THE EFFECT OF ARTIFICIAL INTELLIGENCE TOOLS ON IMPROVING UNDERGRADUATE STUDENTS' ENGLISH-SPEAKING CONFIDENCE

Akhil AUGUSTINE[1] and Kanthathip SINGHANETI[1]
1 Faculty of International Programs, Thongsook College, Thailand;
akhil@thongsook.college (A. A.); dr.kanthathip@thongsook.college (K. S.)

ABSTRACT
This study examined the effects of artificial intelligence tools on undergraduate students' English-speaking confidence. It employed a mixed-methods, four-group pretest-posttest design involving 200 second-year Business English students at Thongsook College, Bangkok, Thailand. Participants were recruited through purposive sampling and randomly assigned to Gemini, ChatGPT, Copilot or non-AI TBLT groups with 50 students per group. The three-week intervention included role-play, opinion-sharing and presentation practice. A 20-item, five-point questionnaire assessed self-reported speaking confidence. The findings showed that the overall mean increased from 2.78 (SD = 0.67) to 3.86 (SD = 0.46), with a significant mean improvement of 1.08 points, $t(199) = 18.77$, $p < .001$, Cohen's $dz = 1.33$. Confidence increased for 84.5% of participants. However, the study identifies its four-group comparisons as hypothetical and its interview accounts as fictional; these cannot establish an advantage of AI-supported instruction or substantiate participants' experiences. The reported overall improvement suggests an association between structured speaking activities and increased perceived confidence. Verified comparative analyses, authentic interview data and longer-term speaking assessments are required to determine the effectiveness of AI-supported TBLT. The study provides pedagogical implications for classroom speaking practice and offers suggestions for further investigations into AI-supported instruction and the development of undergraduate students' English-speaking confidence.
Keywords: Artificial Intelligence Tools, Task-Based Language Teaching, English-Speaking Confidence, Undergraduate Students

CITATION INFORMATION: Augustine, A., & Singhaneti, K. (2026). The Effect of Artificial Intelligence Tools on Improving Undergraduate Students' English-Speaking Confidence. *Procedia of Multidisciplinary Research, 4*(11), 31.

INTRODUCTION

Speaking English is important for undergraduate students' participation in academic and professional communication. For Business English students, classroom speaking activities provide opportunities to explain ideas, discuss business situations recommend products and deliver presentations. However, knowing vocabulary and grammar does not necessarily mean that students feel confident communicating. MacIntyre et al. (1998) proposed a model linking second-language confidence with willingness to communicate, while Peng and Woodrow (2010) states that communication confidence alongside motivation, learner beliefs and classroom environment. Horwitz et al. (1986) state that communication apprehension and fear of negative evaluation as relevant aspects of foreign language classroom anxiety. These perspectives supported investigating students' confidence alongside providing opportunities to use English.

Some students may hesitate to speak because they worry about making mistakes, being misunderstood or receiving negative reactions. Such concerns can make classroom discussions and presentations uncomfortable. MacIntyre and Gardner (1994) examined language anxiety in relation to second-language cognitive processing, while Teimouri et al. (2019) state that concerning anxiety and language achievement. Horwitz et al. (1986) provided a rationale for supportive speaking environments. Nevertheless, speaking anxiety, confidence and willingness to communicate are distinct constructs. The present study therefore measured self-reported speaking confidence directly rather than assuming that greater participation or reduced anxiety necessarily indicated increased confidence.

AI-supported interaction offers additional opportunities to investigate language practice. Dizon (2020) state that intelligent personal assistants for second-language listening and speaking development. Kohnke et al. (2023) reviewed ChatGPT's possibilities for language teaching and learning, while Fathi et al. (2024) examined AI-mediated activities in relation to speaking skills and willingness to communicate. These studies provided a foundation for exploring AI tools as practice partners. However, they addressed different technologies, instructional conditions and outcomes, so they did not directly establish that Gemini, ChatGPT and Copilot improved speaking confidence in the study setting.

