

Enhancing EFL Students' Speaking Performance Through Asynchronous Voice-Recording Tasks

Safrul Muluk¹; Syamsul Bahri²; Suryanto³; Uswatun Hasah⁴; Muhammad Nur Akbar Rasyid⁵; Fakhry Hamdani⁶; Muhammad Safwan Safrul⁷

^{1,2,4,7}Universitas Islam Negeri Ar-Raniry, Banda Aceh, Indonesia

³Universitas Muhammadiyah Yogyakarta, Daerah Istimewa Yogyakarta, Indonesia

⁵Universitas Islam Negeri Alaudin Makassar, Sulawesi Selatan, Indonesia

⁶Universitas Islam Negeri Sunan Gunung Djati Bandung, Jawa Barat, Indonesia

²Correspondence Email: syamsul.bahri@ar-raniry.ac.id

Article Info

Received: March 07, 2026

Accepted: May 15, 2026

Published: June 28, 2026

Online First: June 15, 2026

Keywords

Asynchronous;
Speaking Tasks;
Voice Recording;
EFL;
Communication.

Abstract

Speaking proficiency remains challenging for English as a Foreign Language (EFL) learners, particularly when opportunities for oral practice are limited and communication anxiety is prevalent. This study examined the effectiveness of asynchronous speaking tasks using voice-recording technology in improving EFL students' oral production. A quantitative pre-experimental one-group pretest-posttest design was employed with 23 undergraduate students from the English Language Education Department at Universitas Islam Negeri Ar-Raniry, Banda Aceh. Participants completed a speaking pretest, engaged in three instructional sessions involving asynchronous voice-recording tasks, self-reflection, and instructor feedback, and completed a posttest. Speaking performance was assessed using an analytic rubric covering pronunciation, grammar, vocabulary, and fluency. Data were analyzed using descriptive statistics and a paired-sample t-test. The findings showed substantial improvement, with the mean score increasing from 54.78 in the pretest to 84.13 in the posttest, representing a 29.35-point gain. The improvement was statistically significant ($t = 10.906, p < .05$), indicating that asynchronous voice-recording activities effectively enhanced learners' oral proficiency. The intervention also encouraged repeated practice, self-monitoring, reflection, revision, and greater confidence. These findings support the use of mobile-assisted asynchronous speaking activities to promote learner autonomy and continuous oral practice beyond the classroom. The study recommends voice-recording applications as an accessible and low-cost strategy for technology-enhanced speaking instruction in higher-education EFL contexts.

A. Introduction

English is widely spoken throughout the world, making it a global language. It is used for worldwide communication and applies to many global activities, including educational conferences, athletic events, international trade, business, and daily interactions. These activities require proficiency in verbal communication. English competency encompasses four fundamental skills: listening, reading, writing, and speaking. Articulation, an aspect of human self-expression demonstrated through dialogue, constitutes a talent that requires both a speaker and a listener (Foley, 1999). Speaking is a productive skill that plays a crucial role in communication (Habiburrahim et al., 2020). Oral communication facilitates the improvement of pupils' English skills (Aliakbari & Mardani, 2022). Speaking is a method of communication that conveys thoughts and emotions and facilitates the creation and accumulation of information. Nonetheless, acquiring and practicing it as a foreign language presents considerable challenges for learners. They must traverse several phases and procedures, among which the act of speaking is a crucial element of efficient communication.

Attaining competency in English, especially in oral communication, poses challenges for language learners, particularly when classroom activities include not just textbooks but also strategic actions such as utilizing a computer while recording. In developing oral communication skills, learners may prepare for their speaking activities by composing a conversation or brief talk, which they then record. The educational process promotes student participation, facilitates information sharing, and enables students to take greater ownership of their achievements. Prior studies indicate that self-recording practices can improve students' competence in academic English speaking. This approach enables self-evaluation, encourages cooperative learning, and enhances students' enthusiasm for English-speaking tasks (Kang, 2022; Kim, 2014; Meşe & Mede, 2022). Numerous studies have demonstrated that self-recording enhances academic English-speaking proficiency and significantly mitigates psychological hurdles, including anxiety, while promoting self-confidence in English communication. Bajrami and Ismaili (2016) support this perspective, asserting that video self-recording improves multiple facets of language acquisition, such as listening comprehension, pronunciation, vocabulary, grammar, and nonverbal expression, while also fostering students' autonomy, engagement, and motivation.

The researcher noted a significant event in English-speaking classrooms, specifically with Indonesian EFL students. A deficiency in confidence, increased anxiety, timidity, hesitation, and the weight of cultural expectations lead these pupils to choose their mother language over participating in English conversation (Ur 2012). This complex problem may arise from the insufficiency of efficient teaching techniques utilized by educators and a lack of opportunities for formal English practice.

To address this challenge, it is imperative to identify the root of the problem and devise solutions to enhance students' speaking abilities, confidence, and competence. In light of these challenges, the researcher proposes that utilizing audio recordings as a teaching tool for speaking skills could offer a viable alternative to the traditional methods. This method fosters a sense of security among students, allowing them to practice English confidently after familiarizing themselves with expressions from previous conversations (Whiteside, 2019).

