International Journal of Learning, Teaching and Educational Research
Vol. 23, No. 3, pp. 293-321, March 2024
https://doi.org/10.26803/ijlter.23.3.15
Received Jan 25, 2024; Revised Mar 17, 2024; Accepted Mar 24, 2024

The Effects of an Artificial Intelligence Voice Chatbot on Improving Vietnamese Undergraduate Students' English Speaking Skills

Tam Duong* **and Suksan Suppasetseree**
Suranaree University of Technology
Nakhon Ratchasima, Thailand

Abstract. Speaking English skill is crucial and advantageous for Vietnamese undergraduate students. Speaking English accurately and fluently, however, has proven difficult for these students due to limited English speaking environments, resulting in their insufficient language proficiency and unwillingness to speak. While artificial intelligence (AI) technologies are recommended as effective tools in foreign language education for their flexibility, interactivity, and learner-centeredness, enhancing students' oral communication, little research has considered using an AI voice chatbot to improve students' English speaking skills in Vietnam. Therefore, there are still gaps for this study to be conducted. An 8-week quasi-experiment was conducted with 30 Vietnamese undergraduate students. The participants were informed of research purposes, implementation, and confidentiality. They practiced English speaking with an AI voice chatbot in two class sessions per week, completed a pre-experiment and post-experiment speaking test, responded to a questionnaire, and participated in a semi-structured interview at the end. The results revealed students' significant improvement in English speaking skills after using an AI voice chatbot to practice speaking with $p < 0.05$. The students also agreed that their English speaking skills improved after the intervention because they could speak English better by using suitable hedging words, grammar structures, and vocabulary. The findings pose positive ways for teachers to integrate the AI voice chatbot into their lesson plans for teaching and learning activities. Some limitations were found related to sample size and technological issues. The study brings about a new English learning environment to the Vietnamese EFL undergraduate.

Keywords: artificial intelligence; chatbots; speaking; fluency; accuracy

*Corresponding author: *Tam Duong, dvttam@outlook.com*

1. Introduction

English speaking is a crucial skill EFL learners need to master as it shows learners' knowledge of the English language (Nazara, 2011). Successful English speaking is determined based on how accurately and fluently a speaker uses spoken language during a conversation (Brown & Lee, 2015; Dao, 2017; Manurung, 2015). However, Vietnamese undergraduates have faced difficulties in achieving English speaking skills due to insufficient proficiency, demotivating classroom environments, and limited exposure to English (Dao, 2017; Quyen et al., 2018; Thao & Nguyet, 2019). Similarly, Wang (2014) stated the situation in EFL classrooms where teachers likely spoke in mother tongue language and this situation limited students' English speaking as they did not have many opportunities speaking English in classrooms. The consequence resulted in the students' lack of development and insufficiency in English speaking. Baek and Lee (2018) and Chen and Hwang (2019) pointed out the out-of-date instructional approaches in EFL English classrooms which were not renewed or improved provided insufficient practice, limiting exposure to English speaking. Therefore, overcoming the mentioned problems would be necessary in modern era where a flexible learning and teaching approach might be considered.

According to El Shazly (2021), AI technologies constitute an effective learning tool for EFL learners due to their flexibility, interactivity, and learner-centeredness, they are significant for enhancing oral communication. Among the AI technologies, chatbots are mostly used for education (Colace et al., 2018). With speech recognition and natural language processing algorithms, AI voice chatbots can talk and have discussions as quickly and explicitly as humans (Kaplan, 2016) and are beneficial for classroom applications. While artificial intelligence (AI) is a useable new technology for classrooms applications in terms of improving learners' speaking skills (Ali, 2020), little research has been conducted in Vietnam. Moreover, although the Vietnamese government has encouraged educators to utilise technology in the educational process (Thao et al., 2019), little research in Vietnam has suggested utilising a new technology like artificial intelligence technology to help students improve their English speaking skills (Dao, 2017; Quyen et al., 2018; Thao & Nguyet, 2019). Therefore, this study aimed to fill this gap by exploring the effects of an AI voice chatbot, an application of AI technology, on Vietnamese undergraduate students' English speaking skills. Accordingly, the study's results and findings were expected to enhance learning and teaching experiences by contributing more practical options for classroom application, together with innovative concepts for future research. Therefore, two research questions are posed:

1) What are the effects of an AI voice chatbot on Vietnamese undergraduate students' English- speaking skills?
2) What are the opinions of Vietnamese undergraduate students on using an AI voice chatbot to improve their English speaking skills?

2. Literature Review

2.1 Speaking Skills

Brown and Lee (2015) stated that four crucial skills determining EFL learners' competence in English were reading, listening, writing, and speaking. Among the skills, Nazara (2011) considered speaking as the most vital when EFL students would like to have a successful communication. Theoretically, speaking skill is the learner's ability to communicate comprehensively with listeners in various situations, proving his understanding and acquisition of the language (Brown & Lee, 2015). A successful L2 speaker is evaluated based on fluency and accuracy, demonstrated by smooth and coherent speech with correct pronunciation, grammar, and vocabulary (Walker & White, 2013; Wang, 2014). However, there are many issues regarding students' English speaking skills in the Vietnamese context. Interestingly, Khasawneh (2023) suggested that students should learn English speaking through conversation and authentic contexts so that facilitation could be made optimal through teaching language communication. Within this aspect, using an AI voice chatbot to provide a conversation environment between students and an AI agent is expected to be useful as AI voice chatbots were reported by Kim et al. (2019) that they could speak with users in native English, providing authentic inputs.

2.2 Issues Regarding Students' English Speaking Skills

Dao (2017) implemented a study exploring the barriers preventing Vietnamese undergraduate students from achieving English speaking skills. The study observed 108 undergraduate students learning English speaking in a classroom. A questionnaire was employed to collect data. The findings revealed that the English classrooms in Vietnam provided the students with limited English exposure, which hindered the students' insufficient English language proficiency and demotivated the students' willingness to speak English in class.

Quyen et al. (2018) conducted a study investigating factors that challenged Vietnamese undergraduate students' English speaking skills. The sample included 131 students. A questionnaire, interviews, and classroom observation were used for data collection. The study found that the students' insufficient vocabulary and limited chances to speak English posed difficulties in speaking English successfully. Interestingly, the study reported that the lack of native speakers and opportunities for practicing English speaking in the English classrooms in Vietnam demotivated the students' willingness to speak English.

Thao and Nguyet (2019) examined aspects of undergraduate students' speaking difficulties. There were 150 undergraduates selected. A questionnaire was utilised to collect data. The students were reported to have difficulties in speaking English because they had limited chances to speak English after class and refused to speak with partners. These problems limited the students' exposure to English speaking.

In summary, students had difficulties in English speaking skills because they had limited exposure to English speaking environments (eg. outside classrooms, comprehensible input, time and places), which hindered their language

proficiency. While AI voice chatbots were reported to be flexible in use and provide comprehensible input when speaking as native speakers (Adamopoulou & Moussiades, 2020; Kim et al., 2019), little research reviewed has mentioned using this tool for a solution. Moreover, an Artificial Intelligence (AI) chatbot, one of the widely used technologies, is receiving increasing attention in language learning and it should be implemented to facilitate learning and solve students' oral communication problems.

2.3 Artificial Intelligence Chatbots

Artificial intelligence (AI) chatbots are defined by Haristiani (2019) as the programs featured with AI algorithms which can conduct conversations by audio and text and are able to update knowledge. In exploring the features of AI voice chatbots, Fryer and Carpenter (2006) found that the chatbots could attentively discuss the same material with students through text and speech conversations. Ahmad et al. (2018) reported that AI voice chatbots could understand and answer questions from users all day and support a large number of users at the same time. Çakmak (2022) also mentioned that AI voice chatbots could communicate with students as native speakers, which enhances students' interest in speaking.

