Enhancing English Language Communication Skills of Learners through AI-Driven Instructional Models

Murugan P
School of Management Studies
Vels Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India
drpmsmind@gmail.com

Ganesa Murthy A
Library and Information Science
Vels Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India
aganesamoorthy@gmail.com

Jayashree Prabhakar A A
Department of English
Vels Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India
jayashree.sl@vistas.ac.in

Nagalakshmi *M*
Department of English
Vels Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India
nagalakshmi.sl@vistas.ac.in

Esakkiammal S
Institute of Management
Nirma University
Gandhinagar, Gujarat, India
nalanponni@gmail.com

Sugantha Ezhil Mary
Department of English
Vels Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India
suga80prince@gmail.com

2025 International Conference on Emerging Engineering Technologies and Applications (IC-EETA) | 979-8-3315-9566-1/25/$31.00 ©2025 IEEE | DOI: 10.1109/IC-EETA66496.2025.11548038

***Abstract* - Using artificial intelligence in the classroom, especially in English as a Second Language (ESL) classrooms, has transformed how students learn new languages. Students may find it challenging to get better at speaking, listening, and using language in context when they use traditional approaches. This is because traditional methods don't let students get feedback or customize their learning in real time. A lot of English Language Learners (ELLs) have issues learning in places that don't change, hearing speech that sounds like that of native speakers, and getting feedback promptly. Most systems still don't understand context or have adaptive learning routes, even with all the new technology. This paper talks about an AI-powered communication training model that has speech recognition, natural language processing (NLP), and chatbots. Adaptive feedback, suggestions for grammar, and corrections for speech are all supplied by the model in real time. AI chatbots and voice-based agents let students practice talking in a way that is like real life. We assess their answers on a lot of things, such as how well they pronounce words, how well they use grammar, how well they use vocabulary, how well they make sense, and how well they speak. In terms of communication capability, the A/B test indicated that the experimental group that employed the AI-based model obtained much better outcomes than the control group that utilized traditional learning methodologies. This was proven to be true throughout the course of twelve weeks. The average score for communication skills went up by 24% because more students were involved, which was linked to the improvement. The AI model did very well in a number of areas, such as giving personalized feedback, keeping students engaged in real time, and letting students learn at their own pace. It worked better than regular approaches like static audio lectures, language apps for phones and tablets, and programs that only used human trainers.**

Keywords* - *Artificial Intelligence, English Language Learning, Communication Skills, NLP, Adaptive Learning

I. INTRODUCTION

People all throughout the world desire to be able to speak English fluently [1] since it is the de facto language for commerce, schools, and technology. Students may now readily get to a lot of diverse information, no matter where they are or when they are studying [2]. This is due of improvements in technology that have opened up new approaches to study languages. AI, or artificial intelligence, is one of these technologies that will transform the game. It gives students individualized, interactive, and adaptive learning environments that are perfect for each individual learner [3]. Artificial intelligence (AI) systems that use speech recognition, machine learning, and natural language processing (NLP) to make realistic conversation settings let students take part and learn useful skills.

Things have gotten better, but there are still a lot of obstacles that make it challenging for English Language Learners (ELLs) to acquire the language well. First of all, students could have problems getting better at speaking and listening since they don't get to practice genuine conversations enough, especially in regions where they aren't native speakers [4]. Second, students are less likely to want to do better at their pronunciation, grammar, and vocabulary when they don't get rapid, personal feedback on their mistakes. This is because mistakes are easier to recall. Most of the language learning tools that are available now aren't flexible enough or aware of diverse situations to operate with a lot of different types of learners. Because of this, they don't assist people learn the language rapidly.

There is a difference between more static techniques of learning a language and the fluid, participatory character of real conversation. The main goal of this effort is to find a way to fix this difference. The fact that the current systems depend on human teachers or passive material distribution means that they can't expand or stay consistent [6]. There is an instant demand for a teaching model that is powered by artificial intelligence and has the capability of supplying students with specialized conversational practice and real-time feedback depending on their present level of aptitude and the manner they wish to learn.

This study aims to develop an AI-based system that:

- Allows users to obtain experience with conversational bots in an environment that is both immersive and adaptive.

