

The Effect of AI-Powered Chatbots and Prompt-Engineering Frameworks on English Speaking Proficiency and Speaking Anxiety among Medical Students

Fatema M. Albaqali^{1,*}, Sahar R. Hamzah², Safaa A. Abdulrahim², and Almothana M. Gasaymeh²

¹Department of Education and Curricula, Arabian Gulf University, Manama, Kingdom of Bahrain

²Department of Education and Curricula, Arabian Gulf University, English Language Unit, Manama, Kingdom of Bahrain

Email: 20241727fa@agu.edu.bh (F.M.A.); saharh@agu.edu.bh (S.R.H.); safaaa@agu.edu.bh (S.A.A.), almothanamag@agu.edu.bh (A.M.G.)

*Corresponding author

Abstract—This research investigates the use of Artificial Intelligence (AI)-powered chatbots on English speaking proficiency and speaking anxiety among English as a Foreign Language medical and nursing students in a Gulf Cooperation Council higher education context. Prior studies report positive results from AI-mediated speaking practice but significant research gaps remain. Existing literature primarily focuses on non-medical student populations on ready-made AI language apps that are not easily adaptable by instructors. Existing literature rarely addresses speaking proficiency and speaking anxiety concurrently. The Gulf Cooperation Council is a promising area of research since given that few studies in this region explicitly articulate how large language models are pedagogically guided through structured prompting. This leaves an underexplored gap in prompt engineering for language learning in the region. To address these gaps, this research will employ a pretest-posttest control group design that involves foundation program medical and nursing students at the Arabian Gulf University in the Kingdom of Bahrain. The experimental group will engage in structured speaking practice using Grok Voice Mode, an AI-powered conversational chatbot, while the control group will follow the standard curriculum. Data will be collected using standardized English speaking proficiency assessments and an adapted Foreign Language Speaking Anxiety Scale administered before and after the intervention. In addition to measuring learning outcomes, this research aims to develop and validate a structured prompt engineering framework to guide learners' interactions with Grok. This framework will deploy pedagogical objectives into customizable prompts aligned with speaking components, proficiency levels, and anxiety-reduction strategies. The findings aim to contribute theoretically to AI integration in language learning, English as a Foreign Language curriculum design, and evidence-based prompt engineering practices.

Keywords—Artificial Intelligence (AI) chatbots, English as a Foreign Language (EFL), speaking proficiency, speaking

anxiety, prompt engineering, medical education, Gulf Cooperation Council (GCC) higher education

I. INTRODUCTION

Since the end of World War II, the global use of the English language has increased and become the lingua franca for communication worldwide, including education, healthcare, science, and technology [1, 2]. The number of English speakers in the world is over 1.5 billion, with English as a Foreign Language (EFL) speakers outnumbering native speakers, according to widely quoted linguistic sources [3]. The study of English, speaking ability in particular has become a requirement in higher education institutions and a fundamental skill for various professions [4]. Studies have also repeatedly demonstrated that EFL students experience linguistic barriers such as differences in vocabulary level and mother-tongue dependency, and non-linguistic barriers such as Foreign Language Speaking Anxiety (FLSA). These barriers inhibit EFL speaking, particularly FLSA, as it directly impairs effective oral communication [4, 5]. English is an established academic and professional language in the Gulf Cooperation Council (GCC) region in which cross-border exchanges are common. A lack of English language proficiency and speaking anxiety still hinders effective communication in the region [6]. In the Kingdom of Bahrain, this is reflected in university admission requirements. Many universities have implemented English foundation programs designed to help bridge students' language gap [7]. In spite of this, students at the foundation level in medical schools or nursing schools, still struggle with their language proficiency. Recent developments in Artificial Intelligence (AI) chatbots have, shown strong potential in creating low-stakes, interactive

speaking experiences that are designed to promote speaking development and reduce anxiety [8, 9].