A preliminary investigation by Widyawan and Hartati (2016) examined the potential of audio recordings to enhance oral competency. Their findings validated that student-generated audio recordings enhanced the development of speaking skills. Ramadhan's (2021) following inquiry, "Improving Students' Speaking Skill by Using Audio Recorder for Self-assessment Correction," further explored this issue. This study utilized a descriptive quantitative framework with a pre-experimental design comprising a single group evaluated before and after therapy. Data were gathered through preliminary and concluding evaluations, revealing a considerable enhancement in learners' oral communication abilities after the utilization of audio recordings. This indicates a positive change in student performance following education that includes the use of audio recording equipment.

Another study by Bella (2022) investigated "improving students' speaking ability using podcasts." Employing both quantitative and qualitative methods, including a pre-experimental design of one group pre-test and post-test, and qualitative analysis of the implementation and student responses, the study involved 11th-grade students from a senior high school in Indonesia. The results revealed a discernible increase in students' speaking ability following the incorporation of an audio recorder.

Despite these promising findings, several limitations remain in the existing body of research. Previous studies have primarily focused on the general effectiveness of audio

or video recordings in improving speaking performance, often emphasizing learning outcomes while paying limited attention to the pedagogical processes through which these improvements occur. Moreover, most investigations have examined recording activities as isolated classroom techniques without explicitly conceptualizing them as asynchronous speaking tasks that encourage repeated practice, self-monitoring, and learner autonomy beyond the classroom (Viriya, 2022). Consequently, there is still insufficient empirical evidence regarding how asynchronous voice-recording tasks facilitate sustained oral production and reflective learning in higher-education EFL contexts, particularly among Indonesian university students (Rochmawati et al., 2024).

Another limitation is the contextual diversity of previous research. Many earlier studies were conducted in secondary school settings or employed relatively small-scale classroom interventions, leaving the effectiveness of asynchronous voice-recording activities in tertiary EFL education underexplored. In addition, few studies have simultaneously examined improvements across multiple dimensions of speaking performance, including pronunciation, grammar, vocabulary, and fluency using an analytic assessment framework. Addressing these gaps is important because higher education learners require opportunities for independent speaking practice that extend beyond limited classroom interactions while receiving meaningful opportunities for self-evaluation and continuous improvement.

Building on these research gaps, the present study investigates the effectiveness of asynchronous speaking tasks implemented through voice-recording technology in enhancing EFL students' oral production. Unlike previous studies that primarily examined voice recording as a technological aid, this study conceptualizes voice recording as an asynchronous learning strategy that enables learners to rehearse, record, review, and refine their spoken performances before submitting their work. This iterative learning process is expected to promote self-reflection, increase learner autonomy, reduce anxiety, and ultimately improve speaking proficiency.

Accordingly, this study contributes to the growing literature on technology-enhanced language learning in three ways. First, it extends previous research by examining asynchronous voice-recording tasks within the Indonesian university EFL context, which remains underrepresented in the literature. Second, it provides empirical evidence regarding the effectiveness of asynchronous speaking activities in improving multiple dimensions of speaking performance, namely, pronunciation,

grammar, vocabulary, and fluency. Third, it offers practical pedagogical implications for integrating low-cost mobile-assisted learning technologies into speaking instruction to promote continuous practice beyond the classroom.

Based on the identified research gap, this study addresses the following research question:

Does the implementation of asynchronous voice-recording tasks significantly improve EFL students' speaking performance in an Indonesian higher education EFL context?

Accordingly, the objective of this study is to examine the effectiveness of asynchronous voice-recording tasks in improving students' speaking performance, particularly in pronunciation, grammar, vocabulary, and fluency.

The researchers would investigate the effects of employing voice recordings to enhance the speaking proficiency of English as a Foreign Language (EFL) students. Subsequently, the researchers opted to conduct this study in the Department of English Language Education at Universitas Negeri Islam Ar-Raniry in Banda Aceh. The primary objective was to scrutinize the impact of utilizing voice recordings on the improvement of EFL students' speaking abilities, with a specific focus on speaking classes.

In foreign language education, technology has become an integral component of classroom instruction, transitioning from audiotapes to contemporary CDs. This evolution has eliminated the need for students to carry low-quality microphones and cumbersome cassette players. Modern technological advancements, particularly digital audio recordings, have revolutionized the ease with which student-generated voice recordings can be collected in educational settings. Recognized as an effective tool for foreign language instruction, digital audio recording facilitates the primary objective of enabling students to practice speaking and engage with the auditory nuances of language, thereby enhancing their oral proficiency (Hughes, 2011).

The practice of self-monitoring, considered an indispensable facet of language acquisition at all proficiency levels, is facilitated through the utilization of various free and user-friendly smartphone applications such as voice recording tools (Wijayani, 2010). These applications serve as a means for students to document their language practice sessions, where they engage in spoken interactions in their target languages. By recording these sessions and subsequently reviewing them, students can evaluate

their grammar, fluency, accent, intonation, and other linguistic aspects before sharing the audio lessons with their instructors.

