Applied Research in English Education
Vol.3 No. 1 e-ISSN: 2962-1607
https://jurnal.unimus.ac.id/index.php/AREE/index

Impacting Artificial Intelligent Chatbot In Flipped Classroom To Enhance Students' Speaking Competence

Tri Yuli Setyoningrum [1*]
Dias Andris Susanto [2]
Arso Setiaji [2]

[1]Universitas PGRI Semarang, INDONESIA
[2]SMAN 1 Blora, INDONESIA

Abstract

Traditional teaching methods have changed in recent years due to the integration of technology into education. The use of artificial intelligence (AI) in education has received the greatest attention. The flipped classroom concept is one area where AI has demonstrated promise. However, one challenge teacher's face in language learning contexts is providing opportunities for students to practice speaking skills effectively. The purpose of this study is to examine how AI Chatbots in flipped classrooms can enhance students' speaking competence. This study employed a mixed method research design. Students were split into two groups: the experimental group received instruction via an AI Chatbot before speaking class, and the control group received instruction using a conventional approach. A pre-test and a post-test were administered preceding to and following treatment, separately. Five participants were interviewed and a questionnaire was distributed to all participants in order to measure students' perceptions. The study's findings showed that teaching speaking through AI Chatbots in flipped classrooms was more effective than the conventional approach. Pupils who received instruction from an AI Chatbot in a flipped classroom were much more likely to pass KKM. The benefits of deploying AI Chatbots in flipped classrooms are also evident from students' questionnaire responses, which show that students enhanced their vocabulary, pronunciation and gained confidence. Briefly stated, AI Chatbots in flipped classroom can be utilized as a substitute for traditional classroom instruction in speaking, and they have been shown to boost students' confidence in addition to their speaking competence.

Keywords: AI Chatbot, Flipped Classroom, Speaking Ability.

1. INTRODUCTION

Traditional teaching methods have changed in recent years due to the integration of technology into education. As mentioned by Wang (2019); Wulandari et al.,(2024); Tonapa et al., (2024); Al Aziz et al.,(2023); Retno et al., (2023) that Since there are numerous ways that modern technologies might improve English education, it is necessary to integrate modern technology into English teaching and learning to change the conventional teaching methods. Besides Szymkowiak,

Corresponding Email: trisetyoningrum67@guru.sma.belajar.id

(2021); Solekah et al., (2023); Sari et al., (2023); Aini et al., (2023) conducted a study to identify how technology and the Internet impact Generation Z's knowledge acquisition process as well as the types of information learning that this generation finds attractive. The result showed that the respondents were more likely to choose to study using mobile applications and video content than through traditional methods. It was also found that students tended to follow in the steps of their teachers who utilized modern technologies for learning outside of the classroom andincluded them in their curricula. Besides that, based on the mini research done by the Students Assembly of SMAN 1 Blora in 2024 students express that they want teacher to use digital media and audio visual media on the learning process.

Of these developments, the use of artificial intelligence (AI) in education has received the greatest attention. There are some studies done by researchers on this topic. One of the studies done by Silitonga (2023); Riswalastika et al., (2023); Affini et al., (2023) the findings demonstrates that AI chatbot-based instruction improves students' motivation to study English composition. AI chatbots can give kids more detailed and comprehensive feedback than a teacher can give them alone. Students may be better able to identify their areas of strength and weakness and make more focused progress as a result.

The flipped classroom concept, in which students complete outside speaking and participate in active learning during class, is one area where AI has demonstrated promise. Sudarmaji (2021); Mawarni et al., (2021); Cahyadi et al., (2021) said that his study proved that not only did the fully online flipped classroom design engage the students with the English curriculum, but it also significantly improved their speaking abilities. Additionally, students gained greater confidence when speaking in English in class.

However, one challenge teacher's face in language learning contexts is providing opportunities for students to practice speaking skills effectively. Based on the data taken from preliminary research done in SMAN 1 Blora, it was found that most students feel that speaking English is difficult. The data showed that the reason of their difficulties are their limitation of vocabularies, the lack of partner to practice speaking, and their limited time to practice speaking in English. So, the impact of using AI chatbots in flipped classrooms to improve students' speaking skills is examined in this thesis.

2. LITERATURE REVIEW

2.1. Review of Previous Study

2.1.1. Previous Study on the Use of AI in Education

Artificial Intelligence has been applied in all sectors, including the education sector.There are many studies about it. Zhai (2021) did research on the use of AI in Education from 2010-2020 and stated that Given AI's fast development, it is crucial that teachers understand how educators might best apply AI strategies to students' academic performance.

AI technology is developing quickly, and shortly, it is anticipated that its use in education will increase significantly Zhang & Aslan (2021). Artificial intelligence (AI) has enormous potential in education, especially in the areas of extending learning possibilities, scaling up personalized learning experiences, and optimizing techniques and strategies for targeted learning goals Bromber (2020).

The impact of AI on education has been extraordinary Singh & Hiran (2022). Because of this, artificial intelligence (AI) has had significant effects on education in general and on the use of technology at specialized educational institutions in particular. Teachers may now accomplish administrative duties more quickly and efficiently, such as grading assignments and giving students feedback, due to AI. Using several types of artificial intelligence, such as chatbots (Chat) and cooperative robotic agents (CORA), can enhance the quality of training. Because AI can evaluate students' abilities and requirements and then create and disseminate customized or customized information that promotes improved uptake and retention which subsequently increases learning, students can receive a better and deeper learning experience.