Several research gaps are evident in the existing literature. Previous research relies on standard commercial applications with static speaking activities and minimal personalized features [8–11]. Most studies also center on non-medical students, focusing on general speaking ability but not speaking anxiety [12]. Limited studies exist in the GCC, and few provide explicit articulation of how large language models are pedagogically guided through structured prompts [13–16].

These gaps need to be addressed through theoretical and practical studies. This study seeks to examine the impact of Grok Voice Mode on English speaking proficiency and speaking anxiety levels in foundation EFL medical and nursing students [17]. The research will present a framework for prompt engineering to assist effective AI-mediated speaking practice. The study contributes to digital technology language learning research and offers insights for educators to integrate AI into foundation-level EFL instruction.

II. LITERATURE REVIEW

This section provides reviews of theoretical and empirical studies related to English-speaking proficiency, speaking anxiety, and the use of AI-powered conversational tools for EFL education. The review is organized thematically and discussed chronologically to highlight the evolution of research in language learning theories, technology-enhanced language learning, and AI-mediated speaking practice.

A. Language Learning Theories Related to Speaking Proficiency and Anxiety

Early theories of language learning emphasize behavior that is observable and reinforced. Behaviorism views learning as dependent on stimulus-response interactions shaped through repetition and the utilization of reward structures [18]. From this perspective, language acquisition is accomplished through repeated exposure, immediate feedback and practice. All these components can be assisted by AI-powered tools that provide instant responses and corrective feedback.

Constructivist theory draws attention to learners' responsibility to actively construct their own knowledge through interaction and experience. According to Vygotsky's idea of the Zone of Proximal Development (ZPD), learners require support from a "More Knowledgeable Other (MKO)" to do tasks that are beyond their current ability [19]. For AI-mediated language learning, conversational chatbots have the potential to act as MKOs through scaffolding speaking practices with input from learners to shape feedback as needed.

Krashen's Monitor Model also sheds some light on second language acquisition, especially through the Input Hypothesis and Affective Filter Hypothesis [20]. Krashen explains that language acquisition takes place when learners are presented with comprehensible input slightly higher than their current level and when affective factors such as anxiety are reduced. High speaking anxiety raises

the affective filter which hinders acquisition, whilst low-anxious settings make learning easier. This agrees with Horwitz *et al.*'s [5] concept of Foreign Language Classroom Anxiety such that apprehension surrounding communication and fear of negative assessment serve as drivers of linguistic performance.

B. Technology-Enhanced Language Learning and AI Integration

With the integration of digital technology into education, frameworks such as the Technological Pedagogical Content Knowledge (TPACK) emphasize the importance of aligning technology, pedagogy, and content knowledge [21]. Using conversational AI, Mishra and Koehler contend that successful technology integration necessitates a careful intertwining of these three fields, rather than considering technology as an attachment. For AI-based language acquisition, this framework highlights the importance of using pedagogically guided tools and not just relying on technology.

The swift emergence of generative artificial intelligence started to reshape educational practices in the aftermath of ChatGPT in 2022 [22]. The emergence has led to the creation of more sophisticated conversational tools with voice-based interaction for these purposes. Grok by xAI represents a newer generation of AI-powered chatbots, which provide live transcripts, interactive voice conversations, and real-time feedback, making it suitable for speaking practice in EFL contexts [17].

C. AI Chatbots and Speaking Proficiency

Empirical research is increasingly investigating the contribution of AI chatbots towards EFL learners' speaking proficiency. In their study of the effects on speaking fluency, coherence, lexicon, grammatical accuracy, pronunciation, and willingness to communicate of the implementation of AI chatbot, Fathi *et al.* [11] find that learners in the AI-powered group outperform those in the traditional tutorship group significantly. Qualitative findings reveal that across all three teaching methods, learners value adaptive teaching methods, personalized feedback, and low-pressure speaking environments.

