Journal of English Language and Education
ISSN 2597- 6850 (Online), 2502-4132 (Print)
Journal Homepage: https://jele.or.id/index.php/jele/index

OPEN ACCESS Article

The Correlation of AI Text-to-Speech and Indonesian EFL Students' Confidence in Pronouncing English Words

https://doi.org/10.31004/jele.v%25vi%25i.1787

*Febi Marelia, Reko Hary Putra[ab]
[12]Institut Agama Islam Negeri, Kerinci, Jambi, Indonesia
Corresponding Author: febimarelia18576@gmail.com

A B S T R A C T

This study explores how AI TTS relates to students' confidence in English pronunciation among English as a Foreign Language (EFL) learners in Indonesia. Although previous studies have examined perceptions and the quality of TTS-generated pronunciation, research rarely addresses the affective outcomes of TTS, particularly the development of learners' pronunciation confidence. In the Indonesian context, empirical evidence directly linking AI TTS use to confidence in pronouncing English words remains limited, even though AI-based pronunciation tools are increasingly used in language classrooms. To address this gap, this study uses a quantitative correlational design and involves 38 students from an English language education program. Data were collected using a questionnaire measure of AI TTS use and pronunciation confidence. The results indicate that students have positive perceptions of AI TTS and its role in supporting pronunciation confidence. The two variables show a positive and statistically significant correlation ($r = 0.850$, $p < 0.05$) according to Pearson's Product-Moment correlation test. Put differently, students with increased interaction with AI TTS tend to feel greater comfortable pronouncing English words. These results also imply that AI TTS can assist the growth of self-confidence, promote autonomous learning, and lessen fear.

Keywords: *AI TTS, Pronunciation, Self-Confidence, EFL Learners, Correlational Study*

Article History:
Received 30th November 2025
Accepted 31st December 2025
Published 03rd January 2026

INTRODUCTION

The rapid advancement of Artificial Intelligence (AI) in recent years has brought significant changes to education, especially in foreign language learning. Text-to-Speech (TTS) is one of the innovative AI systems that has received considerable attention. With this technology, written text can be converted into voice output, allowing learners to listen clear pronunciation models and practice independently. Because TTS is easily accessible and can be used anytime, this technology provides students with more opportunities to practice English pronunciation both inside and outside the classroom. However, despite the increasing integration of this technology, empirical evidence regarding its impact on learners' pronunciation confidence remains limited, particularly within EFL contexts where opportunities for authentic pronunciation practice are constrained. This gap underscores the pressing need for a systematic investigation into the extent to which AI TTS contributes to both linguistic development and affective dimensions of language learning.

TTS provides accurate pronunciation models that learners can independently imitate and practice; the application makes a substantial contribution to enhancing language learning outcomes. Empirical evidence demonstrates that TTS use significantly improves learners' speaking and pronunciation skills (Amin, 2024; Pertiwi, 2020). Repeated listening and self-regulated practice enable students to gradually develop greater fluency while reducing anxiety when communicating in English. Moreover, consistent feedback and exposure to native-like pronunciation offered by TTS technologies foster increased comfort and confidence in articulating English words (Svensson et al., 2021).

The present theoretical approach also highlights the connection between affective states like confidence and speaking anxiety and pronunciation correctness. While greater confidence improves verbal expressiveness and pronunciation accuracy, anxiety during second-language performance might hinder oral engagement and reduce speech clarity (MacIntyre & Gardner, 1991). Affective factors also have a reciprocal link with oral competence (Lamb et al., 2020). Students who pronounce words more clearly typically feel more confident, which in turn helps them enhance their speaking ability.

Although many learners of English as a Foreign Language (EFL) have adequate linguistic knowledge, they still show hesitation to speak English in class. This is evident from research findings in Indonesia, which reveal that students often avoid oral tasks due to a lack of confidence and fear of making mistakes when participating (Fitrah et al., 2024). This concern is not only found in the local context but is also supported by recent research showing that EFL learners globally are still reluctant to engage in oral communication (Rotjanawongchai, 2024). In particular, confidence and positive perceptions of communication skills play an important role in determining their level of participation in oral interactions. Thus, students who have higher confidence tend to speak more actively, make fewer mistakes, and remain motivated to improve their English language skills (Destiawati et al., 2024).

Confidence in speaking is closely related to accurate pronunciation, because learners need clear and easily understood articulation in order to communicate effectively. Good pronunciation not only improves comprehension but is also an important factor that supports learners' courage and readiness to engage in oral communication in public settings (Derwing & Munro, 2015). Therefore, a reliable pronunciation model is essential to help students internalize the correct patterns of stress, rhythm, and intonation in English. AI TTS technology enables consistent and repeatable speech input as needed, allowing students to practice independently and gradually improve the clarity of their pronunciation. Thus, the use of TTS has the potential to strengthen pronunciation skills while increasing students' confidence in speaking English. Beyond the linguistic benefits, AI-based learning systems have recently been shown to foster students' socio-emotional development. Integrating AI for oral practice significantly improved learners' speaking confidence by providing risk-free speaking opportunities (Zhang et al., 2024). A similar study also revealed that AI pronunciation tools reduce communication anxiety because learners can practice repeatedly without fear of negative evaluation from peers or teachers (Wang et al., 2025).