2.1.2. Previous Study on the Use of AI in Speaking Class

Some studies have been conducted about the use of AI in speaking classes. Hapsari & Wu (2022) stated that using AI chatbots for casual conversations before previous speaking classes could help the learning process and learning goals more effectively than conventional models. Talking with an AI chatbot beforehand can help students become less nervous while speaking in front of others, enjoy learning more, and develop their critical thinking skills. Madhavi et al. (2023) stated by using appropriate ICT and AI technologies, learners can effectively improve their speaking abilities. All speech test characteristics are performed better by ICT students than by non-ICT students. Technology and AI tools for language learning and teaching greatly enhance the communication abilities of ESL students.

A research done by Hamuddin et al. (2020) mentioned that when it came to speaking English, students who were instructed with the Lyra virtual assistant performed better than those who weren't. Speaking practice is beneficial when using the Lyra virtual assistant. The Lyra virtual assistant helps students improve their speaking abilities.

2.1.3. Previous Study on the Use of Flipped Classroom in Speaking Class

There are many studies about the use of Flipped Classrooms in speaking classes. Pre- and post-oral proficiency tests, observation, and focus group interviews were employed by Abdullah et al.(2019) in his research to investigate the impact of the flipped classroom paradigm on English speaking ability. Pre- and post-oral proficiency exams showed a significant difference between the control class and the experiment class employing a flipped classroom, according to the results of the paired-sample t-test. Additionally, it can be seen that students' commitment level, English-speaking performance, and willingness to participate in conversational English assignments have all gradually increased.

The use of flipped classrooms for speaking classes has been studied at high schools in Vietnam by Hương & Hưng (2023) The majority of students, according to the interview findings, have improved their fluency, pronunciation, understanding, grammar, and vocabulary. They have also gained new knowledge and are able to pronounce the words correctly. The students were all very pleased with their progress. The results of this study showed that students' language acquisition, particularly their speaking abilities, was enhanced utilizing flipped education, a type of blended learning.

Not only flipped classrooms increase the ability in speaking skills, but they also increase students self-regulated learning as stated by Putu Indra Kusuma (2020). According to the findings, students who received instruction in flipped classrooms excelled over their control group peers in speaking and Self-Regulated Learning (SRL). Furthermore, through three stages—forethought, performance control, and self-

reflection—the students established their SLR, which has four potential implications for English language teaching.

2.2. Review of Theoretical Study

2.2.1. Artificial Intelligent

The term artificial intelligence (AI) is typically closely related to computers. However computers may have served as the foundation for the development of artificial intelligence, it is clear—especially in the context of the education sector—that fewer and fewer people now consider computers, their hardware, and software, or their equipment to be artificial intelligence Chen et al. (2020).

Other definition was written by Reddy (2023), he stated that The goal of artificial intelligence (AI) is to create computers that can behave and respond appropriately, changing their course to meet the needs of the moment. These computers ought to behave in a way that is similar to what is thought to require intelligence in humans.

Thus, in addition to helping with decision-making, artificial intelligence can also simplify other elements of our lives. As a result, AI is gradually penetrating every aspect of our daily lives and is a necessary component of all contemporary technology. It is necessary for climate control that is automated. Artificial intelligence is present in everything, whether it be a washing machine, microwave, or automated manufacturing facility in the automotive or industrial sectors. In addition to increasing accuracy, these AI-based tools and methods have made it practicable to complete tasks that would not have been done without AI. (Kulkani, P., & Joshi,P.(2015).

2.2.2. Chatbot

Khan & Das (2018) mentioned that the classic definition of a chatbot is a computer software that interprets user input in natural language and produces perceptive and pertinent responses to be returned to the user. Currently, artificial intelligence (AI) or rules-driven engines that largely communicate with consumers through text-based interfaces enable chatbots. These are stand-alone computer programs that can be integrated with any of the many messaging services—Facebook Messenger, Slack, Skype, Microsoft Teams, and others—that have become available to developers through APIs.

Chatbots, also known as conversational agents, promote human-computer interaction using natural language by utilizing natural language processing (NLP). Chatbots are rapidly becoming common because they replicate human discussions and so automate service. They are now deployed in numerous sectors, such as healthcare, business services, education, and academic advice.(Kuhail et al., 2023)

2.2.3. Speaking

2.2.3.1. Speaking Definition

There are four skills in learning English: Speaking, listening, speaking, and writing. Those four language abilities are most frequently discussed in language instruction according to their modality and direction Bailey & Nunan (2019). Writing or speaking produced by the students is seen as a productive language while speaking or listening to the students is regarded as receptive language. The term "modality" describes the language's medium, which might be written, spoken, or both. Speaking is hence a productive oral skill.

Speaking is simply the physical expression of abstract systems Widdowson,(1978). Stated differently, speech is the external expression of notions, assumptions, and ideas that exist in the human mind. It gives meaning to abstract concepts by transforming them into sounds and utterances.

