Plagiarism Checker X - Report

Originality Assessment

0%

Overall Similarity

Date: Jan 12, 2026 (08:51 AM)
Matches: 28 / 7890 words
Sources: 1

Remarks: No similarity found, your document looks healthy.

Verify Report:
Scan this QR Code

Journal of General Education and Humanities Vol. 5, No. 1, February 2026, pp. 687 – 706, https://doi.org/10.58421/gehu.v5i1.982 ISSN 2963-7147 687 Journal homepage: https://journal-gehu.com/index.php/gehu Enhancing Students' Speaking Self-Efficacy through an AIIntegrated Audio-Lingual Method in English Pronunciation Learning Ari Prasetyo Wibowo1, Mustaji2, Alim Sumarno3 Magister Teknologi Pendidikan, Fakultas Ilmu Pendidikan, Universitas Negeri Surabaya, Indonesia Article Info ABSTRACT Article history: Received 2025-12-06 Revised 2026-01-07 Accepted 2026-01-11 This study aimed to (1) develop an AI-integrated Audio-Lingual Method (ALM) learning media to enhance students' speaking selfefficacy and (2) evaluate its effectiveness in classroom settings. Employing a Research and Development (R&D) approach with the ADDIE model, the media integrated AI Speech Recognition and Textto-Speech (TTS) features within ALM-based learning sequences— dialogue modelling, imitation, repetition drills, and guided practice— across five modules: Materials, Reading Practice, Listening Practice, Speaking Tools, and Chatbot Practice. A quasi-experimental implementation with undergraduate students measured self-efficacy using a 19-item questionnaire adapted from Darmawan et al. (2021) and based on Bandura's self-efficacy theory. The instrument showed high reliability (Cronbach's α = 0.930) and item validity (r > 0.30). Results revealed significant increases in all self-efficacy dimensions—Personal Ability Belief (3.596→4.387), Growth Through Effort (3.620→4.480), and Influencing Factors (3.583→4.361). The Wilcoxon Signed-Rank Test confirmed a significant improvement (p = 0.000). Overall, AI-assisted, repetitionbased learning with instant feedback effectively enhanced learners' confidence, persistence, and emotional regulation in English speaking. Keywords: Artificial Intelligence (AI) Audio-Lingual Method English Language Learning English Pronunciation Speaking Self-Efficacy Corresponding Author: Ari Prasetyo Wibowo Magister Teknologi Pendidikan, Fakultas Ilmu Pendidikan, Universitas Negeri Surabaya, Indonesia Email: ariprasetyo2223@gmail.com 1. INTRODUCTION Artificial Intelligence (AI) has become 1 one of the most talked-about things in the modern world. It is not just for tech geeks

anymore, it is literally everywhere, from phones and business systems to classrooms. In education, AI has started changing the way people learn and teach. It is not perfect, but it is doing things that teachers alone sometimes cannot: giving real-time feedback, repeating drills endlessly without getting tired, and personalising learning in ways that feel, well, kind of human. In English language learning, AI tools like

https://doi.org/10.58421/gehu.v5i1.982 688 Speech Recognition and Text-to-Speech have opened new possibilities for helping students practice pronunciation and speaking [1]. Global reports show that this technology boom is no small thing. AI investment reached around 92 billion U.S. dollars in 2022 [2]. More than half of educational institutions worldwide 1 are expected to use AI tools by 2025 [3], and studies suggest that AI-based systems can improve learning outcomes by up to 30%. Teachers seem to like it too: about 87% of English teachers report using AI to improve classroom interaction [4]. AI-based Speech Recognition can now recognise words with 95% accuracy, while Text-to-Speech provides crystal-clear models of how words should sound [4]. However, despite all this fancy tech, English pronunciation is still a pain point for most Indonesian students. Honestly, it is not because they do not try; it is because their first language keeps interfering with their speech. A study by Kemelbekova et al. [5] found that nearly 69% of vocational students mispronounced English diphthongs, and Sholekhah and Fakhrurriana [6] reported that only about 9% of Javanese learners could pronounce the /ʒ/ sound correctly. It is sad to see students practice for years but still sound unsure when they speak English aloud. The root problem is not just technical; it is also psychological. Many students lack confidence; they fear making mistakes, especially in front of others. That is why methods like the Audio-Lingual Method (ALM) still matter. ALM may be old-school, but it works. It builds habits through mimicry, repetition, and pattern drills. The issue is that it needs tons of practice and feedback, which teachers cannot always provide in large classes. During classroom observation in the Speaking and Listening course at Universitas Negeri Surabaya (UNESA), it was found that pronunciation activities were still done the

“classic” way: students repeat after the lecturer or a recording, and that is it. No interactive feedback, no correction, no motivation. Some students looked bored, others anxious. A few did not even try to speak because they did not want to “sound stupid.” It is an absolute problem; accuracy goes down, and confidence follows right after. The present research was conducted to explore how integrating Artificial Intelligence (AI) with the Audio-Lingual Method (ALM) could enhance learners’ self-efficacy in English speaking. Drawing on Hekmat et al. [7] self-efficacy theory, the study posited that a technology-mediated learning environment equipped with AI Speech Recognition and Textto-Speech (TTS) systems could provide continuous feedback, repeated exposure, and affective support factors that are essential for shaping learners’ beliefs in their speaking competence. The integration of ALM’s repetition-based pedagogy with AI-driven feedback was therefore conceptualised as a dual mechanism to improve both linguistic performance and psychological assurance in oral communication. To address this objective, the study was guided by two primary research questions: RQ1: How can an AI-integrated Audio-Lingual Method (ALM) learning media be systematically designed and developed to enhance students’ self-efficacy in English speaking? RQ2: To what extent does the implementation of the AI-based Audio-Lingual Method media significantly improve students’ self-efficacy in speaking English? In summary, these research questions framed the investigation around both the pedagogical development and empirical validation of an AI-integrated ALM model. The study sought to contribute to the growing body of evidence in Computer-Assisted Language Learning