The efficacy of audio recordings as a pedagogical tool lies in their capacity to create a comprehensive archive that showcases students' progress over time (Aliakbari & Mardani, 2022; Whiteside, 2019). These recordings, saved in easily shareable formats such as downloadable MP3 files, can be exchanged not only with instructors but also among peers. A voice recording app, characterized by its simplicity, empowers language learners to create basic audio recordings effortlessly. By recording their lessons with a single click, students can save the recordings in various formats, including downloadable MP3 files. Subsequently, these audio files can be shared with peers and language instructors via class Telegram and WhatsApp groups, fostering collaborative learning and providing a platform for constructive feedback (Huang, 2008).

According to Pop et al. (2011), voice recording technologies have been incorporated into language education through multiple approaches, creating opportunities for students to generate spoken material. These tools enable learners to refine their verbal abilities beyond the classroom while obtaining input regarding their spoken output. Through voice capture applications available on mobile devices, students can document their speeches for personal evaluation. The increasing prevalence of smartphone and tablet software offers numerous possibilities for learners to explore their vocal delivery by capturing themselves while speaking. Educators can listen to these recordings and provide targeted feedback on oral presentations, or they can ask students to assess or conduct self-reviews of their recorded material.

Utilizing audio recordings proves to be a highly effective technique for enhancing students' learning capabilities (Santika, 2021; Wicaksono, 2017).

For example, it enables students to record their spoken voices and subsequently replay them as often as needed. Audio recordings serve as valuable aids for students in honing their listening skills, particularly in discerning the nuances of sound and tone that may pose challenges. Additionally, the recorded nature of audio allows for repeated playback, affording students the opportunity to distinguish and analyze the sounds they encounter carefully. Moreover, audio recordings enable students to scrutinize their own speech, including pauses and repetitions. Engaging with audio recordings brings students closer to cultivating the confidence and comfort necessary

for language proficiency. By listening to their recorded sessions, students gain a novel method for refining their linguistic expressions. Furthermore, audio recordings function as valuable self-assessment tools, empowering students to enhance their speaking abilities.

B. Method

This study employed a pre-experimental one-group pretest–posttest design to examine the effectiveness of asynchronous speaking tasks using voice-recording technology in improving EFL students' speaking performance. The design involved measuring participants' speaking proficiency before the intervention (pretest), implementing a series of instructional sessions incorporating asynchronous voice-recording activities, and administering a posttest after the intervention. Because the study involved only one experimental group without a comparison or control group, it was appropriately classified as a pre-experimental design rather than a true or quasi-experimental design.

Although this design is appropriate for investigating preliminary instructional effectiveness, it has inherent limitations because improvements cannot be attributed solely to the intervention with the same level of certainty as in controlled experimental studies. Therefore, the findings should be interpreted as evidence of instructional effectiveness within the specific research context rather than as definitive causal proof. The hypothesis of this study is as follows:

Ha: There is a significant difference in using voice records to help students increase their speaking ability.

Ho: There is no significant difference in using voice records to help students increase their speaking ability.

To gather information, a pre-experimental method was employed, with the research subject being one unit selected from several units in batch 2023, specifically those enrolled in a basic speaking or basic language skill class. The participants in this study comprised students from the Department of English Language Education at Ar-Raniry State Islamic University, Banda Aceh, specifically students from batch 2023 (202 students) consisting of five class units. The researchers applied a random

sampling technique to select one class unit. A total of 23 students were recruited as respondents for this study.

To collect the data, the researcher conducted a pre-and post-test. Three teaching interventions were implemented. The investigation commenced with a preliminary assessment conducted during the first gathering. This initial evaluation aimed to establish baseline measurements of learner competency before introducing any instructional intervention, thereby revealing the extent of existing knowledge concerning the upcoming material. In this spoken examination, each participant was required to share details about themselves individually and sequentially. This activity marked the researchers' introductory interaction with the student group. Subsequent to the preliminary assessment, the participants engaged in a series of three structured intervention sessions designed to foster speaking development through voice recording technology.

Prior to data collection, the participants were informed of the study objectives and procedures and voluntarily agreed to participate. All speech recordings were used solely for research purposes, anonymized during data analysis, and treated confidentially. Permission to conduct the study was obtained from the Department of English Language Education at Universitas Islam Negeri Ar-Raniry.

a. Instructional Procedures

Following the administration of the pretest, the participants completed three consecutive instructional sessions over a three-week period. Each session lasted approximately 100 minutes and integrated asynchronous speaking tasks using a voice-recording application available on students' smartphones.

During the first session, the students were introduced to the topic of personal information and received instructions on the communicative purpose, organization, and language features of self-introductions. The instructor demonstrated how to use the voice recording application and explained the assessment criteria. Students then prepared short speaking scripts individually before recording a two- to three-minute oral presentation outside the class. The recordings were submitted through the class WhatsApp group before the next meeting.