Other studies focus on specific parts of speaking. Aly and Alsheibany work mainly on pronunciation quality, and report that AI-enabled tools have been shown to make substantial advances in segmental and suprasegmental pronunciation attributes, and high learner motivation [23]. Together, these findings indicate that AI-mediated instruction can support targeted speaking outcome goals if performed in a systematic manner.

D. DAI Chatbots and Speaking Anxiety

Having speaking anxiety in EFL learning is a major barrier, as it is recognized to hinder effective communication [5]. Fauzi *et al.* [8] show that the use of AI applications significantly reduces speaking anxiety because improved speaking performance leads to reduced fear of negative evaluation whilst gaining increased confidence. Similarly, Ding and Yusof [9] provide evidence that students who practice speaking through AI-powered conversational chatbots exhibit greater

reductions in speaking anxiety and stronger improvements in speaking proficiency compared to those using traditional speaking methods [9].

Studies have also been conducted on younger learners. Tai and Chen [10] show that elementary EFL learners that interact with generative AI chatbots have better speaking performance and lower anxiety levels compared to learners in normal classroom settings. This indicates an increasing trend towards AI-mediated speaking practices as a strong means of overcoming linguistic and affective barriers in EFL learning.

E. Prompt Engineering as Pedagogical Practice

Prompt engineering is increasingly recognized as the *interface layer* that guides how learners and educators access the instructional ability of Large Language Models (LLMs). In education-oriented definitions, a prompt can be understood as a set of instructions that “programs” an LLM tool to unlock its abilities. Prompt engineering refers to crafting prompts that produce the most effective task performance. The content and structure of prompts can significantly influence model output quality. This means that learning outcomes are partially determined by how the human-AI interaction is set up, instead of being driven by technology alone. This agrees with broader prompt-engineering experience, which demonstrates that “designing specific prompts” can introduce high quality task-specific information and improve performance. Prompt engineering is now a rapidly developing field with identifiable methodologies (e.g., manual vs. automated prompting) and observable challenges that shape efficiency and effectiveness across multiple domains [13–16].

In the discipline of language education, the successful implementation of these toolkits is not “straightforward” for non-technical users. It often develops through trial-and-error through developed expertise and practice [13, 14]. Pedagogical perspective tells us that it reinforces the evolving role of the teacher as an educational prompt engineer who must guide the format and sequencing of AI feedback because system responses are driven by the specificity of the inputs or prompts [13]. Empirical EFL research further indicates that there is a gap in understanding how prompting functions as a learning competence. Studies note that learners’ success with ChatGPT depends on their ability to create “purposeful and well-structured prompts”, and examining prompt effectiveness and strategies is quickly becoming a core mechanism that supports successful learning outcomes [15]. Research examining EFL learners’ prompting during human-AI writing tasks also finds that there is “neither established practice nor much instruction” for prompt engineering, and that teachers should provide scaffolded guidance that is responsive to learners’ purposes for prompting [15].

F. Research Gap and Direction

While the value of prompt quality in AI-supported language learning interventions is increasingly recognized, the documentation and operationalization of prompts remain significant gaps. Many studies do not explicitly

document how prompts are designed to operationalize the pedagogical aim systematically. Examples of this could include prompts for gradual task progression, feedback framing, or anxiety-sensitive interactions. This limits the replicability of the studies and limits the ability of instructors to reproduce or adapt reported results in their own teaching contexts.

This study aims to address these interrelated gaps by looking at the use of Grok Voice Mode within a foundation-level EFL medical and nursing context, as well as establishing an explicit prompt engineering framework on this topic adapted to speaking practice. Building on the literature that positions prompt quality as a determinant of interaction quality and learning effectiveness, the framework codifies prompt components including: role assignment, task specification, constraints, scaffolding sequence and feedback rules, into a replicable pedagogical model. This model can be implemented with Grok Voice Mode to support speaking proficiency development while simultaneously reducing speaking anxiety [15].