 (CALL) by illustrating how artificial intelligence can mediate not only pronunciation accuracy but also the affective dimensions of learning, particularly self-efficacy and confidence in oral performance. To answer these questions, this research develops an AI-integrated learning medium that allows students to practice pronunciation independently while receiving instant feedback on their speech accuracy. The goal is not just to make them pronounce words better, but to make them believe that they can. Because in the end, confidence in speaking does not come from

being perfect, it comes from knowing that you are getting better every time you try. In short, this study stands on the idea that a mix of human methods and machine feedback can lead to more than just linguistic accuracy. It can build self-trust. By merging the structured drills of the Audio-Lingual Method with the precision and consistency of AI, students can finally practice pronunciation without fear, improve their accuracy, and grow their confidence in speaking. 2. METHOD This study adopted a Research and Development (R&D) design using the ADDIE model (Analysis, Design, Development, Implementation, and Evaluation) as the methodological framework for systematically developing and validating an AI-integrated Audio-Lingual learning media to enhance students' speaking self-efficacy. The research procedure was structured to align with both technological development processes and pedagogical evaluation principles, ensuring that the resulting media product was not only technically sound but also pedagogically effective in supporting students' self-confidence in oral English communication. Figure 1. Research Framework The research framework, as illustrated in Figure 1, depicts the five interconnected phases of the ADDIE cycle. The Analysis stage involved identifying students' learning challenges and psychological barriers to speaking performance, with a focus on low selfconfidence and speaking anxiety in the Speaking and Listening course. Data were collected through classroom observation, lecturer interviews, and student needs analysis, revealing that learners required an interactive, feedback-oriented media environment to strengthen their belief in their speaking capability.

https://doi.org/10.58421/gehu.v5i1.982 690 In the Design phase, the instructional strategy was developed based on Bandura's [8] Self-Efficacy Theory, integrated with the Audio-Lingual Method (ALM) principles, including repetition, mimicry, and pattern drills. The design blueprint included the development of practice modules that allow continuous repetition and immediate feedback through AI-driven Speech Recognition and Text-to-Speech systems. Learning objectives, performance indicators, and evaluation instruments, including the Speaking Self-Efficacy Questionnaire, were also formulated during this

phase. The Development phase focused on producing the AI-integrated learning media prototype, combining language pedagogy with computational intelligence. Subject-matter experts and media design specialists validated the media to ensure its content relevance, usability, and instructional quality. Subsequent revisions were made based on expert feedback, followed by a pilot test involving ten students to assess practicality and initial learner response. The Implementation phase used a quasi-experimental design, with two groups: an experimental group that used the developed AI-based ALM media and a control group that continued learning through conventional methods. Both groups were drawn from undergraduate students enrolled in the Speaking and Listening course at the English Education Department, Universitas Negeri Surabaya (UNESA). The intervention spanned several sessions, during which the experimental group practised speaking using AI-mediated drills, interactive dialogues, and real-time feedback. Finally, the Evaluation phase comprised both formative and summative assessments. Formative evaluation was conducted during expert validation and pilot testing, while summative evaluation measured the effectiveness of the developed media in improving students' self-efficacy. The Speaking Self-Efficacy Questionnaire, adapted from [8] and grounded in Bandura's theoretical framework, was used as the main instrument. Quantitative data were analysed using Paired Sample t-Tests and Independent Sample t-Tests to determine significant differences between pre-test and post-test results at the 0.05 significance level. Reliability and validity tests confirmed that the questionnaire met the required psychometric standards (Cronbach's $\alpha \geq 0.70$; KMO > 0.6). Overall, this methodological framework ensured that each stage of the research was grounded in both instructional theory and empirical evaluation. The integration of AI-driven language technology into the Audio-Lingual Method created a structured environment for repeated speaking practice, while the feedback mechanism directly strengthened learners' self-efficacy, particularly by building confidence, persistence, and belief in their communicative ability. 3. RESULTS AND DISCUSSION Analysis Phase The Analysis Phase was conducted to identify affective and contextual barriers affecting students' self-efficacy in English speaking within the

Speaking and Listening course. According to Bandura [9], self-efficacy refers to an individual's belief in their ability to perform specific tasks successfully, as reflected in this study by students' confidence in

https://doi.org/10.58421/gehu.v5i1.982 691 their English-speaking ability. This stage involved classroom observation, semi-structured interviews with lecturers and students, and analysis of documentation to determine how far the learning environment supports or hinders students' speaking confidence. Observations during classroom sessions revealed that many students showed hesitation when asked to speak. They frequently avoided oral participation, preferring to respond in writing or remain silent rather than risk making mistakes. These behaviours indicated 1 a lack of belief in their linguistic competence and a fear of negative evaluation. Only a few confident students volunteered to speak, while the majority relied on passive participation. Interviews with lecturers provided more profound insights into the psychological dimension of the problem. The first lecturer remarked: "Students rarely take the initiative to speak unless they are called directly. Even when their answers are correct, they tend to hesitate because they do not believe they can express them well in English." This suggests that the issue is not merely linguistic accuracy, but rather low selfperceived ability. The second lecturer reinforced this observation, stating: "Most of them are afraid of making mistakes. The fear of being corrected in front of others seems stronger than their motivation to practice. They need a space where they can make errors without feeling embarrassed." A third lecturer pointed to the potential role of technology in addressing this affective challenge: "Some students have potential but are too shy to speak up. If they could practice using AI, where no one is judging them, they might gain confidence faster." These statements consistently emphasise that 1 the absence of a safe, supportive learning environment largely contributes to students' low self-efficacy. In a large classroom setting, lecturers cannot provide individualised, immediate feedback, leaving students uncertain about their progress and reinforcing their belief that they "are not good at speaking English." Student interviews further confirmed

these observations. Many expressed frustration and anxiety when asked to perform speaking tasks. One student admitted: “I always feel nervous when I have to speak in English. Sometimes I know what to say, but my voice just will not come out because I am scared of being wrong.” Another added: “When I make mistakes, my friends sometimes laugh, and that makes me more embarrassed. After that, I prefer to stay quiet.” However, several students expressed interest in using AI-based tools for practice, noting that such tools could provide a more private, judgment-free environment. As one student mentioned: “If there is an app that can tell me my mistakes but does not make me feel ashamed, I think I will practice more often. I want to know how I sound, but without the pressure.” These testimonies underline that low self-efficacy among students stems from anxiety, fear of social judgment, and a lack of private opportunities for self-improvement. They also demonstrate the strong perceived need for a learning medium that supports