During the second session, the instructor reviewed common grammatical structures related to personal information, including the simple present, past, and future tenses. Students listened to selected anonymous recordings to identify strengths

and weaknesses in pronunciation, vocabulary, grammar, and fluency. Subsequently, they revised and re-recorded their oral presentations based on the feedback received.

The final instructional session emphasized reflective learning. Students independently listened to their previous recordings using a self-assessment checklist adapted from Brown's (2001) analytic speaking rubric. The instructor then provided individualized formative feedback, highlighting areas that required improvement. Students incorporated the feedback into their final recording before completing the posttest under the same assessment conditions as the pretest.

Unlike conventional classroom speaking activities that require immediate oral performance, asynchronous voice-recording tasks enable students to rehearse, monitor, revise, and refine their spoken production repeatedly before submission, thereby promoting learner autonomy, self-reflection, and reduced speaking anxiety (Kang, 2022).

b. Methods of Data Analysis

In this study, the researchers employed a speaking assessment model utilizing voice recordings to evaluate students' proficiency levels. The success of the study was gauged using a speaking assessment rubric, drawing on the scoring categories proposed by Brown (2001). The rubric outlined aspects crucial for achievement in the speaking skill test, encompassing pronunciation, grammar, vocabulary, and fluency. The analytical score comprised four items, each scored five points. Consequently, the maximum score totaled 20; however, when multiplied by five, the final maximum score reached 100.

Table 1. The Speaking Scoring Rubric

Aspects	Score	Description
Pronunciation	5	The student can pronounce well
	4	The student make mispronunciation in a few words
	3	The student make mispronunciation in some words
	2	The student's pronunciation is hard to understand
	1	The student's pronunciation is possible to understand
Vocabulary	5	The student uses various vocabulary
	4	The student uses the wrong word
	3	The students mix vocabulary with English and Bahasa
	2	The student rarely uses Bahasa
	1	The student use Bahasa or no vocabulary use
Fluency	5	The student speak fluently with no hesitation
	4	The student speak Fluently with occasional hesitation
	3	The student speak Fluently enough with several unnatural

Aspects	Score	Description
Grammar		hesitation
	2	The student speak Fluently not good enough, many unnatural hesitation
	1	The student cannot speak fluently and use complete utterances
	5	The student makes few or (if any) grammatical errors or word order.
	4	The student made few grammatical errors, but they did not change the meaning.
	3	The student frequently makes grammatical errors.
	2	The students' grammar and word order errors made comprehension difficult.
	1	The students' grammar and word order errors made the speech unintelligible.

(Oral Proficiency Scoring Categories (Brown 2001))

To classify students' scores, the researcher used the criteria by Arikunto (2013).

Table 2. The Criteria of Students Score

No	Classification	Score
1	Excellent	81-100
2	Good	61-80
3	Fair	41-60
4	Less	21-40
5	Poor	0-20

c. Instrument Validity and Reliability

Students' speaking performances were assessed using the analytic speaking rubric developed by Brown (2001), which evaluates four dimensions: pronunciation, grammar, vocabulary, and fluency. The rubric was selected because it has been widely adopted in English language assessment research and demonstrates strong content validity in measuring oral proficiency.

All recordings were evaluated by the same trained lecturer using Brown's analytic rubric to maintain scoring consistency. Although inter-rater reliability was not established, the use of standardized analytic criteria helped improve the scoring objectivity.

d. Data Analysis

Students' speaking performances were scored using Brown's (2001) analytic speaking rubric, which assesses four dimensions of oral proficiency: pronunciation,

grammar, vocabulary, and fluency. The scores from the four dimensions were summed to obtain a total speaking score ranging from 0 to 100. To facilitate the interpretation of students' performance, the scores were classified into five proficiency levels adapted from Arikunto (2013): excellent (81–100), good (61–80), fair (41–60), less (21–40), and poor (0–20).

Prior to conducting the paired-samples *t*-test, the normality of the difference scores between the pre- and post-tests was examined using the Shapiro–Wilk test because the sample size was fewer than 50 participants. The Shapiro–Wilk test indicated that the difference scores did not significantly deviate from normality ($W = 0.96, p = .48$).

The data were analyzed using descriptive and inferential statistical methods. Descriptive statistics, including the mean and standard deviation, were calculated to summarize students' speaking performance before and after the intervention. Mean scores were used to indicate the overall level of speaking achievement, whereas standard deviations were computed to examine the variability of students' scores within each testing occasion.

To determine whether the asynchronous voice-recording intervention significantly improved students' speaking performance, a paired-samples *t*-test was conducted by comparing the pre- and post-test scores of the same participants. Statistical significance was set at $< .05$ level. The null hypothesis (H_0) stated that there was no significant difference between the pre- and post-test speaking scores, whereas the alternative hypothesis (H_1) stated that a significant improvement occurred following the implementation of the voice-recording intervention. All statistical analyses were conducted using IBM SPSS Statistics Version 26, with the significance level set at $.05$.