2.2.3.2. The Significance of Speaking in an EFL Context

In the classroom, speaking is important for three basic reasons. First, speaking exercises give teachers insight into their students' proficiency levels and the kind of speaking issues that they most frequently face. Speaking assignments allow students to engage in "real-life speaking" practice. Finally, students develop an automatic usage of language elements (word order, tenses, and parts of speech) the more they speak, activating their understanding of language structures and functions. Speaking thus improves students' accuracy and fluidity Lightbown & Spada (2013).

3. METHODS

3.1. Participants

In this study, a mixed methods research strategy was employed. The researchers developed a hypothesis, gathered quantitative information, and performed data analysis in the first investigation. The type of data gathered in a second study that includes the collection, analysis, and interpretation of qualitative data was decided by the results of the quantitative investigation. To further explain the quantitative findings, the researchers employed qualitative interpretation.

Twelfth graders made up the research's population, and the samples are as follows:

Tabel 1. Sample of study

No	Class	Number of students
1	XII A1	36 students
2	XII A2	36 students
Total		72 students

A few tools were used to gather the data. Students' speaking abilities were assessed by a test, and their perceptions were investigated using questionnaires, interviews, and observations.

3.2. Experimental Procedure

The experimental investigation was conducted over five meetings, each lasting 90 minutes. Every participant completes the pretest during the first meeting. Additionally, they are given an introduction to the course they are enrolled in. An orientation to using an AI chatbot is given to the experimental group. During the first meeting, the control group is introduced to traditional learning. Once every student has completed the pretest, the main experiment begins. The students learnt how to express agreement or disagreement on a particular issue throughout the course of three meetings, from the second to the fourth. The posttest was the focus of the most recent conference. Figure 1 shows the experimental process.

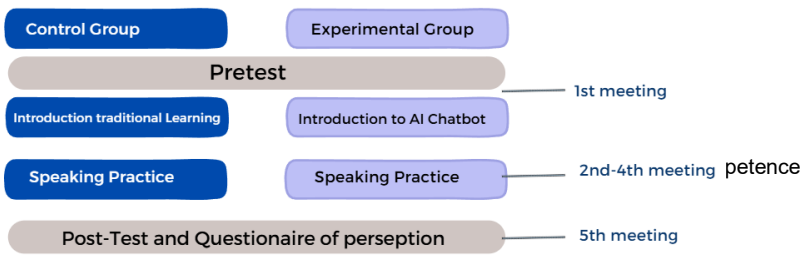

Fig 1. Experimental Procedure

Through real-time feedback, the experimental group's students can engage with the AI chatbot to learn and expand their knowledge. They had fifteen minutes to communicate with the AI chatbot at home before speaking class on their mobile devices. Students in the control group gained knowledge from their instructor's teachers.

Questionnaires, interviews, observations, tests, and documentation can all be used to establish the method's appropriateness. Two observers, both English instructors, participated in the teaching and learning observation. Additionally, each participant received a questionnaire. Five students were interviewed to find out how is their perceptions about the use of AI Chatbots in speaking classes. Because participants in this study were asked the same topics, a structured interview was employed, with the following guidelines for the interview:

1. How do you feel about using an AI chatbot to learn speaking?
2. What are the advantages of using an AI chatbot to learn how to speak?
3. What challenges do you face while using an AI chatbot to learn how to speak?
4. How can you get past the challenges you face when using an AI chatbot to learn to speak?

Quantitative data was collected through speaking (pre-test and post-test). The pretest measures participants' abilities prior to treatment, while the post-test evaluates an attribute or characteristic of experiment participants following treatment. Five criteria—fluency, accuracy, clarity, pronunciation, and performance skill—were employed in the test.

This study employed SPSS, which includes the t-test, homogeneity test, and normality test, to analyze the data. The following formula can be used to determine the normalcy test criteria:

a. If the sig. value (p-value) is more than 0.05, it means the data is normally distributed.

b. If the sig. value (p-value) is less than 0.05, it means the data is not normally distributed.

4. RESULT AND DISCUSSION

As indicated in the methodology, 72 participants were divided into two groups for this study. The experimental group, XII A1, received instruction from an AI chatbot before they had speaking class, whereas the control group, XII A2, received instruction via a traditional way. The researcher starts their investigation by confirming the hypotheses before testing. These were made prior to the hypotheses being verified. The normality and homogeneity tests are the targets of the assumption test. Following the collection of pre-test and post-test data, this analysis was carried out.

Table 2. Normality Test Result

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Class	Statistic	df	Sig.	Statistic	df	Sig.
Score	Experimental pretest	.129	36	.137	.954	36	.138
	Experimental posttest	.150	36	.039	.945	36	.073
	Control pretest	.147	36	.046	.942	36	.057
	Control posttest	.186	36	.003	.964	36	.280

a. Lilliefors Significance Correction

According to the data presented in table 2, where in decision-making the sig. value in the Shapiro-Wilk column must be greater than 0.05 (sig. >0.05). In the experimental class, the sig. value of the pretest was 0.138 and in the posttest was 0.073 the sig. value in the experimental class and the control class >0.05 or it can be concluded that the data is normally distributed. Meanwhile, in the control class, the sig. value for the pretest was 0.057 and for the posttest was 0.280 the sig. value in the control class for the pretest and posttest >0.05 or it can be concluded that the data is normally distributed.

