Scientific Research Publishing

Open Journal of Social Sciences, 2026, 14(11), 404-415
https://www.scirp.org/journal/jss
ISSN Online: 2327-5960
ISSN Print: 2327-5952

AI Speaking Partner's Effect on College English Learners' Speaking Anxiety and Willingness to Communicate: A Mixed-Methods Study

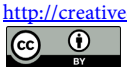
Sha Yang

School of Foreign Languages, Guizhou University of Finance and Economics, Guiyang, China
Email: 916964120@qq.com

How to cite this paper: Yang, S. (2026). AI Speaking Partner's Effect on College English Learners' Speaking Anxiety and Willingness to Communicate: A Mixed-Methods Study. *Open Journal of Social Sciences, 14,* 404-415.
https://doi.org/10.4236/jss.2026.1411022

Received: October 21, 2026
Accepted: November 20, 2026
Published: November 23, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/



Open Access

Abstract

Speaking apprehension and poor willingness-to-communicate (WTC) continue to pose challenges for college students' oral English proficiency in China, where classroom speaking tends to be perceived as face threatening. Artificial intelligence (AI) speaking partners—conversational agents that allow learners to practise spoken English without a human interlocutor—have been proposed as a low-anxiety alternative to classroom interaction, but the empirical evidence on their affective effects is still limited and mixed. This convergent mixed methods study investigated changes in speaking anxiety and willingness to communicate associated with an 8-week AI speaking partner intervention among freshman college ELLs. A total of 48 participants completed adapted versions of the Foreign Language Classroom Anxiety Scale and a willingness-to-communicate scale immediately before and after the intervention, and 12 of them took part in semi-structured interviews. Paired-samples t-tests showed that participation in the intervention was associated with significantly lower speaking anxiety, $t(47) = 7.26$, $p < 0.001$, Cohen's $d = 1.05$, and significantly higher willingness to communicate, $t(47) = -6.81$, $p < 0.001$, Cohen's $d = 0.98$. Thematic coding of the interview data revealed three ways in which participants described the AI partner as operating: being a nonjudgmental presence, eliminating audience-related fear, and providing iterative low-stakes practice that developed communicative confidence. These results indicate that AI speaking partners could be used to supplement, but not replace, interpersonal contact to foster the affective preparedness underlying spoken language production. Implications for pedagogy, the design of AI language tools, and directions for future research are discussed.

Keywords

AI Speaking Partner, Speaking Anxiety, Willingness to Communicate, EFL,

College English, Mixed Methods Study

1. Introduction

Oral communication has been recognized as one of the most stressful skills in foreign language learning, and it is perceived by many Chinese EFL learners as a source of distress rather than of development. Horwitz, Horwitz, and Cope's (1986) concept of foreign language classroom anxiety reflects a sense of stress, worry, and/or anxiety specific to the use of a non-native language, and decades of research have validated its consistently detrimental relation to linguistic performance. Among these constructs, speaking apprehension appears to be of particular significance, since the oral mode of expression is ephemeral, public, and difficult to control or revise being spoken. The possibility for students who fear negative evaluation speaking in front of their peers may provoke an escape reaction that deprives them of what oral ability requires most.

This affective obstacle is closely connected to the notion of willingness to communicate (WTC), which MacIntyre, Dörnyei, Clément, and Noels (1998) define as a readiness to enter into discourse at a particular time with a specific person or persons using a second language. Originally conceptualized as a trait-like predisposition in first-language communication (McCroskey & Baer, 1985), WTC was extended to the second-language context and modeled as the outcome of interacting enduring and situational factors, with state anxiety and perceived communicative competence operating as the most immediate determinants of a learner's decision to speak (MacIntyre et al., 1998). Because learners who are anxious are less willing to initiate and sustain communication, the two constructs jointly constrain oral development: anxiety suppresses WTC, and low WTC deprives learners of the practice that would build confidence and reduce anxiety. In Chinese tertiary contexts, where English classes can be large and opportunities for individual speaking are limited, this cycle is difficult to break, and learners frequently report that they understand the target language well but are simply afraid to speak.