Learners' emotional experiences also deserve attention when designing AI-supported speaking activities. Zhang et al. (2024) state that enjoyment, anxiety and willingness to communicate through the AI-speaking assistant Lora. Their research complemented the attention to language anxiety in Horwitz et al. (1986) and the consideration of confidence and classroom environment (Peng and Woodrow, 2010). Drawing on these perspectives, AI-supported rehearsal was considered a potential setting in which students could prepare ideas and practice responses before classroom interaction. This explanation provided a rationale for the study rather than evidence of an established effect of the three selected tools.

Students used their assigned AI tool for structured speaking practice inside and outside the classroom. Classroom activities included teacher guidance and peer interaction, while independent tasks provided opportunities to rehearse conversations, express opinions and prepare presentations. Research on AI-supported language learning provided a basis for examining these opportunities (Dizon, 2020; Fathi et al., 2024; Kohnke et al., 2023). Topics included university life, everyday situations, travel, technology and business communication. Suitable voice access was checked before implementation, and comparable topics, task requirements and practice time were maintained across groups.

Task-Based Language Teaching (TBLT) provided the common instructional foundation. Prabhu (1987) offered an influential account of task-based instruction, including information-gap, reasoning-gap and opinion-gap activities. Ellis (2009) state that meaning-focused tasks with communicative outcomes, while Nunan (2004) presented a framework for organizing task-based instruction. Bygate et al. (2015) state that communication tasks as fundamental to

TBLT in face-to-face and virtual settings. Students prepared relevant ideas and language, completed purposeful speaking tasks, and reflected on feedback afterward. Activities included exchanging information to plan an event, comparing products to make a recommendation and sharing opinions about technology.

The reviewed literature provided grounds for investigating AI-supported speaking, but it did not directly address the comparison of Gemini, ChatGPT and Copilot within a shared TBLT procedure at Thongsook College (Dizon, 2020; Fathi et al., 2024; Zhang et al., 2024). The study examined changes in confidence from before to after the intervention and compared post-intervention outcomes while controlling for initial confidence.

Research Problem Statement

Undergraduate students' knowledge of English vocabulary and grammar does not necessarily translate into confidence in speaking. Concerns about making mistakes and receiving negative evaluation may limit their confidence during classroom discussions, business communication tasks and presentations. Although previous research provided a rationale for AI-supported speaking practice, the reviewed literature did not directly establish how students' confidence changed following the use of Gemini, ChatGPT or Copilot within a common TBLT framework at Thongsook College. The study therefore investigated changes in undergraduate students' English-speaking confidence following AI-supported TBLT activities inside and outside the classroom. It also compared post-intervention confidence among the three tool groups while controlling for initial confidence and another group practiced without the AI tools. Speaking anxiety provided background motivation for the study, while speaking confidence was its principal measured outcome. Research Questions were RQ1) Did students receiving AI-supported TBLT show greater improvements in English-speaking confidence than students receiving equivalent TBLT instruction without AI? RQ2) Did post-intervention English-speaking confidence differ among the Gemini, ChatGPT, Copilot and non-AI TBLT groups after controlling for initial confidence?

Research Objectives

RO1: To compare improvements in English-speaking confidence between students receiving AI-supported TBLT and students receiving equivalent TBLT instruction without AI.

RO2: To compare post-intervention English-speaking confidence among the Gemini, ChatGPT, Copilot and non-AI TBLT groups while controlling for initial confidence.

Research Hypotheses

The following hypotheses were formulated:

H1: Students receiving AI-supported TBLT were expected to demonstrate significantly greater improvements in English-speaking confidence than students receiving equivalent TBLT instruction without AI.

H2: Post-intervention English-speaking confidence was expected to differ significantly among the four groups after controlling for initial confidence.