Several studies provide strong support for the view that pronunciation skills play an important role in building confidence in speaking for EFL learners. Research indicates that students with improved pronunciation tend to exhibit higher confidence when communicating orally in English (Husna, 2021; Suwari et al., 2023). This indicates that improving pronunciation quality not only has an impact on linguistic aspects but also contributes to the affective aspects of language learning. Furthermore, empirical evidence increasingly confirms that technological support plays a strategic role in the process of pronunciation development. Pronunciation practice activities through digital media have been proven to increase students' motivation and confidence in speaking (Shadiev & Yang, 2020). Specifically, the use of technology, such as Text-to-Speech (TTS), allows learners to access consistent and repeatable pronunciation models, thereby encouraging continuous improvement and strengthening their readiness to participate in oral communication (Duong, 2022). Thus, interactive learning tools, including AI TTS, have significant potential in supporting simultaneous linguistic and affective development in the context of learning English as a foreign language.

Furthermore, Self-Determination Theory (Deci & Ryan, 2000) provides a pertinent theoretical framework for comprehending the possible advantages of AI TTS in pronunciation learning. TTS fosters competence by offering precise models and feedback that improve students' performance, it encourages autonomy by allowing students to practice independently, and it makes students feel more connected and involved by promoting active engagement in oral communication. These psychological experiences can boost learners' confidence to use English in authentic communicative contexts while progressively lowering

speaking anxiety. The ideas of autonomous learning, which emphasize the importance of independence and self-control in language development, are reflected in this system. When students can manage their learning materials and keep an eye on their progress, their autonomy increases (Benson, 2013; Holec, 1981). More recent research demonstrates that by encouraging repeated independent practice at a customized pace, digital pronunciation aids promote autonomous learning habits (Lai, 2022). As a result, AI TTS may promote affective development, which is crucial for EFL learners, in addition to pronunciation improvements.

At an Islamic higher education institution in Jambi, Indonesia, students of the English Language Education Department have demonstrated a growing interest in employing technology-enhanced tools to support their learning, including AI Text-to-Speech. They can effortlessly integrate these tools into their pronunciation practice because of their familiarity with technology. On a broader scale, recent national survey data indicate that the vast majority of university students in Indonesia, approximately 95 % have begun integrating AI technology into their academic activities. Despite this extensive adoption, there remains a lack of quantitative evidence on how English as a Foreign Language (EFL) learners in Indonesia use TTS for pronunciation development and how this use may relate to their confidence in spoken communication. Considering that linguistic and cultural variables can shape learners' perceptions and use of AI tools, examining this correlation within the Indonesian EFL context becomes particularly relevant.

Despite growing interest in the use of technology in language acquisition, research on the subject is still dispersed. Numerous studies concentrate on how TTS affects pronunciation accuracy (Ren et al., 2019), but affective outcomes specifically, pronunciation confidence has not received the same amount of attention. Correlational research also demonstrates a connection between TTS use and perceived speaking confidence (Choi et al., 2024; Khanam et al., 2022), but there is currently little data specifically on Indonesian EFL learners. Because correlational research finds naturally existing associations without experimental manipulation, it offers valuable insights into language instruction (Creswell, 2017). Given that motivation, engagement, and self-confidence are influenced by the learning environment and technological support (Dörnyei, 2014; H. Zhang et al., 2025). Examining how AI-driven pronunciation tools relate to confidence in speaking becomes theoretically and practically significant. Although existing research demonstrates the pedagogical and affective advantages of AI for pronunciation learning, empirical studies examining how AI TTS specifically relates to pronunciation confidence remain limited. Previous correlational studies were conducted outside Indonesia, and contextual factors such as culture, learning environment, and technological readiness can influence learners' confidence in speaking. Therefore, localized evidence involving Indonesian EFL learners is essential to validate whether AI TTS contributes to pronunciation confidence in this specific educational context. Consequently, the present study seeks to address this empirical gap and contribute localized insights into technology-mediated pronunciation learning.

Therefore, the present study aims to determine whether there is a significant correlation between the use of AI Text-to-Speech (TTS) and students' self-confidence in pronouncing English words. The findings are expected to contribute to an understanding of how AI-based learning technologies can support both cognitive and affective aspects of language learning. Specifically, the present study aims to address the research question: Is there a correlation between the use of AI Text-to-Speech (TTS) and the self-confidence of English department students in pronouncing English words?