https://doi.org/10.58421/gehu.v5i1.982 692 independent practice with immediate, non-judgmental feedback. Based on these findings, the development of an AI-based interactive learning medium became necessary to support both learners' pedagogical and psychological needs. The envisioned system should create a private and supportive learning space that encourages repeated practice, provides instant corrective feedback, and reduces performance anxiety. By integrating Speech Recognition technology, students could receive objective and real-time evaluations of their speech, while Text-to-Speech (TTS) technology could model accurate pronunciation patterns for imitation. Unlike traditional classroom settings, this AI-driven medium allows learners to practice autonomously, make mistakes freely, and track their improvement over time, fostering a sense of mastery and strengthening their self-efficacy in speaking English. Such a system directly aligns with the principles of the Audio-Lingual Method (ALM), which emphasises repetition and reinforcement and extends their benefits into the affective domain. Through AI-mediated feedback loops, learners not only improve linguistic accuracy but also develop psychological resilience and self-confidence in oral communication. The analysis phase

revealed that the primary barrier to students' progress in speaking is not a lack of skill but a lack of confidence. Therefore, the proposed AIintegrated Audio-Lingual media is designed not merely as a pronunciation trainer but as an empowering tool that enables students to practice independently, receive constructive feedback, and gradually rebuild their belief in their own ability to speak English effectively. Design Phase The design phase focused on transforming the analytical findings into a structured model for AI-based interactive learning media to enhance students' self-efficacy in English speaking. This stage emphasised how integrating Artificial Intelligence with the AudioLingual Method (ALM) could improve both linguistic performance and psychological confidence. Guided by the ADDIE framework, the design process ensured that the developed system was pedagogically sound, technologically feasible, and emotionally supportive. The Audio-Lingual Method provided the instructional backbone, emphasising repetition, mimicry, and feedback as the foundation of learning, while Bandura's self-efficacy theory served as the affective framework, reinforcing students' belief in their ability to succeed in speaking English. The first design stage, task inventory, identified and organised key learning activities and technological processes within the system. Six core modules were established: (1) Main Page, introducing the platform and providing motivational cues; (2) Learning Materials, containing vocabulary, phrases, and model sentences as linguistic input; (3) Reading Practice, where learners engage in oral reading analyzed by AI Speech Recognition for pronunciation accuracy; (4) Listening Practice, offering auditory exercises that strengthen phonetic perception; (5) Speaking Tools, providing pronunciation modeling for mimicry and repetition; and (6) Chatbot Practice, delivering interactive AI-based dialogues for contextual speaking application. Each module was designed to promote repeated engagement, constructive feedback, and low-anxiety learning conditions essential for fostering mastery and confidence in spoken English.

  Figure 2. Framework of AI-Based Interactive Learning Media Using Audio-Lingual Method to Enhance Students' Self-Efficacy in English

Speaking The subsequent design stage focused on composing performance objectives that define measurable learning outcomes related to self-efficacy development. These objectives were framed using Bandura's four sources of self-efficacy: mastery experience, vicarious experience, verbal persuasion, and affective regulation. The AI-based media supports mastery experience by allowing learners to achieve repeated success in pronunciation through continuous practice. Vicarious experience is facilitated through AI-generated pronunciation models that serve as consistent references. Verbal persuasion is delivered via automated feedback, encouragement messages, and score-based responses. Affective regulation is addressed by providing a private practice space, which reduces social pressure and performance anxiety. All elements were deliberately embedded to form a reinforcement cycle that strengthens confidence, persistence, and learners' perception of their speaking ability [10]. The final stage involved designing the testing strategy to evaluate both the effectiveness and feasibility of the media. A pre-test and post-test design was used to examine changes in students' self-efficacy before and after using the AI platform. Measurement employed a self-efficacy questionnaire adapted from Darmawan et al. [11], grounded in Bandura's Social Cognitive Theory. The instrument assessed beliefs about speaking competence, learning persistence, and emotional control during speaking tasks. Expert validation was also conducted to review pedagogical relevance and technical quality of the media. Data analysis used descriptive and inferential statistics to identify the impact of the media on learner confidence and engagement. Overall, the design phase produced a media framework that is pedagogically grounded and psychologically supportive, yet not overly complicated. The integration of the Audio-Lingual Method with self-efficacy theory allows repetition without boredom and feedback without judgment. Through AI-driven responses, structured learning sequences, and self-paced interaction, the media offers a learning space that feels safe, motivating, and practical. Learners are encouraged to practice more, handle anxiety better, and gradually build confidence in spoken English.

https://doi.org/10.58421/gehu.v5i1.982 694 Develop Phase The Develop Phase focused on turning the design concepts into a fully functional and interactive AI-based learning media that helps students strengthen their self-efficacy in English speaking. At this stage, all components designed earlier, including content structure, practice types, and user interaction flow, were developed and integrated into the application's core system. The final product was realised through five main learning features: Materials, Reading Practice, Listening Practice, Speaking Tools, and Chatbot Practice. These features were not only built to train pronunciation but also to create a safe, encouraging environment where students could practice repeatedly, receive immediate feedback, and gradually build their confidence to speak English without fear of making mistakes [12]. During the content development process, each feature was carefully crafted to reflect both the pedagogical structure of the Audio-Lingual Method (ALM) and the psychological foundation of Bandura's self-efficacy theory. The Materials feature provides structured vocabulary, phrases, and model sentences as the foundation for learners before they begin oral practice. The Reading Practice and Listening Practice modules function as the central audio-based drills, allowing students to practice pronunciation accuracy and comprehension with real-time AI feedback through Speech Recognition and Text-to-Speech systems. The Speaking Tools feature provides pronunciation models that students can repeatedly imitate to form automatic speaking habits, while the Chatbot Practice serves as a digital speaking companion, enabling students to engage in simple, real-time conversations without social pressure. Each of these features was designed to give students consistent opportunities for practice and reinforcement, key elements in developing mastery and confidence. Table1. Media Content Development and Interface Design Overview No Application Feature Content Development Result Mockup Interface Design (UI) Relation to Previous Design Stage 1 Main Page Contains a brief introduction and description of the AI-based learning media, including its purpose and navigation guide. Implements 1 the concept of user introduction and motivation from the design phase.

