



DOI: <https://doi.org/10.46296/yc.v10i19.0947>

TUTORÍA CON APOYO DE IA: PRÁCTICA DE CONVERSACIÓN CON STIMULER EN LA FORMACIÓN DE FUTUROS PROFESORES DE INGLÉS.

AI-SUPPORTED MENTORING: SPEAKING PRACTICE WITH STIMULER IN FUTURE ENGLISH TEACHER TRAINING

Chávez-Zambrano Verónica Vanessa ¹; Vera-Toala Karla Jamileth ²

¹ Docente de la Carrera de Pedagogía los Idiomas Nacionales y Extranjeros de la Universidad Laica Eloy Alfaro De Manabí. Manta, Ecuador. Correo: veronica.chavez@uleam.edu.ec.
ORCID ID: <https://orcid.org/0000-0003-3958-5053>

² Estudiante de la Carrera de Pedagogía los Idiomas Nacionales y Extranjeros de la Universidad Laica Eloy Alfaro De Manabí. Manta, Ecuador.
Correo: e135016878@live.uleam.edu.ec. ORCID ID: <https://orcid.org/0009-0005-3825-1344>

Resumen

La habilidad oral es un requisito esencial en la formación inicial de profesores de inglés como lengua extranjera, ya que los futuros docentes deben ser capaces de comunicarse eficazmente, controlar la interacción en el aula y proporcionar modelos de lenguaje oral adecuado. Este estudio empleó un diseño de métodos mixtos, descriptivo, transversal, de campo y no experimental para evaluar la competencia oral de los estudiantes del programa educativo "Pedagogía de Lenguas Nacionales y Extranjeras: Inglés" de la Universidad Laica Eloy Alfaro de Manabí e investigar las percepciones e ideas de los docentes sobre los factores que determinan el éxito en este campo. Se matricularon 197 estudiantes en el programa, mientras que la muestra analítica consistió en 26 evaluaciones diagnósticas completas de expresión oral, seleccionadas mediante muestreo por conveniencia. Once estudiantes fueron evaluados según los descriptores A1-A2 y quince, según los descriptores B1. Cuatro profesores de inglés participaron en entrevistas semiestructuradas, que se analizaron temáticamente. Se utilizaron las escalas de expresión oral de Cambridge para evaluar la gramática y el vocabulario, la pronunciación, la comunicación interactiva y el logro global, así como la gestión del discurso en el nivel B1 de los estudiantes. En cuanto a la pronunciación, fue el área más fuerte en ambos tipos de estudiantes (A1-A2: M = 3,27, DE = 1,01; B1: M = 2,53, DE = 0,64); los valores medios más bajos se obtuvieron en la evaluación de gramática y vocabulario (M = 2,09, DE = 1,45; M = 1,60, DE = 0,63, respectivamente). Las percepciones de los docentes coinciden con el perfil diagnóstico en cuanto a la limitada recuperación de vocabulario, la ansiedad, la interacción espontánea insuficiente, la baja exposición al inglés y las dificultades de comprensión auditiva. Los hallazgos respaldan las recomendaciones de una interacción guiada más frecuente, práctica integrada de escucha y expresión oral, colaboración entre pares y retroalimentación sistemática dentro del Programa de Lengua Inglesa.

Palabras claves: competencia oral, profesores de inglés en formación, evaluación diagnóstica, métodos mixtos, interacción oral, inteligencia artificial.

Abstract

The speaking skill is an essential requirement in pre-service EFL teacher education as future instructors must be able to communicate effectively, control classroom interaction and provide models of proper spoken language. This study involved a mixed-methods, descriptive, cross-sectional, field-based, and non-experimental design to assess the speaking proficiency of students of the "Pedagogy of National and Foreign Languages: English" educational program at Universidad Laica Eloy Alfaro de Manabí and to investigate teachers' perceptions and ideas regarding the factors determining success in this field. 197 students were enrolled in the program, while the analytical sample consisted of 26 complete diagnostic speaking assessments, which were selected on the basis of convenience sampling. 11 students were assessed according to

Información del manuscrito:

Fecha de recepción: 08 de mayo de 2026.

Fecha de aceptación: 10 de julio de 2026.

Fecha de publicación: 11 de agosto de 2026.





A1-A2 descriptors and 15 students, according to B1 descriptors. Four English language teachers participated in semi-structured interviews, which were analyzed thematically. Cambridge speaking scales were used to evaluate grammar and vocabulary, pronunciation, interactive communication, and global achievement, as well as discourse management at the level of B1 students. As to pronunciation, it was the strongest area in both types of students (A1-A2: $M = 3.27$, $SD = 1.01$; B1: $M = 2.53$, $SD = 0.64$); the lowest mean values were acquired in the assessment of grammar and vocabulary ($M = 2.09$, $SD = 1.45$; $M = 1.60$, $SD = 0.63$, respectively). Teacher perceptions coincide with the diagnostic profile regarding limited vocabulary retrieval, anxiety, insufficient spontaneous interaction, low exposure to English, and listening-comprehension difficulties. The findings support recommendations for more frequent guided interaction, integrated listening-speaking practice, peer collaboration, and systematic feedback within the English Language Program.

Keywords: speaking proficiency, pre-service English teachers, diagnostic assessment, mixed methods, oral interaction, artificial intelligence.