The recent advances in artificial intelligence offer an exciting new perspective on addressing this problem through AI conversation agents—spoken dialogue systems, usually implemented using an LLM and speech technology, through which the learner interacts in spoken English. Dialogue systems for computer-assisted language learning date back some time, and surveys of the field have asserted that automatic dialogue partners are capable of providing rich, patient, and non-judgemental interaction that might reduce the affective barriers to speaking (Bibauw et al., 2022). A systematic review of AI chatbots in university EFL contexts has suggested that they have potential to support a range of interactional competences but also found unevenness in the literature and noted that pedagogical design is far from clear (Zhai & Wibowo, 2023).

Only few empirical papers that have studied the emotional effects of chatbot AI

returned inconsistent but encouraging results. Tai and Chen (2020), for instance, reported that interacting with a smart personal assistant promoted the EFL learners' willingness to communicate, increased their communication self-confidence, reduced their speaking anxiety, and was perceived by learners as less threatening. Zheng, Zhou, Chen, and Ye (2025) found that LLM chatbots used as collaborative dialogue partners significantly improved Chinese first-year college learners' speaking skills, increased their willingness to speak and self-efficacy, and considerably reduced their foreign language anxiety. Wu, Hapsari, and Huang (2025) demonstrated that incorporating AI chatbots into think-pair-share activities alleviated speaking anxiety, enhanced language enjoyment, and improved speaking performance among first-year Indonesian university students. Voice-based chatbots have likewise been associated with greater oral fluency and lower speaking anxiety (Azizimajd, 2023), and chatbot-mediated interaction has been linked to increased willingness to communicate and communicative confidence in young EFL learners (Guo & Yan, 2025). In a survey, perceived usability and sense of presence of generative AI chatbots reported by Chinese university students were positively correlated with a stronger growth mindset toward language ability, with willingness to communicate as the mediating variable. (Huang et al., 2025).

The situation is less encouraging in other respects. Ballıdağ and Aydın (2025), contrasting AI-powered chatbot tasks with conventional peer interactions for Turkish preparatory university students, found that the chatbot tasks had no significant effect on reducing speaking anxiety, while speaking to peers did. This counter-finding matters because it warns us not to assume that every AI interlocutor will be anxiolytic and suggests that both the design of the interaction and the broader pedagogical context are consequential. In addition, much of the prior work is purely quantitative (based on pre- and post-intervention questionnaires) and thus provides little understanding of how AI speaking partners affect learners' affective state. Qualitative and mixed-methods evidence that might shed light on why, how, and under what circumstances these tools help or fail to help remains limited, especially among first-year College English students in China, who face the challenge of adapting to a new mode of university teaching and therefore constitute a special research population.

This study fills the gap by using a convergent mixed methods design that combines quantitative measures of speaking anxiety and WTC with qualitative interviews. The research questions are as follows: (1) To what extent is an eight-week AI speaking partner intervention associated with changes in college English learners' speaking anxiety and willingness to communicate? (2) How do learners make sense of changes in their speaking anxiety and willingness to communicate when using an AI speaking partner? By combining these two strands, this work seeks not only to estimate how large any given affective change is but also to explain what is happening in that process, thereby providing a more comprehensive picture of the pedagogical potential of AI speaking partners than either strand alone.

2. Literature Review

2.1. Speaking Anxiety in EFL

According to Horwitz et al. (1986), foreign language classroom anxiety (FLCA) is a unique, situational form of anxiety comprising three interrelated dimensions: communication apprehension, fear of negative evaluation, and test anxiety. Their Foreign Language Classroom Anxiety Scale (FLCAS) is now the most widely used instrument in the field, and later factor-analytic work has consistently verified that communication apprehension and fear of negative evaluation are especially salient for speaking. The consequences of speaking anxiety have also been well documented: anxious learners speak less, take longer to respond, produce less fluent output, and are more likely to avoid opportunities for oral interaction altogether. In China, research has demonstrated that foreign language anxiety is negatively correlated with achievement and becomes stronger at more advanced educational stages, suggesting that negative emotions do not decrease as students progress through school.