C. Results and Discussion

The data collection process involved administering pre- and post-tests to students within the classroom setting. The researcher sought to evaluate the impact of Voice Record on students' speaking skills by systematically gathering data through pre- and post-tests. This approach involved a comprehensive examination of students' performance before and after the intervention, providing a robust foundation for analyzing the effectiveness of voice recording in the classroom.

1. Results

a) The Analysis of Pre-test

For the initial phase of the study, a group of 23 students actively participated in the pre-test, contributing valuable data that formed the basis for the subsequent analysis. The ensuing table illustrates the specific details and scores obtained by each student during the pre-test, offering a comprehensive overview of their initial performances. The pre-test data are presented in the table below:

Table 3. Students Score of Pre-test

No	Students' Initial	Pronunciation	Grammar	Vocabulary	Fluenc	Total
1	AZA	15	15	10	10	50
2	AR	10	10	10	10	40
3	CNA	10	10	15	15	50
4	DU	10	10	10	10	40
5	DP	10	15	15	15	55
6	DCD	15	10	10	10	45
7	HS	15	10	15	10	50
8	K	15	10	15	15	55
9	MR	10	15	15	15	55
10	MS	20	15	20	20	75
11	MF	15	10	10	15	50
12	MA	15	15	20	20	70
13	MK	15	15	15	20	65
14	NSL	15	10	15	10	50
15	NPR	10	15	15	15	55
16	QA	15	10	15	10	50
17	RNN	10	15	10	20	55
18	SH	10	15	15	15	55
19	SSS	15	15	15	15	60
20	SHA	15	15	20	20	70
21	SRA	15	15	15	15	60
22	SPN	10	10	10	15	45
23	TR	15	15	15	15	60
Total				1.260		

The results from the pre-test score table of the students revealed significant variations in their language abilities. student MS achieved the highest score of 75, making her the top performer in the pretest. However, it is important to acknowledge that even a score of 75, while the highest in this specific pre-test, may still be considered relatively moderate. This suggests that there is still room for improvement, even for the highest-scoring student. Furthermore, when we consider students like

MA and SHA, who scored 70, it becomes evident that these scores, although relatively high, do not necessarily indicate complete mastery. In contrast, the lowest scores were recorded by students AR and DU, each scoring only 40. These scores indicate a substantial need for improvement, particularly in the evaluated areas.

Utilizing the data exhibited in the table, the average score for the pre-test was determined through a computation that involved summing the individual scores and dividing by the total number of participants. This statistical measure, known as the mean score, provides a central point of reference for understanding the overall performance of the students in the initial assessment, we could count the mean score of the pre-test as follows:

The mean score, computed as the sum of individual scores divided by the total number of students, served as a representative measure of the central tendency in the distribution of pre-test scores. In this context, the mean pre-test score of 54.78 suggests that, on average, students demonstrated a proficiency level below the perfect score of 100. This information provides insights into the overall performance trend, highlighting the room for improvement and distance from the maximum attainable score.

b) The Analysis of Post-test

There were also 23 students who participated in the post-test. The ensuing table illustrates the specific details and scores obtained by each student during the pre-test, offering a comprehensive overview of their initial performances. The post-test data are presented in table below.

Table 4. Students Score of Post-test

No	Students' Initial	Pronunciation	Grammar	Vocabulary	Fluency	Total
1	AZA	15	15	20	20	70
2	AR	15	20	20	15	70
3	CNA	15	25	25	25	90
4	DU	20	20	20	15	75
5	DP	25	25	20	25	95
6	DCD	10	20	15	20	65
7	HS	25	20	25	25	95
8	K	15	15	20	25	75
9	MR	25	20	25	25	95
10	MS	20	20	20	25	85

No	Students' Initial	Pronunciation	Grammar	Vocabulary	Fluency	Total
11	MF	20	20	25	25	90
12	MA	25	15	25	25	90
13	MK	20	15	20	25	80
14	NSL	15	20	20	15	70
15	NPR	20	15	20	25	80
16	QA	25	20	25	25	95
17	RNN	20	20	25	25	90
18	SH	15	20	20	25	80
19	SSS	25	20	25	25	95
20	SHA	25	20	25	25	95
21	SRA	25	20	25	25	95
22	SPN	15	20	20	15	70
23	TR	20	20	25	25	90
Total		1.935				

c) Overall Speaking Performance

Descriptive statistics indicated a substantial improvement in students' speaking performance following the implementation of the asynchronous voice-recording activities. The mean pretest score increased from 54.78 (SD = 9.10) to 84.13 (SD = 10.40) in the posttest, representing an average gain of 29.35 points. A paired-samples *t*-test showed that students' post-test scores ($M = 84.13$, $SD = 10.40$) were significantly higher than their pre-test scores ($M = 54.78$, $SD = 9.10$), $t(22) = 10.91$, $p < .001$, Cohen's $d = 2.27$. These findings indicate that students' speaking performance improved significantly following the implementation of the asynchronous voice-recording activities.

In addition to statistical significance, the magnitude of the treatment effect was examined using Cohen's d for paired samples. This very large effect size suggests that the intervention produced not only statistically significant but also educationally meaningful improvements in students' speaking performance.

