nghĩ là ko vì nó ko có giải thích tiếng Việt nên em cũng ko hiểu từ.	I think no because they did not explain in Vietnamese. Therefore, I did not know the words, too.
PAR10	Có.	Yes.
PAR11	Có, nhất là khi AI trả lời lại mình thì mình biết là 1 câu nói có nhiều cách để thể hiện nên em biết nhiều từ vựng hơn.	Yes, particularly when the AI voice chatbots replied to me, I knew that in one sentence, there were many ways to express, which helped me know more vocabulary.
PAR12	Có nhờ AI lặp lại nhiều lần nên em nghe em lặp lại theo từ từ cải thiện.	Yes, because the AI voice chatbots repeated again and again, I could listen and repeat, which helped me

Question 4: Do you think that your vocabulary has been improved after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng từ vựng của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
		gradually improve.
PAR13	Có, như em nói xong thì AI phản hồi lại từ nào ko hiểu thì em dịch ra sẵn em học luôn nên tang từ vựng.	Yes, after my speech, the AI voice chatbots responded. If there were words I did not know, then I have them translated so that I could also learn to improve vocabulary.
PAR14	Có, sau khi luyện nó với AI em thấy từ vựng em tang khá nhiều. Vì khi em thấy từ tiếng Anh em nhìn em hiểu.	Yes, after practicing speaking with the AI voice chatbots, I found that my vocabulary increased quite much because when I saw the words, I could understand.
PAR15	Có một ít nhưng em cần cố gắng thêm để cải thiện thêm.	Yes, but only a little. I think I need to improve more.
PAR16	Có vì mỗi lần AI trả lời có từ mới hoặc cấu trúc mới thì buộc em phải tìm hiểu về nó.	Yes, because whenever the AI voice chatbots responded with new words or new sentence structures, I had to learn about those.
PAR17	Khi mà luyện với AI thì nó cung cấp cho em nhiều từ vựng của 1 chủ đề khi em đặt câu hỏi. Khi AI trả lời thì dùng khá nhiều từ mới nên từ đó em có thể học được thêm từ vựng.	When I practiced with the AI voice chatbots, they provided me with many words related to a topic. When I made a question, the AI voice chatbots answered with many new words. Thanks to that I could learn more vocabulary.
PAR18	Có, khi chat có nhiều cái AI hỏi em ko có từ để nói nên em phải ra tra cứu để có thể vào chat lại với AI.	Yes, when we chatted, the AI voice chatbots asked me many things which I did not have enough words to talk with. Therefore, I had to look up some words to be able to chat with the AI voice chatbots.
PAR19	Có vì khi em học thì có phần thực hành nói với AI có những chữ em ko biết thì em tìm hiểu thêm nên từ vựng được cải thiện.	Yes, because when I studied, there was a part of practicing speaking with the AI voice chatbots. There were words which I did not know, then I had to learn more. Thanks to that my vocabulary has been improved.
PAR20	Từ vựng được cải thiện là tất nhiên do em nhìn thấy mặt chữ và hiểu nghĩa của chữ đó khi nói với AI.	It is certainly that my vocabulary has been improved because I could see the words and understand them when speaking with the AI voice chatbots.
PAR21	Từ vựng được cải thiện vì em được gợi ý khá nhiều từ vựng từ web học và qua chat với AI.	My vocabulary has been improved because I could learn quite a lot of words from the learning websites and from chatting with the AI voice chatbots.
PAR22	Có luôn vì em có làm theo hướng dẫn bài học nên em có thể học từ vựng từ đó và ứng dụng vào nói với AI. Tựa như em có thể làm bài tập cùng với AI vậy.	Yes, because when I followed the instructions of the lessons, I could learn vocabulary and applied those into speaking with the AI voice chatbots. It looks like I can do homework with the AI voice chatbots.
PAR23	Có nhiều từ vựng mới AI nói mà em được biết thêm lắm.	There were many words spoken by the AI voice chatbots I have learnt from.
PAR24	Có những từ em ko biết khi chat với AI thì em tra từ điển và tra luôn cách phát	There were words I did not know when having a chat with the AI voice chatbots. Therefore, I looked them

Question 4: Do you think that your vocabulary has been improved after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng từ vựng của bạn đã được cải thiện sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
	âm cũng như nhớ nghĩa từ đó nên em nâng cao từ vựng.	up on the dictionary and learnt how to pronounce them, too. Therefore, I could remember those words and improve my vocabulary.

In question 5, the participants were asked if they could speak with less hesitations and pauses after speaking with the AI voice chatbots. Most of the participants said that they could speak with less hesitations compared with their first time speaking with the AI voice chatbots. They explained that at the first time, they spoke with a lot of hesitations and pauses, from which the AI voice chatbots failed to recognize what they said. As a result, they had to practice speaking again and again until they could speak quite fluently so that the AI voice chatbots could recognize what they said. Other students further reported that they frequently spent time speaking with the AI voice chatbot. Therefore, they got used to with the AI voice chatbots and felt more comfortable and confident when speaking, which helped them speak with less hesitations and pauses (Table O19).

Table O19. The Participants' Responses to Interview Question 5 in the Pilot Study

Question 5: Do you think that you can speak with fewer hesitations or pauses after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng bạn có thể nói ít vấp hay ít ngập ngừng hơn sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
PAR01	Có, tại vì tập với AI càng ngày càng quen hơn và thấy dễ hơn nên ko vấp nhiều.	Yes, because the more I practiced with the AI voice chatbots, the more familiar I got with them, and I found it easier to speak. Therefore, I could speak with less hesitations.
PAR02	Có 1 phần, 1 phần là do em chưa cố gắng.	I think a little bit because I myself did not try hard much.
PAR03	Em thấy em nói lưu loát hơn sau khi thực hành với AI. Thể hiện rõ nhất là qua bài test em thấy em nói tốt hơn lúc đầu.	I found that I could speak with the AI voice chatbots more fluently. Apparently, I found myself speaking better in the test.
PAR04	Sau lần 3 lần 4 thì em mới tiến bộ hơn vì lần đầu em chưa biết nhiều.	I think I have got advance after three or four tries because I did not know much at first.
PAR05	Có, tại vì tập với AI em biết cách diễn tả ý và sử dụng ah umh để tìm ý.	Yes, because when I practiced with the AI voice chatbots, I could use ah... umh... to find ideas to speak.
PAR06	Em thấy em nói được hơn lần đầu	I found that I could speak better than the first time.