Conceptual Framework

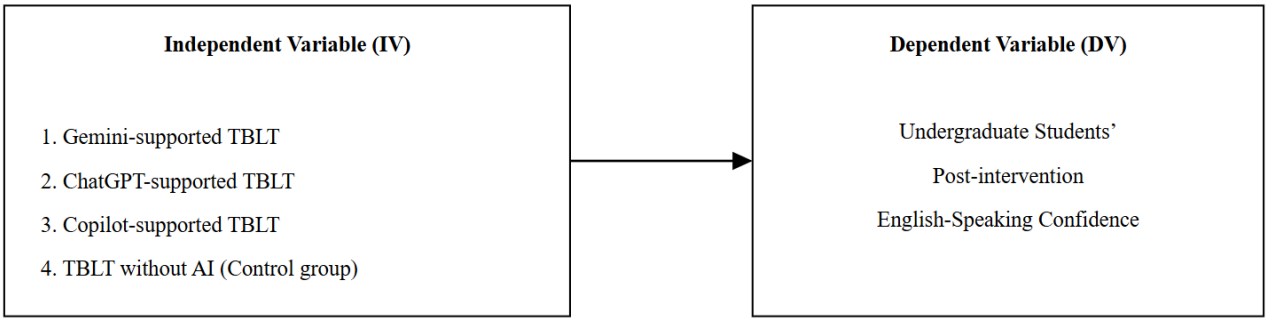

Figure 1 The conceptual framework was adapted from Ellis (2009).

LITERATURE REVIEW

Foreign Language Speaking Anxiety

Foreign language speaking anxiety can influence students' participation in oral communication inside and outside the classroom. Woodrow (2006) stated that speaking anxiety significantly predicted oral achievement among students preparing for university study in Australia. The study distinguished between classroom and external communication situations, indicating that learners' emotional responses could vary across speaking contexts. These findings supported examining students' experiences of both classroom activities and independent speaking practice. Speaking anxiety also varies according to the organization of classroom interaction. Liu (2006) stated that Chinese EFL students experienced greater anxiety when answering teachers' questions or being individually selected to speak, whereas pair work was less anxiety-provoking. Students with higher proficiency generally reported lower anxiety, and increased exposure to oral English was associated with greater comfort. These findings suggested that speaking activities should provide manageable opportunities for participation before requiring public performance.

Individual differences further complicate the relationship between anxiety and speaking performance. Oya et al. (2004) examined personality, anxiety and oral performance among Japanese speakers of English, assessing performance through fluency, accuracy, complexity and overall impressions. Their approach highlighted the importance of considering emotional and personal characteristics alongside linguistic outcomes. For the present study, this perspective supported recognizing that students could respond differently to identical speaking tasks and practice requirements. Speaking confidence was therefore measured directly rather than inferred from anxiety scores. Bandura (1977) stated that self-efficacy concerned individuals' beliefs about their ability to perform particular actions. Applied cautiously to speaking practice, this perspective provided a rationale for examining students' perceived ability to complete specific communication tasks. The present study therefore assessed confidence in speaking activities while treating anxiety as a related emotional experience rather than an interchangeable construct.

AI Tools and Speaking Practice

AI-supported language practice has attracted attention because conversational systems can provide additional opportunities for interaction. Huang et al. (2022) reviewed chatbot-supported language learning and examined its technological, pedagogical and social affordances. They stated that educational usefulness depended on how chatbots were incorporated into learning activities. This perspective supported using AI tools within planned speaking tasks that included clear objectives, appropriate instructions and opportunities for meaningful language use. Empirical evidence also supported investigating automated conversational partners in speaking education. Tai and Chen (2023) studied 112 adolescent EFL learners who participated in Google Assistant activities over two weeks. They stated that participants demonstrated improvements in willingness to communicate and communicative confidence, alongside reduced speaking anxiety. Although these findings concerned adolescents and a different technology, they provided a rationale for examining confidence development among undergraduate learners receiving AI-supported speaking practice.