Literature Review

Self-Confidence in English Pronunciation

Self-confidence is widely recognized as a key psychological construct that influences individual performance in academic and communication tasks. According to theoretical viewpoints, self-belief, perceived competence, and emotional control are the foundations of self-confidence, which progressively grow with experience and social interaction (Jonker et al., 2024). According to Bandura's self-efficacy theory, students' motivation and observed performance outcomes are influenced by their perceptions of their capacity to do particular

tasks (Bandura, 1997). In academic settings, students who have higher levels of self-confidence are more likely to be persistent and involved in reaching their learning objectives (Pajares & Miller, 1994; Zimmerman & Kitsantas, 2005). All things considered, these theoretical stances imply that self-confidence acts as a motivator and a predictor of learning process performance. Overall, these theoretical perspectives suggest that self-confidence serves as a motivational driver and predictor of performance in the learning process.

Confidence is continuously linked to communication willingness and oral contact success in the field of English as a Foreign Language (EFL). Higher-confidence learners typically participate more actively in speaking assignments and do better in oral presentations (Hidayad et al., 2023). On the other hand, those with limited confidence are often reluctant to speak because they worry about making mistakes or receiving negative feedback from classmates and instructors (Fadillah et al., 2025). Because of this avoidance tendency, there are less opportunities to practice pronunciation, which hinders the development of accuracy and fluency. These results highlight a problem that frequently occurs in the setting of EFL: learners may possess sufficient language knowledge but are unwilling to communicate because of emotive rather than cognitive constraints.

Pronunciation is crucial to effective oral communication because it facilitates listeners' comprehension of the speaker's message. According to experts, pronunciation quality should be assessed based on both how effectively the speech may be understood by others in real-world interactions and how properly the sounds are delivered (Derwing & Munro, 2005). In practical communication, students are more likely to feel comfortable and willing to speak when they believe their speech is understandable and accepted by listeners, so improving pronunciation competence becomes essential to lowering communication anxiety and increasing learners' sense of control when speaking English.

The theory that self-confidence directly influences pronunciation development is supported by empirical research. Research demonstrates that students' speaking accuracy and fluency significantly increase when they get systematic pronunciation coaching along with self-confidence-boosting techniques (Chau et al., 2022). However, contextual elements like emotional states, peer support, and content mastery are important determinants of students' confidence when speaking English orally (Asnaini et al., 2025). Furthermore, it has been demonstrated that students' sense of efficacy in oral language output is strengthened by linguistic competence and independent learning activities (Yuehua & Qiumei, 2023). However, although these studies provide valuable insights, many of them focus on isolated variables, limiting our understanding of how psychological, linguistic, and contextual factors interact to shape confidence in speaking.

Overall, the existing literature agrees on the importance of self-confidence in improving students' speaking abilities. However, research remains scattered in its emphasis, and few studies systematically examine pronunciation confidence as a multidimensional construct integrating cognitive, affective, and environmental components. This theoretical and empirical gap highlights the importance of further research, particularly in the context of English as a foreign language (EFL) learning, where pronunciation remains a persistent learning challenge.

Use of text-to-speech AI tools in Language Learning

Text-to-Speech (TTS) technology has emerged as a valuable tool for improving pronunciation in English language learning because it provides consistent exposure to accurate speech models. Previous research has shown that tools such as TTS support the refinement of pronunciation accuracy by allowing learners to repeatedly listen to and imitate clear vocal outputs generated from written text (Ren et al., 2019). Through this repeated exposure, learners can gradually adjust their articulation and improve the clarity of their spoken English.

TTS helps build suprasegmental characteristics including word stress, rhythm, and intonation in addition to segmental accuracy. These characteristics are crucial for natural and comprehensible spoken communication. TTS gives students the chance to more successfully

internalize English prosodic patterns, which promotes fluency and oral expressiveness, by providing consistent and dependable speech models on demand.

More recent research has broadened the focus to encompass learners' emotional responses to TTS use. Because they lessen public speaking anxiety and enable autonomous practice without worrying about being judged, TTS-based pronunciation aids are frequently seen favorably by students (Bione et al., 2016; Khanam et al., 2022). Students who regularly use TTS report better self-monitoring abilities, a greater sense of control over their speech production, and an enhanced willingness to participate in speaking tasks (Choi et al., 2024; Zhao et al., 2023). These results imply that TTS may enhance both pronunciation accuracy and the self-confidence required to actively engage in oral communication.

Nevertheless, current research still reveals a number of shortcomings despite these encouraging findings. While affective consequences, especially confidence in pronunciation, have not received enough attention, the majority of prior research has concentrated on technical outcomes pertaining to the correctness and naturalness of speech synthesis. Additionally, institute-level English as a Foreign Language (EFL) learners in Indonesia have received little attention, even though poor pronunciation of English words continues to be a problem in English language instruction. In order to close this gap, this study examines the correlation between students' confidence in their ability to pronounce English words and the use of AI TTS, establishing TTS as both a linguistic resource and a psychological support aid for pronunciation acquisition.

