

Reducing Speaking Anxiety through AI Interaction: ChatGPT-Assisted English-Speaking Practice and Language Proficiency Development in Chinese Universities

Qingqing Cui¹, Supyan Hussin², Wei Lun Wong¹

¹ Faculty of Education, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

² Institute of Ethnic Studies, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

Correspondence: Qingqing Cui, Faculty of Education, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia. E-mail: p141111@siswa.ukm.edu.my

Received: March 17, 2026

Accepted: May 7, 2026

Online Published: June 17, 2026

doi:10.5430/wjel.v16n6p524

URL: <https://doi.org/10.5430/wjel.v16n6p524>

Abstract

This study aims to explore the correlation between ChatGPT-supported English-speaking practice, speaking anxiety, and students' self-perception of their English-speaking ability improvement at Chinese universities. It also examines speaking anxiety statistically mediates the relationship between ChatGPT-assisted practice and proficiency development and if perceived interaction quality moderates the relationship. This study is a sequential explanatory mixed method design. The quantitative section presents the results of a cross-sectional survey of 412 undergraduate students in six public universities in China. The analysis is conducted according to PLS-SEM reporting logic including reliability, convergent validity, discriminant validity, predictive relevance and bootstrapped path estimates. ChatGPT-assisted English-speaking practice has a positive relationship with the self-reported English-speaking proficiency development and a negative relationship with speaking anxiety. The negative correlation between speaking anxiety and proficiency development and the significant mediating path between ChatGPT-assisted practice and proficiency development suggests that speaking anxiety hinders the utility of ChatGPT to assist in proficiency development. The perceived interaction quality also moderates the practice to proficiency relationship. This paper is an original article that contributes to the existing research on ChatGPT-assisted language learning by incorporating an interaction-based perspective of ChatGPT use, an affective mechanism and a contextual moderator in the Chinese university context. Because the quantitative data are cross-sectional and self-reported, the results are interpreted as evidence of substantial relationships instead of causal relationships.

Keywords: ChatGPT-assisted speaking practice, speaking anxiety, English speaking proficiency development, perceived interaction quality, Chinese universities, AI-assisted language learning

1. Introduction

The entry of generative artificial intelligence into higher education in a short span of time has impacted the way English as a foreign language (EFL) learners prepare, rehearse and review oral tasks. In university language classrooms, students no longer rely on teachers, peers, printed scripts, and occasionally on language-lab sessions alone. They now engage with large language model systems which can create prompts, continue conversations, reformulate learner output and give immediate feedback at any point in the day. Recent reviews confirm the growing use of AI tools for EFL learning since these tools can provide personalization, repetition, and flexibility for learning outside of class through practice (Jiang, 2022; Crompton et al., 2024; Yang & Li, 2024; Kundu & Bej, 2025). Within that larger movement, the practice of speaking has become one of the most promising, yet unevenly understood, areas. A chatbot-based practice will perhaps offer more opportunities for spontaneous use of English, but it is not clear whether it is educationally meaningful in the sense that the learners will feel safe enough to speak up; if the interaction is meaningful in some way; and if repeated practice leads to measurable development in oral performance (Du & Daniel, 2024; Klimova & Ibna Seraj, 2023; Alrajhi, 2024; Wu, 2024).

The latest literature has provided encouraging but still scattered answers. Studies on AI-mediated speaking practice report on gains in enjoyment, willingness to communicate, speaking performance and confidence, along with decreases in anxiety in many settings (Belda-Medina & Calvo-Ferrer, 2022; C. Zhang et al., 2024; D. Zhang et al., 2024; Wang & Wen, 2025; Ebadi et al., 2025). Reviews of AI in language education also indicate that speaking is one of the domains of skills that are the most sensitive to interactive AI support to the extent that learners receive timely feedback and ongoing opportunities for repetition (Du & Daniel, 2024; Yang & Li, 2024; Wu, 2024; Torres & Kahveci, 2025). Yet, the field still does not have a coherent explanatory model of Chinese universities that connects one concrete predictor to one central affective mechanism (in this case, speaking anxiety) and one central education outcome (in this case, proficiency development) - in this case, speaking practice with ChatGPT. It is also not yet clear though if all ChatGPT use is equally effective, or if the benefit is greater if students judge their interaction to be clear, responsive, supportive, and personally relevant.

The present manuscript suggests ChatGPT assisted English speaking practice to be implemented through two related routes. First, it does increase the amount and frequency of oral rehearsal which should directly enhance the developing process of speaking proficiency. Second, it can reduce speaking anxiety by providing learners with a private and nonjudgmental partner for low-stakes practice and that reduced anxiety should also support proficiency gains. The manuscript also suggests that interaction quality is important to be perceived, as students are more likely to benefit from AI speaking practice in response to student stimuli, which are timely, understandable, and adapted to their communicative needs. In this sense, interaction quantity is not sufficient; the learner's experience of the interaction should also be taken into consideration (Tram et al., 2024; Alrajhi, 2024; Muniandy & Selvanathan, 2025).

This research therefore develops and tests a five-hypothesis model that focuses on four constructs, namely, ChatGPT-Assisted English-Speaking Practice, Speaking Anxiety, English Speaking Proficiency Development, and Perceived Interaction Quality. The following sections are used to define the key variables, review the relevant literature, identify the research gap, present the theoretical model, and report the draft empirical results.