Speaking anxiety is theoretically important because it interacts dynamically with performance and practice. Anxiety does not simply accompany poor performance; it can contribute to its development by disrupting the cognitive resources needed for producing speech and by preventing the repetition of utterances necessary to build automatism. In the case of students preparing for university entrance—especially those who have been taught English in school and for whom spoken language is assumed to have already been mastered—there is often a lingering dread of speaking that prevents them from converting receptive competence into productive confidence.

2.2. Willingness to Communicate

Willingness to communicate originated as a first-language construct, defined by McCroskey and Baer (1985) as the probability of initiating communication when free to do so, and was conceived as a stable, trait-like predisposition influenced by personality. MacIntyre et al., and colleagues (1998) re-situated the construct within second-language acquisition, arguing that WTC should be treated as the final, most immediate precursor to actual L2 communication and proposing a pyramid model in which enduring influences, such as personality, inter-group climate, and communicative self-confidence, support situational influences that culminate in a momentary decision to speak. In this model, state anxiety and self-perceived competence are the proximal determinants of communication confidence, and anything that increases anxiety reduces confidence and, in turn, willingness to communicate.

This theoretical framing makes the connection between anxiety and WTC explicit and gives it practical significance. Because WTC is the behavioral gateway to communication, educational interventions that aim to foster oral proficiency must, at some point, address the affective conditions that allow learners to choose to speak. For the present study, the pyramid model provides a useful lens: if AI

speaking partners reduce state anxiety and enhance perceived competence in the speaking situation, they should, in principle, raise WTC. Whether they do so in practice is precisely the empirical question examined here.

2.3. AI Speaking Partners and Affective Outcomes

The theory behind AI speaking partners rests on the argument that automatic interlocutors remove most of what makes classroom speaking anxiety-inducing. Students can talk without fearing judgment by classmates or the presence of a real audience, and without the shame of making errors that peers might observe. Such arguments are perennial in dialog-based CALL, and recent reviews suggest that conversational agents can provide the quantity of conversation, the patience, and the low-stakes feedback that many classroom settings cannot (Bibauw et al., 2022; Zhai & Wibowo, 2023).

The evidence so far, while increasing, is mixed. On the one hand, Tai and Chen (2020) found that interacting with a smart personal assistant promoted willingness to communicate, higher levels of confidence, and lower speaking anxiety among adolescent EFL learners, who explicitly indicated that the tool was not an intimidating environment. Zheng et al. (2025) found that LLM chatbots used as collaborative dialogue partners enhanced Chinese college freshmen's oral ability and decreased their foreign language anxiety. Wu et al. (2025) found that incorporating AI chatbots into think-pair-share activities lowered speaking anxiety and enhanced speaking achievement among first-year students. Voice-based chatbots have been associated with improved oral fluency and lower speaking anxiety (Azizimajd, 2023). Chatbot interactions have increased willingness to communicate in young learners (Guo & Yan, 2025), and perceived usability and presence of generative AI chatbots are related to a stronger growth mindset via willingness to communicate (Huang et al., 2025).

In contrast, Ballıdağ and Aydın (2025) did not observe a decrease in speaking anxiety as a result of the AI chatbot task but did observe lower anxiety levels following peer interaction. This juxtaposition indicates that an AI partner's affective value is contingent rather than intrinsic, depending on factors such as learners' proficiency, the quality of the conversational experience, task design, and the extent to which the tool is embedded in a supportive pedagogical context. This highlights the need for research that moves beyond aggregated questionnaire scores and examines the mechanisms as perceived by learners themselves.

2.4. Research Gap

Two limitations can be figured out in the existing literature. First, most studies are single-method and quantitative, measuring changes in anxiety and WTC without capturing the processes behind those changes; as a result, the mechanisms by which AI speaking partners exert their effects remain largely inferred rather than demonstrated. Second, although several studies involve Chinese learners, few focus specifically on first-year college English students in their second semester, a

population in the midst of the transition to university-level instruction and therefore of particular interest to language educators. The present study addresses both limitations through a convergent mixed methods design that combines a pre- and post-intervention quantitative comparison with semi-structured interviews, thereby attending to both the magnitude and the meaning of any observed change.