Correlation between Use of AI TTS Tools and Self-Confidence

Correlational research seeks to determine statistical correlation between variables, helping educators understand relationships that occur naturally without experimental manipulation (Creswell, 2017). In language learning, technology use is often associated with affective outcomes such as motivation, engagement, and self-confidence. Digital devices and technology-supported learning tools have been shown to reduce student anxiety and promote more positive emotional states during language practice (H. Zhang et al., 2025). Affective factors, including self-confidence, are influenced by the learning environment and the availability of supporting technological devices (Dörnyei, 2014).

Empirical evidence suggests that increased use of AI TTS can be associated with higher pronunciation confidence. Studies by Choi et al. (2024) and Khanam et al. (2022) demonstrated a positive correlation between the frequency of TTS use and learners' perceived confidence in speaking English. Based on this foundation, this study focuses on the correlation between AI TTS use and students' confidence in pronouncing English words in an Indonesian EFL context. Exploring this correlation is crucial to understanding how new technologies can foster language proficiency and self-confidence among language learners.

METHOD

This study employed a quantitative correlational design to investigate the correlation between AI TTS use and students' self-confidence in pronouncing English words using survey techniques to obtain the information the researcher needed. Using a series of questionnaires as an instrument to collect the data. Questionnaires remain a widely utilized research instrument due to their capacity to collect extensive data efficiently and facilitate systematic analysis (Dörnyei & Taguchi, 2009). Data will be collected based on responses from the questionnaires.

Participants

The population of this study was the third and fifth semester students from the English Education Department at an Islamic Institute in Jambi. A total of 38 students, consisting of 12 males and 26 females, with an average age ranging from 19 to 21 years, who demonstrated adequate digital literacy, particularly in operating AI TTS technologies. The participants were selected through purposive sampling. Through the Google Form, participants voluntarily responded to the provided questionnaire. The sampling criteria are that on average, they have used the AI TTS application to learn to pronounce English words both in pronunciation classes

and outside pronunciation classes during their studies in the English department, and all of them have used the tool so that they could use their knowledge to answer the questionnaire well.

Throughout the study, ethical guidelines were strictly followed. Students were informed about the goal of the study, the kinds of data being gathered, and the voluntary nature of their participation before their participation. By indicating that they were willing to participate by completing the online questionnaire, they provided their consent. Responses were handled anonymously for academic purposes only, and no names or personally identifiable information were requested to maintain confidentiality. Prior to data collection, institutional authorities obtained ethical approval, and participants were reassured that their choice to participate or withdraw would not have an impact on their academic status.

Instruments

This study used a modified questionnaire as a data collection tool. The original questionnaire developed by Gagliardi (2016) consisted of 12 statements focusing on students' confidence in speaking English in an ESL context. Although this tool was developed several years ago, it is still frequently cited in recent EFL studies (e.g., Asnaini et al., 2025; Yuehua & Qiumei, 2023), demonstrating its validity and feasibility. To ensure alignment with the current research focus, specifically the correlation between AI TTS tool use and students' confidence in pronouncing English words, several modifications were made. Items addressing topics unrelated to AI-assisted pronunciation (such as familiarity with accents or teacher-dependent performance) were removed. Meanwhile, new items were created to capture students' frequency of TTS use and their perceived benefits, referring to indicators suggested by previous studies on AI-assisted pronunciation learning (Amin, 2024; Duong, 2022).

The revised questionnaire consisted of twelve items, divided into two parts: (1) five items (1,2,3,4, and 5) designed to assess AI TTS usage as the independent variable, and (2) seven items (6,7,8,9,10,11, and 12) intended to measure students' confidence in English pronunciation as the dependent variable. Example statements include: "I often use AI TTS tools to practice my English pronunciation" (usage) and "I feel confident speaking English in front of the whole class" (confidence). A five-point Likert scale was applied, ranging from 1 = Strongly Disagree to 5 = Strongly Agree. Higher scores indicate greater use of AI TTS tools and higher levels of pronunciation confidence. Only the questionnaire section of the original instrument was adapted for this research; no open-ended questions or additional qualitative components from the source study were included.

Data Collection

Before gathering data, the researcher asked the head of the English Education Department for formal permission to conduct the study within the department. To ensure that the survey methods complied with institutional policies, the researchers submitted a request for clearance to the academic authorities responsible for the third- and fifth-semester student groups. The instructors in charge of each class were then asked for permission to verify the appropriateness of the questionnaire items and the viability of conducting the survey.

Following permission, the researchers sent out the link to the questionnaire via the official WhatsApp group that was used for academic correspondence during both semesters. The online distribution strategy was selected because it provided more effective access, allowing students to complete the questionnaire at any time or location, and aligned with the typical use of digital platforms for academic purposes. Students were made aware that their involvement was entirely voluntary and that the information gathered would only be utilized for study.