III. MATERIALS AND METHODS

This research uses a quantitative research design to explore the impact of introducing an AI-driven conversational tool on EFL foundation-level medical and nursing students’ speaking proficiency and speaking anxiety.

A pretest–posttest control group design was employed (no random assignment was used). This design is common among prior studies investigating the impacts of AI-enabled chatbots on both speaking performance and anxiety. This gives an opportunity to compare AI-supported and traditional instructional approaches in classroom conditions [8–11].

Descriptions of the approaches to meet the objectives of the study are presented in detail to facilitate replication.

A. Ethical Approval and Informed Consent

This study will be reviewed and approved by the Arabian Gulf University Ethics Committee before the work is conducted and no data collection will occur until formal ethical clearance is provided. This aligns with ethical standards applied in similar educational studies involving AI-supported learning environments [11]. Informed consent will be obtained from all participants included in the study. Participants will be informed of the study’s purpose, procedures, voluntary nature, and their right to withdraw at any time without suffering any academic penalty. No identifiable data will be published. All data will be anonymized and used solely for research purposes. This study’s strong emphasis on informed consent is important since previous research has highlighted learners’ affective factors, such as anxiety and fear of negative evaluation, as central to language learning experiences [5].

B. Research Design

This research employs a pretest–posttest control group design, utilizing an experimental group and a control group. This design has been commonly adopted in studies

investigating AI-powered conversational tools in EFL contexts, as it allows researchers to measure changes in speaking proficiency and speaking anxiety attributable to the intervention while controlling for baseline differences [11].

The inclusion of both speaking proficiency and speaking anxiety as outcome variables is justified by prior research demonstrating that linguistic and non-linguistic barriers jointly affect EFL learners' oral performance [4]. Studies grounded in Krashen's Affective Filter Hypothesis and Horwitz *et al.*'s [5] work on Foreign Language Classroom Anxiety emphasize that anxiety must be addressed alongside skill development to achieve meaningful improvement.

C. Research Variables and AI Tool Selection

The independent variable featuring in this study is the use of Grok Voice Mode [17]; an AI-powered conversational tool that facilitates learners to be able to practice their English speaking through interactive voice-based dialogue. Grok employs Automatic Speech Recognition (ASR) and Natural Language Processing (NLP), which work together to support real-time conversational interaction. The developers of Grok have trained their own Voice Activity Detection (VAD) system to identify human speech in audio. The developers state that "Grok Voice Agents can speak dozens of languages with native-level proficiency, accurately capturing nuances in dialects and pronunciations" [24].

The selection of Grok Voice Mode is justified by several pedagogical and practical considerations identified in the literature. Prior studies indicate that AI-powered chatbots are most effective when they provide immediate feedback, allow repeated practice, and create a low-pressure, non-judgmental speaking environment [8, 9]. Grok's live transcript feature further supports learning by allowing learners to visually monitor their spoken output, reinforcing form-meaning connections through immediate feedback.

The dependent variables are:

- English speaking anxiety;
- English speaking proficiency.

D. Population, Sample, and Contextual Justification

Foundation program students served as the population for this study. They are currently enrolled at the Arabian Gulf University. The research sample will be students from the second semester, divided into an experimental group and a control group.

This population was selected since previous research highlights that foundation-level EFL students from healthcare disciplines often struggle with speaking proficiency and confidence despite meeting academic admission requirements [4, 6]. Limited research has examined AI-mediated speaking practice among medical and nursing students in GCC contexts as mentioned previously, making this sample particularly relevant [6, 9–11].

Participants are native Arabic speakers from the Kingdom of Bahrain, Saudi Arabia, and Oman. Their estimated English proficiency level is A2 according to the

CEFR. This aligns well with prior research work that demonstrates AI-mediated speaking practice is beneficial for lower-proficiency language learners [10].

E. Intervention Procedure and Prompt Engineering Framework

An orientation session will be conducted to introduce students to the voice tool in order to explain its ethical use and train them to interact with it using guided prompts. Participants in the experimental group will act in structured speaking practice using Grok Voice Mode.