https://doi.org/10.58421/gehu.v5i1.982 695 No Application Feature Content Development Result Mockup Interface Design (UI) Relation to Previous Design Stage 2 Learning Materials Vocabulary, phrases, and English sentences developed as the conceptual foundation for pronunciation learning.) Realises the material design as the initial conceptual framework for pronunciation practice. 3 Reading Practice Oral reading exercises to train pronunciation accuracy, integrated with AIbased Speech Recognition for instant feedback. Implements the design of pronunciation drills based on Speech Recognition technology aligned with the ALM pattern. 4 Listening Practice Listening comprehension exercises requiring learners to transcribe what they hear for accuracy reinforcement. Realises the design of listening practice following the Audio-Lingual Method sequence. 5 Speaking Tools Provides model pronunciations and standard articulation samples for learners to imitate and repeat. Implements the modelling and repetition components of the ALM framework. 6 Chatbot Practice Interactive textbased conversation exercises with AI responses to simulate real communication. (Realises the design of communicative reinforcement and habit formation through interactive AI engagement.

https://doi.org/10.58421/gehu.v5i1.982 696 To support media, the developed application uses a combination of static visual elements and dynamic audio, creating an interactive multimedia learning experience. The interface was designed with simple icons, readable text, and intuitive navigation to help users move naturally through the exercises. However, the real core of the media lies in its audio components. The AI-based Text-to-Speech (TTS) system generates accurate pronunciation models, while the Speech Recognition module captures and analyzes students' spoken input to provide automatic feedback. This two-way interaction between the user and system makes the learning experience active and engaging. Essentially, the application functions as an AIassisted language lab, allowing learners to practice as much as they want, at their own pace, and in

a stress-free setting, an environment proven to support emotional regulation and enhance self-efficacy. The learning guidance for both teachers and students wassalso developed to ensure smooth classroom implementation. For instructors, the guide outlines each ALM stage from dialogue presentation to evaluation and feedback, providing details on how to integrate the AI media into class activities. Teachers act more as facilitators, guiding students to interact with the system, monitor their progress, and provide personalised encouragement based on AI-generated reports. Meanwhile, the student guide focuses on independent practice: listening to model dialogues, repeating aloud, checking their pronunciation score, and improving through feedback loops. The goal was to make learners feel supported yet autonomous, able to take control of their own speaking development. Figure 3. Audio-Lingual Method Flow for Self-Efficacy Development In short, the Develop Phase successfully transformed the conceptual design into a real, usable AI-integrated language learning tool. The final product combines the systematic repetition of the Audio-Lingual Method with real-time AI feedback and self-paced learning design. This integration ensures that students can practice pronunciation intensively while gradually developing confidence, resilience, and a sense of achievement in their speaking performance. The interactive and adaptive nature of the application ultimately supports both linguistic accuracy and psychological empowerment, helping learners not only speak better English but also believe they can.

https://doi.org/10.58421/gehu.v5i1.982 697 Implementation Phase The implementation phase focused on applying the developed AI-based learning media in a real classroom setting to examine its impact on students' self-efficacy in English speaking. This phase was designed as a quasi-experimental study involving two groups: an experimental class using the AI Speech Recognition and Text-to-Speech system integrated with the Audio-Lingual Method (ALM), and a control class that followed conventional pronunciation teaching without technological assistance. Each group consisted of 30 students enrolled in the Speaking and Listening course during the same semester, with the same instructor

teaching both classes to ensure consistency in instructional delivery [13]. Before implementation, the instructor conducted a brief orientation to ensure a uniform understanding of how to integrate AI-based media into the ALM learning cycle. The training emphasised how to guide students during dialogue presentation, repetition drills, and communicative practice while using the AI feedback system as an instructional aid. In the experimental class, students interacted with five primary media features Materials, Reading Practice, Listening Practice, Speaking Tools, and Chatbot Practice to receive instant feedback on their pronunciation and speaking performance. These features not only allowed for repeated practice but also created a psychologically safe environment where students could rehearse speaking without fear of making mistakes in front of peers. This design aligned with the theoretical basis of self-efficacy [9], which emphasises mastery experience, vicarious learning, verbal persuasion, and emotional regulation as core sources of efficacy beliefs. The AI system provided direct feedback that served as verbal persuasion (“Good pronunciation!”), measurable progress scores that strengthened mastery experience, and repeated autonomous practice that supported self-regulation. As a result, students were able to make progress at their own pace and internalise the belief that improvement is attainable through consistent effort, a principle referred to as the growththrough-fort dimension of self-efficacy [11]. Quantitative data from the self-efficacy questionnaire revealed substantial improvement across all three dimensions of speaking selfefficacy: Personal Ability Belief, Growth Through Effort, and Influencing Factors. The descriptive statistics are presented in Table 2 below. Table 2. Descriptive Statistics of Students’ Speaking Self-Efficacy (Pre-Test and PostTest, N = 30) Indicator Time Mean Std. Deviation Minimum Maximum Interpretation Personal Ability Belief (Q1–Q8) Pre-Test 3.596 0.941 1 5 Moderate self-confidence before using AI-based media Growth Through Effort (Q9–Q13) Pre-Test 3.620 0.974 1 5 Belief in improvement through effort was present but not strong. Influencing Factors (Q14– Q19) Pre-Test 3.583 0.968 1 5 Moderate control over anxiety, expression, and gestures

 Indicator Time Mean Std. Deviation Minimum Maximum Interpretation Personal Ability Belief (Q1–Q8) Post-Test 4.387 0.763 2 5 High confidence in speaking ability after AIbased training Growth Through Effort (Q9–Q13) Post-Test 4.480 0.748 1 5 Strong belief that effort and practice lead to improvement Influencing Factors (Q14– Q19) Post-Test 4.361 0.831 2 5 Better control of emotional and expressive aspects of speaking The pre-test results indicated that students had a moderate level of self-efficacy before using AI-assisted learning media. Most students expressed hesitation to participate actively in speaking sessions due to fear of making pronunciation errors and being judged by others. One of the instructors noted, “Some students still feel embarrassed or hesitant to speak in front of the class; they are afraid of being wrong. A private AI-based practice environment can help them gain confidence.” Following implementation, post-test scores increased significantly across all dimensions. The Personal Ability Belief mean rose from 3.596 to 4.387, reflecting an apparent increase in students’ self-confidence in communicating in English. The most notable improvement occurred in the Growth Through Effort dimension, where the mean increased from 3.620 to 4.480, indicating a strong belief among students that consistent practice leads to tangible improvement. The Influencing Factors dimension also improved substantially (from 3.583 to 4.361), suggesting that students improved their emotional regulation and control over the physical and vocal aspects of speaking performance. Overall, the results of this implementation phase show that the AI-integrated AudioLingual Method effectively enhanced speaking self-efficacy by combining structured repetition, automated feedback, and self-paced learning. The students’ ability to see their own progress in real time fostered motivation and perseverance, while the safe, technologymediated learning environment reduced speaking anxiety and boosted their confidence to use English actively. Evaluation Phase The evaluation phase was conducted to assess the validity, reliability, and effectiveness of the developed AI-based learning media in improving students’ speaking self-efficacy. This stage ensured that all research instruments and analytical procedures were statistically valid and reliable, and