Data analysis

The data collected from the questionnaire were coded on a Likert scale from 1 (Strongly Disagree) to 5 (Strongly Agree). The analysis first employed descriptive statistics to summarize student responses and provide an overview of score distributions for both variables. Descriptive statistics helped simplify complex data into easily understandable values (Sugiyono, 2019). To further examine the correlation between the two variables, a correlation test was conducted. Prior to determining whether Pearson's product-moment correlation or

Spearman's rho correlation would be applied, tests for normality, linearity, and outliers were carried out. The normality test was performed using the Shapiro–Wilk test, as the number of participants in this study (N = 38) was below 50, making this test more appropriate and reliable for small sample sizes. Linearity and outliers were assessed using scatter plots to ensure that the data met the assumptions required for correlation analysis.

Once the required assumptions were evaluated, the dataset was exported from Google Forms to Microsoft Excel and subsequently analyzed using SPSS version 21. If the data were normally distributed and demonstrated a linear correlation without significant outliers, Pearson's correlation was used. Otherwise, if the assumptions of normality or linearity were violated, Spearman's rho correlation was applied. The correlation coefficient (r) was then interpreted based on its strength and direction, while the p-value was examined to determine whether the correlation between AI TTS use and students' pronunciation confidence was statistically significant ($p < 0.05$).

Validity and Reliability

To ensure the quality of the instruments, validity and reliability tests were conducted. The validity of the instruments was examined through expert validation and empirical statistical tests. First, the instrument was reviewed by specialists, particularly lecturers with expertise in English language teaching. The expert validation process assessed item relevance to variable indicators, suitability for the research context, and clarity of wording. After revisions based on expert feedback, an empirical validity test was carried out using item–total correlation. All 12 questionnaire items exceeded the minimum acceptance threshold (r-table = 0.320; $p < 0.05$), indicating that every item was statistically valid and contributed meaningfully to the overall construct. Reliability was subsequently tested using Cronbach's Alpha to determine the internal consistency of the instrument. The coefficient obtained was $\alpha = 0.946$ for all 12 items, surpassing the recommended minimum benchmark of 0.70 and confirming that the instrument demonstrated a high level of reliability, consistently measuring the intended constructs across participants.

FINDINGS AND DISCUSSION

Findings

The purpose of this study was to examine the correlation between the use of AI Text-to-Speech (TTS) technology and students' self-confidence in pronouncing English words. The quantitative analysis was conducted using data from 38 English education students. The descriptive analysis showed that the majority of participants reported frequent use of AI TTS tools and relatively high levels of pronunciation confidence. To provide a clearer overview, Table 1 presents the participants' responses to each questionnaire item.

Table 1. Students' Responses to Each Questionnaire Item

Statements	SA	A	N	DA	SD
1. I often use the AI TTS to practice English pronunciation	7.9%	44.7%	36.8%	10.5%	0%
2. I use AI TTS outside of class hours (as self-study).	21.1%	36.8%	31.6%	10.5%	0%
3. I rely on AI TTS to check the pronunciation of new words.	23.7%	44.7%	23.7%	7.6%	0%
4. I use AI TTS to learn intonation, word stress, and rhythm (suprasegmental	10.5%	47.4%	34.2%	7.9%	0%
4. I use AI TTS to learn intonation, word stress, and rhythm (suprasegmental	13.2%	34.2%	36.8%	15.8%	0%
6. I feel confident speaking English in small groups.	18.4%	50%	23.7%	7.9%	0%
7. I feel confident speaking in front of the whole class.	10.5%	42.1%	36.8%	7.9%	2.6%
8. I feel confident pronouncing long or difficult words in English	10.5%	52.6%	23.7%	13.2%	0%
9. I feel my pronunciation is clear enough for the lecturer/teacher to understand	15.8%	44.7%	23.7%	15.8%	0%
10. I speak loudly and clearly when using English in class.	7.9%	31.6%	52.6%	5.3%	2.6%
11. I feel less anxious about speaking after practicing pronunciation.	7.9%	42.1%	39.5%	10.5%	0%
12. After practicing with AI TTS, I feel more confident pronouncing new words	15.8%	34.2%	36.8%	13.2%	0%

Note: SA = Strongly Agree; A = Agree; N = Neutral; D = Disagree; SD = Strongly Disagree.

Based on the table above, the majority of participants showed a positive response toward the use of AI TTS for pronunciation practice. Most students either agreed or strongly agreed with statements indicating that AI TTS helped them improve pronunciation, understand native speakers' accents, and enhance their confidence. For instance, more than half of the participants (52.6%) agreed that AI TTS increased their confidence when pronouncing difficult words (Item 8). Additionally, 44.7% reported frequently using AI TTS to check word pronunciation (Item 3). Overall, these findings suggest that AI TTS plays a supportive role in helping learners gain better pronunciation awareness and practice autonomously. The descriptive data also indicate that learners associate the use of AI TTS with higher levels of speaking confidence, particularly in small group interactions and pronunciation accuracy.

