Journal of Digital Sociohumanities
ISSN: 3032-7865 **Vol. 3 No. 1 2026**
Website : https://jds.unand.ac.id
Email : jds@hum.unand.ac.id | digitalsociohumanities@hum.unand.ac.id

Low-Proficiency ESL Learners' Emotional, Cognitive, and Metacognitive Engagement in Face-to-Face Versus AI-Mediated Spoken English Communications

Sumudu Nisala Embogama
Department of English Language Teaching, University of the Visual and Performing Arts, Colombo, Sri Lanka

Article information

DOI : 10.25077/jds.3.1.1-17.2026
Correspondence : sumudu.e@vpa.ac.lk

Submission Track

Submission : Janury 12, 2026
Final Review : April 15, 2026
Accepted : May 08, 2026
Available online : May 11, 2026

Keywords

AI chatbots, English speaking skills, Low-proficiency ESL learners, multimodality

Abstract

This multimodal phenomenological study investigates the lived experiences of low-proficiency ESL undergraduates' emotional, cognitive, and metacognitive responses during human versus AI chatbot-partnered spoken interactions using the target language of the learners, i.e., English. A multimodal visual analysis approach, grounded in thematic coding, supported the analysis of participant drawings, and the data from the semi-structured focus group discussions were thematically analyzed. The findings reveal a significantly positive response to including voice-enabled AI-chatbots as potential L2 speaking partners due to their non-judgmental, uncritical, accommodating, and facilitative potential, drastically reducing Foreign Language Anxiety (FLA), cognitive overload, fear of making mistakes, being judged, and ridiculed, all of which were present during their spoken interactions in English with a proficient speaker. Situated in a postcolonial context, the implications of reticence in speaking English yield significant findings that highlight nuanced, problematic, and ideologically and socially laden implications as they unravel the internalized power asymmetries and linguistic marginalization rooted in unwarranted colonial legacies. The study's major contribution lies in its depiction of the underrepresented postcolonial perspectives of linguistically disempowered individuals during their L2 speaking skills development. While reflecting on the potential of AI as an enabling resource for supporting the development of spoken English among beginner-level adult ESL learners, this study also draws attention to the structural inequalities within a postcolonial community that should be addressed with the understanding that AI alone cannot mitigate reticence in English speaking.

INTRODUCTION

For centuries, in Sri Lanka's linguistic landscape, the development of English as a second language (ESL) has been marked by a persistent paradox: ESL learners' yearning to be fluent speakers of the target language and their aversion to conversing in English (Parakrama, 1995; Rambukwella, 2018). Within the postcolonial Sri Lankan social milieu, the learning and use of English carries the loaded baggage of British colonial legacy, denoting power imbalances, hierarchies, and exclusion, whereby English, as the language of the colonizers, has long functioned as a marker of elitism and a perpetuator of classism, with proficient users gaining social mobility and class privilege as a result (Rambukwella, 2018). According to Parakrama (1995), the insistence on using the so-called Standard English has continued to sustain social hierarchies resulting from linguistic imperialism, resisting the acknowledgment of the evolution of world Englishes, which have often been delegitimized as somehow inferior and subordinate to British English. In such a complex scenario, where instead of perceiving English as a tool for global communication, its symbolic status as a denominator of upward mobility, elitism, and a marker of intelligence has not just social, economic, and political implications but also educational ones. By attributing linguistic and cultural capital to a language such as English, symbolic power is manifested through its users through covert and/or overt mechanisms (Bourdieu, 1991; Thompson, 1991).

Research conducted within the Sri Lankan ESL context consistently documents the evolving yet problematic/problematized space of English language teaching and learning. The persistent dilemma of Sri Lankan adult ESL learners' desire and, paradoxically, struggle to be competent users of English carries broader implications beyond the personal level (Parakrama, 1995; Rambukwella, 2018). The reticence to communicate in the second language is particularly acute in environments where the primary language of instruction is the ESL learners' first language, either Sinhala or Tamil, resulting from authentic exposure to the L2 (Sawalmeh, 2023. Scholars have noted that in such contexts, the affective and sociopolitical dimensions of learning a target language intersect in a manner that inhibits L2 communication, especially during L2 speaking of beginner-level learners (Horwitz et al., 1986; Teimouri et al., 2019). This persistent gap between the desire to be articulate in the target language and the reluctance to do so in private spaces constitutes a significant pedagogical challenge for ESL educators that warrants empirical studies.

Within the scope of this phenomenological research, the lived emotional, cognitive, and metacognitive experiences of visual arts undergraduates in two distinctive communicative contexts: face-to-face (F2F) and speech-enabled AI-chatbot spoken interactions are the foci for analysis. In spite of the considerable number of studies on AI integration in language education, several noteworthy gaps exist. Firstly, one notable gap in the literature is that most studies have been conducted in the Western world and Central and Southeast Asia, and almost none in Sri Lanka. Secondly, while most related studies tend to consider the emotional and cognitive dimensions of AI integration during the development of ESL, the sociopolitical and postcolonial dimensions have been overlooked (Rambukwella, 2018). In the present study, however, such broader issues are integrated into the interpretations to critically evaluate the findings by moving beyond individual factors to broader issues that impact the acquisition of English-speaking skills among beginner-level L2 learners. Moreover, the study attempts to fill the methodological gap resulting from the lack of studies that delve deeper into low proficient learners' affective and metacognitive states by using an art-based multimodal approach for data gathering to evaluate the multilayered visual and verbal data gathered from students situated within the postcolonial landscape of second language acquisition with the intention of obtaining a more sophisticated understanding of the topic by moving beyond mere linguistic expression.

Considering the existing body of literature, studies have examined how L2 learners' writing skills have been affected by the integration of AI-assisted writing. Abdullaev et al. (2025) engaged in a phenomenological study on how EFL Uzbekistani learners responded to personalized AI feedback for their English writing. The main findings indicate that participants valued the promptness and detailed feedback given by AI, despite some mentioning the need for teacher input so as to further verify the corrections and comments received. Similarly, Azman and Ibrahim (2025) explore the use of AI tools for supporting the academic writing of Malaysian TESL Diploma students, whereby they found that technological support prompted self-correction and developed the participants' writing proficiency. Ahmadi (2025) discovered that introducing ChatGPT as a supportive tool for developing English grammar received a largely positive response from the cohort, as the EFL Iranian undergraduates stated that their L2 grammar knowledge increased due to the AI's ability to provide corrective feedback.