To determine if the data was homogeneous or not, I employed the Levane Test of Homogeneity of Variance, which may be used to calculate the homogeneity test requirements. The next assumption test was homogeneity.

a. If the sig. value (p-value) is more than 0.05, it means the data is homogenous.
b. If the sig. value (p-value) is less than 0.05, it means the data is not homogenous

The analysis of the homogeneity test data is presented in the table 3:

Table 3. Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Score	Based on Mean	1.276	3	140	.285
	Based on Median	1.222	3	140	.304
	Based on Median and with adjusted df	1.222	3	139.010	.304
	Based on trimmed mean	1.300	3	140	.277

According to the data presented in Table 3 and referring to the determination of the significance value, which is >0.05. The significance value in the table refers to the average obtaining a calculated result of 0.285 >0.05. It can be concluded that the obtained data is homogenous data.

According to the homogeneity and normalcy tests. It can be concluded that both experimental classes' pretest and posttest results were homogeneous and regularly distributed. In order to evaluate the hypotheses, those statistical data might then be used to examine the means that differ among the research variables.

To ascertain how independent factors affect the dependent variable, the hypothesis test is employed. The following is an explanation of the research findings:

Table 4 provided an explanation of the outcome before being instructed by an AI chatbot.

Table 4. Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Pretest Experimental Class	36	25	80	58.75	13.222
Pretest Control Class	36	25	80	55.42	14.850
Valid N (listwise)	36				

In Table 4 the number of samples tested is presented as 36 students with a minimum score of 25, a maximum score of 80, and an average score of 58.75. The decision for students to pass is those who have a score > 75. Based on the results of the grade processing, 8 students were declared passed and 28 students were declared not passed in both the experimental and control classes.

In testing the hypothesis "is there a significant difference in students' speaking ability before and after using regular learning" in this section. The researcher used a paired t-test using the SPSS application. In decision-making, it still refers to the sig. (2-tailed) value <0.05, which indicates a significant difference between the pretest and posttest scores before and after being taught using regular learning. Here are the results of the data analysis:

Table 5. Pired Samples Statistic of Control Class

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pretest Control Class	55.42	36	14.850	2.475
	Posttest Control Class	71.94	36	12.322	2.054

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Pretest Control Class-Posttest Control Class	-16.528	7.253	1.209	-18.982	-14.074	-13.673	35	.000

The table 5 shows that the sig. (2-tailed) value of 0.00 < 0.05 indicates a significant improvement in student learning outcomes after the learning process in the control class. Based on these results, it shows that both the pretest and posttest

scores of the control class, which used conventional teaching methods, experienced an improvement.

In testing the hypothesis "is there a significant difference in students' speaking ability before and after using the AI Chatbot" in this section. The researcher used a paired t-test using the SPSS application. In decision-making, it still refers to the sig. (2-tailed) value <0.05, which indicates a significant difference between the pretest and posttest scores before and after being taught using the AI Chatbot. Here are the results of the data analysis:

Table 6. Pired Samples Statistic of Experimental Class

Paired Samples Statistics

			Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pretest Class	Experimental	58.75	36	13.222	2.204
	Posttest Class	Experimental	81.39	36	11.564	1.927

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Pretest Experimental Class - Posttest Experimental Class	-22.639	12.563	2.094	-26.890	-18.388	-10.812	35	.000

sig. value (2-tailed) 0.00 < 0.05 or it can be concluded that there is a significant improvement in student learning outcomes after using the AI chatbot in the experimental class

In table 6, the data processing results show that the mean difference for the pretest and posttest scores in the experimental class is 22.639 with a standard deviation of 12.563, a t-value of 10.812, and a sig. (2-tailed) value of 0.000. From the results of the data analysis, it can be seen from the sig. (2-tailed) value of 0.000<0.05 that there is a significant difference between the pretest and posttest scores before and after using artificial intelligence Chatbot technology to improve students' speaking skills

In testing the hypothesis "is there a significant difference in the speaking ability of students using conventional learning and AI Chatbot" in this section. The researcher used the independent Sample T Test type using the SPSS application. In decision-making, it still refers to the sig. (2-tailed) value <0.05, which indicates a

significant difference between the posttest scores of the experimental class and the control class. Here are the results of the data analysis:

Table 7. Difference between the post test scores of the experimental class and the control class

Group Statistics

	Class	N	Mean	Std. Deviation	Std. Error Mean
Score	Experimental Class	36	81.39	11.564	1.927
	Control Class	36	71.94	12.322	2.054

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Score	Equal variances assumed	.490	.486	3.353	70	.001	9.444	2.816	3.827	15.062
	Equal variances not assumed			3.353	69.720	.001	9.444	2.816	3.827	15.062

At the sig. (2-Tailed) value of 0.001 < 0.05, it can be concluded that there is a significant difference between the scores in the experimental class and the control class.

From the results of the data analysis, it can be seen from the sig. (2-tailed) value of 0.000<0.05, which concludes that there is a significant difference between the pretest and posttest scores before and after using the artificial intelligence Chatbot in students' speaking abilities.