A key methodological contribution of this study is the creation, development and implementation of a structured prompt engineering framework. Previous research work relied on existing 'off the shelf' applications with fixed speaking activities and limited pedagogical customization [8–11]. This study deploys pedagogical objectives through carefully designed prompts that are focused on learners' proficiency level and anxiety reduction whilst addressing: fluency, coherence, vocabulary, grammatical accuracy and pronunciation.

These strategies include increasing gradual task complexity and offering supportive feedback framing. This framework is developed and used after constructivist principles, in which learners actively construct knowledge through interaction. It is also informed by Krashen's Affective Filter Hypothesis, which emphasizes the need for low-anxiety learning environments to facilitate learning [19, 20]. Prompts are scaffolding tools, positioning the AI as a "More Knowledgeable Other" to move learners through progressively challenging speaking tasks, consistent with Vygotskian perspectives on mediated learning [19].

The control group will still receive regular speaking instruction following the existing foundation program curriculum, without AI intervention.

F. Data Collection Instruments and Justification

Two tools will be used:

1) English speaking proficiency test

Speaking proficiency will be assessed before and after the intervention via an oral test graded by qualified examiners. The rubric will be adapted from the Cambridge PET (B1 Preliminary) speaking assessment exercise. This assessment exercise is a widely recognized standard used in EFL speaking research globally [25].

2) English speaking anxiety scale

Speaking anxiety will be measured with an adapted version of Horwitz *et al.*'s [5] Foreign Language Classroom Anxiety Scale (FLCAS). This instrument has been extensively validated and frequently used in studies examining speaking anxiety in EFL contexts, including research involving AI-mediated instruction [5, 8].

G. Data Analysis

Quantitative data will be analyzed with descriptive statistics and inferential tests, such as paired-sample and independent-sample t-tests to examine changes in speaking proficiency and anxiety within and between groups. This approach can be found in prior AI-mediated

EFL studies [8, 9]. This allows for direct comparison of the intervention effects. In addition to the outcome analysis, interaction logs and prompt usage patterns will be analyzed descriptively to inform any refinement required of the proposed prompt framework to support its future applications in educational contexts.

IV. RESULTS AND DISCUSSION

Since this study is in progress, results reported in this section are expected outcomes according to the proposed research design, the reviewed literature, and theoretical frameworks underpinning the study. The discussion integrates these anticipated findings with existing research to show the research's potential contribution and significance.

Given that previous evidence has shown use of AI-based conversational tools in EFL settings, the experimental group with Grok Voice Mode is expected to make a measurable improvement in English speaking proficiency compared to the control group. Enhancements are anticipated in key speaking components (e.g., fluency, coherence, lexical resource, grammatical accuracy, and pronunciation) as measured by the adapted Cambridge PET speaking rubric [25].

It is further expected that after taking the post-test, the scores of speaking proficiency in the experimental group will be statistically greater than pre-test scores, while in the control group, it will show either modest improvement or no statistically significant change. These results are in agreement with previous research that found that AI-mediated speaking practice creates frequent, low-pressure opportunities to practice oral communication [10, 11].

In the context of speaking anxiety, it is anticipated that participants in the experimental condition will experience a decrease in their level of English speaking anxiety upon intervention, as measured by the adapted Foreign Language Classroom Anxiety Scale (FLCAS). Specifically, reductions are anticipated on measures of communication apprehension and fear of negative evaluation, having been identified as core components of FLSA [5]. Control group anxiety is likely to remain flat, or slightly [8, 9].

A descriptive analysis of learner interactions with Grok is expected to demonstrate an increased level of engagement, such as sustained speaking turns and the capability to handle progressively complex tasks.

This study is expected to be an important contribution to the theoretical understanding of how speaking proficiency and speaking anxiety are closely related and should be studied and accounted for at the same time. Improved speaking performance combined with a decrease in anxiety may support Krashen's Affective Filter Hypothesis where lower anxiety promotes language acquisition by allowing learners to process inputs more effectively [20].