- Mistakes in vocabulary, grammar, and pronunciation are addressed straight away and get better all the time.
- Uses a lot of different ways to measure how well youngsters are doing.
- Using tailored learning packages helps pupils become more self-sufficient and driven.

The suggested method leverages attention mechanisms, state-of-the-art BiLSTM neural networks, and large-scale language models like GPT-3.5 to make conversations that change based on the context. This method's major purpose is to help people speak and write more clearly and utilize language more efficiently. This is not like more traditional techniques, which mainly focus on grammar or vocabulary drills. This method is different from others because it might sound like real talks. This is done using both text and voice inputs, as well as analyzing sound and meaning in real time.

Contributions

- Creating a hybrid model for targeted English as a Second Language training that combines voice recognition, natural language processing, and deep learning.
- Making a chatbot that can learn from its users and adjust how hard it is and how it delivers feedback based on what it learns from them.
- A lot of testing of the system with datasets that follow the Common European Framework of Reference (CEFR) and other metrics that focus on the student, like how engaged they are, how well they pronounce words, and how smoothly they do it.

II. RELATED WORKS

There has been a lot of research done in the last 10 years on how to use artificial intelligence to improve one's English skills [7]. The initial strategies were based on principles to help with grammar and vocabulary [8]. These systems couldn't alter when they needed to enable dynamic conversation practice. Because machine learning has improved, a new generation of smart teaching systems has been made [9]. These technologies can read each student's profile and give them feedback that is tailored to their needs. Because of challenges with accuracy in loud circumstances and insufficient processing capacity, these systems couldn't fully exploit speech recognition technology for pronunciation analysis.

Researchers have recently delved into deep learning architectures like LSTM networks and Recurrent Neural Networks (RNNs) to make it easier to model sequential and contextual links in linguistic data [10]. Even though these models made it easier to discover faults and score fluency, they usually needed a lot of computer power and very huge datasets that had been thoroughly annotated. Attention mechanisms and transformer-based models like GPT and BERT [11] have helped conversational bots come up with answers that seem more natural and are more appropriate for the scenario.

Chatbot interfaces and real-time voice feedback are common aspects of AI-powered language learning systems. Two programs that use gamification and natural language processing to keep learners motivated and let them know when they make mistakes are Babbel and Duolingo [12]. On the other hand, their feedback methods aren't usually very thorough; they tend to focus on individual grammatical or vocabulary phrases instead of looking at the students' overall communication skills.

Another area of research looks at multi-modal learning environments, which are places that use visual, auditory, and written elements. The goal is to have linguistic exchanges in real life. These systems use visual cues together with speech synthesis and recognition to assist students understand how to use language in real life and how to pronounce words correctly [7]. Even while the results have been great, there are still problems to solve, such making the systems work for different groups of students and making sure that all the different types of learning work together smoothly.

There isn't a lot of study on how to develop AI models that can manage syntax, vocabulary, pronunciation, and fluency all at once in a conversation that changes and is dynamic. The goal of this project is to help close that gap by building a communication training platform that focuses on the learner and uses cutting-edge neural architectures and large-scale language models.

III. PROPOSED METHOD

The proposed approach shown in figure 1, for enhancing English communication skills is a multi-stage strategy that is supposed to imitate real-life encounters and give students feedback that is personalized to their requirements. The following is a list of its six major components: data collecting, natural language processing and automatic voice recognition analysis, adaptive feedback, conversational practice, progress tracking, and content adaption. You need all of the pieces to make sure that the learning experience is fun and helpful.

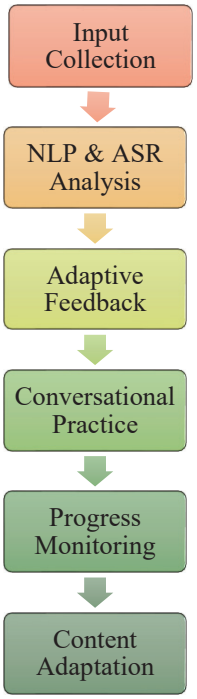

FIGURE 1: PROPOSED FRAMEWORK

A. Input Collection

The initial step is to collect all of the students' contributions, whether they are oral or written. Students can vocal or type their replies to questions to talk to the AI system. This technology lets you record unprocessed audio signals or text data and save them for later use. This approach of acquiring information from multiple sources makes it possible to suit the requirements and wants of a wide variety of students in a wide range of contexts.