Question 5: Do you think that you can speak with fewer hesitations or pauses after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng bạn có thể nói ít vấp hay ít ngập ngừng hơn sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
PAR01	Có, tại vì tập với AI càng ngày càng quen hơn và thấy dễ hơn nên ko vấp nhiều.	Yes, because the more I practiced with the AI voice chatbots, the more familiar I got with them, and I found it easier to speak. Therefore, I could speak with less hesitations.
	vào. Nhất là có ý và em đã nói ra được ý em muốn.	Especially, I could find ideas to speak and I could express what I wanted to say.
PAR07	Có, khi em nói với AI kiểu có người phản hồi cho em thì em tự tin hơn và em có thể học theo cấu trúc của AI nên ko nói vấp nữa và có thể suy nghĩ ý để nói.	Yes, when I spoke with the AI voice chatbots, it looked like there was a person responding to me, which made me more confident, and I could learn the sentence structures spoken by the AI voice chatbots. Therefore, I could speak with less hesitations and I could find ideas to speak.
PAR08	Có, em nói chuẩn hơn.	Yes, I could speak better.
PAR09	Em nghĩ là có.	I think yes.
PAR10	Có, tại vì mấy lúc học với AI xong thì em sai xong em chỉnh lại từ từ nên nhiều lần vậy làm em ít vấp hơn.	Yes, because when I learnt with the AI voice chatbots, I made mistakes. Then I had to fix it gradually for many times, which helped me speak with less hesitations.
PAR11	Có, tương đối. Khi luyện với AI nếu em nói vấp thì AI ko hiểu. Nếu nói vấp nhiều quá thì AI sẽ càng ko hiểu nên em phải nói chậm và nói nhiều lần hơn để AI hiểu.	Yes, a little. When I practiced with the AI voice chatbots, if I spoke with too many hesitations, then the AI voice chatbots would not understand at all. Therefore, I had to speak slowly and repeat many times so that the AI voice chatbots could understand.
PAR12	Có, em thấy ít bị vấp hơn lúc ban đầu khi mới luyện với AI. Từ từ luyện nhiều em nói ít vấp hơn.	Yes, I found that I could speak with less hesitations compared with the first time speaking the AI voice chatbots. With many times practicing, I found that I could speak with less hesitations.
PAR13	Có, em thấy em nói chậm lại nhưng phát âm đúng hơn và ít ah oh hơn.	Yes, I found that I could speak more slowly but pronounce more accurately, less ah... oh..., too.
PAR14	Có, em ít vấp hơn. Tại nhờ em luyện nói thường xuyên với AI nên em thấy em nói trôi chảy hơn.	Yes, I have spoken with less hesitations. It is because I frequently practiced speaking with the AI voice chatbots. I found that I could speak more fluently.
PAR15	Em thấy cũng đỡ vấp hơn trước khi tập. Lúc đó nói còn vấp nhưng hiện tại đã đỡ hơn nhiều.	I found that I could speak with less hesitations compared with my first time. At first, I spoke with quite a lot of hesitations, but now I found less hesitations in my speaking.
PAR16	Có, em phản xạ tốt hơn trước kia.	Yes, I interact better than before.
PAR17	Sau vài tuần học với AI thì em thấy khả năng nói của em lưu loát hơn và nhanh hơn nhờ luyện tập thường xuyên.	After some weeks studying with the AI voice chatbots, I found that I could speak more fluently and faster because I had practiced frequently.
PAR18	Có, tại 1 số ý sau khi nói với AI em đã ghi nhớ sẵn nên khi nói lại em	Yes, because there were some ideas I could remember after speaking with the AI voice chatbots. Therefore, when I spoke

Question 5: Do you think that you can speak with fewer hesitations or pauses after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng bạn có thể nói ít vấp hay ít ngập ngừng hơn sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
PAR01	Có, tại vì tập với AI càng ngày càng quen hơn và thấy dễ hơn nên ko vấp nhiều.	Yes, because the more I practiced with the AI voice chatbots, the more familiar I got with them, and I found it easier to speak. Therefore, I could speak with less hesitations.
	ko thấy bị thiếu ý nữa nên nói ít ngập ngừng hơn.	about that again, I could have more ideas to speak, and I could speak with less hesitations.
PAR19	Ít vấp hơn. Nhờ em thực hành nhiều lần từ từ em quen.	Less hesitations. It is because I had practiced for many times and got acquainted.
PAR20	Ít vấp hơn vì nói với AI em đỡ áp lực hơn nói với người nước ngoài. Em cảm thấy thoải mái hơn rất nhiều so với việc nói tiếng Anh với người nước ngoài kiểu face-to-face.	Less hesitations because I found less stress than speaking with a true native speaker. I felt much more comfortable when I talked online, compared with talking with a native speaker in a face-to-face manner.
PAR21	Có ít vấp hơn.	Yes, less hesitations.
PAR22	Đỡ vấp hơn.	Less hesitations.
PAR23	Em vẫn còn vấp nhưng đỡ hơn nhiều so với lần đầu lúc mới luyện với AI.	I still speak with hesitations but less than my first time practicing with the AI voice chatbots.
PAR24	Có vì khi nói sai hoặc ko được tốt vào lần đầu thì em sẽ cố gắng nói lại và sau nhiều lần thì em quen và cải thiện nhịp nói hơn.	Yes, because when I spoke inaccurately or did not speak better at the first time, I tried to repeat. After many tries, I got acquainted and improved my speaking flow.

In question 6, the participants were asked if they could speak better after speaking with the AI voice chatbots (Table O20). Most of the participants said that the greatest improvement was their grammar and pronunciation while only two participants (PAR13 and PAR17) reported that they could learn more vocabulary and new words to support for their speaking after they practiced speaking with the AI voice chatbots, which resulted in their less hesitations when speaking. Interestingly, there was one participant reporting that she could improve her listening because she could not understand the AI voice chatbots at the first time she listened to them, but later she could listen and understand what the AI voice chatbots said (PAR11). However, for the rest of the students, they did not explain in detail why they found that they could speak better after practicing with the AI voice chatbots.