Different chatbot technologies should not be treated as educationally equivalent. Jeon et al. (2023) presented a conceptual framework and systematic review of speech-recognition chatbots, highlighting the diversification of conversational technologies and their possibilities for language education. Their work supported distinguishing tools by their interaction features and instructional use. Accordingly, the present study documented the interaction modes used with Gemini, ChatGPT and Copilot. Comparable speaking requirements were maintained across the three AI groups and the TBLT group without AI. Learners' experiences were also important when evaluating AI-supported practice. Jeon (2024) stated that chatbot affordances

could operate as opportunities or constraints depending on the learner. This finding suggested that access to a conversational tool did not guarantee a uniform learning experience. In the present study, role-play, opinion-sharing and presentation practice provided a shared structure across all four groups. Structured interviews explored students' experiences of their assigned practice condition, including the opportunities and difficulties encountered.

AI and Learners' Emotional Experiences

Learners' emotional experiences involve both positive and negative responses to language learning. Dewaele et al. (2018) examined foreign language enjoyment and classroom anxiety in relation to learner characteristics and teacher-related variables. Their research supported considering the learning environment alongside individual differences when interpreting participation. In the present study, this perspective supported exploring whether students experienced speaking activities as enjoyable, demanding or uncomfortable rather than assuming that all learners responded positively to AI-supported practice. AI interaction may provide a comparatively comfortable setting for some learners. Tai and Chen (2023) stated that participants enjoyed communicating and playing games with Google Assistant. Their interview findings indicated reduced apprehension and greater motivation to use English. These experiences suggested a possible connection between emotional comfort and participation. However, this connection required investigation within the present setting rather than an assumption that it applied equally to Gemini, ChatGPT and Copilot.

Emotional responses may also depend on whether students perceive a chatbot's features as helpful. Jeon (2024) stated that learners interpreted the same technological affordances differently, creating varied opportunities and constraints. This finding provided a rationale for asking students about successful interactions, communication difficulties and their responses to speaking practice. Such accounts supported examining why students completing comparable tasks could report different levels of confidence or engagement. Successful practice experiences offered a possible explanation for confidence development. Bandura (1977) identified performance accomplishments as a source of self-efficacy beliefs. In the present study, students' perceived success in completing speaking tasks was explored as a possible explanation for changes in confidence. Structured interviews examined whether students believed repeated practice helped them express ideas, manage difficulties and prepare for classroom communication. Experiences across the AI-supported and non-AI TBLT groups were considered without assuming that confidence development resulted exclusively from AI use.

Research Gap and Rationale

The reviewed literature provided a foundation for chatbot-supported language learning but did not directly address comparisons among Gemini, ChatGPT, Copilot and equivalent TBLT instruction without AI in the selected undergraduate setting (Huang et al., 2022; Jeon et al., 2023). Findings from adolescent and primary EFL learners also required further investigation before application to undergraduate students at Thongsook College (Tai & Chen, 2023; Jeon, 2024).

The study therefore examined English-speaking confidence across four groups following a common TBLT procedure. It compared post-intervention confidence while controlling for initial confidence, with a non-AI control group providing a comparison for AI-supported instruction. Structured interviews complemented questionnaire results by exploring students' experiences of classroom and independent speaking practice.

RESEARCH METHODOLOGY

The study employed a mixed-methods approach combining a four-group pretest-posttest design with structured interviews. Participants comprised 200 second-year undergraduate students enrolled in Business English at Thongsook College, Bangkok, Thailand. Purposive sampling

was used to select students whose course enrollment matched the study's focus. Following institutional approval and informed consent, participants were randomly assigned to four groups: Gemini-supported TBLT, ChatGPT-supported TBLT, Copilot-supported TBLT and TBLT without AI, with 50 students in each group. Purposive sampling determined recruitment, while random assignment determined group membership.