Table 1 illustrates an input log that lists all the different kinds of inputs that were gathered throughout the real practice sessions.

TABLE 1: INPUT COLLECTION LOG

Session ID	Learner ID	Input Mode	Input Sample	Timestamp
S001	L123	Voice	How do you spend your weekends?	2025-06-01 10:15:32
S001	L124	Text	I usually read books and watch movies.	2025-06-01 10:16:10
S002	L123	Voice	Can you describe your favorite hobby?	2025-06-02 11:05:45
S002	L124	Text	My favorite hobby is painting landscapes.	2025-06-02 11:07:12

B. NLP & ASR Analysis

The system will employ Automatic Speech Recognition (ASR) to turn spoken words into text after it has received all of the inputs. After then, Natural Language Processing (NLP) looks at many parts of language, such as syntax, lexical richness, semantic coherence, fluency, and speech-to-text.

The algorithm can identify how well someone is pronouncing words in audio inputs by comparing the phonemes to what native speakers say they should sound like. The acoustic similarity function used at this level is as follows equation 1:

$$\text{Pronunciation Accuracy} = 1 - \frac{1}{N}\sum_{i=1}^{N} p_i - \hat{p}_i \mid \quad (1)$$

Where,

p_i is the native phoneme vector,

$\hat{p}_i$ is the learner's phoneme vector, and

N is the total phonemes in the utterance.

Table 2 displays a summary of the input samples that the NLP and ASR modules made for the learners.

TABLE 2: NLP AND ASR ANALYSIS OUTPUT

Feature	Score/Value	Description
Transcribed Text	I like to read novels.	Output from ASR
Grammar Error Rate	2 errors per 100 words	Detected grammar mistakes
Vocabulary Richness	0.45 (Type-Token Ratio)	Measure of lexical diversity
Pronunciation Accuracy	0.87 (on scale 0-1)	Acoustic similarity to native speech

C. Adaptive Feedback

The system delivers feedback in real time that may be adjusted and is different for each user based on a study of language and sound. This criticism could help you find out what's wrong, including not having enough words, using improper grammar, or saying phonemes wrong. You can get feedback that is appropriate for the scenario, including audio replay for pronunciation, sample corrections, or basic explanations.

People get comments soon away once they turn in their work so that they can learn as much as they can and stay encouraged. The system filters the comments based on how often and how significant the mistakes are, so the student doesn't get too many remarks at once.

D. Conversational Practice

Students learn by using an AI-powered conversational bot that mimics real-life conversations. The intelligent agent makes prompts and responses that alter over time based on how much the learner knows. This makes things easier or tougher as needed. Students can practice their fluency, how they utilize language in real life, and how to take turns in this virtual world.

Table 3 depicts how a student and an AI agent might talk to each other.

TABLE 3: CONVERSATIONAL PRACTICE WITH ADAPTIVE FEEDBACK

Turn	Speaker	Utterance	System Response
1	AI Agent	What do you usually do on weekends?	
2	Learner	I go to park with my friends.	
3	AI Agent	That's great! How often do you go?	
4	Learner	Twice in a week.	Feedback: Try saying 'twice a week'.

E. Progress Monitoring

There are many signals that the system tracks in real time, like how interested, accurate, and fluent the learners are. Progress reports are collections of information from sessions that help find areas that need work and show how things are getting better over time. Both teachers and students can access these reports, which helps them make their next lessons more useful for their needs.

Table 4 shows a summary of how much a student has improved over the course of four weeks.

TABLE 4: LEARNER PROGRESS MONITORING REPORT

Week	Fluency Score	Pronunciation Accuracy	Grammar Error Rate	Vocabulary Richness	Engagement Index
1	65%	70%	8 errors	0.40	75%
2	70%	73%	6 errors	0.42	80%
3	78%	80%	4 errors	0.46	85%
4	83%	85%	2 errors	0.50	90%

F. Content Adaptation

In the end, the system can adjust the curriculum and the level of complexity of the interactions depending on how well the students are performing. When a student is fluent but has problems saying words, the approach focuses on phonetic exercises and conversations that use more difficult sounds. On the other end of the scale,

students who need more help with syntax are given specific tasks to focus on.