capable of accurately reflecting the intervention's true impact. The evaluation focused on three aspects: (1) instrument validity and reliability, (2) normality and homogeneity testing, and (3) the effectiveness test of the AI-based media through the Wilcoxon Signed-Rank Test on self-efficacy scores. The self-efficacy questionnaire consisted of 19 items (Q1–Q19) developed based on Bandura's (1997) self-efficacy construct and adapted by Darmawan et al. (2021) for English-

https://doi.org/10.58421/gehu.v5i1.982 699 speaking contexts. The items covered three dimensions: (1) Belief in Speaking Ability, (2) Belief that Ability Can Grow with Effort, and (3) Factors Influencing Speaking Self-Efficacy. The validity and reliability tests were conducted to ensure that the instrument accurately measured these constructs. The validity test used the Corrected Item–Total Correlation method, with a minimum correlation coefficient criterion of $r > 0.30$. The results indicated that all items had correlation coefficients above 0.30, confirming that every statement effectively represented the self-efficacy construct. Meanwhile, the reliability test using Cronbach's Alpha yielded $\alpha = 0.930$, well above the 0.70 threshold. This indicates that the instrument possessed excellent internal consistency and stability. The "Cronbach's Alpha if Item Deleted" values showed that removing any item would not increase overall reliability, proving that all items contributed positively to the instrument's cohesiveness. Table 3. Instrument Validity Results Item Corrected Item–Total Correlation (r) Criterion (r > 0.30) Decision Q1 0.689 ✓ Valid Q2 0.616 ✓ Valid Q3 0.738 ✓ Valid Q4 0.519 ✓ Valid Q5 0.538 ✓ Valid Q6 0.732 ✓ Valid Q7 0.692 ✓ Valid Q8 0.588 ✓ Valid Q9 0.640 ✓ Valid Q10 0.549 ✓ Valid Q11 0.827 ✓ Valid Q12 0.570 ✓ Valid Q13 0.597 ✓ Valid Q14 0.700 ✓ Valid Q15 0.585 ✓ Valid Q16 0.341 ✓ Valid Q17 0.655 ✓ Valid Q18 0.603 ✓ Valid Q19 0.626 ✓ Valid Table 4. Instrument Reliability Results Item Cronbach's Alpha if Item Deleted Decision Q1 0.925 Retained Q2 0.927 Retained Q3 0.924 Retained Q4 0.929 Retained Q5 0.929 Retained Q6 0.925 Retained Q7 0.925 Retained Q8 0.928 Retained Q9 0.927 Retained Q10 0.928 Retained Q11 0.922 Retained

https://doi.org/10.58421/gehu.v5i1.982 700 Item Cronbach's Alpha if Item Deleted Decision Q12 0.928 Retained Q13 0.927 Retained Q14 0.925 Retained Q15 0.928 Retained Q16 0.932 To Be Reviewed Q17 0.926 Retained Q18 0.928 Retained Q19 0.927 Retained The results above confirm that the self-efficacy questionnaire is both valid and reliable, capable of measuring the targeted construct with precision and consistency. Therefore, it was considered appropriate for evaluating the effectiveness of AI-integrated learning media. Before conducting inferential analysis, the normality and homogeneity tests were performed to determine the appropriate statistical approach. Since the self-efficacy data consisted of pre-test and post-test measurements from a single experimental group (N = 30), the Shapiro–Wilk Test was used. The test results revealed that the p-values were below 0.05 for both pre-test (p = 0.020) and post-test (p = 0.000), indicating that the data were not normally distributed. Consequently, a non-parametric Wilcoxon Signed-Rank Test was used to analyse differences in students' self-efficacy before and after the implementation of AIbased media. Table 5. Normality Test Results for Speaking Self-Efficacy (Shapiro–Wilk Test) Test Stage Statistic Sig. (p-value) Criterion (p > 0.05) Decision Pre-test 0.915 0.020 Not Normal Data Not Normal Post-test 0.834 0.000 Not Normal Data Not Normal The findings confirmed that the self-efficacy data did not meet the normality assumptions, necessitating the use of non-parametric methods. The Wilcoxon Signed-Rank Test was applied to examine whether 1 there was a significant difference between students' self-efficacy before and after using the AI-integrated learning media. The results showed that all 30 students increased their self-efficacy scores after the intervention, with no declines (negative ranks = 0). Furthermore, the significance value obtained (p = 0.000) was lower than the 0.05 threshold, indicating a statistically significant difference. Table 6. Wilcoxon Signed-Rank Test Results for Speaking Self-Efficacy Analysis Component Result Positive Ranks (Post-test > Pre-test) 30 Negative Ranks (Post-test < Pre-test) 0 Ties (Post-test = Pre-test) 0 Total Respondents 30 Z 0.000 Asymp. Sig. (2-tailed) 0.000 Decision (α = 0.05) Significant