1. Introduction

The effectiveness of speaking skills determines successful academic participation, professional interaction, and English modeling by teachers. Oral competence goes beyond general interaction for pre-service educators and does cover explanation, question, feedback, classroom management, and mediation of interactions. According to a meta-studies performed by Faez et al. (2021), language proficiency relates to teaching performance but does not fully explain the effectiveness of teaching.

Numerous international studies stress the challenges of speaking development since students have to coordinate grammar, vocabulary, pronunciation, discourse, and the process of interaction. The works like “The Impact of Automatic Speech

Recognition Technology on Second Language Pronunciation and Speaking Skills of EFL Learners” (Sun, 2023), “Artificial Intelligence in EFL Speaking: Impact on Enjoyment, Anxiety, and Willingness to Communicate” (Zhang et al., 2024), and “Investigating the Impact of Online Language Exchanges on Second Language Speaking and Willingness to Communicate” (Zhou, 2023) point out the interplay of linguistic competence, affective states, feedback, and authentic learning practice.

This phenomenon is relevant to Latin America and Ecuador as well. In the paper “An Analysis of the Levels of English of EFL Teachers in Ecuador”, Parra-Gavilánez (2022), stated that teacher proficiency varied across planning areas of the country. Narváez-Cantos (2022), in the work “The Effective English Language



Teacher from the Perspective of Ecuadorian University Students” concluded that university students considered pedagogical knowledge, communicative competence, socio-affective competence, and language proficiency as interrelated characteristics. Moving on in time, Guerrero Rodriguez and Moreira Baquerizo (2025) pointed out that limited instruction time, high emphasis on grammar, and lack of practice are impediments for Ecuadorian EFL learning process.

These findings are especially significant for teacher education. It is expected that future teachers will use the English language as the main means of instruction and will help students learn how to communicate. However, the application of general proficiency definitions provides no information regarding the type of difficulties concerning grammar, pronunciation, and overall speaking ability of individuals.

The current research covers this gap by combining learners’ speaking assessment with opinions of their teacher trainers from one Ecuadorian university. The study does not involve testing teaching methods or their efficiency but regards AI-

supported mentoring as a significant pedagogical matter to be researched once the achievements of students are analyzed.

The purpose of this work is to determine the current English-speaking proficiency of students enrolled in the English teacher education program at Eloy Alfaro University, to inform evidence-based pedagogical recommendations. The research is aimed at answering four questions: (1) What is the current level of English-speaking proficiency of pre-service English teachers enrolled in the English Education Program at Eloy Alfaro University of Manabí? (2) Which components of speaking proficiency (pronunciation, grammar, vocabulary, discourse management and interactive communication) represent the main strengths and weaknesses of pre-service English teachers? (3) How do English teachers assess students’ speaking ability and factors influencing it? (4) What pedagogical strategies could be recommended to improve oral performance in English teacher education based on the findings?

Literature Review

• Previous Research and the Study Gap

Research into the proficiency of teachers in terms of language has established that language knowledge is related to teaching efficiency, but it can't be treated as isolated factor leading to the quality of teaching. In the article, "Connecting Language Proficiency to Teaching Ability: A Meta-Analysis," Faez and his colleagues (2021) claimed that the relationship is positive but varies depending on the context and measurement used. The discovery justifies the necessity of including analytic diagnostics which define certain aspects of speaking in comparison to evaluation of the level of a program or course in terms of the overall level of spoken proficiency.

The studies that have been carried out in Ecuador have provided data that is important for the context, but they still can be elaborated further. Thus, Parra-Gavilánez (2022) used the data coming from standardized tests to define teacher language proficiency by national planning zone, and identified that the zones did not reach the expected what was expected proficiency level. He further

conducted a survey looking into the students and their understanding of the effective teachers' language proficiency in university EFL context and concluded that language proficiency is one of the factors that define the successful teaching along with the pedagogical abilities.

Guerrero Rodriguez and Moreira Baquerizo (2025) conducted a survey of Ecuadorian teachers about the strategies and difficulties faced by teachers in the process of teaching speaking to their students. As a result of their survey they found out a number of limitations such as lack of time, insufficient emphasis on oral practice, lack of motivation, and the effectiveness of communicative activities. On the global level, Zhou (2023) stated that structured online language exchange has a positive influence on speaking skills and willingness to speak, while Sun (2023) and Zhang et al. (2024) provided the connection between the speaking results and technology, anxiety, pleasure and readiness to speak.

Considering the above-mentioned literature, it is possible to identify three gaps present in the research. Firstly, there is not enough



Ecuadorian research concentrating on the direct language performance and interpretation of this performance by teacher educators. Secondly, often wide confusions in speaking proficiency categories conceals the internal differentiation of phoric, grammatical, and vocabulary skills, as well as discourse management, interactive communication and global result. Thirdly, there is lots of research devoted to using technology in raising speaking proficiency where the diagnosis is needed.

- **Speaking Proficiency and Communicative Competence**

Speaking is an active and engaging skill used to find meanings, interact with others, express ideas, and change the language depending on the context of the communication. According to Goh and Burns (2012), speaking development involves the working together of language competence, discourse, fluency, pronunciation, and communication. Thus, oral language skills are more than being able to produce syntactically correct sentences; they also comprise intelligibility, effectiveness, responsiveness, and the given individual's ability to cope

with the spontaneous nature of communication.