Based on the data presented in the table above, we calculated the mean score of the post-test as follows:

The average post-test score of 84.13 reflects a significant improvement compared to the pre-test mean score of 54.78. Students' speaking scores improved by an average of 29.35 points. These findings indicate that students demonstrated considerably higher speaking performance following the intervention. The considerable increase in scores

from the pre-test to the post-test provides compelling evidence of the method's efficacy in fostering skill development and highlights the positive outcomes of incorporating voice records in the instructional process.

d) T-Test and Hypothesis

The researcher computed the sum of squared deviations to determine the noteworthy difference between the pre- and post-tests, employing the t-score formula presented below:

The paired-samples *t*-test indicated that the students' speaking performance improved significantly following the intervention, $t(22) = 10.91, p < .001$. Therefore, the null hypothesis was rejected, indicating that the asynchronous voice-recording activities were associated with a statistically significant improvement in students' speaking scores.

2. Discussion

The present study demonstrates that integrating asynchronous voice-recording tasks into EFL speaking instruction can substantially improve learners' oral performance. The findings further suggest that the improvement was facilitated by repeated opportunities for rehearsal, self-monitoring, revision, and instructor feedback (Santika, 2021; Wicaksono, 2017). The findings demonstrated a statistically significant improvement in students' speaking performance following the intervention, indicating that asynchronous voice-recording activities constitute an effective pedagogical approach for developing oral communication skills in higher-education EFL contexts. Beyond confirming positive learning outcomes, these findings provide insights into the mechanisms through which technology-mediated speaking practice facilitates language development.

In addition to being statistically significant, the intervention produced a **large effect size (Cohen's $d = 2.27$)**. According to conventional benchmarks (Cohen, 1988), this indicates that the asynchronous voice recording activities had a substantial educational impact on students' speaking performance. The large effect size suggests that the observed improvement was not only statistically reliable but also practically meaningful for EFL speaking instruction.

This interpretation is supported by recent research demonstrating that asynchronous digital speaking tasks promote learners' self-regulated learning by encouraging planning, rehearsal, self-evaluation, and reflection throughout the

speaking process. Zhou et al. (2025) found that digital oral dialogue journaling significantly enhanced learners' cognitive and metacognitive speaking strategies because students were able to repeatedly review their recordings and incorporate teacher feedback into subsequent performances. Likewise, Kang (2022) reported that self-recorded speaking tasks foster greater learner autonomy, reflective learning, and sustained engagement in online EFL environments.

The findings may also be interpreted from the perspective of self-monitoring theory in language learning. Listening to one's own recorded speech enables learners to detect pronunciation errors, grammatical inaccuracies, inappropriate vocabulary choices, and disruptions in fluency that often remain unnoticed during spontaneous oral interactions (Wicaksonmo, 2017; Wijayani, 2010). Such reflective practice increases learners' metacognitive awareness of their own language production and encourages deliberate efforts to improve subsequent performances. The opportunity to compare successive recordings further supports continuous language development through repeated cycles of reflection and revisions.

Recent empirical evidence suggests that self-monitoring is a key mechanism underlying improvements in technology-mediated speaking performance. Zhang and Wilson (2023) demonstrated that learners who actively monitored their oral production during computer-assisted speaking tasks achieved higher levels of speaking performance because monitoring facilitated greater metacognitive control over their language production. These findings reinforce the present study by indicating that repeated listening to and evaluation of recorded speech help learners identify linguistic weaknesses and make purposeful revisions.

From the perspective of Swain's Output Hypothesis, the effectiveness of voice-recording activities can be attributed to the requirement that learners actively produce meaningful spoken language rather than merely receiving linguistic input. Producing spoken output encourages learners to notice the gaps between what they intend to express and what they are able to articulate. These moments of noticing stimulate deeper cognitive processing, prompting learners to modify their language use, refine their grammatical structures, and expand their vocabulary. Because students could repeatedly record and revise their oral production, the asynchronous learning environment created multiple opportunities for pushed output, thereby strengthening their speaking proficiency.

The results further support the pedagogical value of asynchronous learning in foreign-language education. Asynchronous learning environments provide flexibility that accommodates differences in learners' proficiency levels, learning pace, and confidence. Students can practice independently without the immediate social pressure often associated with face-to-face speaking activities. This flexibility may reduce communication anxiety while simultaneously increasing learners' willingness to communicate, resulting in more frequent speaking practice and greater confidence in oral production.

The effectiveness of asynchronous speaking activities has also been highlighted in recent studies on digital language learning. Zhou et al. (2025) showed that flexible scheduling, unlimited opportunities to rehearse and review recordings, and dialogic teacher feedback substantially improved learners' speaking strategy use and reduced anxiety associated with oral communication. Similarly, Kang (2022) concluded that asynchronous self-recorded speaking tasks enabled learners to regulate their own learning process while increasing their confidence and engagement during speaking activities. The findings suggest that asynchronous voice-recording activities complement classroom interaction by extending meaningful speaking practice beyond scheduled instructional time.