A variety of AI voice chatbots were mentioned by Kim et al. (2019) that they could be accounted for studying English such as ELIZA, ALICE, Cleverbot, Elbot, Talk to Eve, Replika, Lyra, Andy English Bot, Poket Friend, Mondly, and Duolingual. Accordingly, there were nine criteria to have an AI voice chatbot evaluated. The first criterium was its ability to understand complex user input. The second criterium was the turn-taking scheme which should be sufficient enough to perform turn-taking thorough the communication. Third, it should be able to recall user's name. Next, it should be capable of supporting multilanguage. Fifth, it should support voice input and output. Sixth, it should also feature text input and output. Seventh, it should have the ability for recalling historical conversations. Eighth, it should be able to answer strange questions. Finally, it should be able to overcome typographical errors. Based on the criteria mentioned above, most of the AI voice chatbots could only perform turn-taking and feature text I/O while some of them could support voice I/O, multilanguage, and recalling historical conversations. Few were able to perform other tasks like understanding complex user input, recalling a user's name, answering complicated questions, or overcoming typos except ALICE, Replika, and Andy English Bot which were dominated. Among the three dominants, Andy English Bot was reported to be more able to facilitate English learning as it could make follow-up questions with a user and encourage the user to learn English speaking by teaching new vocabulary and grammar to the user. In addition, Andy English Bot possessed a huge vocabulary bank and various grammar lessons for a user to learn (Figure 1). Therefore, it can facilitate learners as a virtual English tutor.

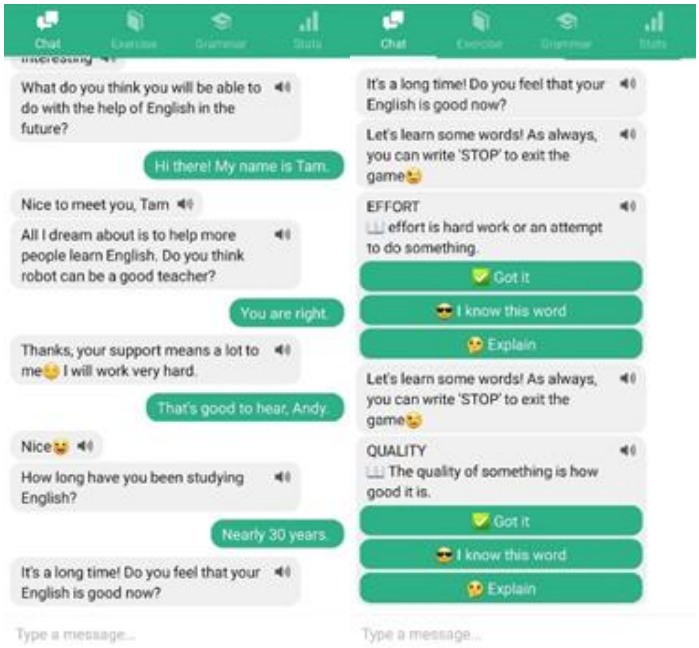

Figure 1: Example chat with Andy English bot

2.4 Features of AI voice chatbots

According to Adamopoulou and Moussiades (2020), AI voice chatbots are featured with three important machine learning techniques including the natural language processing (NLP), the natural language understanding (NLU), and the natural language generation (NLG), illustrated in Figure 2. The NLP was processed to perform analysis process in which the AI voice chatbot classified the message prompted by a user and extracted relevant entities from the built-in responding pattern, allowing the AI chatbot to infer the message before sending responses. The NLU processed an additional function allowing the AI voice chatbot to check spelling, translate, and analyse semantic components of the message. Finally, the NLG generates appropriate responses to the user's prompts.

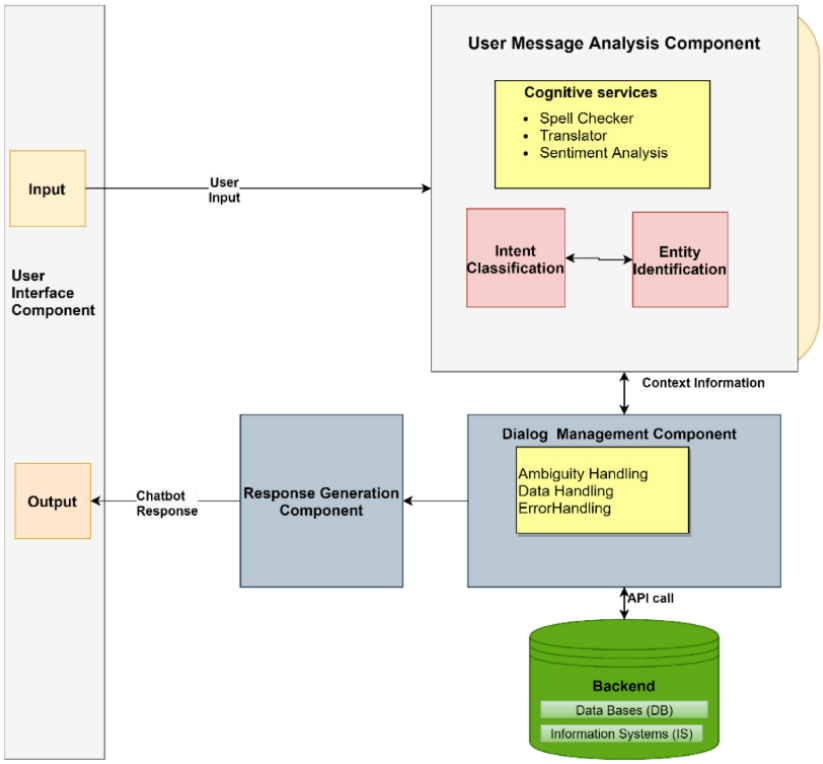

Figure 2: Features of AI voice chatbots (Adamopoulou & Moussiades, 2020, p. 11)

AI voice chatbots were also introduced by previous studies to have additional features such as repeatability, interaction, productivity, and authenticity. Kim (2017) stated that with the process of speech recognition and language processing, the AI voice chatbot could negotiate meaning with learners. Additionally, it could also repeat or rephrase an utterance to support for the negotiating process. Colace et al. (2018) mentioned that using the AI voice chatbot was an innovative approach as it reduced the distance between technology and education, providing interactive and personalized learning experiences to learners and enhancing their language skills. Adamopoulou and Moussiades (2020) mentioned that the AI voice chatbot featured high productivity as it could communicate with a great number of users at the same time on a messenger system. Furthermore, Kim (2018) found that the AI voice chatbot could provide the students with native and human-like English conversation, bringing about authentic input.

2.5 Previous Studies about AI Voice Chatbots in EFL Contexts

AI voice chatbots have been common for their role of assisting learners in EFL educational contexts. Nghi et al. (2019) implemented an experimental study investigating the effectiveness of an AI voice chatbot in facilitating the students' learning of prepositions. There were 200 undergraduate students involving in the study and they were divided into one control group and one experimental group. The experimental instrument was the Facebook chatbot. The instructional instrument was the textbook titled English Pronouns and Prepositions. The learning duration took 15 periods for the control group and 10 periods for the

experimental group. The study provided three useful findings. First, the AI voice chatbot provided the students with new learning experiences during the class meeting, enhancing the students' learning performance. Second, the students showed increasing interests of using the AI voice chatbot for learning because they could feel free to self-practice and self-access their learning. Finally, the AI voice chatbot made a fun learning environment, motivating the students to share their experiences and understanding with their classmates.