The aspects included in the Cambridge speaking scales suggest that such view is multidimensional. Grammar and lexis deal with the range and mastery of language; pronunciation shows how understandable the person is; discourse refers to delivering information; interactive speaking includes initiating, replying, developing, and negotiating; and global achievement assesses how the candidate performs in the task (Cambridge Assessment English, 2020). Each aspect provides their own piece of information meaning that strong pronunciation marks do not mean that all other components will be also good.

Pronunciation is still significant due to the way it influences comprehension, listening effort, and success in conversation. Metruk (2024) said in his review that mobile pronunciation teaching is a newly emerging area of teaching that has new opportunities for practice and feedback. Though intelligibly means different things and it should be combined with knowledge of

vocabulary, grammar, structure, and responsiveness to be fully revealed.

- **Factors Influencing EFL Speaking**

The performance of speaking in a language relies on various elements such as linguistic, cognitive, affective, and contextual aspects. Among the elements in the linguistic dimension, we have vocabulary limitations like the inability to form correct grammatical structures or sentences that would contain inaccurate information and those pertaining to phonetics. The cognitive aspects come into play in situations when speakers have to perform operations such as the processing of information and language retrieval; they have to organize meaning while monitoring its correctness with very little preparation time.

The influence of affective elements appears to be essential when it comes to spontaneous speech events and assessments of speaking proficiency. The findings of Sun (2023) showed that the role of speech recognition has positive implications for the development of speech skills; on the contrary, Zhang et al. (2024) suggested that the

combination of enjoyment, willingness to communicate, and anxiety is significant in AI environments. According to El Shazly's (2021) evidence, anxiety could be lowered through AI-supported experiences.

The opportunities for interaction contribute to the development of speaking. In the study conducted by Zhou (2023), it was determined that learners involved in language exchange had meaningful conversations, which allowed them to practice speaking. The researchers, Ericsson and Johansson (2023), claim that technology enables an endless supply of online practice sessions. Therefore, even though interaction appears to play a vital role, technology does not provide the ultimate solution to the problems of speaking production.

- **Speaking Assessment in Teacher Education**

Diagnostic evaluation of learning indicates the current profile of learners before specific instructional programs are designed. In speaking, analytical scales are important because they distinguish between dimensions that may undergo



uneven development. The descriptors given by Cambridge set out the observable criteria for the different levels and present the outcome of evaluation in the format of signs rather than in the form of a total score (Cambridge Assessment English, 2020a, 2020b). This information is important for the field of teacher education, as it can help make decisions related to oral practices, feedback, peer-mentoring activities, and language-support activities sequence.

Assessment also needs to take into account the conditions under which performance has taken place. For example, one task can be impacted by familiarity with the topic, time pressure, support from interlocutors, listening skills, and anxiety. For this reason, quantitative scores can be interpreted through qualitative evidence to facilitate the contextual understanding of the above scores. Mixed-method integration allows direct scores to be compared with educators' observations to find the areas of similarity and differences instead of regarding both sources as sufficient independently (Creswell & Creswell, 2018).

- **AI-Supported Mentoring as a Pedagogical Implication**

There is an emergence of Artificial intelligence in the field of language instruction. Various applications of AI in education can be found in conversational agents, speech recognition, adaptive practice, and other tools which ease the learning process. According to Alshumaimeri and Alshememry (2024), a number of ways of implementing AI in foreign language teaching has been gaining popularity. Du and Daniel (2024) also supported the idea that chatbots could help students to have more speaking opportunities if tasks were properly designed.

Scholars connect the use of AI tools in speaking practice with the improvement of speaking skills, pronunciation feedback, self-regulation, and anxiety reduction. Qiao and Zhao (2023) presented a study where the authors connected AI-based language training with the development of speaking abilities and self-regulation skills. Ding and Yusof (2025) conducted a mixed-methods study that showed how chatbot technology facilitates the development of speaking skills and diminishes anxiety levels. Finally,

Mingyan et al. (2025) found that the use of an AI-powered mobile application improves students' speaking abilities, although this improvement was not equal in all aspects of speaking. The research results should be further investigated but cannot be utilized in the settings without this type of intervention.

AI-assisted mentoring requires human evaluation as well. Ayanwale et al. (2024) consider the AI literacy of future educators to be crucial, while Peña-Acuña and Corga Fernandes Durão (2024) describe both possibilities brought about by AI technologies and the necessity to investigate AI-assisted English language learning more closely. Köbis and Mehner (2021) draw attention to ethical issues in the use of AI for mentoring. Alghamdy (2023) also stresses the importance of considering pedagogical and ethical aspects of using AI in EFL contexts.

One of the most significant is teacher scaffolding, as automated feedback may be imperfect and hard to read. Ma and Chen (2025) provide evidence explaining that combining AI with teacher support gives better results than letting learners work with AI independently. Liu et al. (2024)

emphasize that informal teaching and learning through AI depends on different factors, such as motivation and context. Thus, the author of the present work would propose using AI-supported mentoring as an addition to human feedback and not as a method that can replace the work done by teachers.

2. Methodology

Research Design

The design for this study was a mixed-methods, descriptive, cross-sectional, non-experimental, and field-based design. The quantitative part of the study focused on how students performed on speaking criteria, and the qualitative part of the study looked into teacher-researchers' perceptions of various linguistic, emotional, contextual, and pedagogical factors that affect oral performance. The study was non-experimental in nature because of not manipulating any variables, the nonexistence of a control group, and the nonuse of post-test and AI intervention (Creswell & Creswell, 2018).