The present study also contributes to the growing literature on technology-enhanced language learning in several important ways: While previous studies have generally demonstrated that audio or video recordings can improve speaking performance, they have tended to evaluate recording technologies primarily as instructional media. The present study extends this body of knowledge by conceptualizing voice recording as an asynchronous pedagogical strategy that promotes self-regulated learning, reflective practice, and sustained oral production (Rochmawati et al., 2024). By emphasizing the learning process rather than the technology itself, this study offers a more comprehensive explanation of how asynchronous speaking tasks facilitate language development.

These findings have important practical implications for English language teaching. For language teachers, asynchronous voice-recording activities provide an accessible and low-cost instructional strategy that increases students' opportunities for speaking practice without requiring additional classroom time. Recordings also allow teachers to provide individualized formative feedback and monitor students' progress

over time (Liao, 2025). Curriculum developers should integrate asynchronous speaking assignments into speaking courses to encourage learner autonomy and continuous practice, while educational institutions are encouraged to integrate asynchronous voice-recording activities into blended learning environments as sustainable, low-cost instructional practices that promote learner autonomy, formative assessment, and continuous speaking practice.

Although this study was conducted in an Indonesian university context, its pedagogical implications extend to broader international EFL settings where opportunities for authentic oral interaction remain limited. Many higher education institutions worldwide continue to adopt blended and technology-supported learning models, making asynchronous speaking activities increasingly relevant in diverse educational contexts. The flexibility, accessibility, and learner-centered characteristics of voice-recording tasks suggest that this approach may be applicable to various EFL classrooms, regardless of geographical location or institutional resources. Consequently, this study contributes to the international discussion on technology-enhanced speaking instruction by demonstrating how simple mobile technologies can support meaningful language learning through reflective and autonomous practice.

The international relevance of these findings is reinforced by growing evidence that technology-supported speaking activities are increasingly being adopted across diverse higher-education settings to enhance learner autonomy and oral proficiency. Recent studies conducted in Hong Kong, South Korea, Türkiye, and other EFL contexts consistently report that asynchronous digital speaking tasks improve students' speaking performance, self-regulation, and confidence, suggesting that the pedagogical benefits observed in the present study are applicable beyond the Indonesian context.

D. Conclusion

The findings indicate that asynchronous voice-recording tasks constitute an effective pedagogical strategy for improving EFL students' speaking performance by promoting repeated practice, self-monitoring, reflection, and learner autonomy. Rather than functioning merely as a technological tool, voice recording serves as a learning strategy that supports sustained oral language development beyond classroom instruction.

This study contributes to the growing body of research on technology-enhanced language learning by conceptualizing voice recording as an asynchronous pedagogical strategy rather than simply an instructional medium. By highlighting the roles of learner autonomy, reflective learning, and repeated oral production, this study extends the current understanding of how simple mobile technologies can facilitate speaking development in higher education EFL contexts.

These findings have important practical implications. English language teachers can integrate asynchronous voice-recording tasks into speaking courses to provide students with additional opportunities for independent practice and individualized formative feedback. Likewise, curriculum developers and educational institutions should incorporate low-cost mobile-assisted learning activities into blended and online learning environments to promote continuous speaking practice and learner engagement.

Despite these contributions, several limitations of this study should be acknowledged. This study employed a pre-experimental one-group pretest-posttest design involving a relatively small sample from a single university class without a control group. Consequently, the findings should be interpreted with caution and cannot be generalized to all EFL learning contexts. In addition, the intervention was conducted over a relatively short period, limiting the ability to examine the long-term effects of the asynchronous voice-recording activities on speaking development.

Future research should employ true or quasi-experimental designs with larger and more diverse samples to improve the generalizability of these findings. Longitudinal studies are also recommended to investigate the sustained impact of asynchronous voice-recording activities on speaking proficiency, learner autonomy, and anxiety. Furthermore, future investigations should compare different forms of digital speaking technologies, such as audio journals, video reflections, artificial intelligence-assisted speaking applications, and automated feedback systems, to determine which approaches most effectively support oral language development across diverse EFL contexts.

Declaration of Generative AI

In the initial stage of writing the manuscript, especially for brainstorming, proofreading, and language polishing, the authors used ChatGPT (GPT-4, OpenAI) to

improve the readability and clarity of the manuscript. After using this tool, to ensure scholarly integrity, the authors thoroughly and carefully reviewed and edited the content as needed and take full responsibility for the final content of the publication.