3. Theoretical Framework

The study is organized around the pyramid model of L2 willingness to communicate proposed by MacIntyre et al., and colleagues (1998). The model holds that the decision to speak is determined at the most immediate level by a desire to communicate with a specific interlocutor and by state communicative self-confidence, which in turn reflects the interplay of state anxiety and self-perceived competence. Because the model locates anxiety and perceived competence as the proximal influences on WTC, it provides a natural framework for predicting and interpreting the effects of an AI speaking partner. If such a partner reduces the audience-related and evaluation-related sources of state anxiety, and if repeated successful interaction raises self-perceived competence, the model predicts an increase in communication confidence and, consequently, in WTC.

The framework is also suited to accommodate the qualitative strand of the study. The interview data can be read in terms of the elements of the model, including the desire to speak, anxiety, perceived competence, and situation-specific characteristics of the conversation, allowing the mechanisms reported by learners to be mapped onto theoretical constructs. This mapping helps us understand not just whether participation in the intervention was associated with change, but how, and provides a theoretically grounded basis on which to discuss pedagogical implications. Thus, the theory is used to bring together the quantitative and qualitative strands of a convergent design, not simply as motivation for hypotheses.

4. Methodology

4.1. Research Design

A mixed methods design was used in which quantitative and qualitative data were collected within the same study period and analyzed independently, with integration occurring at the point of interpretation (Creswell & Plano Clark, 2018). The quantitative strand employed a one-group pre-test and post-test design to estimate changes associated with participation in the intervention, while the qualitative strand used semi-structured interviews, conducted after the post-test, to explore learners' perceptions of the mechanisms at play. Although the two data sets were not collected simultaneously, the design is described as convergent because the qualitative and quantitative strands were developed to answer complementary questions about the same phenomenon and were merged during the interpretation phase, in accordance with recommended practice for mixed methods research in applied linguistics (Dörnyei, 2007).

4.2. Participants and Context

Participants were 48 first-year undergraduate students enrolled in the second semester of a compulsory college English course at a university in southern China. Participants were recruited through convenience sampling from three intact classes and gave informed consent before the study began. Their ages ranged from 18 to 21 years (M = 19.27, SD = 0.82), and 60.4% were female. All participants reported seven to nine years of prior English study and self-assessed their English proficiency as intermediate. Ethical approval for the study was obtained from the University Research Ethics Committee, and participation was voluntary, with no effect on course grades.

4.3. Instruments

Three instruments were employed. First, speaking anxiety was assessed with an adapted 12-item version of the Foreign Language Classroom Anxiety Scale (Horwitz et al., 1986) that focused on speaking-specific situations (e.g., "I start to panic when I have to speak English without preparation"). The scale was translated into Chinese following a translation and back-translation procedure and reviewed by two researchers for content validity. Items were rated on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree), yielding a possible total score ranging from 12 to 60. The adapted scale showed acceptable internal consistency in the present sample (Cronbach's alpha = 0.84).

Second, willingness to communicate in English was measured with a 10-item scale adapted from established willingness-to-communicate measures (McCroskey & Baer, 1985), in which respondents rated how willing they would be to initiate or sustain spoken communication in a range of situations, such as speaking to a stranger, speaking in a small group, or speaking in class. The scale was also administered in Chinese after the same translation and back-translation procedure, with items adapted to the college English classroom context. Responses were given on a five-point Likert scale (1 = definitely unwilling, 5 = definitely willing). Cronbach's alpha for the present sample was 0.81.

Third, a semi-structured interview protocol was developed to explore learners' experiences of the AI speaking partner, including their feelings about speaking with the tool, any changes in anxiety or confidence, and their views on how the experience compared with speaking to a human. The protocol comprised 8 open questions with follow-up probes.