El Shazly (2020) conducted a study exploring the role of AI voice chatbots in foreign language anxiety (FLA) management among the EFL students in Egypt. The research involved 48 undergraduate EFL students during the eight-week quasi-experimental period. Pre-speaking and post-speaking tests were utilised to assess the students speaking proficiency level during the intervention. The research adapted an IELTS speaking test rubric for evaluating the level of anxiety the students might get when speaking. Mondly, Audrey, Charles, Cristal, and Mike were the AI voice chatbots used as the tools for the intervention. The research resulted that the students improved their speaking proficiency after the intervention. However, apart from the positive findings such as increasing the students' efforts in speaking, and enhancing their oral performance with cognitive faculties and linguistic abilities, a negative finding revealed that the students still experienced a slight degree of anxiety even after the intervention. Yet, El Shazly (2020) also explained that unfamiliar and irrelevant topics might trigger the students' anxiety.

Kim et al. (2021) conducted a study aiming at investigating the effects of an AI voice chatbot on the EFL learners' speaking skills. The research was conducted with 49 undergraduate EFL students at two different English proficiency levels. The participants took 14-week experiment having voice chats with one of the three AI voice chatbots namely Replika, Andy, and Google Assistant. The studied utilised questionnaires, pre-speaking and post-speaking tests, and an interview to serve as research instruments. The research findings revealed that the participants improved pronunciation, intonation, and word stress after the intervention. However, their fluency was not found to get improved by low-level students while it was significantly found at intermediate level students. Kim et al. (2021) also found positive and negative perceptions from the students. The study positively reported that the students could get more chances to practice speaking as they did not worry about losing face when making mistakes. Their pronunciation, confidence, activeness, and interests of speaking were enhanced owing to the unlimited time they were given for speaking with the AI voice chatbot. Nevertheless, the study found that the students did not feel very comfortable when having voice chats with the AI voice chatbot because the AI voice chatbot did not recognise their voices correctly.

In investigating the interactive effect of an AI voice chatbot toward the students' speaking performance and anxiety, Çakmak (2022) engaged 90 EFL students from a Turkish university and asked them to practice English speaking with Replika during a 12-week period. The study used a questionnaire to explore the students' perception of practicing English speaking with Replika. The study

found both positive and negative findings. Positively, practicing English speaking with Replika effectively enhanced the students' English speaking performance compared with face-to-face practice. Negatively, the research reported that students' anxiety was not reduced because they worried about how to make Replika understand them correctly in some contexts, Therefore, Çakmak (2022) suggested that using an AI voice chatbot interaction proved to be an effective way to help the students improve their English speaking performance while it was not a useful way to reduce the students' anxiety.

2.6 Technology-Enhanced Language Learning

According to Walker and White (2013), The Technology-Enhanced Language Learning or TELL is the transformation from Computer-Assisted Language Learning or CALL to a broader scope of language learning in which a variety of tools and devices such as mobile phones, tablets, games, and virtual worlds are engaged in the learning process instead of merely using computers for learning. The technology-enhanced language learning (TELL) theory determines technology as the facilitator for learners in achieving learning objectives (Devlin et al., 2013). For a wider view, Walker and White (2013) discussed TELL based on approaches. Within the approaches, technological devices are seen as environmental resources which facilitate the process of English communication and interaction teaching. Learning is guided by connectivism theory (Siemens, 2005) which involves the structural, cognitive, socio-cognitive, and adaptable aspects. Within this perspective, connectivism involved with behaviorism (Skinner, 1957), constructivism (Piaget, 1964), and social constructivism (Vygotsky, 1978). Furthermore, TELL indicates the normalisation of technology in education and views learners' autonomy as a principal learning objective. Table 1 illustrates the TELL approaches.

Table 1: TELL Approaches (Walker & White, 2013)

Approach	**TELL**
Technology	Mobile devices, tablets, multiplayer games, virtual worlds
English-teaching paradigm	Communication, interaction
View of language	Structural, cognitive, socio-cognitive, adaptable
Principal use of technology	Normalised
Principal objective	Autonomy with community
View of learning	Connectivism
Role of technology	Environment resource

2.7 Conceptual Framework

This study was conceptualised with the connectivist paradigm guided by TELL (Walker & White, 2013). Within the TELL approaches, the process of teaching and learning English speaking can be divided into three stages: (1) pre-speaking, (2) while-speaking, and (3) post-speaking. Figure 3 presents the conceptual framework for developing the lessons.

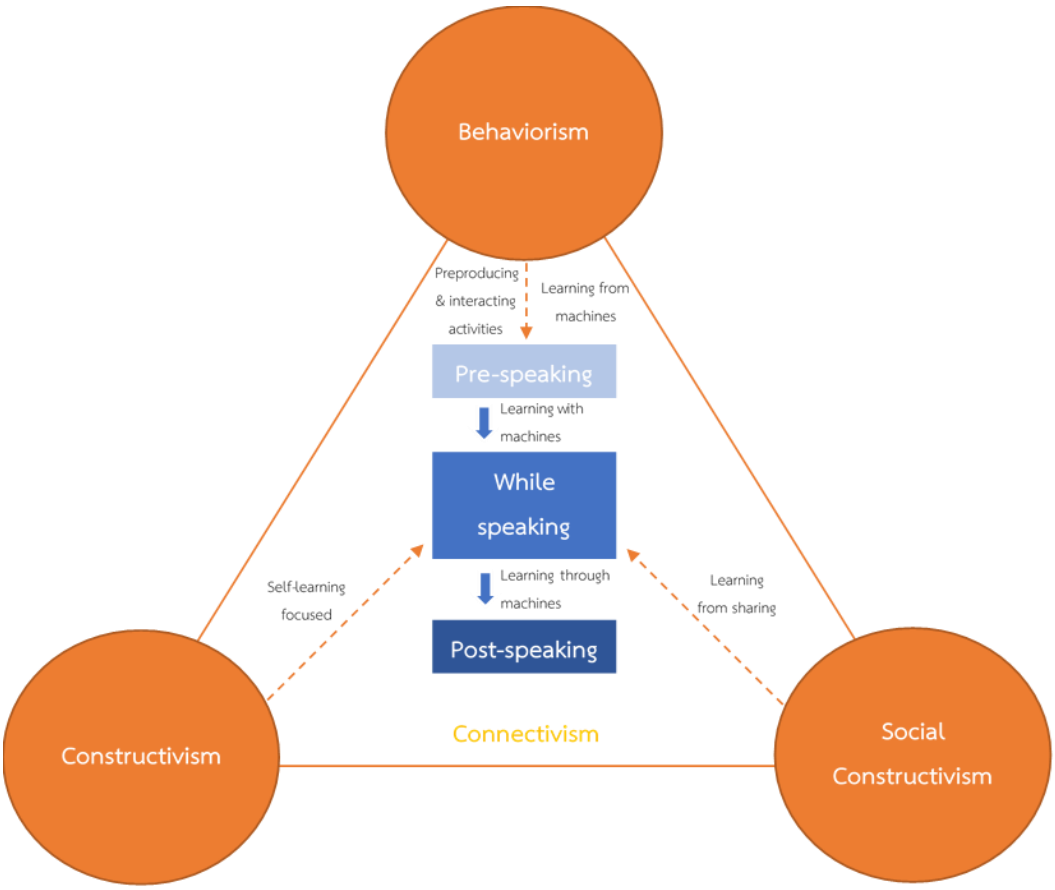

Figure 3: Conceptual framework for developing lessons

The pre-speaking stage is processed under the theories of behaviorism (Skinner, 1957) and connectivism (Siemens, 2005), and allows for the students to learn through interacting in group and pair work activities in which machines play the role of instructors.