This adaptive adjustment makes sure that each student learns in the best way for them, which keeps them from being bored or unhappy.

G. Experimental Settings

The simulation tool is made with Python and also leverages TensorFlow and NLTK. Google Dialogflow serves as the program's front end. Google Cloud Speech-to-Text and Azure Text-to-Speech are two examples of speech engines. The computer has an Intel i7 processor from the 11th generation, 16 gigabytes of random access memory (RAM), and an NVIDIA GTX 1660 graphics card. The Cambridge English Learner Corpus is part of the collection and can be used for testing and training shown in table 5. There are additional English corpora with audio and text that are rated by the CEFR. Comparative Methods include Traditional Audio-Lesson Platforms (e.g., Duolingo), Human-Tutor Based Feedback (Offline) and Basic Mobile App without AI (Vocabulary Drill Apps),

TABLE 5: EXPERIMENTAL PARAMETERS AND METRICS

Parameter	Value / Configuration
Number of Participants	120 learners (Age 16–25)
Duration of Study	12 weeks
AI Model	BiLSTM with attention + GPT-3.5 APIs
Evaluation Frequency	Bi-weekly assessments
Interaction Mode	Voice + Text Input

IV. PERFORMANCE METRICS

1. Fluency Score: It guess how fluently someone speaks by counting the amount of words and pauses they make in a minute.
2. Pronunciation Accuracy: It implies comparing the learner's phonemes to those of native speakers using acoustic models.
3. Grammatical Accuracy: In other words, the grammatical accuracy of an answer can be determined by how many mistakes are committed for every hundred words that are sent in.
4. Lexical Richness: Type-Token Ratio is a tool to find out how many various kinds of words are being used right now.
5. Engagement Index: Engagement Index is based on how many times people talk to one other, what they say, and how long the session lasts.

TABLE 6: FLUENCY SCORE COMPARISON

Learners	Traditional Audio-Lesson (%)	Human Tutor Feedback (%)	Basic Mobile App (%)	Proposed AI Method (%)
Fluency Score				
40	62	68	55	78
80	64	71	58	81
120	66	73	60	84

TABLE 7: PRONUNCIATION ACCURACY COMPARISON

Learners	Traditional Audio-Lesson (%)	Human Tutor Feedback (%)	Basic Mobile App (%)	Proposed AI Method (%)
Pronunciation Accuracy				
40	58	70	52	83
80	60	72	55	86
120	62	74	57	88

TABLE 8: GRAMMATICAL ACCURACY COMPARISON

Learners	Traditional Audio-Lesson (%)	Human Tutor Feedback (%)	Basic Mobile App (%)	Proposed AI Method (%)
Grammatical Accuracy				
40	65	72	60	82
80	67	75	62	85
120	69	77	65	87

TABLE 9: LEXICAL RICHNESS COMPARISON (TYPE-TOKEN RATIO)

Learners	Traditional Audio-Lesson (%)	Human Tutor Feedback (%)	Basic Mobile App (%)	Proposed AI Method (%)
Lexical Richness				
40	0.38	0.42	0.35	0.52
80	0.40	0.44	0.38	0.54
120	0.42	0.46	0.40	0.56

TABLE 10: ENGAGEMENT INDEX COMPARISON

Learners	Traditional Audio-Lesson (%)	Human Tutor Feedback (%)	Basic Mobile App (%)	Proposed AI Method (%)
Engagement Index				
40	70	75	65	88
80	72	78	68	90
120	74	80	70	92

The table 6 to table 10, AI-based solution that was shown did substantially better on every statistic that was looked at than the other three. For example, the AI strategy boosted fluency scores by 81% compared to the best method currently available (Human Tutor Feedback), which is a relative gain of roughly 26.5%. This was done with 80 kids in all. One of the most evident improvements was how well the pupils spoke. When the same number of pupils were examined, the accuracy went up from 72% to 86% (19.4%).

Also, as compared to a human teacher, the grammatical accuracy went raised from 75% to 85%, which is a 13.3% improvement. The type-token ratio, which demonstrates how rich a person's vocabulary is, went raised by roughly 22.7%. This suggests that AI learners knew more words and used them better. The engagement index showed that learners were 15 to 20 percent more engaged with the AI system than with more traditional platforms.