Pearson Correlation between Variables

Before the correlation test was carried out, a linearity check was performed using a scatter plot and the Test for Linearity procedure in SPSS. The scatter plot indicated an upward pattern from the lower left to the upper right, suggesting a positive linear relationship between the two variables. Furthermore, the significance values in the Test for Linearity showed that Sig. Linearity = 0.000 (< 0.05) and Sig. Deviation from Linearity = 0.162 (> 0.05), confirming that the data fulfilled the assumption of linearity required for Pearson's correlation analysis. In addition, the data were tested for normality using the Shapiro–Wilk test, which is considered more suitable for small sample sizes (N = 38). The results showed that both variables had significance values greater than 0.05 (AI TTS usage: Sig. = 0.407; pronunciation confidence: Sig. = 0.269). These results indicate that the data for both variables were normally distributed. Therefore, since the assumptions of both normality and linearity were met, the Pearson Product-Moment correlation test was subsequently performed to examine the relationship between AI TTS usage and students' pronunciation confidence.

The analysis demonstrates that there is a very strong and statistically significant positive correlation between students' use of AI TTS and their confidence in pronouncing English words ($r = 0.850$, $p = 0.000$). This means that the more frequently students use AI TTS technology, the higher their level of pronunciation confidence. The coefficient of determination ($r^2 = 0.7225$) indicates that approximately 72.25% of the variation in students' pronunciation confidence can be explained by their use of AI TTS tools, while the remaining 27.75% may be influenced by other factors such as motivation, classroom environment, or teacher feedback.

Discussion

These results are in line with other research demonstrating that using AI text-to-speech technology increases learners' confidence while pronouncing English sounds in addition to improving pronunciation accuracy. For instance, Asnaini et al. (2025) found that students who used TTS models regularly grew more at ease speaking English in class. In a similar vein, a study by Yuehua and Qiumei (2023) discovered that using TTS to practice pronunciation lessens students' anxiety and increases their readiness to speak, particularly when they come across unfamiliar or challenging words. AI-based practice provides a plausible explanation for these findings. AI TTS offers a setting where students can practice regularly and in private without worrying about being judged, in contrast to traditional pronunciation exercises that may expose students to criticism and humiliation in front of peers or teachers. Students can concentrate more on making appropriate sounds rather than worrying about social pressure in such a non-threatening setting, which seems to lessen emotional barriers. Having unlimited opportunities to repeat and imitate a model also gives learners greater control over the pace and intensity of practice, which can lead to a sense of ownership over their learning.

This ownership aligns with Self-Determination Theory's notion of autonomy. Deci and Ryan (2000) assert that learners become more intrinsically motivated and self-assured when they believe they can autonomously control and steer their learning. AI TTS serves as a personal training partner in this study, allowing students to track their own performance, identify areas for growth, and decide when and how frequently to exercise. The main psychological mechanism underlying improved speech confidence seems to be this cycle of

self-regulation, advancement, and ongoing self-recognition. These results show how useful AI-based devices could be as an additional teaching tool for pronunciation. With little exposure to appropriate speech models, many EFL students find it difficult to internalize phonetic patterns through classroom instruction alone. By providing clear, regular, and repeated auditory input that promotes cognitive processing and articulatory learning outside of the classroom, AI TTS helps overcome these restrictions. This is consistent with Godwin-Jones (2021), who highlights that frequent exposure to phonological patterns through technology-mediated pronunciation practice enables learners to assimilate them more successfully. Overall, these findings suggest that pronunciation instruction is most effective when both technical and affective dimensions are addressed simultaneously. AI TTS contributes to both aspects: AI TTS helps learners improve their English pronunciation accuracy while building the confidence needed to use those sounds in real-life communication situations.

In summary, the correlation results strongly indicate that AI TTS technology contributes significantly to students' self-confidence in English pronunciation. By combining authentic auditory models, accessibility, and self-paced learning, AI TTS becomes a powerful learning tool that fosters both proficiency and confidence among EFL students.

CONCLUSIONS

This study examined the correlation between the use of AI Text-to-Speech (TTS) technology and students' confidence in pronouncing English words. The results indicate a clear positive association: learners who engaged in regular pronunciation practice with AI TTS reported higher confidence when producing English sounds. AI TTS supported both technical development and affective growth by enabling students to repeat pronunciation privately, regulate their pace of practice, and internalize accurate speech patterns without social pressure. These findings suggest that AI-based pronunciation tools can strengthen learners' linguistic accuracy while simultaneously reducing anxiety and increasing their readiness to speak. The implications of this study highlight the potential of AI TTS as a complementary resource in pronunciation instruction. Integrating this technology into teaching provides learners with opportunities for individualized practice beyond classroom hours and fosters greater learner autonomy. Students are encouraged to make use of AI TTS independently as part of their routine language practice, as consistent listening and repetition can contribute to greater fluency and confidence. Future research may extend this study by examining larger and more varied populations or by employing mixed-method designs to gain deeper insights. Further exploration of how AI TTS influences other affective factors, such as anxiety, motivation, or self-efficacy would be valuable. Longitudinal work could also reveal the longer-term impact of sustained TTS use on both pronunciation accuracy and overall speaking performance.