Many studies on AI integration for learning English as a second language (ESL) focused on the impact it has on mitigating the presence of Foreign Language Anxiety (FLA). Similarly, in the case of the current study, attention was paid to the emotional and psychological dimensions of technology integration in ESL contexts. Aly and Didah Nurhamidah (2025) consider Indonesian undergraduates' responses to the use of AI for developing their English and its impact on FLA. Madina (2025) also delves into the emotional engagement of a group of Indonesian EFL learners using a multimodal approach, similar to this study, whereby data were gathered through visual and verbal cues. Similarly, Wang and Meng (2026) engaged in an analysis of the effect of AI chatbot intervention on EFL learners' communication skills. Nguyen and Kim (2025) specifically focused on the feasibility of chatbot-based speaking assessments among low-level EFL (English as a Foreign Language) students in South Korea. Students' perception

regarding the use of AI chatbots during L2 speaking provided similar conclusions in the sense that they were more willing to be assessed by AI interlocutors than human ones.

Despite the vast number of studies concluding that AI chatbots facilitate and enhance ESL (English as a Second Language) learners' spoken English, a few studies have proven otherwise. For example, Çakmak (2022), by contrast, found that after being exposed to the AI chatbot "Replika" during oral communication, the Turkish EFL learners, despite having improved their speaking skills, reported experiencing higher levels of anxiety due to the chatbot's inability to sometimes fully grasp what the speakers were saying. This finding is contradictory to the majority of the study findings claiming a reduction of FLA when AI technologies were integrated.

Moving on to the theoretical framework, Horwitz et al.'s (1986) conceptualization of Foreign Language Anxiety (FLA) is the primary theoretical lens used in this study. According to these scholars, L2 anxiety is seen as detrimental rather than facilitative to second language acquisition. The presence of such multifaceted variables could result in substantial impairment for learning, identified as avoidance, silence, and cognitive blockade of L2 learners' engagement and performance during oral exchanges (Sawalmeh, 2023). Studies conducted by MacIntyre and Gardner (1991), Teimouri et al (2019), and Woodrow (2006) exemplify how manifestations of anxiety negatively impact the process of L2 development in contexts where human interaction occurs. In contrast, many studies, including those conducted by Fathi et al. (2024), Kang (2022), Butarbutar (2024), Hanh (2024), Huraiti (2024), Anh (2024), Patiño (2024), Wang et al. (2024), Celik (2025), Yildiz (2024), Wang et al. (2025), Wibowo and Sumarno (2026) and Zhang (2026) claim that when using AI as their speaking partners in the second language journey, learners demonstrated increased Willingness to Communicate (WTC), took more risks when using the L2, and were more spontaneous in terms of language production.

In order to align the study findings in psycholinguistics, since it deals with affective domains, Self Determination Theory (SDT), developed by Deci and Ryan (1985), was also incorporated into the interpretive lens of the study. SDT refers to three tenets of learner needs, namely, autonomy, competence, and relatedness, and it supports the understanding of how these needs are fulfilled or unmet during spoken L2 exchanges in F2F versus AI-mediated contexts. Moreover, Bandura's (1977) Self-Efficacy Theory is used to substantiate claims related to learners' experiences during the two types of interaction, whereby they express their beliefs in their ability to perform successfully as L2 speakers. Studies have found that when self-efficacy is high, students display increased motivation to learn as they are emotionally receptive to what is being taught. Since anxiety is a main variable of concern, Krashen's (1982) Affective Filter Hypothesis is also crucial in identifying what raises or lowers the presence of anxiety during both human-mediated and AI-mediated spoken exchanges in the second language, i.e., English. According to him, negative affective states such as anxiety, low motivation, and reduced self-confidence raise a metaphorical filter that hinders productive L2 output.

The sociopolitical scaffolding for this study is based on Bourdieu's (1991) conceptualization of symbolic power and linguistic capital, adapted to explain the historical and sociopolitical contexts of these Sri Lankan users of English and their mediating role in spoken English contexts. As explained in the introduction to this paper, the pervasive presence of elitism of English whereby this language is considered a determinant for upward mobility and class privilege in the Sri Lankan postcolonial setting. This phenomenon often leads to the marginalization of non-proficient English users, resulting in resentment, demotivation, and frustration on their part (Rambukwella, 2018; Parakrama, 1995). Such uneven power dynamics appear to result in intense pressure for low-proficient users of English during interactions with proficient users of the L2. In addition, Control Value Theory (CVT) (Pekrun, 2006) looks at the perceived control of learners, whereby those who feel in control in educational contexts experience positive emotions such as pride and enjoyment, whereas those experiencing low levels of control are negatively impacted so that emotions such as anxiety and frustration are generated.

Against the above contextual and theoretical backdrop, the present study is guided by the overarching objective of investigating the emotional, cognitive, and metacognitive responses of low-proficiency Sri Lankan ESL undergraduates to English-speaking contexts during face-to-face communications versus with voice-enabled AI Chatbots and how their cognitive and affective experiences are impacted by postcolonial sociopolitical factors.

Accordingly, the following research questions were formulated:

1. What are low-proficient ESL visual arts undergraduates' emotional, cognitive, and metacognitive responses to face-to-face verbal interactions in English?
2. What are these visual arts undergraduates' emotional, cognitive, and metacognitive responses during English-speaking exchanges with voice-enabled AI chatbots?
3. To what extent do the postcolonial context and sociopolitical factors mediate these ESL learners' affective and cognitive responses during both face-to-face and AI-based verbal interactions?

METHODS

Study Design

To achieve the study objectives, the most appropriate methodological approach was multimodal phenomenology since the study explored subjective human perceptions and lived experiences within an L2 context. As such, qualitative data were gathered through a multimodal elicitation technique whereby secondary use of classroom-generated material in the form of student artifacts created during a reflective learning-oriented session, followed by group discussions involving reflection and analysis.

Participants

Eighty-four convenience-based purposively sampled beginner-level ESL undergraduates majoring in visual arts were the study participants. As is a common practice in qualitative classroom-based exploratory studies focusing on the affective domain, students' authentic and unbiased visual narratives and verbal responses were used as primary data. Participant anonymity was ensured through the removal of their identities from their artifacts, and through voluntary participation, learners were given the option of sharing their drawings via an online platform, in this case, WhatsApp, so that they were not coerced into sharing their illustrations. Verbal data were gathered through in-class semi-structured focus group discussions in the form of a reflective activity, whereby they shared their experiences and perceptions regarding English speaking with another human speaker and with an AI chatbot.

Materials and Instruments

Two primary instruments were used for gathering data. First, the primary data were the students' drawings, collected following an art-based elicitation task. Each student produced two illustrations: one involving their response to F2F L2 spoken interactions and the other during L2 verbal exchanges with AI-chatbots. This type of data is considered highly effective in qualitative studies that focus on mental and cognitive states that elude verbal expression (Madina, 2025). Second, semi-structured focus group discussions were employed in Sinhala to ensure clear articulation of their perceptions, and these utterances were translated into English when reporting the findings. The primary AI platform used by the participants was ChatGPT, although some mentioned the use of Google Gemini's voice-enabled feature.