Table O20. The Participants' Responses to Interview Question 6 in the Pilot Study

Question 6: Do you think that you can speak better after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng bạn có thể nói tốt hơn sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
PAR01	Có, nhìn chung kỹ năng nói cải thiện hơn so với trước khi nói với AI	Yes, in general, my speaking ability has been improved better than before I practiced speaking with the AI voice chatbots.
PAR02	Có 1 phần về ngữ pháp và phát âm.	Yes, particularly in grammar and pronunciation.
PAR03	Em thấy được nâng cao phần nói hơn sau khi luyện với AI.	I found that my speaking ability has been improved after practicing with the AI voice chatbots.
PAR04	Có, em thấy tốt hơn.	Yes, I think I can speak better.
PAR05	Có, em cảm thấy em nói tốt hơn về mặt phát âm và phong thái khi nói.	Yes, I feel that I can speak with better pronunciation and speaking style.
PAR06	Có cải thiện.	Yes, it has been improved.
PAR07	Có.	Yes.
PAR08	Có.	Yes.
PAR09	Có, đặc biệt là về phát âm.	Yes, particularly in pronunciation.
PAR10	Có, ban đầu thì em nói ko tốt nhưng luyện nói với AI nhiều từ từ em cải thiện phần nói hơn.	Yes, at first I did not speak very well, but after practicing with the AI voice chatbots frequently, I could improve my speaking much better.
PAR11	Có, nhất là em nghe được nhiều và hiểu nhiều hơn nên em trả lời lại nhanh hơn. Lý do em ko nói được nhiều là do em nghe ko hiểu hiểu nhưng khi nghe nhiều lần em sẽ hiểu và quen hơn nên nói tốt hơn.	Yes, in particular, I could listen and understand better, so I could reply much faster. The reason why I could not speak much was because I could not listen and understand. But when I listened again and again, I became familiar with the information and could speak better.
PAR12	Có.	Yes.
PAR13	Em thấy em nói tốt hơn được xiu. Thứ nhất là em phát âm nó hiểu, thứ hai là em học được từ vựng và cuối cùng là em nói ít vấp hơn.	I found that I could speak a bit better. First, when I pronounced, they could understand. Second, I could learn vocabulary from that. Finally, I could speak with less hesitations.
PAR14	Có.	Yes.
PAR15	Có.	Yes.
PAR16	Có.	Yes.
PAR17	Em thấy phần nói tốt hơn nhiều. Trước kia em còn ngập ngừng và biết ít từ. Sau khi luyện thì em biết nhiều hơn.	I found my speaking became much better. In the past, I got hesitated to speak and knew only a few words. After practicing, I have known more.
PAR18	Có.	Yes.
PAR19	Cũng có tốt hơn đặc biệt là về phát âm và ngữ pháp.	It got a bit better, particularly in pronunciation and grammar.
PAR20	Theo đánh giá của bản thân thì em thấy em nói tốt hơn khoảng 20 đến 25% so với trước khi luyện với AI mà chỉ nói với người nước ngoài.	For me, I found that I could speak about 20 to 25% better if compared to the time I practiced speaking with native speakers.
PAR21	Nói chung em thấy có tiến bộ hơn.	In general, I found that my speaking skill has got better.

Question 6: Do you think that you can speak better after speaking with the AI voice chatbots? Why do you think so? <i>Bạn có nghĩ rằng bạn có thể nói tốt hơn sau khi nói với AI không? Nhờ đâu bạn nghĩ như vậy?</i>		
Participant ID	VN	EN
PAR22	Thường xuyên luyện tập em thấy cách em hành văn suông hơn và tốt hơn.	By frequent practice, I found that I can use better sentence structures use them more smoothly.
PAR23	Có.	Yes.
PAR24	Em thấy là có.	I think yes.

In question 7, the participants were asked what they liked most about speaking with AI voice chatbots and what they disliked (Table O21). Almost all of the participants reported that they liked to speak with the AI voice chatbots due to many reasons. First, they said that the AI voice chatbots made them feel friendly and gave them the sense of chatting with human, which was not as boring as using Google Translate. Second, they liked to speak with the AI voice chatbots because those chatbots pronounced clearly and loudly enough, which made them feel like they were talking with a native friend who could help them improve their speaking skill in terms of pronunciation, grammar, and vocabulary. Advantageously, some of the participants reported that the AI voice chatbots could speak with them restlessly, which gave them a great chance to practice English speaking with so that they could improve their speaking skill. Moreover, the participants also stated that the chatbots provided various topics for them to speak. However, some participants reported their dislikes. First, they mentioned that the AI voice chatbots could not get their names even though they tried to pronounce their names the best they could. In some cases, they believed that they pronounced a word correctly but the AI voice chatbots still failed to recognize. Especially, PAR05 reported that she disliked the fixed responses from the AI voice chatbots because those responses made her feel bored when they were repeatedly uttered. Moreover, the topics were still not so plentiful for learners to speak. Therefore, she further suggested adding more types of responses as well as more topics so that learners could speak more in the future. Some other participants also shared similar ideas with PAR05 about the topics which they could speak with the AI voice chatbots. In particular, they reported that there were not many topics that they could speak with the AI voice chatbots because they found that the AI voice chatbots failed to respond in some topics they talked about.

Table O21. The Participants' Responses to Interview Question 7 in the Pilot Study