The anticipated increase in speaking skills can be accounted for by the numerous characteristics of Grok Voice Mode mentioned in the proposed scenario, such as real-time conversational interaction, immediate feedback, live transcripts, and repeated practice opportunities. Such

aspects are consistent with the behaviorist theories of reinforcement and repetition effects as well as constructivist perspectives of active knowledge construction through interaction [18, 19]. It is anticipated that the ability to access live transcripts will further support learners' metalinguistic awareness by promoting the immediate noticing and self-monitoring of spoken output, a concept that has been emphasized in technology-enhanced language learning research [26].

The resulting reduction in speaking anxiety should be achieved through Grok's non-judgmental and self-paced teaching and learning atmosphere which decreases the apprehension when learners encounter negative evaluation, an important aspect of the effect of FLSA [5]. This expected effect aligns with earlier work that shows that conversational tools embedded with AI enable psychologically safe contexts for oral practice, especially in the case of less proficient learners [10].

This study is expected to extend existing research by demonstrating that AI-powered chatbots work not only according to their technology but also in how they are pedagogically managed. The prompt engineering framework presented here represents a major methodological addition that addresses the observation that there is a gap in previous research where commercial applications were utilized. Learner-AI interactions were never explicitly structured [8–11].

This article makes a strong contribution to methodology in this field and utilizes AI to aid in pedagogical design using a structured prompt engineering framework to establish methodology as an influential research target.

This framework is based on well-known and widely accepted language learning theories. With an understanding of Krashen's Affective Filter Hypothesis, prompts are designed with the intent to reduce anxiety through supportive feedback, gradual task progression, and learner control over the pace of interaction [20]. As scaffolding tools; and, in contrast, from a constructivist framework, prompts serve as scaffolding aids by placing AI as a "More Knowledgeable Other" that facilitates learning within the learners' Zone of Proximal Development [19]. Repetition, reinforcement, and immediate corrective feedback reflect behaviorist principles [18].

The expected results will show that increased speaking proficiency and decreased speaking anxiety may result not only from the AI technology, but also from the orchestrated execution of learner-AI interaction. In this regard, the methodology serves as both an investigative mechanism and a transferable pedagogical model.

Finally, through applying this methodological approach within the context of a foundation medical and nursing practice in the GCC context, the study extends local evidence in a literature dominated by non-GCC contexts [6]. Accordingly, the anticipated findings contribute to existing literature by providing a transferable methodological framework for implementing large language models in EFL speaking instruction, with implications for curriculum development, teacher training, and responsible adoption of AI in higher education.

V. CONCLUSION

The purpose of this study is to explore the impact of conversational chatbots powered by AI on English speaking proficiency and speaking anxiety among foundation-level EFL medical and nursing students, as well as to address a critical gap in methodological approaches to the pedagogical use of such a tool. Based in the higher education environment of the GCC, the study aligns with the increasing necessity of teaching students academic and professional proficiency in using English, particularly in healthcare disciplines, which are primarily the domain of oral communication.

Instead of treating AI chatbots as isolated technologies, this study suggests that pedagogical design mediates their effectiveness. Building upon behaviorist, constructivist and affective understanding of language learning, this study integrates Grok Voice Mode from the learning environment in structured speaking practice. The anticipated findings, grounded in the reviewed literature and theoretical frameworks, highlight that improvements in speaking proficiency and reductions in speaking anxiety are closely interconnected and must be addressed concurrently. This also supports the notion that both linguistic and nonlinguistic factors mediate EFL learners' oral performances.

A central idea of this research is the framing methodology itself. The proposed prompt engineering framework enacts pedagogical objectives in terms of structured and customizable prompts that mirror learners' proficiency levels, speaking components, and anxiety-reduction strategies. By doing so, the study overcomes the dependence on pre-defined commercial applications that consist of fixed speaking actions with little tailoring, creating a replicable and reusable model for AI-supported speaking teaching.