Procedure

Both types of data were gathered following an initial intervention shortly after the undergraduates' first semester of study. The intervention was in the form of a seminar on autonomous learning practices to develop English through ICTs. During the session, students were introduced to various technologies for enhancing their English language skills. Specific attention was given to explaining how Generative

AI platforms can support this process. Towards the end of the semester, at which point students had adequate experience using AI interlocutors to develop their English speaking, the students created art-based visual narratives by doing two drawings capturing their experiences; firstly, while speaking with an L2 proficient interlocutor, and secondly, when speaking to a voice-enabled AI agent. Subsequently, during the discussion phase, the participants vocally expressed their feelings when speaking in English with both types of interlocutors.

Data Analysis

The data gathered during the discussions were thematically coded using the inductive approach, whereby codes emerged iteratively from the data. Thematic analysis followed the six-phase framework (Braun & Clarke, 2006) involving data familiarization, generation of initial codes, searching for and reviewing themes, defining and naming them, and producing the findings. Visual data were analyzed through thematic coding and multimodal discourse analysis to unravel the explicit and implicit meanings conveyed through their artistic expressions.

Ethical Considerations

The study adhered to established ethical principles involving human participants, as participation was purely voluntary, and their anonymity was ensured through the removal of any identifying information from either the visual or the verbal data. It is acknowledged that the researcher's role as the educator of these participants could have led to relational influence; however, by getting them to share their illustrations via WhatsApp anonymously and with volition and not involving them in individual interviews, the researcher intended to reduce any negative impact.

Methodological Limitations and Delimitations

As far as the study limitations are concerned, since the participants are limited to a particular institutional context, generalizability should be made with caution, although the findings may apply to similar sociolinguistic and educational contexts. In addition, the researcher's position as the ESL educator of this study sample could have potential for relational influence on the data produced. Nonetheless, the art-based methodology contributed to generating rich and nuanced data capturing affective states that elude verbal articulation.

RESULTS

When considering the participants' art-based narratives in the form of drawings, three prominent themes emerged regarding their reactions to F2F L2 speaking contexts.

Negative Evaluation, Fear, and L2 Anxiety

The main findings of the visual data depict the recurring iconographic representations of these low-proficient English speakers' expressions of fear and anxiety through overt and covert artistic expressions. Across the corpus of illustrations, participants depicted themselves with downcast eyes, sealed lips, hands that reach out as if to cover their mouths, and fear-stricken eyes. These visual narratives encoded with ESL learners' obvious desire for invisibility and withdrawal from the spoken encounter. The pervasive use of achromatic and darkly saturated palettes, jagged lines, and distorted/disproportionate compositional figures presents powerful images of psychological distress.

As with Figure 1, these drawings often use no colour, or dark colours indicating the level of darkness in their emotional experiences, while the use of chaotic forms and jagged lines also provides a visual manifestation of the internal struggle during spoken exchanges with proficient English speakers. Figure 2 depicts a resonant symbolic construction as the low-proficient L2 speaker's face, totally obliterated by the paper bag bearing the inscription: "I can't," which indicates the inability of the speaker to produce

English verbal utterances. The speech bubble pointed towards this interlocutee is even more disturbing, as it directly claims that he is "afraid to speak English". The unspoken reason indicated by the thought bubble is significant since the drawing is that of a thought bubble and not a direct utterance, indicating the main reason for the inability, which is identified by the speaker as fear of speaking in English.

Here, the paper bag acts as a polysemous symbol connotating, on the one hand, the participant's sense of suffocation and the acute anxiety. On the other hand, the image of the paper bag, used in instances of hyperventilation for supporting breathing, highlights the tangible panic-stricken state of the low-proficient speaker resulting from the fear of being judged and ridiculed when speaking in English. What is more interesting is the conscious decision by the student to draw himself wearing a tie. It can be interpreted as a sartorial symbol of professional aspiration and social mobility. In the symbolic economy of this drawing, there is an irreparable disconnect between social performance and linguistic performance.

Fig 1. Fear-stricken reaction of a low-proficient speaker to face-to-face spoken English interactions

Fig 2. Expression of linguistic inadequacy, feeling of suffocation, and desire for concealment

The verbal data subsequently elicited during the focus group discussions corroborate and extend on the visual findings as participants describe feelings of desperation, sinking, and raised anxiety levels. One participant invoked the metaphor of a trapped animal, exemplifying the sense of entrapment in L2 communicative contexts; a sentiment that resonated with the others in the group. These students reiterated that their silence is not simply due to their limited proficiency in the L2 but as a result of the social stigma invariably associated with incorrect use of English.

Power Imbalance and Linguistic Hierarchies

Yet another prominent theme that emerged from the analysis of the visual data was that of explicit and implicit manifestations of asymmetrical power relations and linguistic hierarchies arising during the verbal interactions involving English-speaking and non-English-speaking participants.

Figure 3 encodes the unequal relationship between the proficient English speaker and the struggling L2 user through a consciously presented semiotic system of meaning-making through sartorial symbolism. Here, the symbolic synonymizing of the phrase "knowledge of English" with the formal coat worn by the L2 speaker seems to indicate the monitory supremacy, arrogance, and Western imposition of language, signaling the broader historical implications of unwarranted colonial legacies. The proficient speaker is fully clothed in Western attire, whereas the less proficient L2 speaker is incompletely dressed as he is conspicuously barefoot and in shorts. This compositional choice eloquently depicts the impossibility of complete cultural and linguistic acculturation within the postcolonial Sri Lankan milieu. The symbolic deconstruction of the expected norm, of wearing trousers with a coat, is indicative of the fact that English proficiency alone cannot fully compensate for socioeconomic conditions of poverty and inequality, such as in education and job placement.

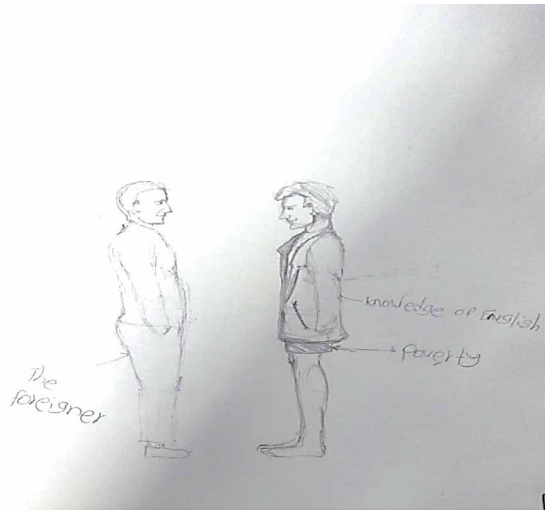

Fig 3. Sartorial symbolism and colonial linguistic hierarchy reflecting asymmetrical power relationships

Fig 4. Manifestation of linguistic subjugation through disproportionate stature between interlocutor and interlocutee

Figure 4 intensifies the above interpretations through the technique of differential scale, whereby the student has drawn himself of smaller stature to his interlocutor, who is a large, overbearing figure pointing to his necktie as if to indicate his superior status acquired through linguistic capital. The light bulb drawn near his head represents the general perception that by virtue of knowing the L2, he is regarded as cognitively more sophisticated as a knowledge generator/creator, whereas the non-proficient English speaker, proportionately less than half the others' height, is surrounded by question marks and a speech bubble directed to his head, reading "my mind". At a superficial interpretation of the constant presence of such question marks in the illustrations drawn by these ESL undergraduate students can be interpreted as indicative of the sense of confusion they experience during interactions with profieicent English speakers and the incomprehensibilities that arise; however, at a more nuanced level, the interrogative marks symbolize a radical questioning of the ideological conditions that have created such profound epistemic and social disempowerment. Thus, these visual depictions correspond well with Bourdieu's (1991) definition of symbolic capital as being embodied and reproduced through social interactions. Moreover, the findings correlate with Parakrama's (1995) stipulation of the gatekeeping function of perpetuating questionable linguistic norms in Sri Lankan society.