https://doi.org/10.58421/gehu.v5i1.982 701 The evaluation results clearly demonstrate that the AI-based Audio-Lingual Method media significantly improved students' self-efficacy in speaking English. The absence of negative ranks indicates that every student benefited from the implementation, while the perfect significance value (p = 0.000) confirms the robustness of this improvement. From a psychological perspective, the findings align with Getman et al. [9], who identified verbal persuasion and emotional regulation as primary sources of efficacy beliefs. The AI-based system, by providing immediate feedback, opportunities for repetition, and a non-judgmental practice space, contributed to stronger beliefs in speaking ability and greater persistence in practice. Students reported feeling more confident and less anxious when speaking English after practising with the AI system. As one participant expressed, this evidence confirms that the developed media not only functions effectively as a learning tool but also acts as a psychological scaffold that enhances motivation, self-belief, and communication readiness. Therefore, the evaluation phase demonstrates that the AI-integrated Audio-Lingual Method is an effective, empirically supported approach to enhancing speaking self-efficacy among university students. Discussion RQ:1 Development and Feasibility of the AI-Based Learning Media The development of the AI-integrated Audio-Lingual Method (ALM) media in this study has shown that combining traditional repetition-based pedagogy with artificial intelligence feedback mechanisms effectively improves both pronunciation accuracy and students' self-efficacy in speaking English. The design followed the ADDIE framework, resulting in a pedagogically structured and technologically feasible product consisting of five interactive modules Materials, Reading Practice, Listening Practice, Speaking Tools, and Chatbot Practice. Each component was carefully designed to reinforce mastery experience, verbal persuasion, and affective regulation, reflecting Bandura's principles of self-efficacy. The findings from the implementation phase demonstrated significant gains across all self-efficacy dimensions Personal Ability Belief, Growth Through Effort, and Influencing Factors, with statistical significance (p = 0.000). This shows that consistent AI-based

repetition and feedback cycles provided learners with tangible mastery experiences, leading to increased confidence and emotional stability during speaking tasks. The use of AI Speech Recognition and Text-to-Speech (TTS) features allowed learners to receive instant, objective, and non-judgmental feedback, a key factor in reducing language anxiety and fostering persistence. These outcomes align closely with those of Yuri Lolita[14], who found that the use of Computer-Assisted Language Learning (CALL) significantly improved students' English vocabulary comprehension and motivation. In their study, students who used CALL-based software were more engaged, autonomous, and confident than those in traditional classrooms. The researchers emphasised that CALL's interactive design promotes active participation, repeated exposure, and reduced fear of failure, which are also mirrored in this study's AI-integrated ALM model [12]. Figure SEQ Figure * ARABIC 1. Title of picture or chart

https://doi.org/10.58421/gehu.v5i1.982 702 Furthermore, the integration of AI-based pronunciation training systems in recent studies supports these findings. For instance, Lolita et al. [15] demonstrated that AI-assisted pronunciation tools like Listnr and Murf AI significantly improved learners' pronunciation accuracy, engagement, and confidence while providing a supportive and anxiety-free environment [15]. Similarly, Getman et al. [16] reported that AI-powered speech recognition integrated into CALL systems for children enhanced both pronunciation quality and learning motivation through gamified, feedback-rich interfaces [16]. The present study's AI-integrated ALM framework also shows strong parallels with Bogach et al. [17], who proposed a CAPT (Computer-Assisted Pronunciation Training) system that uses ASR-based rhythm and intonation feedback. Their findings emphasised that combining feedback on sound, rhythm, and intonation in a single environment leads to better phonological learning and more active engagement. Similarly, Bogach et al. [18] confirmed that AI-based pronunciation tools not only enhance intelligibility but also increase motivation and alleviate speaking anxiety, findings that echo the psychological outcomes observed in this study. From a pedagogical

perspective, the feasibility of integrating AI within the AudioLingual Method underscores the effectiveness of combining repetition, mimicry, and feedback with adaptive learning technologies. The AI-assisted model serves as a selfregulating environment in which learners engage in continuous self-assessment and improvement, aligning with the principles of learner autonomy and self-directed practice identified in CALL research. As Vančová [18] and Abimanto and W. Sumarsono [19] observed, computer-assisted instruction enhances self-efficacy by allowing learners to take greater control of their learning process, leading to increased motivation and persistence. Technologically, this study extends the functionality of traditional CALL frameworks by embedding AI-driven feedback and self-paced interactivity into pronunciation training. Similar systems that employ automatic pronunciation scoring and visual-speech feedback have been shown to correct learners' mispronunciations and foster phonetic awareness effectively. The design of this study's AI media, combining speech recognition, text-to-speech, and motivational feedback, therefore represents a convergence of AI-based CALL innovation and psychologically grounded language pedagogy. Overall, the AI-integrated ALM model developed and validated in this research is both pedagogically sound and technologically feasible. It not only improves pronunciation learning outcomes but also enhances learners' self-efficacy, motivation, and emotional regulation. By aligning with evidence from CALL research and recent AI-based pronunciation studies, this framework confirms that artificial intelligence can effectively extend the Audio-Lingual Method beyond mechanical drills to foster meaningful, confident, and self-determined language learning. RQ:2 Effectiveness of the AI-Integrated Audio-Lingual Method in Enhancing Students' Self-Efficacy The findings of this study clearly demonstrate that integrating AI-based interactive media using the Audio-Lingual Method (ALM) effectively improved students' pronunciation accuracy and overall self-efficacy in English speaking. Based on the

 quantitative results, students in the experimental group (N = 30) showed substantial gains across all three self-efficacy

dimensions after using the AI-integrated media. The mean score for Personal Ability Belief increased from 3.596 to 4.387, indicating a marked improvement in learners' confidence to speak English. Similarly, Growth Through Effort rose from 3.620 to 4.480, reflecting stronger beliefs that consistent practice leads to real improvement. The Influencing Factors dimension also improved from 3.583 to 4.361, suggesting that students became more capable of managing anxiety, regulating emotions, and exercising expressive control during speech. These numerical improvements were further supported by the Wilcoxon SignedRank Test, which showed a significant difference ($p = 0.000 < 0.05$), confirming that AIbased media use had a statistically significant effect on learners' self-efficacy and pronunciation performance. Additionally, the reliability coefficient (Cronbach's Alpha = 0.930) confirmed that the measurement instrument was highly consistent and dependable, ensuring the credibility of these results. The validity test also showed that all 19 questionnaire items had corrected item–total correlation coefficients above 0.30, confirming that each item was valid in measuring self-efficacy. From a pedagogical standpoint, these results reveal how the AI-integrated AudioLingual Method supports both linguistic and psychological dimensions of learning. The systematic repetition characteristic of ALM enabled students to internalise pronunciation patterns through imitation and correction, while AI-driven Speech Recognition and Text-toSpeech (TTS) features provided instant, individualised feedback. This dual mechanism, continuous repetition, and immediate feedback helped students self-correct their errors in real time, thus enhancing both accuracy and confidence. These findings align with the study of Abimanto and W. Sumarsono [20], who reported that using Mobile Assisted Language Learning (MALL) through ELSA Speak significantly improved students' pronunciation and speaking confidence. They emphasised that digital tools such as ELSA help learners practice independently in a low-anxiety, personalised environmenthere errors are treated as opportunities for growth rather than failure [14]. Similarly, in the current study, students expressed that AI feedback felt "encouraging and private," motivating them to repeat pronunciation drills without fear of peer judgment. Moreover, comparable evidence is found