Question 7: What do you like most about speaking with the AI voice chatbots? And what do you dislike? <i>Bạn thích điều gì nhất khi nói chuyện với AI và điều gì bạn không thích?</i>		
Participant ID	VN	EN
PAR01	Khi nói với AI thì tạo cảm giác thân quen, AI phát âm to rõ, có thể đọc lại nhiều lần. Chủ đề cũng khá đa dạng. Điều không thích thì có 1 vài chủ đề em nói AI nó ko hiểu. Có thể là do phát âm sai nên nó ko hiểu.	When I spoke with the AI voice chatbots, this made me feel friendly. The AI voice chatbots pronounced clearly and loudly enough, and they could also repeat many times. The topics were quite various. The thing I dislike is that there are some topics I spoke but they did not understand. Maybe it was because I pronounced inaccurately, which prevented them from understanding me.
PAR02	Em thích nhất là chat với AI giống như chat có cảm xúc và tương tác cũng như thông minh hơn chứ ko chán như Google dịch. Ko có gì không thích.	The thing I like best is that chatting with the AI voice chatbots gives the sense of chatting with human. The AI voice chatbots are also more intelligent. They are not as boring as Google Translate. Nothing I dislike.
PAR03	Em thích nhất là được cải thiện tiếng Anh, nói tốt hơn. Nhưng có 1 số chỗ em nói hoài mà AI vẫn ko hiểu.	The thing I like best is that I can improve my English and speak better. However, there are some sections which I speak again and again but the AI voice chatbots still could not get it.
PAR04	Em thích là AI nó làm cho em tiến bộ hơn phần ngữ pháp. Không có phần nào ko thích.	The thing I like is that the AI voice chatbots help me improve my grammar. Nothing I dislike.
PAR05	Khi thực hành với AI thì em thích nhất là phần hỏi và trả lời được lâu. Điều em ko thích là tại vì AI là robot thì khi nghe câu hỏi chỉ trả lời có 1 dạng câu, ko đa dạng cách trả lời. Em mong là có nhiều chủ đề và đa dạng chủ đề hơn để người ta có nhiều lựa chọn vì có khi em hỏi AI cứ kêu em lặp lại chứ ko trả lời được câu hỏi.	When I practiced with the AI voice chatbots, the thing I like best is that the rest time we could take for asking and answering. The thing I dislike is that because the AI voice chatbots are robots, they can only reply to me with fixed sentences. Their responses are not plentiful. I hope that there will be more topics and a variety of responses so that people can have more choices. Sometimes, the AI voice chatbots only asked for my repetition because they failed to answer me.
PAR06	Em thích nhất là cảm giác như nói trực tiếp giữa người với người. Điều ko thích là em thấy có vài từ em phát âm đúng rồi mà AI vẫn ko nhận diện được.	The thing I like best is that the feeling of speaking. It seems like I am speaking with human. The thing I dislike is that there were some words I pronounced correctly but the AI voice chatbots could not recognize.
PAR07	Điều em thích nhất có lẽ là nó như người bạn bản xứ vậy và có phần hỏi với em. Em ko thích thì nó giống như bạn ảo chứ ko phải bạn thật.	The thing I like best is that they look like my native friends and they reply to me. The thing I dislike is that they are like virtual friends, not real.
PAR08	Em thích là AI giúp phát âm chuẩn. Ko có nội dung nào em ko thích.	What I like is that the AI voice chatbots help me to pronounce accurately. Nothing I dislike.
PAR09	Em thích là AI chỉnh phát âm cho em, rèn kỹ năng nói tốt hơn. KO thích là còn hạn chế nhiều nội dung và 1 số	What I like is that the AI voice chatbots help me fix my pronunciation and speak better. I dislike the limitation of speaking topics because some topics were not

Question 7: What do you like most about speaking with the AI voice chatbots? And what do you dislike? <i>Bạn thích điều gì nhất khi nói chuyện với AI và điều gì bạn không thích?</i>		
Participant ID	VN	EN
	chủ đề chưa có. Cũng như chưa có giải thích tiếng Việt vì có 1 số người như em ko biết từ vựng nhiều nên cũng ko hiểu được khi ko có giải thích tiếng Việt.	available. Moreover, explanations in Vietnamese were not found because there will be some people like me whose vocabulary is too limited, will fail to understand anything without Vietnamese explanation.
PAR10	Em ko có gì ko thích. Em thích là nó giúp em cải thiện hơn tiếng Anh.	Nothing I dislike. What I like is that I can improve my English much.
PAR11	Em thích là cách phát âm của AI rất tốt và dễ nghe. Dễ giao tiếp hơn bình thường. Em ko thích ở chỗ nó chưa đủ nội dung nói và chủ đề để AI có thể hiểu thêm.	What I like is that the AI voice chatbots can pronounce perfectly and easily to hear. It is easier for me to communicate rather than in real life communication. What I dislike is that the contents and topics for speaking are still limited. There are not many topics that the AI voice chatbots can understand.
PAR12	Em thích nhất là tăng vốn từ vựng và ngữ pháp. Ko có điều gì ko thích.	What I like best is that I can improve my vocabulary and grammar. Nothing I dislike.
PAR13	Em thích nhất là em nói xong có người đáp lại. Ko thích là lần đầu tiên trải nghiệm AI ko hiểu lại nói Goodbye em luôn nên làm em hơi sốc.	What I like best is that after I speak, I can hear them respond. What I dislike is that for the first time I tried speaking, the AI voice chatbots did not understand, and they even said goodbye to me, which shocked me.
PAR14	Khi em nói tên em tiếng Việt thì AI ko hiểu nên ko thích lắm. Còn điều em thích là khi nói đúng thì AI có phản hồi lại và cho lời khen nên em thấy thích lắm.	When I told my name in Vietnamese, the AI voice chatbots did not understand, so I somehow dislike. What I like is that when I speak accurately, then AI voice chatbots can respond and compliment, which encourages me a lot.
PAR15	Ko thích thì có nhiều lúc em nói tên nhưng AI ko nhận diện được tên tiếng Việt. Em thích ở chỗ dễ dàng giao tiếp, nói chuyện khi ko có ai để nói cùng. Em cũng có thể luyện nói tiếng Anh mỗi ngày dễ hơn bình thường.	What I dislike is that there are times I told my name, but the AI voice chatbots could not get Vietnamese names. What I like is about the ease of communication. It is easy for me to speak in case I cannot find anyone to talk with. I can also practice English speaking everyday more easily than before.
PAR16	Em thích nhất ở AI là em có thể giao tiếp với AI như người bạn bản xứ khi mà em ko có ai để tập nói cùng. Em ko thích ở chỗ có những câu dù tìm hiểu cấu trúc và phát âm đúng thế nào thì AI vẫn ko hiểu mặc dù em đã chỉnh lại ngữ pháp và từ vựng. Chỉ một vài trường hợp AI ko hiểu thôi còn đa số thì AI đều hiểu em.	What I like best about the AI voice chatbots is that I can communicate with them like with the native speakers when I cannot find anyone to speak with. What I dislike is that there are some sections I used correct grammatical structure and pronunciation but the AI voice chatbots still failed to answer, even though I revised the vocabulary and grammar. However, only some situations the AI voice chatbots did not get me. Most of the time they could.
PAR17	Em thích nhất là AI có thể giúp em học được phát âm, em nghe và phát âm theo từ từ sẽ đúng và AI cũng cung cấp nhiều từ vựng hơn. Em ko thích là phần mềm AI trên máy hay xuất hiện để nghị đánh giá (cho sao) hơi làm	What I like best is that the AI voice chatbots can help me learn pronunciation. I can listen and repeat, and gradually I can pronounce correctly and learn more vocabulary from the AI voice chatbots. What I dislike is that the app on the smartphone usually popped up with recommending for evaluation (giving stars), which