To answer the hypothesis "how students perceive the use of AI Chatbot to improve speaking skills," the researcher used a data collection method in the form of a questionnaire distributed to students who have used the AI Chatbot. The decision-making technique uses the scale range in table 8:

Table 8. Scale range of students perception

Range of Average	Predicate
<1	Strongly Disagree
1-2	Disagree
2-3	Undecided
3-4	Agree
4-5	Strongly Disagree

Here are the results from the questionnaire distribution obtained by the researcher:

Table 9. Result of Questionnaire Distributions

No	Average	Note
1	3,56	Agree
2	4,06	Strongly Agree
3	4,22	Strongly Agree
4	4,33	Strongly Agree
5	4,11	Strongly Agree
6	4,03	Strongly Agree
7	4,14	Strongly Agree
8	4,03	Strongly Agree
9	4,22	Strongly Agree
10	4,19	Strongly Agree

According to the table 9, for question point no. 1, an average score of 3.56 was obtained, which falls under the "agree" category, while for question points 2-10, the "strongly agree" category was achieved. In addition, the average score from all questions is 4.09 with the predicate of strongly agree. It can be concluded that students' perceptions of learning using AI Chatbot can improve students' speaking skills.

Furthermore, the following results of this study focused on how students perceived the use of AI chatbots to enhance their speaking skills. The following are the findings from an interview regarding students' opinions of using AI chatbots to enhance their speaking abilities:

Using AI Chatbot in flipped Classroom helped students perform better in speaking class.The benefits using AI Chatbot in Flipped Classroom were students get used to with the vocabulary used in the speaking class, increased pronunciation and increased their confidence.

The obstacles of using AI Chatbot in flipped classroom is that the students need more internet quota when using the AI Chatbot in the Flipped Classroom, the language used is too robotic, the time is limited only 10 minutes because they used free AI Chatbot. The solution given by students in accrodance to those obstacles are they use AI Chatbot daily for a week it means they practice 7 times in 10 minutes every day. Besides, they paraphrase the response of AI Chatbot use their own words so it is no longer robotic.

Speaking is the ability to communicate using spoken language. Speaking in English is one of the activities that will serve as a foundation for students' language skills in the future. Krisita Omega Putri et al (2023) in their research also concluded that speaking is the ability to articulate words in order to convey or express intentions, ideas, thoughts, and feelings that are organized and developed according to the needs of the audience so that what is conveyed can be understood by the audience. It cannot be denied that for students, boredom during the learning process results in a lack of interest in the media used by the teacher. Strategies or methods that are still conventional and not varied have not yet been able to improve students' speaking skills in the learning process.

Moreover, the implementation of the Merdeka curriculum requires teachers to be more creative and innovative in their teaching. Students can be given a variety of learning activities in new ways that have not been used by them before, making them more interested in using the media provided. This research is in line with the study by Krisita Omega Putri et al (2023), which showed that before the introduction of variations, students' speaking skills were very low. The same was also investigated by Marzuki et al (2021) through their observations, which indicated that students' speaking skills in English language learning were still very low.

Students' speaking abilities before the use of AI chatbots generally reflect the limitations of traditional school learning, everyday experiences, and limited access to modern learning technology. In the context of formal education, students typically learn speaking skills through methods that focus on grammar, sentence structure, and the use of vocabulary determined by the curriculum. As a result, their speaking abilities are often limited to the topics taught in class, with little opportunity to improvise or adjust conversations according to real-life situations. Most students also tend to only be able to speak in formal contexts or highly structured situations, such as reading dialogues from texts or giving presentations based on scripts, so their ability to speak spontaneously and naturally is undertrained.

The test results show a sig. (2-tailed) value of 0.00 < 0.05, which means there is a significant improvement in student learning outcomes after the lesson in the control class. Based on these results, it shows that both the pretest and posttest scores of the control class using regular teaching methods showed an improvement.

The improvement in students' speaking skills in the control class through conventional learning was evident, although with relatively limited results compared to other innovative methods. At the pretest stage, students' speaking abilities were generally low due to a lack of active speaking experience in class and anxiety about making mistakes. Many students show limitations in fluency, vocabulary, and sentence structure. However, after conventional learning that includes methods such as group discussions, simple presentations, and dialogues, there was an improvement in speaking skills. This is due to the targeted practice provided during the learning process, the repetition of material that reinforces the mastery of grammar and vocabulary, and exposure to language input through teacher lectures or simple learning media.

Nevertheless, the level of improvement is not significant. Conventional learning often lacks a variety of strategies and adequate speaking practice time for each student, as well as in-depth real-time feedback. Several studies, such as those conducted by Nasiri & Gilakjani (2016), show that conventional teaching methods can still improve speaking skills, particularly in basic aspects like pronunciation and vocabulary, but the results are more optimal when combined with interactive approaches. Another study by Dewi (2016) also supports that structured practice in conventional methods can help students understand important elements of speaking skills, although the limited variety of activities poses a barrier to significant progress. Thus, conventional learning is capable of providing a foundation for speaking skills, but the integration of more innovative strategies is necessary to maximize the results.