During the discussion, students expressed a similar sense of indignation and critique of being unjustly penalized for their lack of English proficiency, which they reiterate is not due to a fault of their own, but due to structural inequalities in the social and education systems that adversely affect their English development.

"English was a horrible period for us. The teacher would just go through the textbook and make us do the exercises, even though most of the time we did not know how and ended up copying the answers she finally shared." (Student A).

"It is so unfair that job placements often prioritize English skills over other skills, including academic prowess. That makes us hate English even more." (Student B)

Cognitive Paralysis and Linguistic Blockage

The third thematic configuration centres on the psychosomatic and cognitive dimensions of L2 communication. The visual data often depict the less proficient English speakers as demonstrating physiological manifestations of fear and anxiety as a result of being coerced into speaking English. Moreover, these illustrations depict the presence of cognitive paralysis as a result of the unsolicited interactions with fluent English speakers.

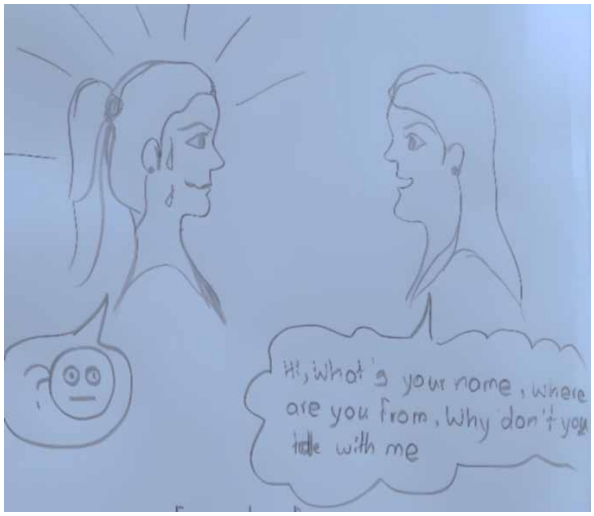

Fig 5. Somatic manifestations of L2 anxiety during face-to-face communication

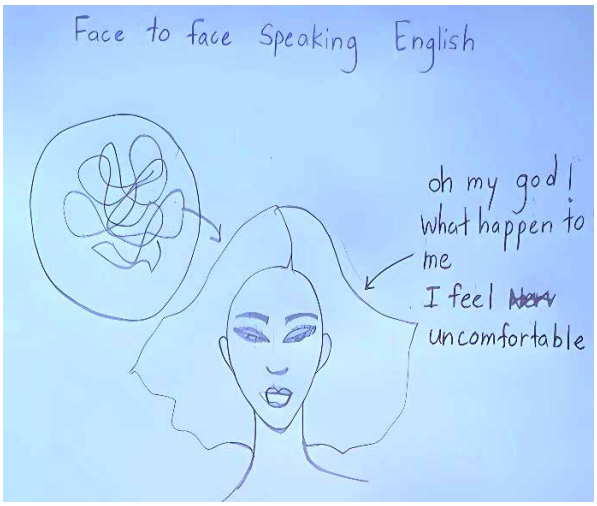

Fig 6. Brain fog and cognitive paralysis during spoken exchanges using English

Figure 5 is an instance of the physiological manifestations of inward psychological renderings, revealing what the presence of stress and unpleasant negative experiences during face-to-face encounters with a proficient English speaker entails. As in this illustration, often participants drew perspiration marks, trembling lines, tightly sealed lips, and downcast eyes suggestive of the seemingly traumatic experience of human-mediated oral exchanges using English. Figure 6 presents a striking visual manifestation of cognitive paralysis. The circle, with an arrow pointing towards the speaker's head, is filled with densely tangled, irrational lines, representing the state of her mind and its internal chaos during English-speaking encounters. The findings further validate these assertions as students claim to "go blank", feel overwhelmed, and have overcrowded thoughts that preclude coherent and comprehensible language output.

Reduced Anxiety and Psychologically Safe Spaces

In marked contrast to the participants' visual and verbal narratives of their lived experience during face-to-face spoken interactions in English, their responses to speaking with voice-enabled AI chatbots, mainly ChatGPT, involved positive and hope-filled illustrations and perceptions. Here, participants tend to depict calm skies, open landscapes, and expansive backgrounds, denoting the later interaction as one that is in a psychologically safe and non-judgmental space.

Fig 7. AI as a non-threatening companion, granting a sense of affective safety

Fig 8. Collaborative empowerment and equal stature between the ESL learner and the AI chatbot

Across the corpus of drawings on AI and human spoken exchanges, a noteworthy representational technique adopted by these undergraduates was the anthropomorphizing of the AI interlocutors. As in the two figures above, almost all participants depicted the AI as smiling, humanized figures with bodily forms and facial expressions communicating warmth and approachability. Such attributions of human-like or, in some cases, ethereal and angel-like qualities to the AI present it as a friendly and emotionally available companion. For example, in Figure 7, the heart icon strips the AI of its mechanical, emotionless status to an entity that feels compassion and even love towards the human speaker. Here, the low-proficient English speaker is depicted as smiling, relaxed, and with open expressions and gestures.

Figure 8 depicts an interesting symbolic construction resulting from the joining of hands between the human speaker and the AI chatbot, as if together they can conquer the world, since the drawing of the world is placed in the central position where the two hands appear to meet. It captures the participant's perceptions of the AI chatbot as her English-speaking partner to be an enabling, collaborative, and globally connected encounter. The drawing of the internal working of the brain with foggy, tangled, or jagged lines and question marks during F2F spoken exchanges is in stark contrast to how it is configured as embodying ordered molecular structures with embedded alphabetical letters masterfully illustrating the integration of English into the ESL learners' cognitive architecture.

During the verbal exchanges, there was a recurring insistence on how when communicating with the AI, these L2 learners felt that they no longer experience the same level of anxiety, embracement of making mistakes and the fear of being judged negatively as the AI felt like a reliable, friendly and non-judgmental confidante always willing to listen, repeat, correct and facilitate any request made unlike the patronizing gaze of a human speaker.