in international studies. Haliza and Lolita [21] proved that AI-powered speech recognition tools significantly enhanced ESL learners' pronunciation accuracy and engagement, underlining that immediate, objective feedback is crucial for improvement [22]. Likewise, Bogach et al. [12] found that students who practised using AI pronunciation tools such as Listnr and Murf AI achieved significant gains in accuracy and greater speaking confidence. In a related study, Wang and Shi [23] developed an AI-based pronunciation correction model using Hidden Markov Models (HMM), which effectively enhanced learners' pronunciation precision by guiding them through automated correction and acoustic feedback loops. This mirrors the technological design of the current study's AI system, which combines phonetic modelling with real-time corrective scoring.

https://doi.org/10.58421/gehu.v5i1.982 704 In addition, Vančová [24] emphasised that AI-driven pronunciation tools not only improve speech intelligibility but also reduce speaking anxiety, allowing learners to practice pronunciation more confidently. This psychological comfort was also observed in the present study. Students reported that the AI platform felt "safe and supportive," helping them to overcome shyness and hesitation when practising English aloud. The trend of improvement is also consistent with Abimanto and Sumarsono [25], who found that integrating Google Read Along, an AI-powered pronunciation app, led to a 65.73% increase in pronunciation accuracy among EFL learners. Their findings reinforced that instant, gamified feedback promotes better articulation and learner motivation. Furthermore, Nguyen and Vo [26] demonstrated that students who used SpeechAce, an AI-driven pronunciation evaluation system, achieved significantly higher post-test scores than those using traditional methods. Learners particularly valued the instant feedback and interactive interface, which fostered greater engagement and self-awareness. Overall, the results of this study reinforce that the AI-integrated ALM model not only improves pronunciation accuracy but also strengthens self-belief, persistence, and emotional regulation. The mean score increases across all three self-efficacy dimensions, Personal Ability Belief (+0.791), Growth Through Effort (+0.860),

and Influencing Factors (+0.778), indicating that learners are not just speaking more accurately, but doing so with greater confidence and control. The Wilcoxon test significance (p = 0.000) further confirms that these improvements were not due to chance, but to the real pedagogical and psychological benefits of AI-assisted practice. In simpler terms, the integration of AI speech recognition and feedback mechanisms provided students with a “personal tutor” that offered instant, tireless, and non-judgmental correction, a learning experience rarely achievable in conventional classrooms. This reflects a new paradigm in pronunciation pedagogy where technology not only teaches pronunciation, but also empowers learners emotionally, encouraging them to believe that “they can improve if they try.” 4. CONCLUSION This study aimed to (1) develop an AI-integrated interactive learning medium based on the Audio-Lingual Method (ALM) for the Speaking and Listening course, and (2) examine its effectiveness in improving students’ speaking self-efficacy. The development process followed the ADDIE model, resulting in a functional and validated learning media that integrates Speech Recognition and Text-to-Speech technologies to facilitate pronunciation practice and self-directed speaking exercises. The design was evaluated by media and pedagogical experts and found to be highly feasible for classroom implementation, both in terms of technical quality and instructional alignment with ALM principles such as mimicry, repetition, and pattern practice. The implementation results indicated that the use of the AI-based learning media led to a significant improvement in students’ self-efficacy in English speaking. Statistical analysis using the Wilcoxon Signed-Rank Test showed that all participants experienced positive growth in confidence, belief in their speaking ability, and persistence in practising English after using the media. These findings suggest that integrating AI feedback

 mechanisms, repetitive practice, and error correction contributed to students’ stronger sense of control, reduced speaking anxiety, and increased motivation to communicate in English. Thus, the developed AI-integrated ALM media not only enhances technical pronunciation skills but also supports affective

growth by fostering learners' confidence and belief in their speaking potential. Despite its promising results, this research has several limitations. The study was conducted with a relatively small sample from a single institution and within a limited duration of classroom application. Future research could involve larger, more diverse participant groups, longer implementation periods, and additional features, such as adaptive learning analytics, to further personalise feedback. It is also recommended that educators continue to integrate AI-based tools as supplementary resources rather than replacements for human interaction, ensuring that technology remains aligned with pedagogical goals and learner needs. Overall, this study contributes to the growing body of research on AIenhanced language learning by demonstrating that AI-integrated Audio-Lingual learning environments can effectively improve students' speaking self-efficacy and engagement in English language education.

ACKNOWLEDGEMENTS We would like to express our deepest gratitude to all parties who have contributed to this research. Thank you to our colleagues who have provided advice, support, and inspiration throughout the research process. We also wish to extend our appreciation to everyone who took the time to participate in this study. Additionally, we are grateful to the institutions that have provided support and facilities for the conduct of this research. All contributions and assistance have been invaluable to the smooth progress and success of this study. Thank you for all the hard work and collaboration that has been established.

REFERENCES

[1] N. V. Nguyen, T. T. Vo, and V. D. T. Tran, "AI-Driven Pronunciation Assessment: The Impact of SpeechAce on EFL Learners' Pronunciation Competency," Computer-Assisted Language Learning Electronic Journal, vol. 26, no. 3, pp. 84–106, May 2025, doi: 10.54855/callej.252635. [2] H. Vančová, "AI and AI-powered tools for pronunciation training," Journal of Language and Cultural Education, vol. 11, no. 3, pp. 12–24, Dec. 2023, doi: 10.2478/jolace-2023-0022. [3] E. Mohammadkarimi, "Exploring The Use Of Artificial Intelligence In Promoting English Language Pronunciation Skills," LLT Journal: A Journal on Language and Language Teaching, vol. 27, no. 1, pp. 98–115, Apr. 2024, doi: 10.24071/llt.v27i1.8151. [4] Y. Lolita, E. Boeriswati, and N. Lustyantie, "The Impact of