4.4. Procedure

The intervention ran for eight consecutive weeks. During the first week, participants were introduced to the AI speaking partner, iTest, a conversational agent that allowed them to hold spoken English dialogues, and were asked to complete two speaking sessions per week, each of approximately 15 minutes, on topics linked to their course themes. iTest is a web-based speaking practice platform that relies on automatic speech recognition, large language models, and text-to-speech technology to generate context-appropriate responses and it provided immediate,

non-evaluative feedback on pronunciation and grammar [Author to specify the developer or provider, system version, and feedback settings]. A representative task was a two-minute open-ended conversation on a course theme, such as describing a personal experience or responding to a situational prompt. To monitor completion of the prescribed sessions, the platform recorded each session's date, duration, and topic in a usage log that the instructor reviewed weekly. The quantitative instruments were administered in class immediately before the intervention (pre-test) and again in the final week (post-test). Following the post-test, 12 participants were invited to take part in semi-structured interviews, which were conducted individually in a quiet room, audio-recorded with consent, and later transcribed verbatim. Each interview lasted approximately 22 minutes.

4.5. Data Analysis

Quantitative data were analyzed with SPSS (version 28). Paired-samples t-tests were used to compare pre-test and post-test scores for speaking anxiety and WTC, with effect sizes reported as Cohen's d. Because the paired-samples t-test assumes that the pre-post difference scores are approximately normally distributed and free of extreme outliers, these assumptions were checked before the analyses. Shapiro-Wilk tests on the difference scores indicated no significant deviation from normality ($p > 0.05$), and inspection of boxplots revealed no extreme outliers in either difference-score distribution.

The qualitative data were analyzed with thematic analysis following the six-phase procedure described by Braun and Clarke (2006). Two researchers independently coded a subset (25%) of the transcripts, discussed and reconciled discrepancies, and then coded all data. Inter-coder reliability (Cohen's kappa) was 0.83. Themes were identified on a semantic level and checked against the theoretical framework, with particular attention to learners' descriptions of the mechanisms underpinning any perceived change. The two strands were brought together during the interpretation phase, where the qualitative themes served as an explanation and context for the quantitative results.

5. Results

5.1. Quantitative Findings

Descriptive statistics for speaking anxiety and WTC before and after the intervention are presented in Table 1.

Table 1. Descriptive statistics for speaking anxiety and willingness to communicate before and after the intervention (N = 48).

Variable	Pre-test *M* (*SD*)	Post-test *M* (*SD*)	Mean difference
Speaking anxiety	38.71 (5.24)	31.44 (4.87)	7.27
Willingness to communicate	26.35 (4.18)	32.08 (4.43)	−5.73

Note: *M* = mean; *SD* = standard deviation. Higher anxiety scores represent greater speaking anxiety; higher WTC scores represent greater willingness to communicate.

The first paired-samples t-test showed that speaking anxiety was significantly lower after the intervention than before, $t(47) = 7.26$, $p < 0.001$, Cohen's $d = 1.05$, suggesting that students' speaking anxiety declined following participation. The second paired-samples t-test indicated a significant increase in willingness to communicate, $t(47) = -6.81$, $p < 0.001$, Cohen's $d = 0.98$, meaning that participants were more willing to speak in English after the intervention. Both effect sizes are conventionally considered large.

5.2. Qualitative Findings

The thematic analysis on the interview data produced three themes which provided insight into how this quantitative change occurred.

The first theme concerned the nonjudgmental nature of the interaction. Participants repeatedly contrasted the AI partner with human interlocutors, emphasizing that they did not fear being laughed at or criticized when they made mistakes. One participant observed that "there is no one to judge you, so you do not feel embarrassed when you cannot find the right word" [Participant 07]. This perception of safety appeared to lower the evaluation related anxiety that the literature identifies as a core component of speaking apprehension.

The second topic involved alleviating the fear regarding the audience. When speaking with a computer, they said, this interaction removed the stress of being seen by classmates, which was the most stressful part of classroom speaking. Without an audience, they could concentrate on what they wanted to say without worrying about how they would be perceived. This result is in line with an opinion that it is not usually the language but the audience which is responsible for speaker's nervousness.

Theme three was related to repeated low-stakes practice and increasing confidence: participants described how repeating, low-stakes interactions with the AI partner allowed them to practise much more frequently than they could during class and that this process also made speaking feel more natural over time. Some participants even described a transfer of confidence from the AI conversation to other situations, reporting feeling more willing to volunteer to speak in class. One user mentioned that "after practising with the machine many times, I do not feel as frightened of speaking English, and I am more willing to answer the teacher's questions." [User 11]

The three themes map directly onto the theoretical framework: the nonjudgmental and audience-free nature of the interaction appears to reduce state anxiety, while repeated successful practice appears to raise perceived competence; both, in turn, support the increase in WTC observed in the quantitative data.