"I feel that with ChatGPT, I can make any mistake, hesitate, and pause without stressing about what it will think. I know it is a machine, but gradually I am beginning to think of it as a friend. Using AI is such a great way to practice my speaking." (Student C)

This mental transformation for cognitive blockage regarding L2 production to mental clarity, leading to enhanced language output, is evident as per the verbal and visual narratives.

Reduced Cognitive Load and Reflective Pausing

Closely related to the theme of affective safety, though analytically distinguishable, is the reported reduction of cognitive load during spoken interactions with the AI as a consequence of the reduction of anxiety and fear. The illustrations often portray a cognitively enabling rather than disabling presence attributed to the AI chatbots.

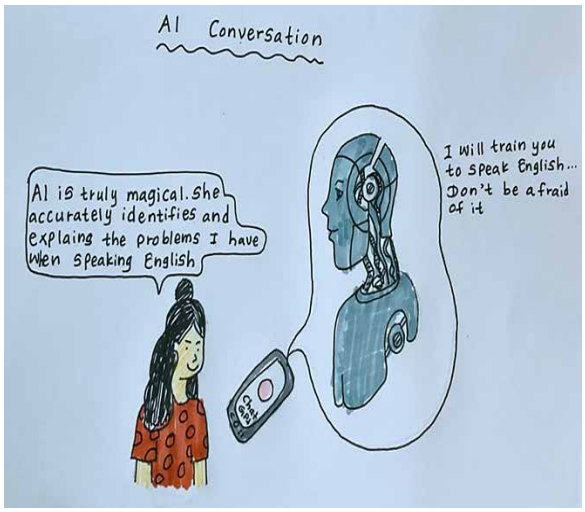

Fig 9. AI chatbot as an ideal speaking partner, granting affective safety

Fig 10. Nonintimidating presence of AI interlocutor affording cognitive flow of language output

The cognitive responsiveness of the AI chatbot appears to be a primary reason for the reduction of cognitive paralysis arising during face-to-face English communication. Illustrations of personified AI interlocutors as speaking facilitators are frequently present in the visual data. The relaxed postures and expressions of the human speaker, in addition to the absence of physiological markers of distress index the significant difference in affective quality between the two communicative contexts.

Similarly, as evidenced in Figure 10, the metaphorical alignment of the AI chatbot with the natural environment suggests that for these ESL learners, AI-mediated spoken interactions feel more organic and natural instead of as a mechanistic process. According to the participants, the AI speaking partners, unlike humans, grant space for the low-proficient English speakers to pause, reflect, correct, reformulate, and repeat utterances. Moreover, they verbalized the convenience granted to them to translanguage as facilitating language output.

Empowerment, Agency, and Willingness to Communicate

This theme highlights the participants' sense of agency and self-efficacy during dialogues with an AI English speaker. The relationship they have with the, more often than not, anthropomorphized AI is presented as one of symmetry, collaboration, and mutual recognition, a stark departure from the hierarchical, power-laden interactions with other human interlocutors.

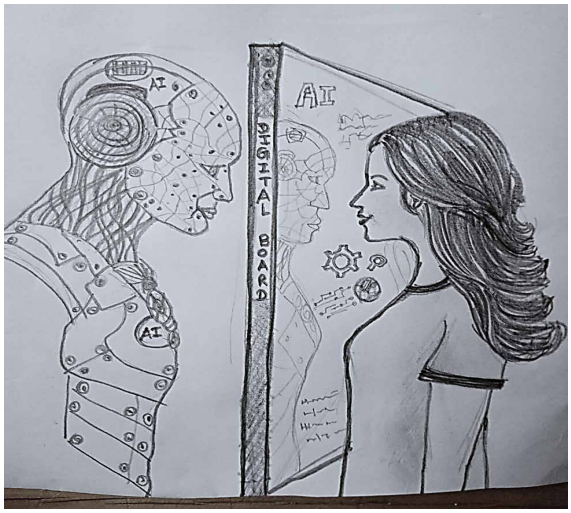

Figure 11. Egalitarian relationship between the AI interlocutor and the human interlocutee

Figure 12. AI as an omnipresent, divine entity constantly available for a language assistant

In Figure 11, direct eye contact and equal stature illustrate the compatibility of the relationship between the two speaking partners, i.e., the AI chatbot and human speaker. The absence of hierarchical cues dominating the visual narratives representing F2F spoken English interactions constitutes a restructuring of the ESL learners' self-perceptions resonant of Bandura's (1977) notion of self-efficacy enhancement through feelings of empowerment and reassurance of mastery of spoken English. As per the Self-Determination Theory, students in this study reported feelings of autonomy and competence, which support the motivational impetus to develop English.

A complex relationship register is depicted in Figure 12, whereby the AI is presented as a god-like figure. This may be due to its ubiquitousness and omnipresence, in addition to its capacity to provide unlimited and unconditional guidance during the acquisition of oral English skills. The L2 learners verbally expressed feeling empowered due to the AI speakers' availability that surpasses temporal and spatial boundaries. This inadvertently grants the L2 learners' agency as opposed to sentiments of being agentless and powerless during actual human spoken interactions.

Interestingly, some participants included cartographic imagery by drawing maps with marked destinations and navigational routes illustrating their English language learning journey when in partnership

with AI. For them, the journey is now made fathomable through the enabling presence of AI. This positive aspiration is of stark contrast to the iconography of entrapment and impossibility characterizing these ESL undergraduates' experiences during speech encounters with fluent human interlocutors, contributing to metaphorically endless journeys filled with humiliation, failure, and negative evaluation.

During the focus group discussions, a comparable notion was shared whereby participants highlighted the transformative nature of AI-mediated spoken interactions, granting self-efficacy, learner autonomy, and agency.

"Until now, I always thought that being able to speak in English fluently would always remain a dream in my life. Now that I've been introduced to ChatGPT and Gemini, I actually feel that it can become a reality one day." (Student D)

DISCUSSION

As per the study findings, for these low-proficient ESL learners in post-colonial Sri Lanka, emotional, cognitive, and metacognitive factors that come into play during communicative face-to-face exchanges with an English speaker are shaped by individually-driven factors as well as historical, sociopolitical, and economic forces affecting authentic communicative interactions. On the contrary, when initiating dialogues with voice-enabled AI chatbots, the evaluative and hierarchical structures diminish, resulting in lowered levels of L2 anxiety, reduced cognitive load, and resistance to the use of the colonizers' language. In this scenario, English is seen as a neutral tool of communication used to convey ideas in the target language as effectively as possible.