The independent variable comprised the four groups, and the dependent variable was post-intervention English-speaking confidence. Pre-intervention confidence was included as a covariate. All groups completed a three-week intervention with equivalent topics, teacher guidance, task requirements and practice time. Activities included role-play, opinion-sharing and presentation practice inside and outside the classroom. Each activity followed the TBLT stages of pre-task preparation, task cycle and post-task reflection. The AI groups used their assigned tools for speaking practice, while the control group completed equivalent activities using peer interaction and non-AI materials. A researcher-developed, 20-item, five-point Likert-scale questionnaire measured speaking confidence before and after the intervention. Experts reviewed its content and clarity, and a pilot study assessed internal consistency using Cronbach's alpha. Structured interviews explored the experiences of eight purposively selected students, two from each group.

Following institutional approval and informed consent, participants completed the pretest, three-week intervention and posttest. Practice records documented participation and practice time. Interviews were recorded with permission and transcribed. Participant codes linked responses, while missing data and withdrawals were documented. Frequencies, percentages, means and standard deviations summarized the data. ANCOVA compared posttest confidence across the four groups while controlling for pretest confidence. A planned contrast compared the three AI groups' average adjusted score with the non-AI control group for H1, while the overall group test addressed H2. Follow-up comparisons used Holm adjustment. Assumptions were checked, and effect sizes and 95% confidence intervals were reported at a .05 significance level. Interview transcripts were analyzed thematically and interpreted alongside quantitative findings.

RESULTS

Respondent Demographic Characteristics

The study included 200 students: 66 females (33.0%), 67 males (33.5%) and 67 who preferred not to disclose their gender (33.5%). The largest age category was under 20 years (39.0%), and most respondents had studied English for 5-10 years (55.0%). Previous speaking practice varied: 26.5% practiced less than once a week, while 26.0% practiced 5-7 days a week. Previous experience using AI for speaking practice was reported by 110 respondents (55.0%).

Descriptive Statistics and Internal Consistency

Table 1 Descriptive Statistics and Internal Consistency of Speaking-Confidence Scores

Assessment	N	Items	Mean	SD	Median	Cronbach's α
Pretest	200	20	2.78	0.67	2.95	.915
Posttest	200	20	3.86	0.46	4.10	.844

Table 1. show that the mean speaking-confidence scores increased from 2.78 (*SD* = 0.67) at pretest to 3.86 (*SD* = 0.46) at posttest. Median scores also increased, from 2.95 to 4.10. The questionnaire demonstrated high internal consistency at both assessments, with Cronbach's alpha coefficients of .915 and .844, respectively.

Overall Pretest-Posttest Comparison

Table 2 Paired-Samples Comparison of Overall Speaking Confidence

Mean difference	SD of differences	95% CI	*t*	*df*	*p*	Cohen's *dz*
1.08	0.82	[0.97, 1.20]	18.77	199	< .001	1.33

Table 2 shows that posttest speaking-confidence scores were significantly higher than pretest scores, $t(199) = 18.77$, $p < .001$. The mean increase was 1.08 points, 95% CI [0.97, 1.20]. Cohen's *dz* of 1.33 indicated a large standardized within-person change. This effect size describes the overall change across participants and does not estimate the advantage of AI-supported instruction over instruction without AI.

Assumption Check and Supplementary Analyses

The Shapiro-Wilk test indicated that the difference scores departed from normality, $W = .952$, $p < .001$. Supplementary Wilcoxon signed-rank and two-sided sign tests were therefore examined. Both tests were statistically significant ($p < .001$), providing additional evidence of an overall increase in speaking-confidence scores.

Table 3 Direction of Individual Speaking-Confidence Changes

Direction of change	Frequency (*n*)	Percentage (%)
Increased confidence	169	84.5
Unchanged confidence	17	8.5
Decreased confidence	14	7.0
Total	200	100.0

Table 3 shows that confidence scores increased for 169 respondents (84.5%), remained unchanged for 17 (8.5%) and decreased for 14 (7.0%). The sign test excluded the 17 tied pairs and indicated that increases predominated among the 183 respondents whose scores changed.