6. Discussion

Taken together, these results indicate that the participation in the eight-week AI speaking partner intervention was associated with significantly lower speaking anxiety and significantly higher willingness to communicate among a sample of

first-year university EFL learners. These findings are consistent with studies that have reported positive affective consequences of using an AI chatbot as a conversational partner (Azizimajd, 2023; Guo & Yan, 2025; Tai & Chen, 2020; Wu et al., 2025; Zheng et al., 2025), and they extend this work by providing qualitative evidence of how such changes are perceived by learners.

The three interview-derived themes—non-judgemental interaction, elimination of audience anxiety, and iterative low-risk practice—correspond well with the more immediate antecedents of willingness to communicate from the pyramid model, namely state anxiety and perceived competence (MacIntyre et al., 1998). We interpret the two properties of no audience and no judgment in terms of their effect on state anxiety: they reduce the evaluation-related and audience-related causes of state anxiety, while the amount of practice is theorized to raise perceived competence. Convergence between the two strands thus provides a theoretically coherent account of the changes associated with participation: the AI partner appears to address both sides of the confidence equation that drives the decision to speak.

However, we should also temper this encouraging result with the counter-evidence in the literature. Ballıdağ and Aydın (2025), for instance, did not find any significant reduction in speaking anxiety through an AI chatbot task while peer interaction did. The difference between their results and ours could be due to a different student population, the specific tool, the length of the intervention, and whether the tool was incorporated into a formal course or not. This might also reflect a design feature emphasized by recent work, namely that the benefits of AI partners depend on the perceived usability and sense of presence of the system (Huang et al., 2025). These factors indicate that AI speaking partners are not a silver bullet but rather instruments which need to be carefully designed with regard to their emotional appeal and embedded in instruction.

The results should thus also be read as evidence that AI speaking partners can supplement, but not replace, human contact. The learners in our study appreciated having an AI partner specifically because it was not a human audience, which is also what makes it a poor substitute for face-to-face interaction. Human communication involves pragmatic, socio-linguistic, and affective aspects—such as turn-taking, negotiation of meaning, and the benefits of authentic interaction—that a chatbot cannot fully replicate (Bibauw et al., 2022). The pedagogical takeaway, therefore, is that AI speaking partners are most usefully deployed as a rehearsal space that lowers the threshold for speaking and builds the confidence that learners can bring to human interaction, rather than as a stand-alone replacement for in-classroom communication.

Several limitations should be acknowledged. Because the study used a one-group pre-test and post-test design without a control group, caution is needed when making any causal claims; the observed changes may be due, at least in part, to maturation, repeated exposure to the instruments, or the overall impact of additional speaking practice rather than to any unique property of the AI partner.

The sample was drawn through convenience sampling within one institution, which limits the generalizability of the results. Furthermore, all data are self-reported and no objective measure of oral performance was included; it is therefore unclear if the reported affective changes translate into better speaking performance. Last but not least, the lack of a delayed follow-up assessment leaves the durability of the observed changes in question.

In future work, these limitations could be addressed in a randomized controlled trial that compares conditions matched on the quantity of speaking practice, adds objective measures of speaking performance, and includes delayed post-tests to determine whether affective gains are sustained. Individual differences (e.g., initial anxiety level, proficiency, and attitude toward technology) should be examined to understand how they moderate the effects of AI speaking partners, and the quality of the conversational experience deserves closer scrutiny, because the mixed results in the literature imply that the design of the interaction is an important factor.

7. Conclusion

In sum, the present mixed-methods study provided some evidence of a positive association between the eight-week AI speaking partner intervention and both lower speaking anxiety and higher willingness to communicate among first-year college English learners in China. The qualitative results also indicated that the observed affective changes were associated with the non-judgmental, audience-free, and practice-rich nature of the experience, which maps onto the state anxiety and perceived competence that the pyramid model identifies as the proximal determinants of willingness to communicate (MacIntyre et al., 1998). This paper contributes to this growing, if mixed, literature by demonstrating both the magnitude and the perceived process of the observed changes within a single coherent design.