The mixed-methods design of the research is called convergent design in the interpretation phase; first, qualitative and quantitative data were analyzed separately and then connected through the joint display showing the principal diagnostic patterns as well as emergent themes based on interviews.

Population, Participants, and Analytical Sample

The population that was part of the study was 197 students enrolled in the Pedagogy of National and Foreign Languages: English program at Universidad Laica Eloy Alfaro de Manabí, Ecuador. It was made of the quantitative analytical sample which included 26 students with completeness of their diagnostic records. The sample was selected by convenience and from the aspect of the availability and engagement of students in the speaking assessment; they were selected through non-probability sampling. It needs to be mentioned that 11 students were evaluated using A1-A2 scale and 15 students by B1 level.

Four English language teachers from the same program were purposively selected as key informants due to

their experience teaching English and their knowledge of students' speaking performance. Their interviews constituted the qualitative data analyzed in this study.

Data Collection Instruments

The instrument of measurement type was composed of speaking assessment scales aligned with Cambridge and appropriate levels. The rubric for the A1-A2 level indicates evaluation of grammar and vocabulary, pronunciation, interactive communication, and achievements across Parts 1 and 2 of Speaking. The B1 rubric indicates grammar and vocabulary, discourse management, pronunciation, interactive communication, and achievements throughout Parts 1 to 4 of Speaking. The ratings were given for the above-mentioned aspects from Band 0 to Band 5 depending on the series of descriptors provided by Cambridge (Cambridge Assessment English, 2020a, 2020b).

The research qualitative tool of choice was a semi-structured interview guide with open questions. Ten major questions of the tool were about the overall level of speaking

ability, spontaneous usage of English, fulfilment of the set goals, conditions for doing good performance, different factors such as linguistic and affective ones, teaching practices and methods of teaching, and recommendations for the program as a whole. The open-ended nature of the questions allowed the interviewers to put forward their comments on some usual things without being bound to particular types of answers.

Procedure and Data Analysis

In this study, the participants performed speaking tasks at their respective levels and their performances were recorded for further processing. The two evaluators participated in this review and one of them evaluated the analytic parameters of the performers while the other i.e., the global score. No statistical coefficients such as inter-rater reliability were used in the analysis of the data that is treated as the limitation (Braun & Clarke, 2022).

In this research, quantitative analysis was conducted for every group separately. The means and standard deviations provided general

quantitative information while frequencies and percentages showed how many performers scored below Band 3. The use of Band 3 was for descriptive purposes only as it was characterized as the first middle descriptor. The researcher did not create an overall score combining A1-A2 and B1 means due to the instruments used for measurement.

The interview transcriptions were analyzed using thematic analysis whereby meaning units were extracted and coded into themes of proficiency variability, spontaneous communication, linguistic barriers, emotional and performance factors, teaching methods used, listening-speaking partnership, and computer-aided instruction. The themes were then generalized up to broader parameters and compared to quantitative findings from the quantitative data (Braun & Clarke, 2022).

Ethical and Reporting Considerations

Student and teacher names have not been disclosed in either the analytical dataset or the article. The results were provided in an



aggregated manner and interviews were summarized thematically. The report clearly distinguishes among the institutional population, the 26 complete diagnostic speaking assessments included in the quantitative analysis, and the four English teachers who participated in the qualitative interviews. The report also explicitly states that Stimuler and other AI applications were not implemented or evaluated, allowing for the elimination of causal claims that are unsupported by the research design.

3. Results and discussion

Analytical Sample and Mean Speaking Bands

The quantitative analysis covered 26 total recordings: 11 learners (42.3%) in the A1-A2 assessment category and 15 students (57.7%) in the B1 class. The means and standard deviations for each criterion are shown in Table 1. The results are analyzed separately by group because the scales for the various levels utilized different descriptors, and the A1-A2 rubric did not include discourse management among the analytical categories so that was never a part of the analysis.

Table 1. Mean speaking bands by assessed level

Group	n	Grammar and vocabulary	Discourse management	Pronunciation	Interactive communication	Global achievement
A1-A2	11	2.09 (1.45)	Not assessed	3.27 (1.01)	2.18 (1.89)	1.18 (0.60)
B1	15	1.60 (0.63)	2.20 (0.56)	2.53 (0.64)	2.13 (0.92)	2.47 (0.64)

A1-A2 Diagnostic Profile

Pronunciation in the A1-A2 group has been comparatively the strongest dimension ($M = 3.27$, $SD = 1.01$), as only two distraught individuals (18.2%) scored under Band 3, meaning the overwhelming majority of participants were

intelligible according to the undertaken descriptors. This capacity should not be considered as sufficient evidence of oral proficiency as performance at other levels was considerably weaker.

With a mean score of 2.09 ($SD = 1.45$), grammar and vocabulary

constitute the second dimension of effectiveness; seven subjects (63.6%) received lower than Band 3. Functions in interactive communication resulted in a similar mean of 2.18, but the standard deviation was the biggest one ($SD = 1.89$); six students (54.5%) demonstrated a level lesser than Band 3 and three of them belonged to Band 0. The results suggest that some of the learners performed locally in difficult communication situations, whereas the majority were not able to respond.

Finally, the overall score was the weakest among all groups ($M = 1.18$, $SD = 0.60$), with 90.9% of respondents being below Band 3. Different outcomes in pronunciation performance and overall achievement indicate being understandable on the level of words or phrases has not always led to the completion of tasks.