Four-Group Hypothesis Testing

Table 4 Research Hypothesis Testing

Hypothesis	Analysis	Result	Decision
H1: Students receiving AI-supported TBLT demonstrated greater improvement than students receiving TBLT without AI.	Planned ANCOVA contrasts comparing the average adjusted posttest score of the three AI groups with the non-AI group.	Adjusted mean difference = 0.32, 95% CI [0.18, 0.46], $p < .001$.	Supported
H2: Post-intervention confidence differed among the four groups after controlling for initial confidence.	Four-group ANCOVA with pretest confidence as the covariate.	$F(3, 195) = 7.20$, $p < .001$, partial $\eta^2 = .100$.	Supported

Table 4. The planned contrast indicated that the three AI-supported groups had an average adjusted posttest confidence score 0.32 points higher than the non-AI TBLT group, 95% CI [0.18, 0.46], p < .001. This hypothetical result supported H1. The ANCOVA also indicated a significant difference among the four groups after controlling for pretest confidence, $F(3, 195) = 7.20$, $p < .001$, partial $\eta^2 = .100$, supporting H2.

Thematic Analysis of Interview

Table 5 Themes Identified in the Fictional Interview

Theme	Codes	Summary
Greater comfort in speaking	Reduced hesitation; willingness to express ideas; persistence after mistakes; familiarity with peers	The accounts portrayed greater willingness to speak and continue after mistakes.
Opportunities for rehearsal and preparation	Repeated practice; organizing ideas; presentation rehearsal; peer practice; feedback application	The examples described how rehearsal and peer practice helped organize ideas and prepare responses.
Challenges requiring instructional support	Speech-recognition problems; unfamiliar vocabulary; unsuitable responses; limited preparation time	The examples highlighted difficulties that required clearer guidance and teacher feedback.

Table 5. The analysis of the fictional interview identified three themes: greater comfort in speaking, opportunities for rehearsal and preparation, and challenges requiring instructional support.

DISSCUSSION AND CONCLUSION

The results showed an increase in undergraduate students' English-speaking confidence following the three-week intervention. Mean confidence increased from 2.78 to 3.86, with a mean difference of 1.08 points, 95% CI [0.97, 1.20]. The paired-samples comparison was statistically significant, $t(199) = 18.77$, $p < .001$, and the standardized within-person change was large (Cohen's $dz = 1.33$). Confidence increased for 169 students (84.5%), indicating that the overall improvement extended across most participants. The significant Wilcoxon signed-rank and sign tests provided additional support for the direction of this change despite the non-normal distribution of difference scores. These findings suggest that structured speaking practice within the TBLT framework was associated with greater perceived confidence. Pre-task preparation, purposeful interaction and post-task reflection may have helped students organize ideas, attempt communication and apply feedback. Role-play, opinion-sharing and presentation practice offered repeated opportunities to complete meaningful speaking tasks. However, the pooled pretest-posttest improvement represents change across all four groups and cannot independently establish the additional contribution of AI. The hypothetical planned contrast presented in the manuscript showed an average adjusted posttest advantage of 0.32 points for the AI-supported groups over the non-AI TBLT group, 95% CI [0.18, 0.46], $p < .001$. If confirmed through verified participant data, this result would support H1 and suggest that AI-supported practice provided an additional benefit beyond the common instructional framework. A possible explanation is that conversational tools offered further opportunities to rehearse responses, repeat difficult exchanges and prepare for classroom communication. Nevertheless, the adjusted between-group difference should be distinguished from the substantially larger overall within-person change.