Bibliography

- Aliakbari, M., & Mardani, M. (2022). Mobile-assisted Language Learning and its Effects on Learners' Speaking Development. *Education Research International*, 2022, Article 9043326. <https://doi.org/10.1155/2022/9043326>
- Arikunto, S. (2013). *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Bajrami, L., & Ismaili, M. (2016). The Role of Video Materials in EFL Classrooms. *Procedia-Social and Behavioral Sciences*, 2(32), 502–506. <https://doi.org/10.1016/j.sbspro.2016.10.068>
- Bella, C. C. O. (2022). Improving Students' Speaking Ability Using Podcast. *Retain: Journal of Research in English Language Teaching*, 10(01), 25–31. <https://doi.org/10.26740/rt.v10i01.41041>
- Brown, H. D., & Abeywickrama, P. (2019). *Teaching by Principles: An Interactive Approach to Language Pedagogy* (5th ed.). Pearson.
- Foley, J. A. Ed. (1999). *English in New Cultural Contexts*. New York: Oxford University Press.
- Habiburrahim, H., Risdaneva, R., Putri, G., Dahliana, S., & Muluk, S. (2020). The Effects of Anxiety Toward Acehnese Students' English speaking ability. *The Qualitative Report*, 25(1), 254–270. <https://doi.org/10.46743/2160-3715/2020.3639>
- Harmer, J. (2007). *How to teach English (New ed.)*. England: Pearson Longman.
- Huang, S. C. (2008). Raising Learner-initiated Attention to The Formal Aspects of Their Oral Production Through Transcription and Stimulated Reflection. *IRAL-International Review of Applied Linguistics in Language Teaching*, 46(4). <https://doi.org/10.1515/IRAL.2008.016>
- Kang, N. (2022). Promoting EFL Learners' Self-Regulated Learning Using a Self-Recorded Video-Speaking Task In an Online ESP course. *Journal of Asia TEFL*, 19(4), 1141–1162. <https://doi.org/10.18823/asiatefl.2022.19.4.1.1141>
- Kim, S. (2014). Review of Digital literacies, by Julia Gillen. *Learning, Media and Technology*. *Learning, Media and Technology*, 41(3). <https://doi.org/10.1080/17439884.2014.959449>

- Liao, M.-H. (2025). Cultivating Proficient and Efficacious L2 English Speakers Via Voice Thread-mediated Self- and Peer Assessments. *Humanities and Social Sciences Communications*, 12, Article 1277. <https://doi.org/10.1057/s41599-025-05674-2>
- Meşe, E., & Mede, E. (2022). Using Digital Differentiation to Improve EFL Achievement and Self-regulation of Tertiary Learners: The Turkish context. *Innovation in Language Learning and Teaching*, 17(2), 340–353. <https://doi.org/10.1080/17501229.2022.2043872>
- Pop, A., Tomuletiu, E. A., & David, D. (2011). EFL Speaking Communication With Asynchronous Voice Tools For Adult Students. *Procedia-Social and Behavioral Sciences*, 15, 1199-1203.
- Ramadhan, Y. F. (2021). *Improving Students' Speaking Skill By Using Audio Recorder for Self-Assessment Correction: At MAN 1 Bekasi, West Java, Indonesia* (Sarjana thesis). UIN Sunan Gunung Djati Bandung.
- Rochmawati, D., Nurkamto, J., Nizam, M., Rochsantiningsih, D., et al. (2024). A systematic Review Of Self-Regulated Learning Approach Through Digital Learning Media In Enhancing Students' EFL speaking competences. *English Education Journal*, 14(4), 779–794. <https://doi.org/10.24815/eej.v14i4.37800>
- Santika, N. (2021). The Application of Audio Recording in a Smartphone to Assist Students' Speaking Skills in Senior High School. *Retain: Journal of Research in English Language Teaching*, 9(2), 102–111. <https://doi.org/10.26740/rt.v9i2.40241>
- Viriya, C. (2022). Exploring The Impact of Synchronous, Asynchronous, and Bichronous Online Learning modes on EFL Students' Self-Regulated and Perceived English Language Learning. *rEFLections*, 29(1), 88–111. <https://doi.org/10.61508/refl.v29i1.260525>
- Whiteside, R. (2019, February 1). *Getting smart with speaking: Voice Recording In Language Learning*. Cambridge English. <https://www.cambridge.org/elt/blog/2019/02/01/getting-smart-with-speaking/>
- Wicaksono, A. (2017). Using Video Recording to Improve Students' Speaking Ability. In *Proceedings of The International Conference on English Language Teaching (ICONELT 2017)* (pp. 21–24). Atlantis Press. <https://doi.org/10.2991/iconelt-17.2018.5>
- Widyawan, K., & Hartati, E. (2016). Improving Students' Speaking Skill by Using Their Spoken Audio Recording in The Middle School. *JELE (Journal of English Language and Education)*, 2(1), 26–32. <https://doi.org/10.26486/jele.v2i1.216>

- Zhang, W., & Wilson, A. (2023). From Self-Regulated Learning to Computer-Delivered Integrated Speaking Testing: Does Monitoring Always Monitor? *Frontiers in Psychology*, 14, 1028754. <https://doi.org/10.3389/fpsyg.2023.1028754>
- Zhou, S., Tso, W. S., & Aubrey, S. (2025). Enhancing L2 learners' Self-Regulated Speaking Through Digital Oral Dialogue Journaling: An EFL Classroom-Based Study. *RELC Journal*. Advance Online Publication. <https://doi.org/10.1177/00336882251313704>