B1 Diagnostic Profile

On the B1 rating scale, pronunciation achieved the highest measure ($M = 2.53$; $SD = 0.64$) and global achievement was in close second place ($M = 2.47$; $SD = 0.64$). However, a little more than half of the

students (eight students or 53.3%) fell below Band 3 in pronunciation, while nine (60.0%) of them failed to meet Band 3 in global achievement, thus indicating that though the strongest domains in the B1 band provided some indicators of performance, there is still substantial underperformance relative to what is considered acceptable performance.

Both grammar and vocabulary had the lowest performance ($M = 1.60$; $SD = 0.63$) – out of 15 students, only one-three scored above Band 3, with no student achieving Band 4 or 5. The same concern is relevant to discourse management ($M = 2.20$; $SD = 0.56$) as its mean result indicates that 11 students were below Band 3, which means that there was difficulty with producing long answers and keeping the speech organized.

In addition, concerning interactive communication, its mean is 2.13 ($SD = 0.92$); exactly 10 students scored below Band 3. Therefore, all the B1 results indicate a profile where pronunciation was relatively strong versus other aspects like linguistic knowledge, organizational skills, or interactive communication but was

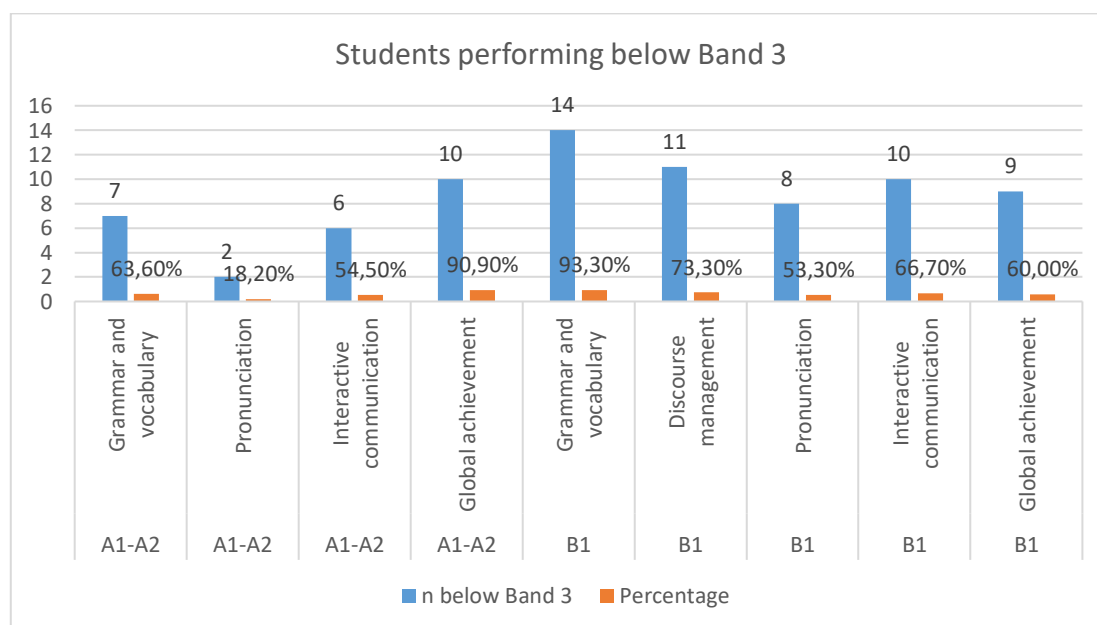


not able to offset the failures in using an appropriate language.

Table 2. Students performing below Band 3

Group	Dimension	n below Band 3	Percentage
A1-A2	Grammar and vocabulary	7	63.6%
A1-A2	Pronunciation	2	18.2%
A1-A2	Interactive communication	6	54.5%
A1-A2	Global achievement	10	90.9%
B1	Grammar and vocabulary	14	93.3%
B1	Discourse management	11	73.3%
B1	Pronunciation	8	53.3%
B1	Interactive communication	10	66.7%
B1	Global achievement	9	60.0%

Figure 1. Percentage of students below Band 3 by assessed dimension



Teacher Interview Findings

The student teachers' statements reflected a variety of proficiency levels and speakers. The majority of

learners fell within the basic and intermediate level, while a smaller group managed to communicate more autonomously. The teachers made a clear distinction between

performance on structured tasks and spontaneous use of the foreign language in classroom conditions: given the appropriate context, students participated more actively when they knew the topic, the tasks were straightforward and group work was used as a way to reduce pressure.

The main linguistic elements of speech were found to be the issues of vocabulary knowledge and grammar, as well as the fact that speaking and listening are interconnected. Teachers reported that students may sometimes know the right words but are unable to respond promptly when it comes to listening to the message of an interlocutor. This gives an additional meaning to the quantitative data acquired.

The reasons which influenced students' performance were divided into affective reasons (such as anxiety and low self-esteem) and performance reasons (such as poor listening and speaking skills and lack of preparation). The teachers laid emphasis on positive feedback and made recommendations to use known contexts for practicing speaking skills.

The classroom activities that had been proposed included communicative tasks, interviews, presentations, discussions, different types of kinesthetic activities, pair work and delayed feedback, etc. The teachers considered technology to be a good method of practice with the help of the recording and recognition.