REFERENCES

Abdel-Reheem Amin, E. (2024). EFL Students' Perception of Using AI Text-to-Speech Apps in Learning Pronunciation. *SSRN Electronic Journal*, *8984*, 887–903. https://doi.org/10.2139/ssrn.4746800

Asnaini, A., Tohamba, C. P. P., & Rusli, T. I. (2025). Self-Confidence Factors of Efl Students in Speaking Class: a Case Study in Higher Education. *Klasikal : Journal of Education, Language Teaching and Science*, *7*(1), 57–72. https://doi.org/10.52208/klasikal.v7i1.1263

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman and Company.

Benson, P. (2013). *Teaching and Researching: Autonomy in Language Learning* (2nd ed.). Routledge.

Bione, T., Grimshaw, J., & Cardoso, W. (2016). An evaluation of text-to-speech synthesizers in the foreign language classroom: learners' perceptions. *CALL Communities and Culture – Short Papers from EUROCALL*, 50–54. https://doi.org/10.14705/rpnet.2016.eurocall2016.537

Chau, T., Huensch, A., Hoang, Y. K., & Chau, H. T. (2022). The Effects of L2 Pronunciation Instruction on EFL Learners' Intelligibility and Fluency in Spontaneous Speech. *Tesl-Ej*, *25*(4), 1–28. https://doi.org/10.55593/ej.25100a7

Choi, H., Bae, J. S., Lee, J. Y., Mun, S., Lee, J., Cho, H. Y., & Kim, C. (2024). MELS-TTS: Multi-Emotion Multi-Lingual Multi-Speaker Text-To-Speech System Via Disentangled Style Tokens. *ICASSP 2024 – IEEE International Conference on Acoustics, Speech and Signal Processing*, 12682–12686.

Creswell, J. W., & Creswell, J. D. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th Editio). SAGE Publications, Inc.

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, *11*(4). https://doi.org/10.1207/S15327965PLI1104_01

Derwing, T. M., & Munro, M. J. (2005). Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, *39*(3), 379. https://doi.org/10.2307/3588486

Derwing, T., & Munro, M. (2015). *Pronunciation Fundamentals*.

Destiawati, L., Zulfa, V., & Marsita, J. (2024). The Effect of Self-Confidence on Students' English Speaking Ability in Front Office Subjects at State Vocational School 32 Jakarta. *International Education Trend Issues*, *2*(2), 225–238. https://doi.org/10.56442/ieti.v2i2.761

Dörnyei, Z. (2014). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Routledge (Taylor & Francis Group).

Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in Second Language Research: Construction, Administration, and Processing* (2nd Editio). Routledge (Taylor & Francis Group).

Duong, T. Van. (2022). The Effects of using Online Text-To Speech Tools on EFL Students' Perceptions in Learning Pronunciation. *International Journal of Science and Management Studies (IJSMS)*, *August*, 270–276. https://doi.org/10.51386/25815946/ijsms-v5i4p129

Fitrah, A. I., Korompot, C. A., Sakkir, G., & Fauzan, M. M. (2024). Assessing Self-Confidence Levels in EFL Students' Speaking Performance. *Celebes Journal of Language Studies*, *4*(2), 215–224. https://doi.org/10.51629/cjls.v4i2.200

Hidayad, F., Lestari, D., Laoli, A., Anam, S., & Despita. (2023). Academic Success in the Oral Presentation Course and Overall Self-Confidence. *Asian Journal of Applied Education (AJAE)*, *2*(3), 293–302. https://doi.org/10.55927/ajae.v2i3.4189

Holec, H. (1981). *Autonomy is the ability to take charge of one's own learning.*". https://www.responsiveclassroom.org/wp-content/uploads/2018/10/Whos-in-Charge-of-Learning-in-Your-Elementary-Classroom-School-Handout.pdf

Husna, A. (2021). *The correlation between students' pronunciation mastery and their self-confidence in speaking English at the eleventh grade of SMAN 1 Seunagan*. UIN Ar-Raniry Banda Aceh.

Jonker, C., Graham, M. A., & Ebersöhn, L. (2024). Measuring dimensions of teacher resilience in Africa: Self-efficacy and teacher efficacy. *South African Journal of Education*, *44*(3). https://doi.org/10.15700/saje.v44n3a2481

Khanam, F., Munmun, F. A., Ritu, N. A., Saha, A. K., & Mridha, M. F. (2022). Text to Speech Synthesis: A Systematic Review, Deep Learning Based Architecture and Future Research Direction. *Journal of Advances in Information Technology*, *13*(5), 398–412. https://doi.org/10.12720/jait.13.5.398-412

Lai, C. (2022). *Learner Autonomy and Technology* (1st ed.). The Routledge Handbook of Second Language Acquisition and Technology.