Question 7: What do you like most about speaking with the AI voice chatbots? And what do you dislike? <i>Bạn thích điều gì nhất khi nói chuyện với AI và điều gì bạn không thích?</i>		
Participant ID	VN	EN
	phiền người dùng.	quite annoyed users.
PAR18	Thích là AI biết nhiều cái để nói với em giống như em đang nói với người thật. Chưa thích lắm là có nhiều chữ AI ko nhận được dù em đã nói rất nhiều lần.	What I like is that the AI voice chatbots know many things to speak with me like I am speaking with real human. What I dislike is that there are many words the AI voice chatbots could not recognize although I repeated many times.
PAR19	Em thích nhất là được thực hành nói nhiều làm em nói tốt hơn và phát âm tốt hơn. Em ko thích nhất là nhiều khi em nói hoài AI vẫn ko hiểu ở một số chủ đề.	What I like best is that I can practice speaking many times, which helps me speak and pronounce better. What I dislike best is that sometimes I spoke again and again in some topics but the AI voice chatbots still could not get me.
PAR20	Về cơ bản thì em ko có ý kiến về thích hay ko thích vì đây cũng là chương trình AI đang được cải thiện. Nhưng mà có phần em đánh giá cao là AI có nhiều chủ đề và ngôn ngữ AI thân thiện hơn tuy là ko nhiều chủ đề như nói giữa người với người.	Basically, I have no ideas of what I like or dislike because this is an AI app which is under development process. However, I appreciate that the AI voice chatbots have quite a lot of topics and the language used by the AI voice chatbots are friendly even though the topics are not as many as speaking with human.
PAR21	Thích nhất là mạng nói được phần hồi nhanh. Ko thích ở chỗ AI chưa được hoàn thiện lắm do có 1 vài chủ đề em hỏi mà AI ko hiểu được.	What I like best is the quick of responses. What I dislike is that the AI voice chatbots have not been fully developed because in some topics I spoke a lot but the could not understand.
PAR22	Thích nhất là em có người bạn để luyện tập thường xuyên, em nói thì AI trả lời lại vì ở Việt Nam tụi em ko tìm được người bạn có thể trò chuyện tiếng Anh yên suốt cùng để học. Còn chưa thích thì có nhiều câu em nói AI trả lời lại chưa khớp với câu em nói. Đôi khi em hỏi một đấng AI trả lời một nẻo.	What I like best is that I have a friend to practice speaking frequently because when I say, the AI voice chatbots reply immediately. In Vietnam, it is hard for us to find a friend to speak English with when we study English speaking. What I dislike is that there are many sentences I uttered but the AI voice chatbots could not respond appropriately. Sometimes, I asked one thing but they answered another thing.
PAR23	Thích nhất là em có người bạn để thực hành và học tiếng Anh mọi lúc. Ko thích ở chỗ có những phần AI chưa hoàn thiện nên chưa hiểu ý em hỏi và trả lời còn chưa liên quan đến câu em hỏi.	What I like best is that I have a friend to practice and learn English everytime everywhere. What I dislike is that the AI voice chatbots have not been fully developed. Therefore, they could not understand what I said and replied to me irrelevantly.
PAR24	Thích nhất là mức độ tương tác của AI nhanh với cách AI nói và sử dụng câu từ giúp em học được nhiều từ vựng và cấu trúc ngữ pháp hay. Ko có ko thích.	What I like is the fast interaction of the AI voice chatbots, they way they use language, and the way they use sentences and words, which helps me a lot when learning vocabulary and grammatical structures. Nothing I dislike.

APPENDIX P
INTERVIEW QUESTIONS FOR THE MAIN STUDY

THE PARTICIPANTS INFORMATION SHEET

	<p>Suranaree University of Technology Institutional Ethics Committee</p>	<p>Information Sheet for Participants and Informed consent Form</p>
---	--	---

This informed consent form is for students at Can Tho University (CTU) who are invited to participate in research, titled *"The Effects of An Artificial Intelligence Voice Chatbot on English Speaking Skill of Vietnamese Undergraduate EFL Students"*.

Principle Investigator

Name: Duong Vo Thanh Tam

Email: dvttam@outlook.com

Co-Investigator

Name: Dr. Suksan Suppasetserree

Email: suksan@sut.ac.th

Organization

School of Foreign Languages, Institute of Social Technology, Suranaree University of Technology

Sponsor

Suranaree University of Technology

Project

The Effects of An Artificial Intelligence Voice Chatbot on English Speaking Skill of Vietnamese Undergraduate EFL Students

This Informed Consent Form has two parts:

- Information Sheet (to share information about the study with you)
- Certificate of Consent (for signatures 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 Duong Vo Thanh Tam, a graduate student in the School of Foreign Languages, Institute of Social Technology, Suranaree University of Technology. I am doing research on investigating the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL 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

Speaking skills is very crucial for Vietnamese undergraduate EFL students in learning English. However, previous studies found that it is hard for the Vietnamese undergraduate EFL students to achieve English speaking skill successfully due to insufficient English language proficiency, their limited exposure to English speaking environments, and their unwillingness to speak in English classrooms which have an intertwined relationship with each other. While using technologies may be helpful solutions, the technology like AI voice chatbots for solving the problems was not found in previous studies. Therefore, the main purpose of this study is to investigate and evaluate the effects of an artificial intelligence voice chatbot on English speaking skill of Vietnamese undergraduate EFL students and explore their opinions on using AI voice chatbots to improve English speaking skill.

3. Type of Research Intervention

This research will involve your participation in an orientation and training on speaking English with the AI voice chatbots, joining an in-class English speaking course, practicing English speaking with the AI voice chatbots or with human partners, taking two (online) speaking tests, answering an online questionnaire, and taking an online interview.

4. Participant Selection

You are being invited to take part in this research because you are a non-English majored student studying English 2 at Can Tho University, and volunteer 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 recordings of the pre-/posttests, pre-/post-survey questionnaires, and an interview. The table below provides an overall timeline of the study. It is informed that you can feel free to take part in these meetings if you are willing to. If you decide not to come to the meetings, you can also refuse joining. Your refusal will not cause any effects and threats on your lives or learning grades.

Meeting with researcher	Activity	Duration	Time	Place
First time	Orientation	20 minutes	Week 1	Online (Zoom)
	Pre-speaking test	10 minutes		
Second time	Teaching sessions	3 hours/week	Weeks 2-9	Classroom / Online (Zoom)
Third time	Post-speaking test	10 minutes	Week 10	Online (Zoom)
	Questionnaire	5 minutes	Week 10	Google Form
	Interview	15 minutes	Week 10	Online (Zoom)

*The semi-structured interview will be conducted in Vietnamese by the researcher. The participants will practice English speaking with a partner or an AI voice chatbot for 45 minutes each week.

7. Duration

The experiment takes 10 weeks, starting with the orientation session and then the pretest in week 1. Next, the experiment will continue until week 9. After that, the experiment ends with a posttest in week 10. During week 10, you will also response to the questionnaire and take an individual interview with the researcher.

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 want.