Table 3. *Integration of diagnostic and interview findings*

Theme	Synthesized teacher evidence	Relationship to diagnostic results
Heterogeneous proficiency	Students were divided into two diametrically opposed groups, with the first group consisting of highly supported speakers and the second group being less dependent on teacher assistance	There was a high variance in the usage of grammar and vocabulary at A1-A2 level (SD=1.45) and interactive speaking (SD=1.89)
Limited spontaneous interaction	The rate of the use of English increased with the delivery of interesting topics, clear communicative tasks, and pair/group work, but did not become autonomous	Interactive speaking was below band 3 in 54.5% of A1-A2 level students and 66.7% of B1 level students



Theme	Synthesized teacher evidence	Relationship to diagnostic results
Linguistic constraints	The participant was faced with a number of linguistic problems, such as vocabulary retrieval, grammar, sentence construction, coherence, pronunciation, and the use of meaningful phrases	Grammar and vocabulary scored the lowest average in B1 and average in A1-A2; discourse competency was poor in B1
Affective and performance factors	Anxiety, low self-esteem, fear of making mistakes, pressure from being evaluated, and feeling ashamed hampered students' participation	Low performance globally and in supported communication might suggest a case of interference under communicative pressures
Listening-speaking relationship	Another barrier to giving appropriate and timely answers was the difficulty of understanding questions spoken by the teachers	This finding may be the reason for low interactive communication and higher pronunciation scores
Pedagogical and technological response	The teachers suggested using authentic communication, peer work, delayed correction, integrated skills, recording, shadowing, and AI-supported practice	The profile indicates the necessity of transferring some elements of controlled knowledge into properly established and independent speech production

Discussion

The analysis provides evidence of an unbalanced speaking profile rather than a comprehensive level of speaking. Pronunciation was recognized as the strongest category of measurement in the two analyzed groups, however, language achievement, grammar and vocabulary capabilities, interactive discourse and management of interactions remained low. This finding indicates that the multi-component view of the Cambridge scale is justified as it distinguishes between intelligibility and language use (Cambridge Assessment

English, 2020a; 2020b). This fact is in agreement with the ideas of Goh and Burns (2012), stating that speaking involves coordination of various processes rather than isolated mastery of sounds or structures.

Grammar and vocabulary were found to be the most consistent weakness of students, particularly in the B1 group, with as many as 93.3% of learners' performance being below Band 3. Teacher trainers also referred to vocabulary issues, problems of construction and errors in grammar as the most frequent difficulties encountered by students.

Guerrero Rodriguez and Moreira Baquerizo (2025) indicate that speaking teaching process in Ecuador is mainly restricted by time limits and absence of sufficient oral practice in instruction. The current study brings direct proof of how these limitations affect the performance in analytical speaking performance.

The relatively better results in pronunciation should be interpreted with caution. Although A1-A2 students demonstrated a considerable level of intelligibility, their overall performance was still below the required level; in B1 group, pronunciation was assumed to be the strongest criterion but at the same time over a half of the group failed to reach Band 3 level. The authors of Metruk (2024) and Abdelhalim and Alsehibany (2025) agree that pronunciation is improved through a combination of practice and feedback provided through technology. However, the profile in this paper shows that the understood pronunciation alone does not guarantee good vocabulary, coherent speech, and the ability to communicate independently.

The level of interactive communication in both groups was

below Band 3 in about half of the cases. The interviews revealed that students were able to speak in English only when discussing familiar topics or following clear instructions, working in pairs, or receiving other types of help from teachers. Zhou (2023) said that communicating online positively influenced speaking and willingness to speak because of real interaction. However, the authors of the study did not test the proposed method and, likewise, did not seek to adopt this method in their research, even though they agree that the activities in the study should include tasks that make students take turns, ask questions, clarify, answer without preparing and communicate more than giving simple answers.

The discoveries relating the state of knowledge to reasons for the failure of this knowledge to become spontaneous communication are explained with such factors as anxiety and fear of making mistakes, insecurity, and evaluation pressure shown by the teachers' comments. This explanation is supported by studies conducted by El Shazly (2021), Sun (2023), Zhang et al. (2024), who note that oral



performance is affected by anxiety and willingness to communicate. Thus, it may be concluded that the anxiety-friendly environment is created through the use of low-stress speaking activities, peer learning, clear rules for evaluation, delay in feedback.

In addition, the importance of listening comprehension mentioned by the teacher educators may provide an interesting insight regarding the interactive issue. Learners may know many words and speak with perfect articulation but be unable to answer the question they heard and understood late. The solution of this problem may be found in integrated listening-speaking task, where students have to find important information and give answers to questions that require negotiating meaning. This approach to teaching corresponds to the whole approach suggested by Goh and Burn (2012), allowing teachers to move from rote learning to real communication and production of language.

Limitations and Future Research

The study has several limitations. Out of the total population of 197

students, only 26 complete diagnostic speaking assessments were available for analysis, while the nature of convenience sampling undermines the possibility of generalizing the results of the study. Furthermore, the classification of the A1-A2 and B1 groups led to an evaluation based on level-specific descriptors, which does not allow direct comparisons across levels. Besides, no inter-rater reliability coefficient is available for the data.

The cross-sectional design only reflects students' speaking proficiency at one point in time, thus not assessing the causes of the phenomena in question or the changes in performance over time. Future studies should include larger samples of students, conduct additional interviews, and analyze inter-rater agreement among evaluators to strengthen the methodological rigor of the study.