The results of the data analysis show that the average difference for the pretest and posttest scores in the experimental class is 22.639, with a standard

deviation of 12.563, a t-value of 10.812, and a 2-tailed significance value of 0.000. From the results of the data analysis, it can be seen that the sig. (2-tailed) value is 0.000<0.05, which can be concluded that there is a significant difference between the pretest and posttest scores before and after using artificial intelligence chatbot technology to improve students' speaking skills.

Based on the results, there is a difference in student learning outcomes after the use of Artificial Intelligence chatbot technology in speaking lessons. The passing rate before the use of AI in the experimental class was 8 students, whereas the passing rate after the use of AI increased to 30, a drastic improvement from the passing rate before the implementation of speaking lessons with the help of artificial intelligence.

According to Magdalena et al (2020) this increased ability is due to several factors, including the fact that students' speaking skills are generally influenced by a lack of practice, low self-confidence, and limitations in vocabulary and grammar. They often feel anxious when speaking, especially in a foreign language, because they are afraid of making mistakes. The available speaking exercises are usually structured and provide little room for improvisation. Students are also less exposed to real-time conversation simulations, so their skills in responding or adapting to real conversations are still low. Therefore, there is new technology to help students learn to speak foreign languages, namely through the use of AI Chatbots. This is in line with research by Mananay (2024) which shows that AI successfully enhances student engagement, proficiency, and autonomy in language skills.

The results of the data analysis show that the sig. (2-tailed) value of 0.000<0.05 indicates a significant difference between the pretest and posttest scores before and after using the artificial intelligence chatbot in students' speaking abilities. Speaking is one of the most important skills to develop as a means of effective communication. Teaching and learning speaking skills are important components in language education classes. Through speaking learning with artificial intelligence, this can become an effective and efficient learning method in improving student learning outcomes, as shown in the research by Nasiri & Gilakjani (2016), which indicates that the findings from this review paper demonstrate that by using the right strategies, EFL teachers can provide a friendly environment for EFL learners to effectively and efficiently improve their English speaking skills.

Based on these results, it can be interpreted that the use of AI Chatbot-based learning has shown significant effectiveness in improving students' speaking skills compared to conventional learning methods. In conventional learning, students receive direct instructions from the teacher through face-to-face interactions, which are effective in providing guidance but less flexible in meeting individual student needs. On the other hand, AI Chatbots offer a more personalized learning experience by providing instant feedback, repetitive exercises, and conversation simulations that boost students' confidence. Studies show that students who use chatbots record an average speaking ability improvement of 20-25% higher after post-tests compared to students who use traditional methods (Yang et al., 2024). Using advanced artificial intelligence (AI) technology in the learning environment is one of the latest challenges for educators and education policymakers. AI chatbots bring new possibilities for alternative and innovative Information and Communication Technology (ICT) tools, such as the AI chatbot by Mageira et al. (2022).

Additionally, chatbots allow students to learn anytime and provide immediate

feedback on mistakes, which not only accelerates learning but also creates a supportive learning environment without the fear of judgment.

According to the result of interview, all students felt satisfied using AI Chatbot in flipped classroom. According to their perception, the use of AI technology before learning greatly helps students in boosting their confidence, especially in their English speaking skills. However, direct interaction and correction with a teacher are necessary to improve students' speaking abilities. From the interviews with the students, almost all of them said that the use or practice of speaking using AI technology is very beneficial, especially for training students in speaking skills. The use of AI can be utilized anywhere and anytime, so students can not only learn in class with their teacher but also use it when studying with friends or at home. The advantages felt by students are the increase in new vocabulary, pronunciation, intonation, and correction of words or mispronunciations, making it easier for students to understand speaking skills and practice them in front of the class.

Teaching with the use of artificial intelligence chatbot technology also introduces students to technology, practices things that have not been understood, especially in mastering English vocabulary, boosts confidence, and of course, speaking skills.

Positive perception of using AI Chatbot to improve their speaking skills. An interactive and flexible learning experience allows students to practice anytime without feeling awkward or afraid of being judged. Immediate feedback from the chatbot, such as pronunciation and grammar corrections, helps students correct mistakes more quickly compared to conventional methods. Additionally, students feel more confident speaking after using the chatbot due to regular practice and a focus on individual weaknesses. However, several challenges were identified, such as the chatbot's limitations in understanding complex conversation contexts and the lack of interactions that feel natural. However, features like conversation simulations are considered very engaging and increase students' interest in learning. Students also suggested the development of additional features, such as voice-based conversations, for a more realistic experience. Thus, the AI Chatbot becomes an innovative learning tool with great potential to support more personalized and efficient language learning. This research aligns with the statement by Chen et al. (2020) that artificial intelligence chatbots can enhance the effectiveness and efficiency of teachers, thereby resulting in richer or better teaching quality.

Students' perceptions regarding speaking lessons using an artificial intelligence chatbot indicate that it can help students understand speaking lessons more deeply (compared to using only conventional learning) because the use of this technology makes it easier for learners to understand vocabulary and word corrections. In addition, the advantages gained by students when engaging in speaking lessons with an artificial intelligence chatbot are that students can directly contribute to learning through the use of technology, improve their literacy skills, and of course, enhance their creativity. Therefore, the use of artificial intelligence chatbots also has an impact on improving students' speaking skills because students find it easier to understand mistakes and can practice directly.