The while-speaking stage is processed under constructivism (Piaget, 1964), social constructivism (Vygotsky, 1978) and connectivism (Siemens, 2005), and is reflected in individual practicing of activities and pair work, encouraging students to learn from a variety of ideas and develop communication with non-human appliances.

The post-speaking stage is processed through individual tasks guided by behaviorism (Skinner, 1957), constructivism (Piaget, 1964), social constructivism (Vygotsky, 1978), and connectivism (Siemens, 2005) to promote students' self-study ability. Machines can be used by students to facilitate the process of solving tasks.

In summary, the concept of learning English speaking skills follows the connectivist paradigm in which learning is a connecting process through behaviorist to constructivist and social constructivist perspectives. Technological devices are utilised as facilitators thorough the process. Therefore, this conceptual framework is considered the blueprint in this study.

3. Research Methodology

3.1 Research Design

In designing the research, a quasi-experiment was selected and a mixed-method design was applied for answering two research questions because the study contained one independent variable which was the use of an AI voice chatbot in teaching English speaking skill and two dependent variables including the students' speaking skills in the speaking tests and their opinions on practicing English speaking with an AI voice chatbot. The experimental period took place for eight weeks as to avoid Hawthorne effect which might probably lead to participants' over achievement when they knew that they were involved in the research (Kim, 2018). The pre-speaking test, the post-speaking test, and a questionnaire were utilised to collect quantitative data while a semi-structured interview was conducted to collect qualitative data. The research was implemented within three steps: (1) Instrumental design, (2) data collection, and (3) data analysis. Figure 4 outlines the research implementation process. At the beginning, the research instruments and instructional instruments were designed and tested for their accuracy and validity through a pilot study. In the second step, the data collection procedure started with the employment of the pre-speaking tests to pre-evaluate the students' English speaking ability before they took the intervention. Next, online teaching sessions were conducted with the intervention in which students were allowed to learn and practice English speaking with the AI voice chatbot during their learning process. Within the online teaching sessions, the participants were requested to practice English speaking with Andy English Bot thorough two class sessions every week, lasting for six weeks. After that, the participants were asked to take a post-speaking test to evaluate their English speaking skills after the intervention. Finally in step two, the participants further responded to the questionnaire and the interview to share their opinions after the intervention. After the data collection procedure, the final step was processed with data analysis in which the data from the speaking tests and responses of the students to the questionnaire and interview were analysed.

Figure 4: Research implementation

3.2 Participants

There were 30 non-English major Vietnamese undergraduate students participating in the research. The students were enrolled in an English 2 course at Can Tho University, Vietnam and their homogeneity was assured based on their placement scores in the TOEIC Bridge test from the screening process conducted by the university before enrolling in the course. After the screening process, only the students who got the overall band score from 43 to 75 (equivalent to A2 level) qualified for English 2. The participants included 13 males and 17 females aged from 19 to 23 and they initially possessed technological experience before taking the intervention. In compliance with ethical issues, the participants were informed of the research purpose, implementation, and were ensured their participation would remain confidential. Consent forms were administered to all participants to confirm their willingness to participate in the research.

3.3 Research Instruments

The pre-speaking test and the post-speaking test were developed to evaluate the students' English- speaking skills in terms of speaking fluency and accuracy. This study adopted the British Council's IELTS speaking test structure which contains three parts. In part 1, the students orally answered questions about food, study, their hometown, daily activities, or family and friends. Part 2 encouraged the students to orally share opinions on a given topic about books, people, countries, exercise, or festivals. In part 3, students were required to give further explanations of their opinions about the topic they discussed in part 2 by answering some questions asked by the examiner. The IELTS speaking test

structure was adopted in this present study because the test's validity and reliability were confirmed by Li (2019) that it can be used for assessing students' speaking skills through various criteria such as fluency, coherence, lexical resources, grammatical range, accuracy, and pronunciation. In designing the speaking tests, the CEFR's principles for assessment were applied to serve evaluating purposes. The CEFR's principles suggested five aspects to consider designing tests including context, purpose, linkability, production, and standards. Accordingly, the speaking tests were designed to match with level A2 of the non-English majored undergraduate EFL students who was learning English 2. The testing purpose was to evaluate the students' speaking skill in terms of fluency and accuracy. The tests were designed to provide linkable results to help with examiners' assessment. The questions and topics of the tests could be used again and again for English speaking assessment, which followed the principle of production. Finally, the tests' structure was standardized with the IELTS speaking test structure developed by the British Council. In addition, the test structure included three speaking parts as to ensure that test takers' scores might demonstrate their true competence as multiple measures might provide reliable and valid assessment than a single measure.

A questionnaire was developed to explore the students' opinions on using the AI voice chatbot and contained three parts: general information, close-ended statements, and an open-ended statement (see Appendix 1). The first part included six items to identify the students' gender, age, major, year of study, English placement score, and technological experience. The second part included five items to explore the students' opinions on using the AI voice chatbot. The final part was to allow the students to share further opinions on learning English speaking skills by using the AI voice chatbot. A five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) following the framework of Cohen et al. (2007) was adopted to measure the students' statements on the questionnaire. For instrument validity, the questionnaire items were checked and verified by three experts in the field of English language teaching by using the index of item objective congruence (IOC) introduced by Turner and Carlson (2003), ranging from -1 (Incongruent), 0 (Questionable), to +1 (Congruent). All items of the questionnaire were rated +1 by all three experts and the average IOC score was 1.00. Therefore, all items were employed.

In investigating the students' opinions on using the AI voice chatbot, a semi-structured interview constructed with the non-directive framework suggested by Cohen et al. (2007) was utilised and contained seven guided questions arranged thematically, referring to the practicing process, speaking accuracy improvement, speaking fluency improvement, and the students' feelings (see Appendix 1). The interview questions were also checked and verified by three experts in the field of English Language Teaching based on the IOC. All guided questions were rated +1 by the three experts and the average IOC score was 1.00. Therefore, all items were employed.

3.4 Instructional Instruments

The study used Andy - English Speaking Bot downloadable on Google Play or App Store through https://andychatbot.com/ as the tool for the students to practice English speaking skills. A learning website was created to facilitate the students' learning processes, featuring five lessons developed based on the lesson plans. Specific instructions were provided on the website to guide the students' self-study in improving their English speaking skills.

A speaking assessment rubric was designed by using the analytic method directed by Ulker (2017) to grade the students' English speaking in the pre-test and the post-test. The rubric included five columns presenting five levels of speaking achievement and two rows representing two criteria of accuracy and fluency assessments.

The lesson plans were designed to encourage the students' self-study ability. The CEFR's Can-Do Statements suggested by Cambridge (2011) were the basis for designing the lesson plans as they provide a description of what language learners can do at different learning stages and determine how learners can engage in a discussion or a conversation through interaction, production, listening, reading, and writing.

3.5 Data Collection

Both quantitative and qualitative data were collected to find answers to the research questions. Accordingly, the results from the speaking tests were used to answer research question 1 while the questionnaire and interview responses were used to answer research question 2. Figure 5 presents the outline for the data collection procedure.