4. Conclusions

The diagnostic assessment indicated that pre-service English language teachers who were in the records showed heterogenous oral proficiency profiles. Pronunciation

was found to be the strongest component of speech output; however, it was not leading to a globally high speech performance. Grammar and vocabulary represented the major weaknesses, especially because of the B1 level; however, interactive speaking skills and discourse management skills also needed attention. The results show that using separate criteria for assessment of speaking is of much more importance than simple identification of speaking proficiency level.

Teacher educators' perceptions were similar with the results of the diagnostic assessment with limited spontaneity of English use and poor vocabulary retrieval, difficulties in grammar and discourse, lack of experience, issues with listening comprehension, anxiety, fear of making mistakes and pressure during assessment. The coincidence of direct assessment data and qualitative information means that oral development requires linguistic support as well as frequent communication practice, integrated listening and speaking practice, and constructive feedback in a classroom where risk is encouraged.

Mentoring with the use of AI technology can provide more options for individual practice, feedback, and self-assessment, however, it has not been investigated so far owing to the absence of the AI application in the study. There was no intervention applied to the study, therefore, the results presented here do not claim that the performance of the students was improved with the help of Stimuler or any other tool; thus, future projects should integrate mediation, ethical issues, transparent assessment, and evidence showing the results of the project.

Recommendations

The English language program ought to encompass speech tasks that will be performed often but in short periods. Such activities like conversation clubs, cross-semester collaboration, role plays, information gap activities, problem-solving tasks, debates, interviews, and English-only activity moments should enforce engagement, meaning making, and immediate responses instead of relying on memorized speeches.

Teachers' ought to combine speaking and listening through the



use of listening exercises and speaking. Analytic criteria for these types of activities ought to be presented beforehand, while feedback should be focused on a small number of issues in the field of grammar, vocabulary, discourse, pronunciation, and interaction.

The program should introduce supportive norms that would help diminish fear of errors and negative peer response towards mistakes. Low-pressure practice, pairing up students in different combinations, gradually increasing difficulty of tasks, self-evaluation, and peer feedback should help increase participation of students while creating clear requirements for participation in speaking activities.

Students experiencing permanent obstacles in speaking should be supported according to their individual problems instead of using the same speaking practice aimed at all students.

Any future use of Stimuler, conversational AI, or pronunciation applications should be built upon established learning goals, teacher education, privacy check, access process, and instructions for how to

interpret machine-generated suggestions. Evaluation about the use of this technology should be based on keeping a complete database of students who participated in the experiment, engagement of trained evaluators, documentation of reliability, checking using pre-test/post-test evaluation, and qualitative research.

References

- Abdelhalim, S. M., & Alsehibany, R. A. (2025). Integrating AI-powered tools in EFL pronunciation instruction: Effects on accuracy and L2 motivation. *Computer Assisted Language Learning*, 1–25. <https://doi.org/10.1080/09588221.2025.2534015>
- Alghamdy, R. Z. (2023). Pedagogical and ethical implications of artificial intelligence in EFL context: A review study. *English Language Teaching*, 16(10), 87–98. <https://doi.org/10.5539/elt.v16n10p87>
- Alshumaimeri, Y. A., & Alshememry, A. K. (2024). The extent of AI applications in EFL learning and teaching. *IEEE Transactions on Learning Technologies*, 17, 653–663. <https://doi.org/10.1109/TLT.2023.3322128>

- Ayanwale, M. A., Adelana, O. P., Molefi, R. R., Adeeko, O., & Ishola, A. M. (2024). Examining artificial intelligence literacy among pre-service teachers for future classrooms. *Computers and Education Open*, 6, 100179. <https://doi.org/10.1016/j.caeo.2024.100179>
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.
- Cambridge Assessment English. (2020a). *A2 Key for Schools handbook for teachers for exams from 2020*. Cambridge University Press & Assessment.
- Cambridge Assessment English. (2020b). *B1 Preliminary for Schools handbook for teachers for exams from 2020*. Cambridge University Press & Assessment.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.
- Ding, D., & Yusof, A. M. B. (2025). Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: A mixed methods study. *Humanities and Social Sciences Communications*, 12, 1223. <https://doi.org/10.1057/s41599-025-05550-z>
- Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, 6, 100230. <https://doi.org/10.1016/j.caeai.2024.100230>
- El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3), e12667. <https://doi.org/10.1111/exsy.12667>
- Ericsson, E., & Johansson, S. (2023). English speaking practice with conversational AI: Lower secondary students' educational experiences over time. *Computers and Education: Artificial Intelligence*, 5, 100164. <https://doi.org/10.1016/j.caeai.2023.100164>
- Faez, F., Karas, M., & Uchihara, T. (2021). Connecting language proficiency to teaching ability: A meta-analysis. *Language Teaching Research*, 25(5), 754–777. <https://doi.org/10.1177/1362168819868667>