The implications of these results for foreign language instructors are that AI conversation partners can be useful tools for developing learners' readiness to speak by reducing the affective barriers that prevent them from participating orally. If used as part of an organized curriculum paired with opportunities for face-to-face contact, such tools could help students develop the self-confidence needed to speak up in class, or even outside it. For developers of AI-based language applications, these findings suggest that systems should include elements that help users feel safe and competent in their use. With more data available, continued mixed-methods research is necessary to understand for whom, under what circumstances, and in what ways AI speaking partners may best serve the affective dimension of second-language speaking.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Azizimajd, H. (2023). Investigating the Impacts of Voice-Based Student-Chatbot Interactions in the Classroom on EFL Learners' Oral Fluency and Foreign Language Speaking Anxiety. *Technology Assisted Language Education, 1,* 61-83. https://tale.razi.ac.ir/article_2732.html

Ballıdağ, M., & Aydın, S. (2025). A Comparison of the Effects of AI-Based Chatbots and Peer Interactions on Speaking Anxiety among EFL Learners. *Future in Educational Research, 3,* 224-238. https://doi.org/10.1002/fer3.70005

Bibauw, S., François, T., & Desmet, P. (2022). Dialogue Systems for Language Learning. In H. Mohebbi, & C. Coombe (Eds.), *Research Questions in Language Education and Applied Linguistics* (pp. 121-135). Routledge. https://doi.org/10.4324/9781351117586-12

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology, 3,* 77-101. https://doi.org/10.1191/1478088706qp063oa

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). SAGE Publications.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics.* Oxford University Press.

Guo, Z. W., & Yan, H. W. (2025). The Impact of Artificial Intelligence Chatbots on Young EFL Learners' Willingness to Communicate. *Curriculum and Teaching Methodology, 8,* 112-119. https://www.clausiuspress.com/article/14941.html

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign Language Classroom Anxiety. *The Modern Language Journal, 70,* 125-132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Huang, Y., Chen, H., & Hu, C. (2025). L2 Growth Mindset in Ai-Mediated Language Learning: Effects of Perceived Usability and Presence of Generative AI Chatbots. *Frontiers in Psychology, 16,* Article 1700117. https://doi.org/10.3389/fpsyg.2025.1700117

Macintyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing Willingness to Communicate in a L2: A Situational Model of L2 Confidence and Affiliation. *The Modern Language Journal, 82,* 545-562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

McCroskey, J. C., & Baer, J. E. (1985). Willingness to Communicate: The Construct and Its Measurement. In *The Annual Meeting of the Speech Communication Association*. ERIC Document Reproduction Service. https://eric.ed.gov/?id=ED265604

Tai, T. Y., & Chen, H. H. (2020). The Impact of Google Assistant on Adolescent EFL Learners' Willingness to Communicate. *Interactive Learning Environments, 31,* 1485-1502. https://doi.org/10.1080/10494820.2020.1841801

Wu, T. T., Hapsari, I. P., & Huang, Y. (2025). Effects of Incorporating AI Chatbots into Think-Pair-Share Activities on EFL Speaking Anxiety, Language Enjoyment, and Speaking Performance. *Computer Assisted Language Learning, 39,* 1408-1446. https://doi.org/10.1080/09588221.2025.2478271

Zhai, C., & Wibowo, S. (2023). A Systematic Review on Artificial Intelligence Dialogue Systems for Enhancing English as Foreign Language Students' Interactional Competence in the University. *Computers and Education: Artificial Intelligence, 4,* Article ID: 100134. https://doi.org/10.1016/j.caeai.2023.100134

Zheng, Y. B., Zhou, Y. X., Chen, X. D., & Ye, X. D. (2025). The Influence of Large Language Models as Collaborative Dialogue Partners on EFL English Oral Proficiency and Foreign Language Anxiety. *Computer Assisted Language Learning, 39,* 1092-1118. https://doi.org/10.1080/09588221.2025.2453191