Emotional Responses to Face-to-Face and Oral English Production

Based on the analysis of the qualitative data, the affective domain covering fear, anxiety, and shame, predominates during F2F spoken interactions, whereas relief, confidence, and acceptance can be identified as the positive aspects of their verbal interactions with AI chatbots. As per Horwitz et al.'s (1986) conceptualization of Foreign Language Anxiety (FLA), the data demonstrate the presence of communication apprehension and fear of negative evaluation among the study participants. This is symbolized in the visual data related to spoken human communication in English, predominated by interlocutees with sealed lips, downcast eyes, and distraught faces. Evidently, the lack of perceived control, as conceptualized by Pekrun (2006), during human spoken interactions results in negative emotions of shame, hopelessness, and anxiety. In contrast, during verbal exchanges with AI agents, positive emotions such as relief, hope, and enjoyment denote a restored sense of control in the process of acquiring English. These findings align closely with Abdullaev et al. (2025), Madina (2025), Aly and Didah Nurhamidah (2025), Wang et al. (2024), Nguyen and Kim (2025), Fathi et al. (2024), and Teimouri et al. (2019), indicative of positive emotional engagement when speaking occurs between AI and humans as they document significant reductions in FLA in such contexts.

Moreover, Krashen's (1982) Affective Filter Hypothesis can be used to explain how emotional distress blocks oral L2 output during F2F interactions with English speakers, whereas the lowered filter during AI-mediated speech production reduces the presence of FLA among these undergraduate ESL learners. As per their visual and verbal narratives, this is because AI chatbots provide nonjudgmental learning spaces, devoid of status attribution and condescending attitudes of certain proficient English speakers.

Although most related studies substantiate these claims, a contradictory finding in a study conducted by Çakmak (2022) reports the target Turkish EFL learners experiencing increased FLA levels during chatbot-led spoken interactions. This paradoxical finding could be a result of the contextual, cultural, and technology usage differences acting as a mediating variable concerning the anxiety-reducing or increasing potential of using AI chatbots to develop spoken proficiency in the L2. Interestingly, the Chatbot named "Replica" used in Çakmak's (2022) study has been designed as a social companionship tool, and there

is a possibility that such a design feature could have contributed to increased anxiety as opposed to the neutral general-purpose AI models like ChatGPT and Gemini.

Cognitive Responses to Face-to-Face and AI-mediated Spoken Interactions

This section interprets the findings related to the cognitive dimension involved during the two methods of spoken communicative exchanges, as it considers the participants' mental functioning to identify any differences or similarities in language processing and production. As with the other major themes, based on the findings of the current study, there is an explicit dichotomy between the participants' cognitive responses to AI-mediated spoken exchanges in English as opposed to F2F spoken exchanges. The former interaction evidently lowers the affective filter, thereby minimizing cognitive overload and cognitive paralysis resulting from FLA, a phenomenon that specifically occurs when using the target language in classroom and social contexts. The resultative "brain fog" that hinders effective spoken output (Figure 6), as opposed to the molecular brain drawing in Figure 8, provides visual evidence of how the two modes of communication can either facilitate or obstruct language input and output processing. Moreover, this exemplifies how the affective filter can impact not just emotional engagement but also cognitive processing during language learning.

The findings corroborate those of Kim et al. (2025) and Stadler et al. (2024), who also consider reduced cognitive load during chatbot-based speaking assessments among low proficiency EFL learners. One main reason for this, in both studies, is identified as speech-enabled AI chatbots' ability to grant L2 learners time for reflective pausing often not available during spontaneous human spoken exchanges. In addition, AI speech assistants can be patient, provide explicit feedback, and give non-intimidating error correction, evidenced in the exalted expressions and relaxed postures in the participants' visual narratives symbolizing the mentally receptive state of beginner L2 users.

According to Bandura's (1977) self-efficacy theory, the understanding here would be that low self-efficacy in the F2F context has resulted in cognitive avoidance and the inability to retrieve and produce language in this high-pressure and high-stakes interaction. Moreover, the cartographic images in the visual corpus indicate metaphorical journeys towards a destination, i.e., that of English proficiency, providing direct visual evidence of self-efficacy restoration in AI interactions. Similarly, drawing on Pekrun's (2006) Control Value Theory, we can account for how these students' low perceived control during oral exchanges with proficient English speakers results in interference to the working memory connecting their claims of "going blank' and having "overcrowded thoughts." To anchor the claim regarding cognitive performance in L2 contexts, similar findings (Stadler et al. 2024, Wang & Meng, 2026, Ahmadi, 2025) can be referred to, whereby references to enhanced language output and cognitive facilitation provide an empirical basis.

Metacognitive Responses to Face-to-Face and AI-mediated Spoken Interactions

To fully comprehend how these low-proficient ESL learners conceptualize their L2 speaking development in the two different contexts involving human versus AI chatbot-mediated spoken interactions, metacognitive processes also play a crucial role in this estimation. In the context of this study, metacognition is understood as learners' awareness and regulation of their own emotional and cognitive processes during the acquisition of the target language. It involves their capacity to monitor, evaluate, and strategically adjust their use of English, reflecting their potential for self-regulated learning that supports autonomous L2 development.

When interpreting the findings related to participants' metacognitive responses across the two contexts, Self-Determination Theory (Deci & Ryan, 1985) provides a useful analytical lens, particularly through its emphasis on learners' basic psychological needs of autonomy, competence, and relatedness. Tabulation of the visual and verbal data clearly indicates that these needs are more readily fulfilled in AI-mediated communication contexts. Participants frequently reported the ability to pause, reformulate,

clarify uncertainties, and engage in translanguaging practices, all of which suggest enhanced autonomy, a form of mediated relatedness, and a developing sense of competence. Conversely, in F2F spoken interactions in the target language, they seem to feel disempowered, incompetent, and relationally unsafe. As such, learner autonomy operates as a foundation for the development of self-efficacy, which in turn supports self-regulation among these ESL learners. This interpretation is consistent with Bandura's (1977) Self-Efficacy Theory. Similar to this study's finding, Abdullaev et al. (2025) found that Uzbekistani EFL learners felt that AI facilitates their reflective practices before production through the space it provides for self-correction. Moreover, Lim et al. (2026) and Stadler et al. (2024) found that AI enhanced learner motivation and confidence relating to the metacognitive dimension of self-regulation.

Subsequently, it can be argued that the participants' expression of confidence during AI-mediated speech interactions facilitates the deployment of self-regulatory strategies such as planning, monitoring, and evaluation. This suggests a heightened sense of control (Pekrun, 2006) over their willingness to communicate. Conversely, their metacognitive responses to human interactions indicate a tendency toward diminished self-regulation, as reflected in their reported silence, unease, anxiety, and disengagement. This pattern aligns with the findings of Wang et al. (2026) and Abdallah (2025), who similarly demonstrate that AI chatbot interventions foster EFL learners' self-regulation and metacognitive engagement.