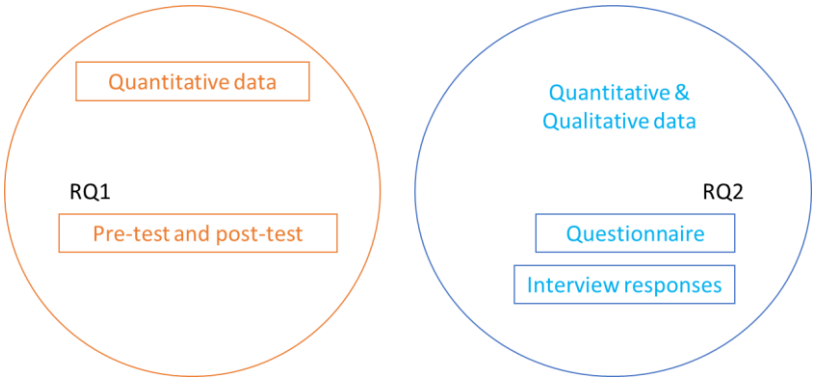

Figure 5: Outline of data collection procedure

The data collection procedure was processed with four steps. First, the participants took a pre-speaking test for input assessment. Second, they learned five English lessons on the website and practiced English speaking with Andy. After that, they took a post-speaking test for output assessment. Finally, they all answered the questionnaire and twelve of them attended the semi-structured interview. The selection process followed the informant sampling method suggested by Schreiber and Asner-Self (2011) with the aim of purposively selecting and talking with the participants. Accordingly, six students with the

highest scores and six with the lowest scores were selected from the post-speaking test to participate in the interview. The interview took around 10-15 minutes in Vietnamese so that the participants could easily find ideas and use their mother tongue language to share opinions about using an AI voice chatbot to practice their English speaking skills. The interview sessions were recorded into audio files to serve for transcription purposes. The transcription was made bilingually in tables including three main columns containing participant code, their responses in Vietnamese, and their responses translated into English.

3.6 Data Analysis

For quantitative data, the paired t-test guided by Schreiber and Asner-Self (2011) was employed to compare the results of the pre-test and post-test to investigate the effects of an AI voice chatbot on improving Vietnamese undergraduate students' English speaking skills and investigate the statistical significance of the intervention based on the p-values. Descriptive and frequency statistical methods were used to find the mean scores (M) of the questionnaire responses in each item. This was done to explore significant variables related to the students' opinions on improving English speaking accuracy and fluency after using an AI voice chatbot. SPSS version 20 was employed to analyse the data.

For qualitative data, the thematic analysis method suggested by Schreiber and Asner-Self (2011) was applied because this method was reported to be appropriate for grouping data within specific themes. Consequently, it would allow researchers to easily interpret data following themes. Under the thematic analysis method, the study classified the interpretation into four themes to consider the relevance of each theme to the research objectives. Accordingly, the interview responses were classified into four themes: practicing process, speaking accuracy improvement, speaking fluency improvement, and the students' feelings. The aim was to get more in-depth information from the students about using an AI voice chatbot to improve their English speaking skills.

4. Results

4.1 Results of Quantitative Data Collection

In investigating the effects of an AI voice chatbot on improving the students' English speaking skills, the results of the pre-test and post-test were collected and analysed. The paired t-test in Table 2 showed that the overall p-value of Pair 3 (pre-test and post-test) was 0.000 ($p < 0.05$; statistical significance), which meant that the participants significantly improved their English speaking skills after the intervention. In addition, the p-values of Pair 1 (fluency in pre-test and post-test) and Pair 2 (accuracy in pre-test and post-test) were 0.000 ($p < 0.05$; statistical significance), which indicated that the participants significantly improved speaking fluency and accuracy after the intervention.

Table 2: The results of paired t-test

Paired t-tests	Paired differences					T	df	Sig. (2-tailed)
	Mean (M)	Std. deviation	Std. error mean	95% confidence interval of the difference				
				Lower	Upper			
Pair 1 (FluencyPre-FluencyPost)	-.4333	.3198	.0584	-.5528	-.3139	-7.421	29	.000 *p<0.05*
Pair 2 (AccuracyPre-AccuracyPost)	-.2833	.2780	.0508	-.3872	-.1795	-5.582	29	.000 *p<0.05*
Pair 3 (Pretest-Posttest)	-.8167	.4676	.0854	-.9913	-.6421	-9.565	29	.000 *p<0.05*

4.2 Results of Qualitative Data Collection

In exploring the students' opinions on using the AI voice chatbot, their responses to the questionnaire were collected and analysed with descriptive and frequency statistical methods. Table 3 presents the students' demographic information. There were 30 students participating in the research and the majority were females with 17 respondents (56.7%) while 13 respondents were males (43.3%). The participants aged from 19 to 23 (M = 20.67) and were mostly in the second year of study (M = 2.73). Their average English placement test score in the TOEIC Bridge test was M = 53.8 which was equivalent to A2 level (43-75). All participants came from non-English majors and had technological experience before taking the intervention.

Table 3: The students' demographic 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%

Table 4 illustrates the descriptive statistics result which revealed that the participants agreed about the improvement of their English speaking skills after using the AI voice chatbot (M = 3.90, SD = 0.77). The mean score of the students' opinions on using an AI voice chatbot to help them speak English with appropriate hedging words was the highest (M = 4.07, SD = 0.74) and the lowest mean score was that after practicing with the AI voice chatbot, the students thought that they could speak English without making too many pauses and hesitations (M = 3.73, SD = 0.74).

Table 4: The students' opinions on using an ai voice chatbot

Students' opinions	N	Mean (M)	Std. Deviation (SD)	Interpretation
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.	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
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
	Total	3.90	0.77	Agree

The mean scores reported in Table 4 were interpreted by adapting the interpretation framework suggested by Banditvilai (2016), illustrated in Table 5.

Table 5: The interpretation of mean scores in the questionnaire responses

Mean scores (M)	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

In investigating the students' opinions on using the AI voice chatbot in detail, a semi-structured interview was conducted with twelve students. The students' responses to the interview questions were interpreted by employing four themes: practicing process, speaking accuracy improvement, speaking fluency improvement, and the students' feelings respectively.

4.2.1 Students' Practicing Process

When asking about the students' practicing process, this study found that the students talked with the AI voice chatbot for around 5-10 minutes when they were free. Participants 2 and 5 said:

> *"Normally I practiced English speaking with the AI voice chatbot for around 10-15 minutes when I was free."*

> *"Normally I spoke around 3-4 times a week when I was free. I usually spoke for around 10-15 minutes each time with the AI voice chatbot."*

4.2.2 Speaking Accuracy Improvement

When asked about opinions on improving speaking accuracy in terms of grammar, most of the students reported that they could use grammar more accurately after speaking with the AI voice chatbot because it used correct grammar structures, helping them learn and follow. Moreover, they could see their uttered sentences on the screen to check their grammar. Participants 1 and 5 shared:

> *"When I spoke, the screen displayed the chat texts so that I could observe and correct my grammar. The AI voice chatbot also used correct and comprehensive grammar structures."*

> *"I had improved in grammar a lot because the AI voice chatbot used correct grammar when speaking so that I could see it on the chat screen to learn and improve myself."*

Regarding pronunciation, most of the students reported that they could speak with better pronunciation because of two reasons. First, they could listen to the AI voice chatbot's speaking voice multiple times and practice pronunciation because it sounded like a native speaker. Second, when they uttered a sentence with the wrong pronunciation, they could see the incorrect words appear on the screen, and then they had to try to speak more carefully with the correct pronunciation. Participants 5 and 6 reported:

> *"I could listen to the AI voice chatbot for many times and repeat after it to practice my pronunciation. It spoke like a native speaker."*

> *"I could look at my words on the chat screen to know if I had pronounced the words correctly or not so that I could correct myself."*