9. Benefits

You will have a chance to join a free English speaking course and more opportunities to practice English speaking with human partners or with the AI voice chatbots to overcome English speaking problems and improve your English speaking skills, which benefits for your English learning process.

10. Reimbursements

You will be able to attend an English speaking class and use the learning website to learn and practice English speaking in class. Moreover, you will be given one more year to use the website for your own learning purposes after you finish the research.

11. Confidentiality

All information will be kept confidential and used for research purposes only. Your identity will be anonymous. Every effort will be made to preserve your confidentiality as follow:

- Assigning code names for participants used on all research notes and documents.
- Keeping notes and interview in a locked file cabinet in the personal possession of the researcher
- Storing electronic data on a password protected computer, 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 result of this study will be published in the researcher's master 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 participant 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 Can Tho 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: Duong Vo Thanh Tam (principal investigator)

Address (in Vietnam): 14/16 Nguyen Thi Minh Khai St., Tan An Ward, Ninh Kieu District, Can Tho City, Vietnam.

Phone (in Vietnam): 0819 759388

Email: dvttam@outlook.com

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 been 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: DUONG VO THANH TAM

Signature of Researcher _____

Date _____

Day/month/year

	<p style="text-align: center;">Đại học Công nghệ Suranaree Ủy ban Đạo đức Cơ sở</p>	<p style="text-align: center;">Thông tin dành cho tình nguyện viên và Đơn đồng ý tham gia nghiên cứu</p>
---	---	--

Đơn đồng ý nghiên cứu này dành cho những sinh viên Trường Đại học Cần Thơ (CTU) được mời tham gia vào đề tài nghiên cứu “Hiệu quả của AI Hội thoại đối với Kỹ năng nói tiếng Anh của Sinh viên Ngoại ngữ bậc Cử nhân đại học Việt Nam”.

Chủ nhiệm đề tài nghiên cứu

Họ tên: Dương Võ Thành Tâm

Email: dvttam@outlook.com

Những người đồng nghiên cứu

Họ tên: Tiến sĩ Suksan Suppasetseree

Email: suksan@sut.ac.th

Đơn vị quản lý

Khoa Ngoại ngữ, Viện Công nghệ Xã hội, Đại học Công nghệ Suranaree

Đơn vị tài trợ

Trường Đại học Công nghệ Suranaree

Dự án

Hiệu quả của AI Hội thoại đối với Kỹ năng nói tiếng Anh của Sinh viên Ngoại ngữ bậc Cử nhân đại học Việt Nam.

Đơn đồng ý tham gia này gồm 2 phần:

- Phần Thông tin (cung cấp thông tin cho bạn về dự án nghiên cứu này)
- Phần xác nhận đồng ý (chữ ký của bạn nếu bạn đồng ý tham gia)

Bạn sẽ nhận được một bản sao y đầy đủ của Đơn đồng ý tham gia nghiên cứu

Phần I: Thông tin

1. Giới thiệu

Tôi là Dương Võ Thành Tâm, nghiên cứu sinh tại Khoa Ngoại ngữ, Viện Công nghệ Xã hội, Đại học Công nghệ Suranaree. Tôi đang thực hiện một nghiên cứu về hiệu quả của AI hội thoại đối với kỹ năng nói tiếng anh của sinh viên ngoại ngữ bậc cử nhân đại học Việt Nam. Tôi sẽ cung cấp các thông tin cho bạn và mời bạn tham gia vào nghiên cứu này. Vui lòng đọc kỹ những thông tin dưới đây. Vui lòng đặt câu hỏi nếu bạn thấy điểm nào không rõ hoặc cần thêm thông tin.

2. Mục đích nghiên cứu

Kỹ năng nói là rất cần thiết đối với sinh viên khi học tiếng Anh. Tuy nhiên, các nghiên cứu trước đây tìm ra rằng sinh viên ngoại ngữ bậc cử nhân đại học Việt Nam khó đạt được kỹ năng nói tiếng Anh do các vấn đề có mối tương quan bao gồm thiếu kỹ năng tiếng Anh, hạn chế môi trường nói tiếng

Anh, và không sẵn sàng nói tiếng Anh trong lớp học. Tuy sử dụng công nghệ là giải pháp hiệu quả cho các vấn đề trên nhưng công nghệ về AI hội thoại chưa được nhắc đến trong các nghiên cứu trước kia. Do đó, mục tiêu chính của nghiên cứu này là tìm hiểu và đánh giá hiệu quả của AI hội thoại đối với kỹ năng nói tiếng Anh của sinh viên ngoại ngữ bậc cử nhân đại học Việt Nam và khảo sát ý kiến của sinh viên về việc sử dụng AI hội thoại để nâng cao kỹ năng nói tiếng Anh.

3. Hình thức nghiên cứu

Trong nghiên cứu này, bạn sẽ tham gia vào một buổi giới thiệu và hướng dẫn về thực hành nói tiếng Anh với AI hội thoại, khóa học Anh văn giao tiếp, thực hành nói tiếng Anh với AI hội thoại hoặc với bạn cùng học, làm hai bài thi nói (trực tuyến), trả lời một bảng câu hỏi khảo sát trực tuyến, và tham gia một buổi phỏng vấn trực tuyến.

4. Lựa chọn tình nguyện viên

Bạn được mời tham gia vào nghiên cứu này vì bạn là sinh viên ngoại ngữ không chuyên tham gia khóa học Anh Văn Căn Bản 2 tại trường Đại học Cần Thơ, và bạn là những tình nguyện viên tham gia nghiên cứu.

5. Tự nguyện tham gia

Việc tham gia vào nghiên cứu này là hoàn toàn tự nguyện. Bạn có thể chọn tham gia hoặc không. Việc lựa chọn của bạn sẽ không có bất kỳ ảnh hưởng nào đến việc học hoặc các đánh giá liên quan trong quá trình học của bạn. Bạn có thể thay đổi quyết định và dừng việc tham gia dù bạn đã đồng ý trước đó.

6. Quy trình

Dữ liệu sẽ được thu thập thông qua các ghi nhận về bài thi đầu/cuối khóa, khảo sát đầu vào đầu ra, và phỏng vấn. Bảng dưới đây cung cấp thông tin về thời gian thực hiện của nghiên cứu này. Theo đó, bạn hoàn toàn có thể tự do tham gia các buổi gặp mặt và thu dữ liệu nếu bạn muốn. Nếu bạn quyết định không tham gia thì bạn cũng có thể từ chối tham gia. Việc từ chối này sẽ không gây bất kỳ ảnh hưởng hay nguy cơ nào đối với cuộc sống hay điểm học tập của bạn.