- Goh, C. C. M., & Burns, A. (2012). Teaching speaking: A holistic approach. Cambridge University Press.
- Guerrero Rodriguez, S. E., & Moreira Baquerizo, A. S. (2025). Ecuadorian EFL teachers' experiences in fostering students' English-speaking skills: Insights into strategies and challenges in public and private schools. UNESUM-Ciencias. Revista Científica Multidisciplinaria, 9(2), 124–136.
<https://doi.org/10.47230/unsum-ciencias.v9.n2.2025.124-136>
- Köbis, L., & Mehner, C. (2021). Ethical questions raised by AI-supported mentoring in higher education. Frontiers in Artificial Intelligence, 4, 624050.
<https://doi.org/10.3389/frai.2021.624050>
- Liu, G. L., Darvin, R., & Ma, C. (2024). Exploring AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation of Chinese EFL learners' AI adoption and experiences. Computer Assisted Language Learning, 1–29.
<https://doi.org/10.1080/09588221.2024.2310288>
- Ma, Y., & Chen, M. (2025). The human touch in AI: Optimizing language learning through self-determination theory and teacher scaffolding. Frontiers in Psychology, 16, 1568239.
<https://doi.org/10.3389/fpsyg.2025.1568239>
- Metruk, R. (2024). Mobile-assisted language learning and pronunciation instruction: A systematic literature review. Education and Information Technologies, 29, 16255–16282.
<https://doi.org/10.1007/s10639-024-12453-0>
- Mingyan, M., Noordin, N., & Razali, A. B. (2025). Improving EFL speaking performance among undergraduate students with an AI-powered mobile app in after-class assignments: An empirical investigation. Humanities and Social Sciences Communications, 12, 370.
<https://doi.org/10.1057/s41599-025-04688-0>
- Narváez-Cantos, D. (2022). The effective English language teacher from the perspective of Ecuadorian university students. Maskana, 13(2), 5–14.
<https://doi.org/10.18537/mskn.13.02.01>
- Parra-Gaviláñez, L. (2022). An analysis of the levels of English of EFL teachers in Ecuador. Medwave, 22(S2), eUTA153.
<https://doi.org/10.5867/medwave.2022.S2.UTA153>

- Peña-Acuña, B., & Corga Fernandes Durão, R. (2024). Learning English as a second language with artificial intelligence for prospective teachers: A systematic review. *Frontiers in Education*, 9, 1490067. <https://doi.org/10.3389/feduc.2024.1490067>
- Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, 14, 1255594. <https://doi.org/10.3389/fpsyg.2023.1255594>
- Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in Psychology*, 14, 1210187. <https://doi.org/10.3389/fpsyg.2023.1210187>
- Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System*, 121, 103259. <https://doi.org/10.1016/j.system.2024.103259>
- Zhou, A. (2023). Investigating the impact of online language exchanges on second language speaking and willingness to communicate of Chinese EFL learners: A mixed methods study. *Frontiers in Psychology*, 14, 1177922. <https://doi.org/10.3389/fpsyg.2023.1177922>

Appendix A. Diagnostic Speaking Assessment Summary

Table A1. A1-A2 speaking descriptors

Band	Grammar and vocabulary	Pronunciation	Interactive communication
5	Good control of simple forms and a range of appropriate everyday vocabulary.	Mostly intelligible, with some control of phonological features.	Maintains simple exchanges with very little prompting or support.
3	Sufficient control of simple forms and appropriate everyday vocabulary.	Mostly intelligible despite limited phonological control.	Maintains simple exchanges with some difficulty and requires prompting.
1	Limited control of a few forms and isolated words or phrases.	Very limited phonological control and often unintelligible.	Considerable difficulty maintaining exchanges; requires additional support.



Band	Grammar and vocabulary	Pronunciation	Interactive communication
0	Performance below Band 1.	Performance below Band 1.	Performance below Band 1.

Table A2. B1 speaking descriptors

Band	Grammar and vocabulary	Discourse management	Pronunciation	Interactive communication
5	Good control of simple forms, attempts complex forms, and uses a range of vocabulary.	Produces extended, relevant speech with cohesive devices despite some hesitation.	Intelligible, with generally appropriate intonation, stress, and clear sounds.	Initiates, responds, develops interaction, and negotiates with very little support.
3	Good control of simple forms and an appropriate range for familiar topics.	Extends beyond short phrases, remains mostly relevant, and uses basic cohesive devices.	Mostly intelligible, with some control at utterance and word level.	Initiates and responds appropriately with very little prompting.
1	Sufficient control of simple forms and a limited appropriate vocabulary.	Produces short phrases with frequent hesitation, repetition, or digression.	Mostly intelligible despite limited control of phonological features.	Maintains simple exchanges with difficulty and requires support.
0	Performance below Band 1.	Performance below Band 1.	Performance below Band 1.	Performance below Band 1.

Appendix B. Interview Guide for English Teacher Educators

Purpose: To collect English teachers' perceptions regarding the current speaking proficiency of students in the English Language Program.

Instructions: Please answer honestly. Information will be used only for academic research and kept confidential.

1. How would you rate the overall speaking level of the students you currently teach in the English Language Program?
2. How often do your students use English spontaneously during class?
3. Do your students reach the expected speaking level according to the semester they are in?



4. In your opinion, what performance factors do you believe might affect the students' speaking skills?
5. In your opinion, what linguistic factors do you believe might affect the students' speaking skills?
6. From your view, what affective factors do you believe might affect the students' speaking skills?
7. In your opinion, what are the main methodological factors that affect students' speaking performance?
8. Which teaching strategies do you usually use to help students improve their speaking skills?
9. Do you think technology can support the development of students' speaking skills? If so, which tools?
10. Which recommendations would you give to improve speaking skills in the English Language Program?