In terms of vocabulary, the students reported that their vocabulary improved because the AI voice chatbot used many new words during conversations and it could suggest and explain new vocabulary for them to learn during conversations, which helped them improve their vocabulary. Participants 1 and 3 reported as follows:

> *"I could ask the AI voice chatbot for an explanation to a new word."*

> *"The AI voice chatbot usually suggested more vocabulary related to the topics we were talking about, which inspired me to learn more to be able to chat with it."*

4.2.3 Speaking Fluency Improvement

When asked about speaking fluency improvement, the students answered that they could speak with fewer hesitations compared with their first time speaking with the AI voice chatbot. They explained that they spoke with a lot of hesitations and pauses for the first time, which made the AI voice chatbot misrecognise what they said. Therefore, they had to practice speaking again until they could speak fluently enough so that the AI voice chatbot could recognise their sentences correctly. Other students reported that the AI voice chatbot could wait for them for a long time, which make ease for them to find

ideas and speak. Therefore, they felt more comfortable and confident and could speak with fewer hesitations and pauses. Participants 6 and 7 stated:

> *"I could ask the AI voice chatbot for an explanation to a new word."*

> *"The AI voice chatbot usually suggested more vocabulary related to the topics we were talking about, which inspired me to learn more to be able to chat with it."*

4.2.4 Students Feelings

When asked about their feelings after using the AI voice chatbot, the students reported that they liked to speak with the AI voice chatbot for some reason. First, they had more opportunities to speak English as they could continuously speak with them at anytime and anywhere. Second, they were familiarised with real-life situations as the AI voice chatbot spoke like a native speaker. As Participants 6 and 17 reported:

> *"The AI voice chatbot frequently spoke, asked, and changed topics to keep the communication unstoppable."*

> *"I could talk and practice speaking with it anytime and anywhere. Its voice is like a native speaker, which familiarises me when speaking with a real native speaker."*

However, some participants reported their dislikes. First, the AI voice chatbot could not get Vietnamese names even though they tried to pronounce the names the best they could. Second, there were not many topics to speak about with the AI voice chatbot. In such cases, they found that the AI voice chatbot failed to respond to some topics. Participants 3 and 27 informed:

> *"The AI voice chatbot was still limited in terms of speaking topics. I wanted to speak about more topics but it still could not do that."*

> *"The AI voice chatbot was still limited in speaking topics and it failed to recognise the names of some places in Vietnam."*

5. Discussion

This study investigated the effects of an AI voice chatbot on Vietnamese undergraduate students' English speaking skills and their opinions about using an AI voice chatbot to practice English speaking skills. Both quantitative and qualitative data were collected and analysed. The following discussion elaborates on the results reported previously.

5.1 The Effects of an AI Voice Chatbot on Students' English Speaking Skills

The results from the paired t-test showed statistical significance between the pre-test and post-test ($p < 0.05$), which indicated the students' improvement in English speaking skills. This finding is in line with Çakmak (2022) saying that EFL students could improve their speaking performance after practicing English speaking with an AI voice chatbot.

The study found improvements in speaking accuracy as the students increased their scores in accuracy in terms of grammar, vocabulary, and pronunciation

after the intervention. This was because the AI voice chatbot used a variety of grammar structures and vocabulary (Figure 6) to speak with the students, and it could also speak with a native voice.

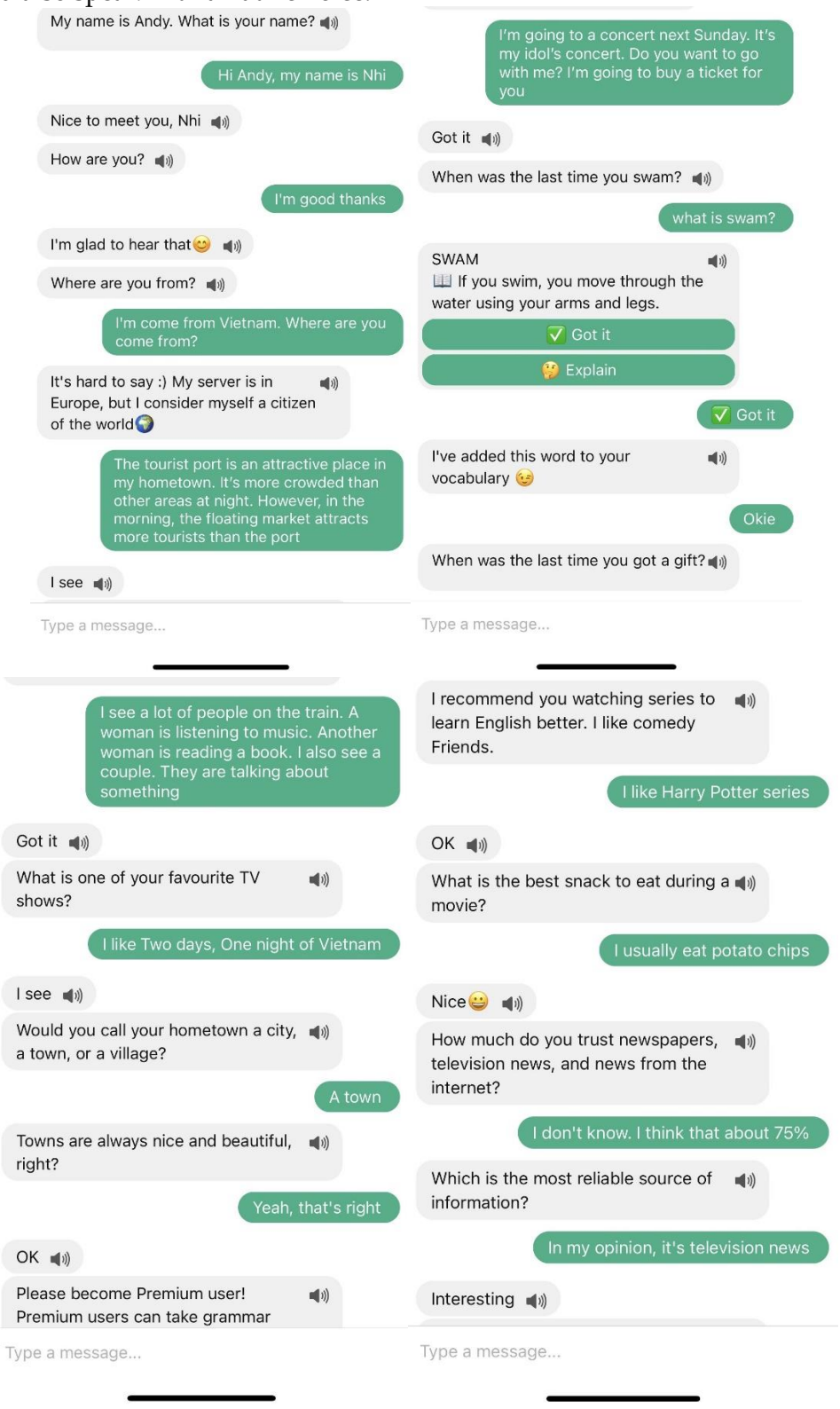

Figure 6: Illustrations for the conversation between Andy English bot and some participants

This finding is in line with Siemens (2005) stating that an important learning objective is to know more rather than knowing what. Within the Connectivist paradigm, students were connected with an expert source of language knowledge like the AI voice chatbot so that they could learn more vocabulary and grammar while speaking. In addition, the chatbot could speak again and again with a native voice to help students expose to a virtual English language environment. This environment might motivate the students to speak more and help them overcome speaking difficulties (Dao, 2017; Quyen et al., 2018; Thao & Nguyet, 2019). The finding is also in line with Çakmak (2022) that AI voice chatbots could communicate with students as native speakers, which ensures authentic language input for English learners. Interestingly, the AI voice chatbot could recognise the students' grammar mistakes for using quantifiers with different types of nouns and for using verb tenses. This finding is interesting because previous studies did not mention any findings alike (Kim, 2017; Kim, 2018; Ahmad et al, 2018; Colace et al., 2018; Adamopoulou & Moussiades, 2020).