Gặp gỡ nghiên cứu viên	Hoạt động	Thời lượng	Thời gian	Địa điểm
Lần đầu	Hướng dẫn Kiểm tra nói đầu vào	20 phút 10 phút	Tuần 1	Trực tuyến (Zoom)
Lần 2	Các buổi giảng dạy	3 giờ/tuần	Tuần 2-9	Tại lớp / Trực tuyến (Zoom)
Lần 3	Kiểm tra nói đầu ra	10 phút	Tuần 10	Online (Zoom)
	Trả lời bảng câu hỏi	5 phút	Tuần 10	Google Form
	Phỏng vấn	15 phút	Tuần 10	Online (Zoom)

* Buổi phỏng vấn bằng tiếng Việt được nghiên cứu viên trực tiếp thực hiện. Người tham gia sẽ thực hành nói tiếng Anh với bạn học cùng hoặc với AI chatbot 45 phút mỗi tuần.

7. Thời gian nghiên cứu

Nghiên cứu thực nghiệm kéo dài 10 tuần, bắt đầu với buổi hướng dẫn và sau đó là kiểm tra nói đầu vào. Tiếp theo thời gian thực nghiệm sẽ tiếp tục cho đến tuần 9. Sau đó thời gian thực nghiệm kết thúc với bài kiểm tra nói đầu ra vào tuần 10. Trong tuần 10, bạn cũng sẽ trả lời bảng câu hỏi và tham gia phỏng vấn với nghiên cứu viên.

8. Các rủi ro

Không có rủi ro nào được dự đoán trước cho các quy trình của nghiên cứu này. Bạn có thể từ chối trả lời bất kỳ hoặc tất cả các câu hỏi, và bạn có thể chấm dứt việc tham gia vào nghiên cứu bất kỳ lúc nào nếu bạn muốn.

9. Lợi ích

Bạn sẽ có cơ hội tham gia khóa học nói tiếng Anh miễn phí và có thêm cơ hội thực hành nói tiếng Anh với bạn học chung hoặc AI hội thoại để vượt qua các trở ngại về nói tiếng Anh và nâng cao khả năng cũng như kỹ năng nói tiếng Anh của bạn, điều này sẽ giúp ích cho việc học tiếng Anh của bạn..

10. Bồi dưỡng

Bạn sẽ được tham gia vào lớp học nói tiếng Anh và sử dụng trang web học tiếng Anh để học và thực hành nói tiếng Anh. Đồng thời, bạn sẽ được phép truy cập vào trang web học tập miễn phí trong vòng một năm để tự học nói tiếng Anh.

11. Bảo mật

Mọi thông tin thu thập được từ nghiên cứu này sẽ được bảo mật và chỉ được dùng cho mục đích nghiên cứu. Danh tính của bạn sẽ được ẩn. Người nghiên cứu sẽ nỗ lực bảo mật thông tin của bạn, bao gồm những điều sau:

- Mã hóa tên của tình nguyện viên khi dùng ở tất cả các ghi chú, tài liệu nghiên cứu.
- cất giữ ghi chú và nội dung cuộc phỏng vấn trong một ngăn hồ sơ cá nhân có ổ khóa của người nghiên cứu.
- Lưu trữ các dữ liệu điện tử trong máy tính cá nhân có bảo mật bằng mật mã, và có thể truy cập được bởi người nghiên cứu.

Mọi dữ liệu của tình nguyện viên sẽ được bảo mật, ngoại trừ những trường hợp người nghiên cứu theo pháp lý phải xuất trình trong những tình huống cụ thể.

12. Chia sẻ kết quả

Kết quả của nghiên cứu này sẽ được công bố trong luận văn thạc sĩ của người nghiên cứu và trong các công bố tương lai. Bạn có thể đọc luận văn này thông qua một liên kết tại Thư viện của Trường sau khi luận văn được chấp thuận. Danh tính của bạn sẽ không bị phát hiện trong bất kỳ báo cáo/công bố nào.

13. Quyền từ chối tham gia hoặc hủy tham gia

Việc tham gia vào nghiên cứu này là hoàn toàn tự nguyện. Bạn có quyền chọn tham gia hoặc không nếu bạn đồng ý tham gia, bạn sẽ được yêu cầu ký vào Đơn đồng ý tham gia. Sau khi bạn ký vào Đơn đồng ý tham gia, bạn vẫn có quyền hủy việc tham gia bất kỳ lúc nào mà không cần đưa ra lý do. Việc hủy tham gia sẽ không ảnh hưởng đến việc học tập của bạn tại trường Đại học Cần Thơ. Nếu bạn hủy tham gia trước khi quá trình thu dữ liệu kết thúc, dữ liệu từ bạn sẽ được trả lại hoặc tiêu hủy.

14. Người liên lạc

Họ và tên: Dương Võ Thành Tâm (chủ nhiệm đề tài nghiên cứu)

Địa chỉ (tại Việt Nam): 14/16 Nguyễn Thị Minh Khai, P. Tân An, Q. Ninh Kiều, TP. Cần Thơ, Việt Nam

Điện thoại (tại Việt Nam): 0819759388

Email: dvttam@outlook.com

Đề cương nghiên cứu này đã được xét duyệt và chấp thuận bởi Ủy ban Đạo đức Trong Nghiên cứu có liên quan đến Con người. Trường Đại học Công nghệ Suranaree – một Ủy ban với nhiệm vụ bảo vệ tình nguyện viên khỏi các tổn thương trong nghiên cứu. Nếu bạn muốn tìm hiểu về Ủy ban Đạo đức Nghiên cứu, xin liên hệ với Ủy viên của Ủy ban Đạo đức, Viện Nghiên cứu và Phát triển, Đại học Công nghệ Suranaree, điện thoại: 044-224-757.

Phần II: Xác nhận đồng ý tham gia

Tôi đã đọc các thông tin phía trên. Tôi đã có cơ hội được hỏi và các câu hỏi, nếu có, đều được giải đáp thỏa đáng. Tôi đồng ý tự nguyện tham gia vào nghiên cứu này.

Tên viết in hoa của Tình nguyện viên _____

Chữ ký của Tình nguyện viên _____

Ngày tháng _____

Ngày/tháng/năm

Tuyên bố của người nghiên cứu

Tôi đã đưa/đọc các thông tin cho tình nguyện viên và bằng khả năng tốt nhất của mình, tôi đã đảm bảo tình nguyện viên hiểu sẽ phải làm gì. Tôi xác nhận rằng tình nguyện viên đã có cơ hội để hỏi về nghiên cứu, và mỗi câu hỏi do tình nguyện viên đặt ra đã được giải đáp chính xác bằng khả năng tốt nhất của tôi. Tôi xác nhận rằng tình nguyện viên không bị đe dọa để đồng ý tham gia, và sự đồng ý này là tự do và tự nguyện.