Computer Assisted Language Learning (CALL) Use of English Vocabulary Enhancement," Linguistic, English Education and Art (LEEA) Journal, vol. 4, no. 1, pp. 206–221, Dec. 2020, doi: 10.31539/leea.v4i1.1896. [5] Z. Kemelbekova, X. Degtyareva, S. Yessenaman, D. Ismailova, and G. Seidaliyeva, "AI in teaching English as a foreign language: Effectiveness and prospects in Kazakh higher education," XLinguae, vol. 17, no. 1, pp. 69–83, Jan. 2024, doi: 10.18355/XL.2024.17.01.05. [6] M. F. Sholekhah and R. Fakhrurriana, "The Use of ELSA Speak as a Mobile-Assisted Language Learning (MALL) towards EFL Students Pronunciation," JELITA: Journal of Education, Language Innovation, and Applied Linguistics, vol. 2, no. 2, pp. 93–100, Jul. 2023, doi: 10.37058/jelita.v2i2.7596. [7] H. Hekmat, P. Balakrishnan, and T. Arumugam, "The Efficacy of AI-Powered Speech Recognition in Enhancing English Pronunciation for Second Language Learners," International Journal on Science and Technology, vol. 16, no. 2, May 2025, doi: 10.71097/IJSAT.v16.i2.5170.

https://doi.org/10.58421/gehu.v5i1.982 706 [8] Y. Getman et al., "Developing an AI-Assisted Low-Resource Spoken Language Learning App for Children," IEEE Access, vol. 11, pp. 86025–86037, 2023, doi: 10.1109/ACCESS.2023.3304274. [9] A. Bandura, Self-efficacy: The exercise of control. W. H. Freeman, 1997. [10] Darmawan, S. P. Alam, and O. N. Nirma, "Speaking Self-Efficacy of EFL Students of Pre-Service Teaching Program in EFL Classroom Setting," JET (Journal of English Teaching), vol. 7, no. 2, pp. 150–162, Jun. 2021, doi: 10.33541/jet.v7i2.2582. [11] W. Darmawan, S. P. Alam, and O. N. Nirma, "Speaking self-efficacy of EFL students of pre-service teaching program in EFL classroom setting," Journal of English Teaching, vol. 7, no. 2, pp. 150–162, 2021, doi: 10.33541/jet.v7i2.2582. [12] N. Bogach et al., "Speech Processing for Language Learning: A Practical Approach to ComputerAssisted Pronunciation Teaching," Electronics (Basel), vol. 10, no. 3, p. 235, Jan. 2021, doi: 10.3390/electronics10030235. [13] Y. Anwar and P. Kalisa, "Students' Problems in Pronouncing Non-Existing English Diphthongs in Indonesian Language," Language Circle: Journal of Language and Literature, vol. 15, no. 1, pp. 17–

25, Oct. 2020, doi: 10.15294/lc.v15i1.26261. [14] Y. Lolita, E. Boeriswati, and N. Lustyantie, “The impact of Computer-Assisted Language Learning (CALL) use of English vocabulary enhancement,” Linguistic, English Education and Art (LEEA) Journal, vol. 4, no. 1, pp. 1–10, 2020, doi: 10.31539/leea.v4i1.1896. [15] E. Mohammadkarimi, “Exploring the use of artificial intelligence in promoting English language pronunciation skills,” LLT Journal: A Journal on Language and Language Teaching, vol. 27, no. 1, pp. 45–60, 2024, doi: 10.24071/llt.v27i1.8151. [16] Y. Getman et al., “Developing an AI-assisted low-resource spoken language learning app for children,” IEEE Access, vol. 11, pp. 86025–86037, 2023, doi: 10.1109/ACCESS.2023.3304274. [17] N. Bogach et al., “Speech processing for language learning: A practical approach to computer-assisted pronunciation teaching,” Electronics (Basel), vol. 10, no. 3, p. 235, 2021, doi: 10.3390/electronics10030235. [18] D. Abimanto and W. Sumarsono, “Improving English Pronunciation with AI Speech-Recognition Technology,” Acitya: Journal of Teaching and Education, vol. 6, no. 1, pp. 146–156, Jan. 2024, doi: 10.30650/ajte.v6i1.3810. [19] A. Zarei and M. Hashemipour, “Computer-assisted language instruction,” Procedia Soc Behav Sci, vol. 192, pp. 591–598, 2015, doi: 10.1016/j.sbspro.2015.06.094. [20] H. P. M. Haliza and Y. Lolita, “The use of MALL (Mobile Assisted Language Learning) by using ELSA Speak for students’ speaking skills,” JADEs Journal of Academia in English Education, vol. 5, no. 2, pp. 224–239, 2024, doi: 10.58926/jades.v5i2.334. [21] R. Hekmat and K. Balakrishnan, “The efficacy of AI-powered speech recognition in enhancing English pronunciation,” International Journal on Science and Technology, vol. 16, no. 2, 2025, doi: 10.71097/ijsat.v16.i2.5170. [22] G. Chhetri et al., “Model Context Protocols in Adaptive Transport Systems: A Survey,” Article, 2025. doi: XXXXXXX.XXXXXXX. [23] S. Wang and X. Shi, “Research on correction method of spoken pronunciation accuracy of AI virtual English reading,” Advances in Multimedia, 2021, doi: 10.1155/2021/6783205. [24] H. Vančová, “AI and AI-powered tools for pronunciation training,” Journal of Language and Cultural Education, vol. 11, no. 2, pp. 12–24, 2023, doi: 10.2478/jolace-2023-0022. [25] D. Abimanto and W. Sumarsono, “Improving English pronunciation with AI speech-recognition

technology,” Acitya: Journal of Teaching and Education, vol. 6, no. 1, 2024, doi: 10.30650/ajte.v6i1.3810. [26] T. Nguyen and L. Vo, “AI-driven pronunciation assessment: The impact of SpeechAce on EFL learners’ pronunciation competency,” Computer-Assisted Language Learning Electronic Journal, 2025, doi: 10.54855/callej.252635.

Sources

1	https://dictionary.cambridge.org/dictionary/english/absence INTERNET <1%	

EXCLUDE CUSTOM MATCHES	OFF
EXCLUDE QUOTES	OFF
EXCLUDE BIBLIOGRAPHY	ON