In terms of fluency improvement, the students could speak more fluently after the intervention as the AI voice chatbot could allow unlimited time for students to think and speak, promoting a relaxing atmosphere during a conversation so that the students could feel more comfortable, confident, and motivated to speak English. This stimulation improved the students' speaking fluency. This finding is also in line with Dao (2017), Quyen et al. (2018), and Thao and Nguyet (2019) suggested that when students are provided with more English speaking opportunities and exposure, they can overcome speaking difficulties and feel more confident in speaking English.

5.2 Students' Opinions on Using an AI Voice Chatbot

The findings of both the questionnaire and interview showed the students' agreement on their improvement in English speaking skills by using the AI voice chatbot. Çakmak (2022) similarly shared this finding as the students agreed that the AI voice chatbot helped them improve their English speaking skills.

The exploration of the students' opinions delves into four themes: practicing process, speaking accuracy improvement, speaking fluency improvement, and students' feelings about using the AI voice chatbot.

5.3 Practicing Process

The study found that the AI voice chatbot could talk with the students without time and place limitations. The students could talk with it anytime and anywhere whenever they were free. Therefore, they could feel increasingly free to arrange their time to practice English speaking, when compared to speaking with a friend, which relied mostly on each other's availability. This finding agrees with Ahmad et al. (2018), mentioning that features of AI voice chatbots are that they can speak with a user all day and be of service to many users at the same time. This has led to the AI voice chatbot becoming highly productive when applied for teaching and learning purposes.

5.4 Speaking Accuracy Improvement

In investigating the students' opinions on improving speaking accuracy, the study recognised grammar, pronunciation, and vocabulary as three improving aspects. All students reported on their improvement in English speaking skills because they could speak with accurate grammar, vocabulary, and pronunciation after the intervention.

First, the students improved grammar because the AI voice chatbot featured native voice and never used wrong grammar in conversations. Therefore, students could learn correct grammar structures while speaking. Moreover, the chatbots' use of various grammar structures encouraged the students to learn more and enriched their linguistic proficiency. This enforcement crucially enhanced the students' speaking grammar. This finding agrees with Walker and White (2013) suggesting that expert instructors who possess expertise language knowledge are crucial to support learners.

Second, the students improved their pronunciation because the AI voice chatbot encouraged them to speak again and again to improve their pronunciation. Actually, the chatbot could not respond appropriately to the students' utterances if they talked with incorrect pronunciation. In such cases, the AI voice chatbot misrecognised the students' uttered words with other unrelated words and responded to the students inappropriately. This forced students to try to practice pronunciation many times to communicate more effectively. Moreover, while practicing pronunciation, students could also listen to the AI voice chatbot's native voice many times and imitate pronouncing the English utterances. This finding is connected with the findings of Ahmad et al. (2018), Fryer and Carpenter (2006), and Çakmak (2022) that AI voice chatbots could talk like native speakers and discuss with students on the same material repeatedly.

Finally, the students improved their vocabulary because the AI voice chatbot could suggest and explain new words to them during conversations. In fact, the AI voice chatbot is specifically featured for teaching vocabulary and grammar to students. Therefore, while talking with students, it can automatically suggest new words related to the topic of the conversation and explain the meanings of the words to students, from which students can learn, listen, and repeat them to improve both vocabulary knowledge and pronunciation. This finding seems to solve students' difficulties in English speaking caused by insufficient vocabulary knowledge reported by Quyen et al. (2018) because the students can learn and update their vocabulary knowledge by usually talking with the AI voice chatbot and then learning vocabulary thorough conversations.

Interestingly, this study found that the students became autonomous when speaking with an AI voice chatbot thanks to the function of the chat screen and vocabulary suggestions and explanations. When talking with the AI voice chatbot, the students could look at the chat screen to check if they used the correct grammar and pronunciation or not so that they could make necessary revisions in subsequent conversations. Furthermore, they could also proactively

learn vocabulary while having conversations with the AI voice chatbot. This sheds light on the learner's autonomy which is the principal objective of TELL.

5.5 Speaking Fluency Improvement

In terms of speaking fluency improvement, the study found that the students could improve their speaking fluency because the AI voice chatbot encouraged them to speak again and again until they could make the smooth flow of speech. It could also provide the students with a friendly atmosphere for communication, which reduces their stress and strengthens their confidence and motivation to speak English. This finding is in line with Fryer and Carpenter (2006) stating that AI voice chatbots can create a relaxing environment for EFL learners when they try to speak English.

5.6 Students' Feelings

This study further explored students' feelings about talking with the AI voice chatbot. The students reported with positive and negative feelings. On the one hand, the AI voice chatbot piqued the students' interest in communicating because of the native voice function. Quyen et al. (2018) reported that most of the students could not speak English well because they could not find a native speaker to practice speaking English with them, which limited their speaking chances after class. While they might speak with their friends or teachers who were non-native speakers, they could not speak at anytime and anywhere like speaking with the AI voice chatbot. Therefore, the AI voice chatbot might bring more speaking chances to the students and encourage them to speak more frequently. Furthermore, thanks to the AI voice chatbot's compliments during the conversations, the students might feel motivated and unpressured, resulting in them speaking more.

On the other hand, the AI voice chatbot was limited to communication topics and voice recognition. The limitation in speaking topics is due to the database of the AI voice chatbot. According to Adamopoulou and Moussiades (2020), the conversational database is stored in the backend which is responsible for suggesting responses for AI voice chatbots to answer user's questions. Therefore, if the database is not updated continuously, then an AI voice chatbot can only afford limited topics. However, with AI algorithms and AI developers' assistance, an AI voice chatbot can definitely learn and update its knowledge to overcome this problem itself. For the limitation in voice recognition technology, the AI voice chatbot could not understand Vietnamese names because it was designed and integrated with American or English native voices. Therefore, the articulation of the native voices might be very different from the articulation of the Vietnamese voice, thus, preventing the AI voice chatbot from recognising Vietnamese names by using the English nicknames instead might be a useful solution for the students.

6. Implications

Based on the findings, this study confirms that the AI voice chatbot has improved students' English speaking skills in terms of accuracy and fluency and provided the students with a relaxing English speaking environment. The

research suggests some implications for the use of the AI voice chatbot in the classroom as well as suggestions for future research.

For pedagogical implications, it is suggested that teachers integrate the AI voice chatbot into their lesson plans for teaching and learning activities. However, the designed activities should follow clear objectives and purposes. For example, teachers may set specific speaking goals for students when speaking with the AI voice chatbot (eg. How many times should they take to repeat their utterance to get the AI voice chatbot understand and respond correctly to their utterance because wrong pronunciation might lead to misunderstanding from the AI voice chatbot). Clear instructions on how to use the AI voice chatbot should be provided to students so that they can be ready for the learning activities. For instance, teachers may have to teach students how to register for a free account to practice speaking with the AI voice chatbot, how to turn on their microphone from their smartphone, or how to make it speak in English with students. Notably, a ready-made AI voice chatbot like Andy is highly recommended in case teachers are new to script writing. In applying this process, teachers should allow students to talk with the AI voice chatbot while learning vocabulary thorough conversations to enhance students' vocabulary knowledge and autonomous learning. Also, if teachers would like to use the AI voice chatbot, they should carefully investigate its functions to select the most suitable one to serve their specific educational purposes. Likewise, teachers can find support from a developer who can intervene in the architecture and update the knowledge database of the AI voice chatbot so that it can afford more speaking topics.