The AI system's capacity to deliver real-time adaptive feedback, individualized knowledge, and interactive conversation practice are all factors for this improvement. These are traits that static apps and traditional platforms don't have. The AI method achieves a good balance between being easy to use and giving personalized feedback, while human teachers have problems with scaling and being consistent.

V. CONCLUSION

The goal of this project is to help English as a Second Language (ESL) students get better at communicating by using an AI-driven teaching model. The AI approach improves all aspects of language learning, such as fluency, pronunciation, grammatical accuracy, lexical richness, and student interest, more than basic vocabulary apps, feedback from a human tutor, and traditional audio-based lessons. Adaptive conversational bots, speech recognition, and excellent natural language processing all help students learn faster and get better. One of the other good things about them is that they move faster. The results demonstrate that AI could improve or perhaps take the place of traditional language teaching methods when there aren't enough trained teachers available. The system can keep track of each learner's progress and update the content as needed. This makes it feasible to build a tailored learning experience that focuses on each learner's weaknesses and encourages them to improve their skills on their own. This allows the system detect and fix the problems with each person.

REFERENCES

[1] J. Fathi, M. Rahimi, and A. Derakhshan, "Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions," *System*, vol. 121, p. 103254, 2024.

[2] E. Madhavi, L. Sivapurapu, V. Koppula, P. E. Rani, and V. Sreehari, "Developing learners' English-speaking skills using ICT and AI tools," *J. Adv. Res. Appl. Sci. Eng. Technol.*, vol. 32, no. 2, pp. 142–153, 2023.

[3] W. Y. Hwang, M. I. Luthfi, and Y. F. Liu, "AI-enhanced video drama-making for improving writing and speaking skills of students learning English as a foreign language," *Innov. Lang. Learn. Teach.*, pp. 1–18, 2024.

[4] S. M. R. Leba and R. B. B. Butar-Butar, "The function of artificial intelligence (AI) in developing English language learner's speaking performance," in *Proc. Int. Conf. Sci. Technol. Internet Things*, vol. 1, no. 01, pp. 240–249, Dec. 2024.

[5] J. Chen, P. Lai, A. Chan, V. Man, and C. H. Chan, "AI-assisted enhancement of student presentation skills: Challenges and opportunities," *Sustainability*, vol. 15, no. 1, p. 196, 2022.

[6] J. Yiling, M. Omar, and F. M. Kamaruzaman, "Exploring the AI-enhanced project-based learning for English language acquisition: A systematic review of the key elements and emerging technology trends," *Int. J. Learn. Teach. Educ. Res.*, vol. 24, no. 2, pp. 636–652, 2025.

[7] A. S. A. E. AbuSahyon, A. Alzyoud, O. Alshorman, and B. Al-Absi, "AI-driven technology and chatbots as tools for enhancing English language learning in the context of second language acquisition: A review study," *Int. J. Membr. Sci. Technol.*, vol. 10, no. 1, pp. 1209–1223, 2023.

[8] N. Sari, "The role of artificial intelligence (AI) in developing English language learner's communication skills," *J. Educ.*, vol. 6, no. 01, pp. 750–757, 2023.

[9] D. M. A. S. K. Khasawneh, "Improving the learning of language proficiency at tertiary education level through AI-driven assessment models and automated feedback systems," *Migr. Lett.*, vol. 21, no. 2, pp. 712–726, 2024.

[10] F. R. Baskara, "AI-driven dynamics: ChatGPT transforming ELT teacher-student interactions," *Lensa: Kajian Kebahasaan, Kesusastraan, dan Budaya*, vol. 13, no. 2, pp. 261–275, 2023.

[11] Y. Chen, "Enhancing language acquisition: The role of AI in facilitating effective language learning," in *Proc. 2024 3rd Int. Conf. Humanities, Wisdom Educ. Serv. Manag. (HWESM 2024)*, pp. 593–600, May 2024.

[12] A. Tahir and A. Tahir, "AI-driven advancements in ESL learner autonomy: Investigating student attitudes towards virtual assistant usability," in *Linguist. Forum - J. Linguist.*, vol. 5, no. 2, pp. 50–56, Jun. 2023.