5. CONCLUSION

This study concludes with some conclusions after a number of analysis and discussions of the data. It is evident from the results and discussion in the preceding

chapter that the use of AI chatbots in flipped classrooms had an impact on students' speaking, as evidenced by the fact that the mean score of these chatbots was greater than that of the traditional approach. Strong evidence of a statistically significant difference between pre-test and post-test scores is shown by these studies. The null hypothesis can therefore be disproved. Students has a good perception about using AI Chatbot in flipped classroom because this technology makes them easier to understand vocabulary, pronunciation and increase their confidence. Therefore, for futher researchers needs to conduct reseach about teachers and parents perception of using AI Chatbot.

REFERENCES

Abdullah, M. Y., Hussin, S., & Ismail, K. (2019). Implementation of flipped classroom model and its effectiveness on English speaking performance. *International Journal of Emerging Technologies in Learning*, *14*(9). https://doi.org/10.3991/IJET.V14I09.10348

Affini, L. N., & Susanto, D. A. (2022). An Analysis Of Students' Speaking Descriptive Ability Between Photo And Short Video, 1(2), 47-56.

Aini, E. N., Susanto, D. A., & Rahayu, D. B. (2023, July). 115. Peningkatan Hasil Belajar Siswa Menggunakan TGT Pada Materi Explanation Text Kelas XI IPS 4. In Prosiding Seminar Nasional Pendidikan Profesi Guru (pp. 1050-1062).

Al Aziz, M. S., Susanto, D. A., & Nugrahani, D. (2023). The Influence Of Using Animaker Media In Improve To Teach Recount Text. Jurnal Penelitian Multidisiplin Ilmu, 2(4), 2125-2132

Bailey, K., & Nunan, D. (2019). *Practical english language teaching speaking*. 104.

Bromber, R. (2020). MOOCs and open education in the Global South: Challenges, successes, and opportunities. *Distance Education*, *41*(1). https://doi.org/10.1080/01587919.2020.1727290

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, *8*. https://doi.org/10.1109/ACCESS.2020.2988510

Creswell, J. W. (2012). Educational research: Planning, conducting, and evaluating quantitative and qualitative research. In *Educational Research* (Vol. 4).

Estevez, J., Garate, G., & Grana, M. (2019). Gentle Introduction to Artificial Intelligence for High-School Students Using Scratch. *IEEE Access*, *7*. https://doi.org/10.1109/ACCESS.2019.2956136

Hamuddin, B., Julita, K., Rahman, F., & Derin, T. (2020). Artificial intelligence in EFL context: rising students' speaking performance with Lyra virtual assistance. *International Journal of Advanced Science and Technology*, *29*(5), 6735–6741.

Hapsari, I. P., & Wu, T. T. (2022). AI Chatbots Learning Model in English Speaking Skill: Alleviating Speaking Anxiety, Boosting Enjoyment, and Fostering Critical Thinking. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *13449 LNCS*. https://doi.org/10.1007/978-3-031-15273-3_49

Hương, N. T. H., & Hưng, T. N. (2023). Applying the flipped classroom to improve speaking skills for high school students in vietnam. *TNU Journal of Science and Technology*, *228*(03). https://doi.org/10.34238/tnu-jst.7303

Khan, R., & Das, A. (2018). Build Better Chatbots. In *Build Better Chatbots*. https://doi.org/10.1007/978-1-4842-3111-1

Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: perspectives of leading teachers for AI in education. *Education and Information Technologies*, *27*(5). https://doi.org/10.1007/s10639-021-10831-6

Krisita Omega Putri, M., Anwar Korompot, C., Lidiyawati Kaharuddin, A., Negeri Makassar, U., & Negeri, S. (2023). Peningkatan Keterampilan Berbicara Bahasa Inggris Melalui Model Pembelajaran Proyek. *Indonesian Journal of Pedagogical and Social Sciences*, *3*(1), 35–44.

Kuhail, M. A., Abu Shawar, B., & Hammad, R. (2023). Trends, applications, and challenges of chatbot technology. In *Trends, Applications, and Challenges of Chatbot Technology*. https://doi.org/10.4018/978-1-6684-6234-8

Lightbown, P. M., & Spada, N. (2013). How Languages Are Learned 4th edition. *TESOL Quarterly*, *29*(1).

Madhavi, E., Sivapurapu, L., Koppula, V., Esther Rani, P. B., & Sreehari, V. (2023). Developing Learners' English Speaking Skills using ICT and AI Tools. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *32*(2). https://doi.org/10.37934/ARASET.32.2.142153

Mageira, K., Pittou, D., Papasalouros, A., Kotis, K., Zangogianni, P., & Daradoumis, A. (2022). Educational AI Chatbots for Content and Language Integrated Learning. *Applied Sciences (Switzerland)*, *12*(7). https://doi.org/10.3390/app12073239

Mananay, J. A. (2024). Integrating Artificial Intelligence (AI) in Language Teaching: Effectiveness, Challenges, and Strategies. *International Journal of Learning, Teaching and Educational Research*, *23*(9), 361–382. https://doi.org/10.26803/ijlter.23.9.19

Marzuki, Samnia, & Ratna. (2021). Meningkatkan Kemampuan Berbicara Bahasa Inggris Siswa Melalui Metode Cerita Bebas dengan Media Kartu Bergambar. *Amanah: Jurnal Amanah Pendidikan Dan Pengajaran*, *2*(2), 95–105. https://jurnal.pgrisultra.or.id/ojs/

Mawarni, A. M., & Susanto, D. A. (2021). The Effectiveness Of Fun Game Kahoot As A Media In The Teaching Of Vocabulary. Journal of English Language Teaching, Linguistics, and Literature Studies, 1(2).