For learning purposes, this study suggested that students at intermediate level learn English speaking with Andy English Bot so that they can practice speaking more easily. This also benefits students' confidence in speaking because it may lower students' risk of feeling inferior when they make mistakes. In order to self-assess, students should also use Andy English Bot when practicing English speaking because it can help them spot their grammar errors more easily, learn more vocabulary and check for pronunciation mistakes. This also improve students' self-awareness of learning and practicing. In addition, students should take advantages of the user interface when speaking with Andy English Bot because it not only communicates with the native voice, but it can also display the messages as texts, facilitating students' learning process with visual aids. Finally, students should spend as much time speaking English with Andy English Bot because the more time they spent, the competence they become.

For research implications, the effects of an AI voice chatbot on other language skills such as reading, writing, or listening should be considered because students' English competence relies not only on productive skills but also on receptive skills. Besides, future research may focus on finding the effects of an AI voice chatbot on different learning contexts (other school levels, different proficiency levels, or specific groups of learners) and other subject areas to bring new learning experiences to students as well as to evaluate if an AI voice chatbot is effective in these areas.

7. Limitations

Within the scope of this study, there were still limitations found in terms of sample size and unexpected issues related to technology. For the sample size, the number of 30 students were quite objectively limited in this study as the time the study was conducted was during summer time. As a result, there were only two classes of English 2 opened by the university during the conducting time. Ideologically, a greater sample size would bring a greater chance for the study to get more identical results in the quasi-experiment. In terms of technological defects, the participants using out-of-dated Android operating system were not successful installing and using Andy English Bot on their smartphone as only Android OS 13.0 or above was featured by the app. These unavoidable limitations should be carefully concerned for future research in this field of interest.

8. Conclusion

The present study was conducted to investigate the effects of an AI voice chatbot on Vietnamese undergraduate students' English speaking skills and to examine the students' opinions on using an AI voice chatbot to improve their English speaking skills. The results of the speaking tests revealed that the students improved their English speaking skills after the intervention. The questionnaire responses showed that the students agreed on their improvement in English speaking after using an AI voice chatbot, and the semi-structured interview responses elaborated on how the students improved English speaking accuracy and fluency. For pedagogy, an AI voice chatbot may provide a creative setting for Vietnamese education. If English speaking classrooms are integrated with the AI tool, a relaxing atmosphere and more chances may be experienced by students, increasing exposure to improve English speaking skills. If it is skillfully applied by teachers, it is possible to increase students' interests and excitement in English classrooms, motivating them to speak more. If it is carefully used by students, it provides a stress-free learning environment, bringing more confidence to students to learn and improve English speaking skills. Careful instruction should be given in the AI-assisted learning environment. Finally, comparison studies should be conducted to compare students' improvement in other productive skills as they may provide in-depth scientific evidence to broad readers. Furthermore, the effectiveness of AI-assisted learning toward students' enhancement in language skills would also be potential for a broader research area. In addition, other EFL contexts and AI-assisted learning tools may also be suggested to future researchers to bring a broader picture of how AI technologies can do in such contexts to support educators and learners' methods during teaching and learning processes.

9. References

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, *2*, 100006. 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), 7-10. https://doi.org/10.5120/ijca2018917606

Ali, S. G. A. (2020). Using an artificial intelligence application for developing primary school pupils' oral language skills. *Journal of Education, 75*(75), 67-110. https://doi.org/10.12816/EDUSOHAG

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.

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.

Cambridge ESOL. (2011). *Using the CEFR: Principles of good practice*. University of Cambridge.

Chen, M. R. A., & Hwang, G. J. (2019). Effects of a concept mapping-based flipped learning approach on EFL students' English speaking performance, critical thinking awareness and speaking anxiety. *British Journal of Educational Technology*, *51*(3), 817-834. https://doi.org/10.1111/bjet.12887

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge. https://doi.org/10.4324/9780203029053

Colace, F., Santo, M. D., Lombardi, M., Pascale, F., Pietrosanto, A., & Lemma, 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://article.sciencepublishinggroup.com/html/10.11648.j.her.20170202.12.html

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

El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems, 38*(3), e12667. https://doi.org/10.1111/exsy.12667

Fryer, L., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology*, *10*(3), 8-14. http://dx.doi.org/10125/44068

Haristiani, N. (2019). Artificial intelligence (AI) chatbot as language learning medium: An inquiry. *International Conference on Education, Science and Technology, 1387(*1), 012020. https://doi.org/10.1088/1742-6596/1387/1/012020

Kaplan, J. (2016). *Artificial intelligence what everyone needs to know*. Oxford University Press. https://doi.org/10.1093/wentk/9780190602383.001.0001

Khasawneh, M. A. S., (2023). Factors affecting the improvement of speaking skills among Jordanian EFL learners. *Journal of Language Teaching and Research, 14*(6), 1559-1568. https://doi.org/10.17507/jltr.1406.14

Kim, H.-S., Cha, Y., & Kim, N. Y. (2021). Effects of AI voice 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. (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

Li, J. (2019). An evaluation of IELTS speaking test. *Open Access Library Journal, 06*, e5935. https://doi.org/10.4236/oalib.1105935

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

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

Nghi, T. T., Phuc, T. H., & Thang, N. T. (2019). Applying AI voice chatbot for teaching a foreign language: An empirical research. *International Journal of Scientific & Technology Research*, *8*(11), 897-902.

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

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

Schreiber, J., & Asner-Self, K. (2011). *Educational research*. 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

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

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), 581-594. https://doi.org/10.33100/jossh5.5.TranQuocThao.etal

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

Ulker, V. (2017). The design and use of speaking assessment rubrics. *Journal of Education and Practice, 8*(32), 135-141.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Havard 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), 110-118. https://doi.org/10.5539/elt.v7n2p110

Appendix 1

QUESTIONNAIRE

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:	☐Male ☐Female
2. Age:	☐18 ☐19 ☐20 ☐21 ☐22 ☐23
3. Major:	☐Medicinal Chemistry ☐Primary Education ☐Other *(please specify):* ....
4. Year of study:	☐1st year ☐2nd year ☐3rd year ☐4th year
5. English placement test score:	
6. What technology have you ever used for learning English speaking? *You can select more than one option listed in this section.*	☐ Podcast ☐ Voice chat apps (i.e. Zoom, Zalo, LINE, etc.) ☐ Chatbots (i.e. Andy English Bot, ELIZA, Siri, Google Assistant, etc.) ☐ Learning websites (i.e. SpeechAce.com, IELTS Liz.com, etc.) ☐ Youtube ☐ Social network (i.e. Facebook, Twitter, Reddit, etc.) ☐ 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.	○	○	○	○	○
8. After practicing with the AI voice chatbot, I think I could speak English with appropriate hedging words such as uhm… ah… oh....	○	○	○	○	○
9. After practicing with the AI voice chatbot, I think I could produce correct pronunciation when I spoke English.	○	○	○	○	○
10. After practicing with the AI voice chatbot, I think I could use appropriate sentence structures when I spoke English.	○	○	○	○	○
11. After practicing with the AI voice chatbot, I think I could use appropriate words and vocabulary when I spoke English.	○	○	○	○	○

Part 3: Additional suggestions or comments on learning English speaking skill by using an artificial intelligence voice chatbot *(if any):*

.....

.....

.....

.....

.....

THE INTERVIEW QUESTIONS

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
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?