Despite the overwhelming positive responses to AI as a potential partner for developing oral English proficiency among these low-proficient English users, several studies have presented findings that raise concerns regarding AI in language learning. Hamid Aly and Nushamidha (2025) refer to the phrase "Technological pressure" as being a sentiment experienced by the participants in their study, whereby the metacognitive aspect is the conscious reflection of the learners regarding the worrisome nature of figuring out the credibility of the AI-generated output and the feeling of being overwhelmed with the information received. Moreover, it reflects the absence of emotional responsiveness in human interactions, which a mechanized system cannot grant. Nonetheless, as per the visual and verbal data gathered in this study, technostress (Brod, 1984), summed up as the inability of a person to cope with new information communication technologies, is only minimally present. This is because these learners have already been exposed to ICTs extensively during the COVID-19 pandemic, specifically for educational purposes, and subsequently to Artificial Intelligence during their compulsory Information Technology module, as well as during their English module, whereby they were trained to use AI platforms for developing their language skills. As such, the study participants mostly indicated that they were comfortable using AI and that it was not the same experience as F2F communication with an English speaker, which caused negative feelings that were not welcomed by the learners.

At a deeper level, the students appear metacognitively aware of their reticence to speak English with human interlocutors as something that was not ingrained in their individualities but indicative of larger, complex, and multilayered sociopolitical and linguistic matters. In fact, these learners' struggles in speaking English extend beyond emotional and cognitive factors, reflecting a legacy of linguistic subordination shaped by the enduring influence of British colonialism in Sri Lanka. From the lens of Bourdieu's (1991) conceptualization of symbolic power and language capital, the intensity of negative emotions arising during F2F spoken English encounters can be better understood. These fear- and anxiety-laden sentiments are not merely person-specific traits but are historically and socially constituted. Participants' visual and verbal narratives further reflect this, depicting asymmetrical power relations between speakers through markers such as sartorial symbolism and the attribution of exaggerated or even demonic traits to English speakers, as well as lowered gazes and distraught facial and bodily expressions.

Unlike in Çakmak's (2025) study, where students experienced increased anxiety levels during AI interactions, participants of the present study responded differently. One possible explanation for this divergence lies in the context-specific nature of FLA, particularly within postcolonial sociolinguistic settings marked by power imbalances and unequal distributions of linguistic capital in human-mediated

interactions. In such contexts, human interlocutors may embody evaluative authority and social hierarchy, thereby intensifying L2 anxiety. Conversely, AI functions as a non-human, non-judgmental interlocutor, offering a neutral platform for spoken exchanges that can reduce perceived social threat and facilitate stress-free oral communication in English.

This premise is further validated through Rambukwella's (2018) and Parakrama's (1995) insistence that the development and use of the English language brings with it connotations of historically and ideologically laden power dynamics, reinforcing colonial structures, class hierarchies, and linguistic dominance within educational and social contexts. This leads to what Bourdieu (1991) refers to as linguistic capital, whereby, in the context of Sri Lanka, English is not merely a communicative tool but also a type of capital permeating social, educational, economic, and political spheres, resulting in the positionality of the non-English users as those belonging to a lesser stature.

Based on the data gathered, the undergraduates of this study appear to be aware of symbolic power and linguistic capital, revealing a sophisticated understanding of the structural and historical dimensions of their linguistic marginalization. Their metacognitive awareness of the impact of the Swa Basha Policy, the imperial legacy of English, and the structurally inegalitarian distribution of linguistic capital in the Sri Lankan educational landscape is reflected in the symbolic expressionism in their drawings, and their explicit renderings of the consequences of the injustices that have rendered them literally speechless to use such a socially loaded and problematic tool; i.e., English. As per the findings, it is evident that Parakrama's (1995) premise regarding the insistence of standard English as a gatekeeping mechanism preventing low or non-English users from accessing the same privileges as their English speaking counterparts, and Rambukwella's (2018) identification of the divided subjectivities of the postcolonial Sri Lankan English users are further substantiated since these abstract claims are validated in the lived realities of the participants of this study.

The study findings should be generalized with caution as it is limited to a single institutional context with the participants being visual arts undergraduates; however, given that the sample belongs to beginner-level ESL learners situated within the sociopolitical, postcolonial scenario, they may be applied to comparable contexts. In addition, the study is a cross-sectional study and does not engage in a longitudinal evaluation to monitor perception changes over time.

CONCLUSION

This study explored the lived emotional, cognitive, and metacognitive experiences of a group of low-proficiency ESL learners during face-to-face and AI-mediated oral communication in English. The main findings of the study are explicit in the sense that oral English speaking with human interlocutors results in psychological agitation, cognitive overload, feelings of marginalization, and unwarranted exposure to criticism shaped by the postcolonial politics of linguistic capital in the country. Paradoxically, speaking in English with voice-enabled AI chatbots is both described and visualized as a psychologically safe space, diminishing anxiety and fear of being negatively judged, resulting in augmented self-efficacy and willingness to communicate in the target language.

The most significant aspect of this research is that, unlike the existing literature, which tends to focus on the emotional and cognitive effects of human versus AI-agent spoken interactions in an English as a second language context, it moves a step further by advancing the theoretical underpinning of FLA as being impacted by sociopolitical factors and linguistic realities in postcolonial contexts. When acquiring the colonizers' language carries with it the adverse effects of colonialism, unjust social stratification, linguistic marginalization, and symbolic violence. Therefore, a study on the development of English speaking skills cannot be limited to individualized emotional and cognitive factors; rather, it should focus on the dynamics of linguistic capital and symbolic power, especially in the context of ESL learners situated in a postcolonial landscape. Within this communicative space, these Sri Lankan ESL learners

harness AI as a collaborative partner in their second language acquisition journey, a journey they perceive as having a final destination through sustained interaction with such technologies.

Future studies can also extend their scope to encapsulate the problematic sociopolitical and linguistic dimensions of language anxiety in comparable contexts. They should also move beyond the cognitive to the metacognitive levels of investigation. More studies should consider phenomenologically oriented multimodality as an effective approach to conducting studies that consider affective and cognitive expressions. Moreover, expanding on the positive findings to delve further into the complex dynamics and ethical dimensions of the dynamics that AI usage affords in second language acquisition, including data privacy issues, AI misuse, mechanization, misinformation, and human alienation.

Based on the study implications, it is evident that the use of AI as language practice partners could be particularly beneficial for low-proficiency L2 learners. In fact, AI platforms can serve as empathetic interlocutors, offering safe spaces for expression and confidence-building. As such, administrators and educators should consider structured integration of such technologies into formal curricula. Educators should also be aware of balancing AI tools, as authentic and effective communication involving human participants should be the ultimate goal.