Lamb, M., Csizér, K., Henry, A., & Ryan, S. (2020). The palgrave handbook of motivation for language learning. In *The Palgrave Handbook of Motivation for Language Learning*. https://doi.org/10.1007/978-3-030-28380-3

MacIntyre, P. D., & Gardner, R. C. (1991). Language Anxiety: Its Relationship to Other Anxieties and to Processing in Native and Second Languages. *A Journal of Research in Language Studies*, *41*(4), 513–534.

Pajares, F., & Miller, M. D. (1994). Role of Self-Efficacy and Self-Concept Beliefs in Mathematical Problem Solving: A Path Analysis. *Journal of Educational Psychology*, *86*(2),

193–203. https://doi.org/10.1037/0022-0663.86.2.193

Ren, Y., Tan, X., Qin, T., Zhao, S., Zhao, Z., & Liu, T. Y. (2019). Almost unsupervised text to speech and automatic speech recognition. *36th International Conference on Machine Learning, ICML 2019*, *2019-June*, 9483–9492.

Rotjanawongchai, S. (2024). Developing EFL University Students' Willingness to Communicate and Self-Perceived Communicative Confidence through Self-Assessment of Self-Recorded Conversations. *REFLections*, *31*(2), 639–669. https://doi.org/10.61508/refl.v31i2.275092

Shabrina Nur Fadillah, Ade Ismail Ramadhan Hamid, & Abdul Halim. (2025). The Influence of Fear of Making Mistakes on Students' Self-Confidence and Engagement in English Classroom Presentations at MTs Muhammadiyah 1 Samarinda. *Journal of English Language Learning*, *9*(1), 825–829. https://doi.org/10.31949/jell.v9i1.13972

Shadiev, R., & Yang, M. (2020). Review of studies on technology-enhanced language learning and teaching. *Sustainability (Switzerland)*, *12*(2). https://doi.org/10.3390/su12020524

Sugiyono, A. (2019). *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif, dan R&D* (Issue January).

Suwari, I. G. A. A. R. P., Joni, D. A. A. W., & Wirastuti, I. G. A. P. (2023). the Correlation Between Self-Confidence and Speaking Skill of the Pre-Intermediate Efl Students. *JOSELT (Journal on Studies in English Language Teaching)*, *4*(1), 41–48.

Suwartono, T., & Eka Pertiwi, W. (2020). How Tertiary EFL Learners Dealt with Problems in Oral Communication in the Target Language. *International Journal of Advanced Science and Technology*, *29*(4), 2958–2966.

Svensson, I., Nordström, T., Lindeblad, E., Gustafson, S., Björn, M., Sand, C., Almgren/Bäck, G., & Nilsson, S. (2021). Effects of assistive technology for students with reading and writing disabilities. *Disability and Rehabilitation: Assistive Technology*, *16*(2), 196–208. https://doi.org/10.1080/17483107.2019.1646821

Wang, X., Gao, Y., Reynolds, B. L., & Wang, Q. (2025). Exploring Chinese EFL learners' beliefs about AI-mediated informal digital learning of English: Insights from Q Methodology. *Porta Linguarum*, *2025-September*(Special Issue 13), 131–146. https://doi.org/10.30827/portalin.viXIII.31925

Yuehua, H., & Qiumei, W. (2023). Assessing the English Speaking Self-Efficacy of Translation-Major Undergraduates: A Current Perspective. *International Journal of English Literature and Social Sciences*, *8*(3), 269–277. https://doi.org/10.22161/ijels.83.45

Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System*, *121*, Article 103259. https://doi.org/10.1016/j.system.2024.103259

Zhang, H., Du, L., Luo, S., Xun, H., Yang, A., & Tang, Y. (2025). Unlocking EFL achievement through informal digital learning: the mediating roles of anxiety and enjoyment among school-age students in the Chinese mainland. *Asian-Pacific Journal of Second and Foreign Language Education*, *10*(1). https://doi.org/10.1186/s40862-025-00352-x

Zhao, X.-J., Hwang, W.-Y., Hoang, A., & Su, Y.-C. (2023). Investigation of Facilitating English Speaking of Elementary School Students in Authentic Contexts with UEnglish. *Human-Centric Intelligent Systems*, *3*(3), 244–262. https://doi.org/10.1007/s44230-023-00031-5

Zimmerman, B. ., & Kitsantas, A. (2005). Homework practices and academic achievement: The mediating role of self-efficacy and perceived responsibility beliefs. *Contemporary Educational Psychology*, *Volume 30*(4), 397–417. https://doi.org/10.1016/j.cedpsych.2005.05.003