Một bản sao y của đơn đồng ý tham gia này được cung cấp cho tình nguyện viên.

Tên viết hoa của người nghiên cứu: DƯƠNG VÕ THÀNH TÂM


Chữ ký của người nghiên cứu _____

Ngày tháng _____

Ngày/tháng/năm

APPENDIX Q
THE INFORMED CONSENT FORMS

AF/23-08/01.0

	Human research Ethic Committee Suranaree University of Technology	Informed Consent Form
---	--	------------------------------

My name is agedyears old, now living at the address no
 at Street Sub-District/ Ward District.....
 Province..... Postal code: Tel:

I, the signatory of this letter, wish to consent to take part in this research project.

Title of the research project is *“The Effects of An Artificial Intelligence Voice Chatbot on English Speaking Skill of Vietnamese Undergraduate EFL Students”*.

Name of the principal researcher: Duong Vo Thanh Tam

Contact address: 14/16 Nguyen Thi Minh Khai St., Tan An Ward, Ninh Kieu District, Can Tho City, Vietnam.

Telephone number: 0819759388

I **have been notified** of the details of the rationale and the research objectives, details of the activities that I must go through, as well as the benefits to be obtained from this research. I have thoroughly read the details in the information sheet for the research participants and **have received explanations from the researcher so that I clearly understand the information.**

I therefore agree to take part in this research project, as specified in the information sheet for research participants. Concerning this, I consent to do the pre-speaking test, post-speaking test, answer the questionnaires, and join the interview.

I **have the right** to withdraw from the research at any time **without having to state the reason.** This withdrawal will in no way negatively affect my study at my university.

I have been assured that the researcher will treat me in accordance with what is specified in the information sheet for the research participants and any information about me **will be treated by the researcher as confidential.** The research findings will be presented as collective data. No information in the report will lead to identifying me as an individual.

If I am not treated according to what is specified in the information sheet for the research participants, I have the right to file a complaint to The Human Research Ethics Committee, Suranaree University of Technology. Institute of Research and Development

Suranaree University of Technology, 1 1 1 University Avenue, Sub District Suranaree, Muang District, Nakhon Ratchasima 30000, Telephone 044 224757, E-mail: ecsut@sut.ac.th

I have signed my name hereto in the presence of a witness. I have also received a copy of the information sheet for the research participants and a copy of the letter of consent.

(Signature).....

(Duong Vo Thanh Tam)

Principal researcher

(Signature).....


(.....)

Research participant

(Signature).....

(.....)

Witness

	Ủy Ban Đạo Đức Nghiên Cứu Con Người Đại học Công nghệ Suranaree	Mẫu chấp thuận tham gia
---	--	--------------------------------

Tôi tên tuổi:, hiện đang sống tại địa chỉ số tại đường phố
 KV/ Phường Huyện..... Tỉnh..... Mã bưu chính:
 ĐT:.....

Tôi, người ký vào bức thư này, đồng ý tham gia vào dự án nghiên cứu này. Tiêu đề của dự án nghiên cứu là “*Hiệu quả của AI Hội thoại đối với Kỹ năng nói tiếng Anh của Sinh viên Ngoại ngữ bậc Cử nhân đại học Việt Nam*”.

Tên nghiên cứu viên chính: Dương Võ Thành Tâm

Địa chỉ liên hệ: 14/16 Nguyễn Thị Minh Khai, P. Tân An, Q. Ninh Kiều, TP. Cần Thơ, Việt Nam.

Số điện thoại: 0819759388

Tôi **đã được thông báo** về các chi tiết về lý do và mục tiêu nghiên cứu, chi tiết về các công việc mà tôi phải thực hiện, cũng như những lợi ích thu được từ nghiên cứu này. Tôi đã đọc kỹ các chi tiết trong bảng thông tin cho những người tham gia nghiên cứu và **đã nhận được lời giải thích từ nhà nghiên cứu để tôi hiểu rõ thông tin**.

Do đó, tôi đồng ý tham gia vào dự án nghiên cứu này, như được liệt kê trong bảng thông tin cho những người tham gia nghiên cứu. Liên quan đến điều này, tôi đồng ý làm bài kiểm tra nói trước và sau khi học, trả lời bảng câu hỏi và tham gia cuộc phỏng vấn.

Tôi **có quyền** rút khỏi nghiên cứu bất cứ lúc nào mà **không cần phải nêu lý do**. Việc rút tên này sẽ không ảnh hưởng gì đến việc học của tôi tại trường đại học.

Tôi đã được đảm bảo rằng nhà nghiên cứu sẽ đối xử với tôi theo những gì được chỉ định trong bảng thông tin cho những người tham gia nghiên cứu và **bất kỳ thông tin nào về tôi sẽ được nhà nghiên cứu coi là bí mật**. Các kết quả nghiên cứu sẽ được trình bày dưới dạng dữ liệu tập thể. Không có thông tin nào trong báo cáo sẽ dẫn đến việc xác định danh tính của tôi.

Nếu tôi không được đối xử theo những gì được quy định trong bảng thông tin cho những người tham gia nghiên cứu, tôi có quyền nộp đơn khiếu nại lên Ủy ban Đạo đức Nghiên cứu Con người, Đại học Công nghệ Suranaree. Viện Nghiên cứu và Phát triển Đại học Công nghệ Suranaree, 111 Đại lộ Đại học, Tiểu khu Suranaree, Huyện Muang, Nakhon Ratchasima 30000, Điện thoại 044 224757, E-mail: ecsut@sut.ac.th

Tôi đã ký tên của tôi ở đây trước sự hiện diện của một nhân chứng. Tôi cũng đã nhận được một bản sao của bảng thông tin cho những người tham gia nghiên cứu và một bản sao của thư đồng ý.

(Chữ ký)..... (Chữ ký).....

(Dương Võ Thành Tâm) (.....)

Người nghiên cứu chính Người tham gia nghiên cứu

(Chữ ký).....

(.....)

Người làm chứng

CURRICULUM VITAE

Tam Duong received his Bachelor's degree in English Studies from Can Tho University, Vietnam in 2013. Then, in 2021, he started his M.A. study English Language Studies (ELS) in the School of Foreign Languages, Institute of Social Technology, Suranaree University of Technology, Thailand. His research interests include artificial intelligence in education, English speaking skills, technology-enhanced language learning, and educational technologies.