There are quite a few future aspirations to this research in addition to the current goals. Once data collection and analysis are done, the study endeavors to make an empirical assertion and enhance the proposed prompt engineering framework based on the results of learner performance data, interaction behaviors, and speaking anxiety. It will examine which prompt structures, sequencing strategies and feedback framing techniques are the most effective for foundation-level EFL learners and particularly in high-pressure disciplines like medicine and nursing. This research was also used to refine the conceptual framework and to consolidate it as a tangible model and embed it into foundation programs such as the curricula and training for instructors.

However, this study is a preliminary step toward a larger-scale evaluation of the effectiveness of structured prompt engineering by conducting a cross-faceted analysis of how well it is scalable and transferable across different proficiency levels and disciplines. Longitudinal studies are necessary to investigate the longer-term effects of AI-mediated speaking practice on the academic performance, professional expression ability and retention of reductions in speaking anxiety of students. Further investigation is also required into the effect of teacher-mediated prompt

design, instructor vs. learner prompts and the consequences for speaking outcomes of learner freedom.

Finally, as few studies have been carried out on the GCC, more regionally-situated research is needed to inform evidence-based policies concerning AI in higher education. This kind of research will serve to contribute to ensuring that the deployment of AI-empowered conversational tools is in line with local learning goals, linguistic realities and the broader drive towards producing graduates with the skills, knowledge, and confidence needed to thrive in the global knowledge economy.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Fatema M. Albaqali led the conceptualization of the research, conducted the literature review, and developed the proposed framework. Dr. Sahar R. Hamzah contributed to the study's research design, methodological development, and academic co-supervision. Dr. Safaa A. Abdulrahim provided academic co-supervision in regards to the English as a Foreign Language methodology. Dr. Almothana M. Gasaymeh provided primary supervision, guidance on aligning project management competencies, and validation of the research approach. All authors critically reviewed the manuscript and approved the final version.

ACKNOWLEDGMENT

This research was supported and created through the initiatives of His Majesty King Hamad Academic Chair Scholarship in Digital Education, aiming to advance digital learning and foster innovative educational practices. Our commitment to enhancing educational accessibility and quality aligns with the Chair's mission to leverage technology for transformative learning experiences.

REFERENCES

- [1] Merriam-Webster. (Jan. 18, 2025). *Lingua franca*. *Merriam-Webster.com Dictionary*. [Online]. Available: <https://www.merriam-webster.com/dictionary/lingua%20franca>
- [2] Cambridge Dictionary. (Jan. 18, 2025). *Lingua franca*. *Cambridge Dictionary*. [Online]. Available: <https://dictionary.cambridge.org/dictionary/english/lingua-franca>
- [3] SIL International. (Jan. 18, 2025). What is the most spoken language? *Ethnologue*. [Online]. Available: <https://www.ethnologue.com/insights/most-spoken-language>
- [4] B. Ross and D. Stuckler, "Barriers to learning English as a Foreign Language (EFL) for university students: A systematic review," *Int. J. Instr.*, vol. 18, no. 4, pp. 395–420, 2025.
- [5] E. K. Horwitz, M. B. Horwitz, and J. Cope, "Foreign language classroom anxiety," *Mod. Lang. J.*, vol. 70, no. 2, pp. 125–132, 1986.
- [6] F. A. M. Yapar and O. E. Yapar, "Language anxiety as a barrier to effective communication in the GCC region," in *Proc. Conceptual Change Theories and Practice Examples in Education*, A. K. and M. Eraslan, Eds. Lyon, France: Livre de Lyon, 2023.
- [7] Arabian Gulf University. (Jan. 19, 2026). Undergraduate admission. [Online]. Available: <https://www.agu.edu.bh/en/page/undergraduate-admission>
- [8] I. Fauzi, R. Hartono, D. Rukmini, and H. Pratama, "AI applications for EFL learners: Enhancing speaking performance and reducing