Nasiri, A., & Gilakjani, A. P. (2016). A Review of EFL Learners' Speaking Skill and the Strategies for Improvement. Modern Journal of Language Teaching Methods (MJLTM). *Modern Journal of Language Teaching Methods*, *6*(9), 53–59.

Putu Indra Kusuma, I. (2020). The investigation of flipped classrooms on students' speaking performance and self-regulated learning. *Pertanika Journal of Social Sciences and Humanities*, *28*(3).

Reddy, S. (2023). An Introduction to Artificial Intelligence. *Translational Application of Artificial Intelligence in Healthcare: -A Textbook*, 1–18. https://doi.org/10.1201/9781003262152-1

Retno, R. R. S., Susanto, D. A., & Setyorini, A. (2023). The Use Of Flashcard Game In Improving Students' Reading Comprehension Of Tenth Grade Students. Didaktik: Jurnal Ilmiah PGSD STKIP Subang, 9(5), 290-299

Riswalastika, E., Susanto, D. A., & Suwarti, T. S. (2023). Using Pop-Up Book In E-Learning To Improve Student's Vocabulary Mastery Student. Edusia: Jurnal Ilmiah Pendidikan Asia, 3(1), 15-27

Sari, A. R., Susanto, D. A., & Rahayu, D. B. (2023). Peningkatan Keterampilan Menulis Teks Explanation Pada Pelajaran Bahasa Inggris Menggunakan Media Gambar Pada Kelas XI IPS 1 SMA Negeri 6 Semarang tahun pelajaran 2022/2023. MARAS: Jurnal Penelitian Multidisiplin, 1(2), 93-106.

Seo, K., Tang, J., Roll, I., Fels, S., & Yoon, D. (2021). The impact of artificial intelligence on learner–instructor interaction in online learning. *International Journal of Educational Technology in Higher Education*, *18*(1). https://doi.org/10.1186/s41239-021-00292-9

Silitonga, L. M., Hawanti, S., Aziez, F., Furqon, M., Zain, D. S. M., Anjarani, S., & Wu, T. T. (2023). The Impact of AI Chatbot-Based Learning on Students' Motivation in English Writing Classroom. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *14099 LNCS*. https://doi.org/10.1007/978-3-031-40113-8_53

Singh, S. V., & Hiran, K. K. (2022). The Impact of AI on Teaching and Learning in Higher Education Technology. *Journal of Higher Education Theory and Practice*, *22*(13). https://doi.org/10.33423/jhetp.v22i13.5514

Solekah, S., Susanto, D. A., & Rahayu, D. B. (2023). Peningkatan Motivasi Belajar Dan Kemampuan Reading Comprehension Materi Explanation Text Melalui Media Mystery Box. MARAS: Jurnal Penelitian Multidisiplin, 1(2), 165-179

Steinbauer, G., Kandlhofer, M., Chklovski, T., Heintz, F., & Koenig, S. (2021). Education in Artificial Intelligence K-12. *KI - Kunstliche Intelligenz*, *35*(2). https://doi.org/10.1007/s13218-021-00734-6

Sudarmaji, I., Amaliyah Anwar, A. A., & Mulyana, A. (2021). Developing Students' Speaking Skills through Flipped Classroom Model. *Journal of English Education and Teaching*, *5*(2). https://doi.org/10.33369/jeet.5.2.188-200

Szymkowiak, A., Melović, B., Dabić, M., Jeganathan, K., & Kundi, G. S. (2021). Information technology and Gen Z: The role of teachers, the internet, and technology in the education of young people. *Technology in Society*, *65*. https://doi.org/10.1016/j.techsoc.2021.101565.

Tonapa, Y. D., Susanto, D. A., & Sukmaningrum, R. (2024). The Application Of Flashcards To Use English As A Medium Of Instruction To Enhance Pre-Kindergarten Learners' English Language Proficiency. Innovative: Journal of Social Science Research, 4(2), 1405-1415

Wang, X., Dostál, J., & Kamal Aziz, A. (2019). Using modern technology to improve English teaching and learning*12*. *th International Conference Innovation in Language Learning Proceedings*, *12*(1), 188–195.

Widdowson, H. G. (1978). *Teaching language as communication*. Oxford University Press.

Wulandari, W., Susanto, D. A., & Hawa, F. (2024). Exploring Bamboozle As Games-Based Learning Media To Support Students' Vocabulary: Its Application In The Classroom. Innovative: Journal Of Social Science Research, 4(2), 6542-6551

Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL*, *34*(3), 327-343

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J. B., Yuan, J., & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. In *Complexity* (Vol. 2021). https://doi.org/10.1155/2021/8812542

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. In *Computers and Education: Artificial Intelligence* (Vol. 2). https://doi.org/10.1016/j.caeai.2021.100025