The hypothetical ANCOVA also indicated differences among the four groups after controlling for initial confidence, $F(3, 195) = 7.20$, $p < .001$, partial $\eta^2 = .100$. If verified, this finding would support H2. However, the overall group test does not establish which individual groups differed. Without adjusted group means and the specified follow-up comparisons, the results cannot demonstrate that Gemini, ChatGPT or Copilot was more effective than the other tools. The fictional interview examples presented three possible explanations and implementation concerns: greater comfort in speaking, opportunities for rehearsal and preparation, and

challenges requiring instructional support. These examples suggest useful areas for investigating how learners experience repeated practice, feedback and communication difficulties. They cannot serve as evidence of participants' experiences or corroborate the quantitative findings. Such interpretations require analysis of the actual interview transcripts. The results have potential implications for classroom practice. Teachers could incorporate AI-supported rehearsal into clearly structured speaking tasks, maintain comparable practice requirements and provide guidance on evaluating responses. Speech-recognition difficulties, unfamiliar vocabulary and unsuitable responses require teacher support. The finding that 7.0% of participants reported lower confidence also indicates that learners may respond differently to the same instructional activities. Teachers should therefore monitor individual experiences and adapt support where necessary.

In Conclusion, the pretest-posttest results indicated increased English-speaking confidence among the 200 participants following three weeks of structured speaking activities. Most students reported improvement, suggesting that repeated opportunities to prepare ideas, practice communication and reflect on feedback may support confidence development. This interpretation is consistent with Bandura's (1977) view that successful performance experiences can strengthen beliefs in one's capabilities. AI-supported rehearsal may provide additional opportunities for students to practice before classroom communication. Tai and Chen (2023) reported improvements in communicative confidence following Google Assistant activities among EFL learners. The fictional interviews cannot verify participants' experiences.

Recommendations

Teachers should integrate AI-supported speaking practice into TBLT classrooms through role-play, opinion-sharing and presentation rehearsal, with clear instructions and feedback. Peer interaction and individual support should remain available to help students develop English-speaking confidence.

Future Research

Future studies should use larger samples and longer interventions, combining confidence questionnaires with speaking assessments. Verified comparisons across AI and non-AI groups, supported by authentic interviews, could clarify their effectiveness.

REFERENCES

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

Bygate, M., Norris, J. M., & Van den Branden, K. (2015). Task-based language teaching. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Wiley. https://doi.org/10.1002/9781405198431.wbeal1467

Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language Teaching Research, 22*(6), 676-697. https://doi.org/10.1177/1362168817692161

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology, 24*(1), 16-26. Journal article

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics, 19*(3), 221-246. https://doi.org/10.1111/j.1473-4192.2009.00231.x

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

Google. (2025). *Gemini* [Generative AI tool]. https://gemini.google.com/

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125-132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning, 38*(1), 237-257. https://doi.org/10.1111/jcal.12610

Jeon, J. (2024). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning, 37*(1-2), 1-26. https://doi.org/10.1080/09588221.2021.2021241

Jeon, J., Lee, S., & Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education, 206*, Article 104898. https://doi.org/10.1016/j.compedu.2023.104898

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537-550. https://doi.org/10.1177/00336882231162868

Liu, M. (2006). Anxiety in Chinese EFL students at different proficiency levels. *System, 34*(3), 301-316. https://doi.org/10.1016/j.system.2006.04.004

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*(2), 283-305. https://doi.org/10.1111/j.1467-1770.1994.tb01103.x

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*(4), 545-562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

Microsoft. (2025). *Copilot* [Generative AI tool]. https://copilot.microsoft.com/

Nunan, D. (2004). *Task-based language teaching*. Cambridge University Press.

OpenAI. (2025). *ChatGPT* [Large language model]. https://chatgpt.com/

Oya, T., Manalo, E., & Greenwood, J. (2004). The influence of personality and anxiety on the oral performance of Japanese speakers of English. *Applied Cognitive Psychology, 18*(7), 841-855. https://doi.org/10.1002/acp.1063

Peng, J.-E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language Learning, 60*(4), 834-876. https://doi.org/10.1111/j.1467-9922.2010.00576.x

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.

Tai, T.-Y., & Chen, H. H.-J. (2023). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments, 31*(3), 1485-1502. https://doi.org/10.1080/10494820.2020.1841801

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

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

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

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



Copyright: © 2026 by the authors. This is a fully open-access article distributed under the terms of the Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).