ACKNOWLEDGEMENTS

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

DECLARATION OF COMPETING INTERESTS

The author declares that she has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

Abdallah, M. M. S. (2025). *How conversational AI chatbots support and reinforce self-regulated language learning (SRLL)*. ERIC. https://files.eric.ed.gov/fulltext/ED673689.pdf

Abdullaev, D., Abdazimov, A., Khusainova, Z., Makhmudov, S., & Sagdullaeva, D. (2025). A phenomenological study on the impact of personalized AI feedback on writing quality: Insights from EFL students in Uzbekistan. *Computer-Assisted Language Learning Electronic Journal: Special Issue AIIDLE, 26*(7), 162–186. https://callej.org/index.php/journal/article/view/816/501

Ahmadi, R. (2025). A qualitative investigation of ChatGPT's impact on grammar proficiency in Iranian EFL students' academic writing. *Applied Linguistics Compass, 1*(3), 65–74. https://doi.org/10.22034/alc.2026.570346.1029

Aly, A. H., & Nurhamidah, D. (2025). Exploring the impact of artificial intelligence on foreign language anxiety in language learning (Indonesia and English). *Ndrumi: Jurnal Ilmu Pendidikan Dan Humaniora, 8*(1), 28–45. https://doi.org/10.57094/ndrumi.v8i1.2907

Anh, N. H. (2024). Harnessing AI-based tools for enhancing English speaking proficiency: Impacts, challenges, and long-term engagement. *International Journal of AI in Language Education, 1*(2), 18–29. https://doi.org/10.54855/IJAILE.24122

Azman, M. A. B., & Ibrahim, A. B. (2025). Investigating AI writing tools in enhancing academic writing skills of Diploma in TESL students of a private college in Malaysia: The moderating role of students' perceptions. *International Journal of Academic Research in Progressive Education and Development, 14*(4), 1307–1314. https://doi.org/10.6007/IJARPED/v14-i4/26844

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bourdieu, P. (1991). *Language and symbolic power* (G. Raymond & M. Adamson, Trans.). Polity Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brod, C. (1984). *Technostress: The human cost of the computer revolution*. Addison-Wesley.

Broom, C. (2015). Empowering students: Pedagogy that benefits educators and learners. *Citizenship, Social and Economics Education, 14*(2), 79–86. https://doi.org/10.1177/2047173415597142

Butarbutar, R. (2024). *The function of artificial intelligence (AI) in developing English language learner's speaking performance* [Doctoral dissertation or Master's thesis, Universitas Sarjana Wiyata Taman Siswa].

Çakmak, F. (2022). Chatbot-human interaction and its effects on EFL students' L2 speaking performance and anxiety. *Novitas-ROYAL (Research on Youth and Language), 16*(2), 113–131. https://doi.org/10.3969/j.issn.1307-4733.2022.16.002

Celik, B. (2025). Using ChatGPT as a virtual speaking tutor to boost EFL learners' speaking self-efficacy. *Australian Journal of Applied Linguistics, 8*(1). https://doi.org/10.29140/ajal.v8n1.102418

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.

Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System, 121*, Article 103254. https://doi.org/10.1016/j.system.2024.103254

Hạnh, N. T. T. (2024). Students' attitudes toward utilizing AI-based technologies to improve their speaking proficiency. *Ho Chi Minh City University of Education Journal of Science, 21*(5). https://doi.org/10.54607/HCMUE.JS.21.5.4270

Horwitz, E. K., Horwitz, M. B., & Cope, J. A. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132. http://dx.doi.org/10.2307/327317

Huriati, N. (2023). The role of artificial intelligence (AI) in developing English language learner's communication skills. *Journal on Education, 6*(1). https://doi.org/10.31004/JOE.V6I1.2990

Kang, H. (2022). Effects of artificial intelligence (AI) and native speaker interlocutors on ESL learners' speaking ability and affective aspects. *Multimedia-Assisted Language Learning, 25*(2), 9–43. https://doi.org/10.15702/MALL.2022.25.2.09

Kim, J. (2025). Voice interaction with conversational AI could facilitate thoughtful reflection and substantive revision in writing. *arXiv*. https://arxiv.org/abs/2504.08687

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

MacIntyre, P. D., & Gardner, R. C. (1991a). Language anxiety: Its relationship to other anxieties and to processing in native and second languages. *Language Learning, 41*(4), 513–539. https://doi.org/10.1111/j.1467-1770.1991.tb00691.x

MacIntyre, P. D., & Gardner, R. C. (1991b). Methods and results in the study of anxiety and language learning: A review of the literature. *Language Learning, 41*(1), 85–117. https://doi.org/10.1111/j.1467-1770.1991.tb00677.x

Madina, G. (2025). Emotional engagement in AI-supported online English learning: A mixed-methods study among Indonesian EFL students. *Advances in Education Journal, 2*(3), 814–823. https://journal.al-afif.org/index.php/aej/article/view/329

Nguyen, T. T., & Kim, J. (2025). Feasibility of administering chatbot-based speaking assessment for low-level EFL students. *Language Learning & Technology, 29*(1), 1–26. https://doi.org/10.64152/10125/73660

Parakrama, A. (1995). *De-hegemonizing language standards: Learning from (post)colonial Englishes about "English"*. Palgrave Macmillan. https://doi.org/10.1057/9780230371309

Patiño, A. E. C. (2024). The improvement of oral communicative competence in English through the

artificial intelligence: Desarrollo de la competencia comunicativa oral en inglés mediada por la inteligencia artificial. *Latam, 5*(1), Article 1850. https://doi.org/10.56712/LATAM.V5I1.1850

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Rambukwella, H. (2018). *The politics and poetics of authenticity: A cultural genealogy of Sinhala nationalism*. UCL Press. https://doi.org/10.14324/111.9781787351288

Sawalmeh, M. H. (2023). Globalization and the increasing demand for spoken English teachers. *Research Journal in Advanced Humanities, 4*(2). https://doi.org/10.58256/rjah.v4i2.1097

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior, 160*, Article 108386. https://doi.org/10.1016/j.chb.2024.108386

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition, 41*(2), 363–387. https://doi.org/10.1017/S0272263118000311

Thompson, J. B. (1991). Editor's introduction. In P. Bourdieu, *Language and symbolic power* (G. Raymond & M. Adamson, Trans.; pp. 1–31). Harvard University Press.

Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: An analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System, 127*, Article 103533. https://doi.org/10.1016/j.system.2024.103533

Wang, F., Zhou, X., Li, K., Cheung, A. C. K., & Tian, M. (2025). The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English. *Interactive Learning Environments, 33*(7), 4633–4652. https://doi.org/10.1080/10494820.2025.2470319

Wang, X., & Meng, Y. (2026). AI chatbots and EFL learners' risk-taking in oral communication: A scoping review. *Eduvelop: Journal of English Education and Development, 9*(1), 144–166. https://doi.org/10.55150/v9i1.6112

Wibowo, A. P., & Sumarno, A. (2026). Enhancing students' speaking self-efficacy through an AI-integrated audio-lingual method in English pronunciation learning. *Journal of General Education and Humanities, 5*(1). https://journal-gehu.com/index.php/gehu/article/view/982

Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal, 37*(3), 308–328. https://doi.org/10.1177/0033688206071315

Yildiz, C. (2024). A path to enhanced speaking self-efficacy: The effect of ChatGPT on EFL learners. *English Language Education Journal*.

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