- anxiety with gender-based analysis,” *Forum Linguist. Stud.*, vol. 7, no. 9, pp. 282–301, 2025.
- [9] D. Ding and M. B. Yusof, “Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: A mixed methods study,” *Humanit. Soc. Sci. Commun.*, vol. 12, no. 1, pp. 1–16, 2025.
- [10] T. Y. Tai and H. H. J. Chen, “Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions,” *Comput. Educ.*, vol. 220, 105112, 2024.
- [11] J. Fathi, M. Rahimi, and A. Derakhshan, “Improving EFL learners’ speaking skills and willingness to communicate via artificial intelligence-mediated interactions,” *System*, vol. 121, 103254, 2024.
- [12] N. Vy and V. P. H. Pham, “AI chatbots for language practices,” *Int. J. AI Lang. Educ.*, vol. 1, no. 1, 10.54855, 2024.
- [13] I. Isemonger, “Generative language models in education: Foreign language learning and the teacher as prompt engineer,” *TEFL Praxis J.*, vol. 2, pp. 3–17, 2023.
- [14] D. J. Woo, K. Guo, and H. Susanto, “Exploring EFL students’ prompt engineering in human–AI story writing: An activity theory perspective,” *Interact. Learn. Environ.*, vol. 33, no. 1, pp. 863–882, 2025.
- [15] W. Andewi, W. Waziana, D. Wibisono, K. A. Putra, T. Hastomo, and I. B. Oktarin, “From prompting to proficiency: A mixed-methods analysis of prompting with ChatGPT versus lecturer interaction in an EFL classroom,” *J. Stud. Engl. Lang.*, vol. 20, no. 2, pp. 210–238, 2025.
- [16] J. Wang *et al.*, “Prompt engineering for healthcare: Methodologies and applications,” *Meta-Radiology*, 100190, 2025.
- [17] xAI. (Nov. 3, 2023). Announcing Grok. [Online]. Available: <https://x.ai/news/grok-1.5>
- [18] B. F. Skinner, “Behaviorism at fifty: The rapid growth of a scientific analysis of behavior calls for a restatement of the philosophy of psychology,” *Science*, vol. 140, no. 3570, pp. 951–958, 1963.
- [19] L. S. Vygotsky, “Mind in Society: The Development of Higher Psychological Processes,” Cambridge, MA, USA: Harvard University Press, 1978.
- [20] S. Krashen, “Principles and Practice in Second Language Acquisition,” Oxford, U.K.: Pergamon Press, 1982.
- [21] P. Mishra and M. J. Koehler, “Technological pedagogical content knowledge: A framework for teacher knowledge,” *Teach. Coll. Rec.*, vol. 108, no. 6, pp. 1017–1054, 2006.
- [22] OpenAI. (Jan. 19, 2026). *Introducing ChatGPT*. [Online]. Available: <https://openai.com/index/chatgpt/>
- [23] S. M. Abdelhalim and R. A. Alsehibany, “Integrating AI-powered tools in EFL pronunciation instruction: Effects on accuracy and L2 motivation,” *Comput. Assist. Lang. Learn.*, pp. 1–25, 2025.
- [24] xAI. (Jan. 19, 2026). *Grok Voice Agent API*. [Online]. Available: <https://x.ai/news/grok-voice-agent-api>
- [25] Cambridge Assessment English, “B1 Preliminary (PET) speaking assessment: Assessment criteria,” *Cambridge English*. 2026. <https://www.cambridgeenglish.org/images/563276-b1-preliminary-assessing-speaking.pdf>
- [26] W. C. Lin, Y. C. Yeh, and C. K. Chang, “Investigating the impact of an English meeting transcription tool on student speaking skills and confidence in EMI contexts,” *Int. J. Learn. Technol. Learn. Environ.*, vol. 8, no. 1, 2025.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).