At the same time, claims related to AI in language education must be carefully qualified. Not every student has the same benefit, and not every use of ChatGPT is pedagogically sound. Some of the learners use it purely for short question-answer interactions but some use it as sustained role-play, pronunciation support or rehearsal for a presentation. These differences point to the importance of considering the value of ChatGPT as not just determined by access, but also by the ways in which the learners engage with the interaction, and whether the interaction leads to meaningful speaking practice. For this reason, the use of AI is not discussed in the present manuscript as a single undifferentiated behavior. It focuses specifically on ChatGPT-assisted English-speaking practice and tests the relationship between the practice and anxiety and proficiency development in Chinese universities.

1.1 ChatGPT-Assisted English-Speaking Practice

How often and how deeply learners engage with ChatGPT or a similar large language model as a conversation partner, rehearsal tool, and source of feedback for spoken English tasks. In practical terms this encompasses such activities as role-play dialogues, oral brainstorming, mock interviews, pronunciation rehearsal, presentation practice, impromptu response to prompts, and follow up clarification after errors. Recent reviews have suggested that AI chatbots are especially useful in the language learning process when they provide students with repeated interaction, immediate response creation and individualized feedback, which students can access outside of teacher-controlled schedules (Belda-Medina & Calvo-Ferrer, 2022; Klimova & Ibna Seraj, 2023; Du & Daniel, 2024; Yang & Li, 2024). In the case of EFL situations where authentic speaking opportunities are limited, a conversational AI system can serve as an always-available practice partner as opposed to a reference tool.

The theoretical base for this construct is mainly interaction-based views of language learning and recent research on AI-supported self-regulation. Interaction-oriented accounts argue that language development occurs through learners being stimulated to notice gaps, produce output, adjust meaning and try again in response to feedback. Recent AI-EFL reviews have approached the practice of chatbot based on this perspective, emphasizing that conversational AI provides more exposure to negotiated meaning, and learners can repeat their conversation until they are satisfied with their output (Jiang, 2022; Du & Daniel, 2024; Yang & Li, 2024). Although self-regulated learning may shape how students choose and sustain AI-based speaking tasks, CAESP in this study is limited to reported oral practice behavior rather than planning, monitoring or motivational regulation

1.2 Speaking Anxiety

Speaking anxiety involves tension, worry, physiological discomfort, and anticipatory fear of using English in speaking. In Chinese University contexts, anxiety seems to be the norm when students are asked to speak without preparation, to respond in front of their peers, to present formal content or to have the fear that errors will be publicly noticed. Recent studies conducted in China indicate that speaking anxiety relations with the expectancy beliefs, public speaking performance, and emotional experience in oral English tasks remain strong (Chen, 2024; Gao & Tay, 2023; Jin & Qin, 2024). Even if learners have adequate knowledge of vocabulary or grammar, fluency can be interrupted by anxiety, and learners will be less willing to communicate and their attention will be narrower when performing speaking tasks.

The most pertinent theoretical background is the affective filter perspective supplemented with recent research on achievement emotions. Recent AI talking studies still view anxiety as an obstruction that prevents access to the knowledge of language which has already been acquired by the learner. When anxiety is high, learners monitor themselves too much, are risk avoidant, are more concise in their response, and focus more on avoiding error than on meaning. In Chinese EFL contexts, this dynamic has been found not only in the classroom performance, but also in their beliefs about their own oral competence (Chen, 2024; Shang & Ma, 2024). Work on English public speaking has also revealed that self-efficacy and anxiety are strongly related, with increased levels of anxiety restricting performance and levels of confidence in oral tasks (Zhang & Ardasheva, 2019; Zhang et al., 2020).

1.3 Perceived Interaction Quality

Perceived interaction quality refers to the learner's judgment that AI mediated interaction is responsive, understandable, personalized and useful for getting the job done. In this manuscript, it does not tell how ChatGPT is technically designed but how students experience the interaction. A learner might interact with ChatGPT many times and still consider the interaction as weak when the responses are generic, confusing, repetitive, or lack context of the speaking task. By contrast, the same amount of practice may be more productive when the AI

provides the learner with clear hints, follows the topic in a logical way, responds rapidly to the learner, and adjusts feedback to the learner's level. This distinction is important in the context of language learning as the educational impact of a tool is not only influenced by access, but also the way in which the learners perceive the quality of the interaction they receive (Belda-Medina & Calvo-Ferrer, 2022; Alrajhi, 2024; Tram et al., 2024).

2. Literature Review

The recent literature can be summed up in three strands. The initial strand is related to AI and speaking practice. Research in this field suggests that conversational AI enhances opportunities for output, eliminating time and turn constraints found in classroom interactions, allowing for immediate feedback and repetition (Du & Daniel, 2024; Wu, 2024). These affordances are helpful in the context of speaking development in the sense that repeated production, reformulation and opportunities to continue speaking despite errors are all necessary to become proficient. The second strand is about AI and affective experience. The studies conducted in the Chinese context and other EFL contexts show that fear of negative evaluation, low confidence, and speaking anxiety are still significant obstacles to oral participation, and that AI-assisted rehearsal could reduce the interpersonal risk of practice (Chen, 2024; Gao & Tay, 2023; Wang & Wen, 2025). The third strand is related to AI and learner engagement. The line of work demonstrates that the access to AI alone does not benefit learners, but the access to AI that is perceived as coherent, responsive, supportive, and relevant to the task (Alrajhi, 2024; Tram et al., 2024).

There are a number of strands interwoven, but each have been studied individualistically. Some studies are concerned with the gains in performance without specifying the affective pathway. Others emphasize anxiety/willingness to communicate, but do not link these to an actual oral practice model involving ChatGPT. Another set of studies focuses on acceptance or perception of AI without modelling the quality change of the practice to outcome. The present model combines these strands by taking the behavior to be predicted, namely ChatGPT-assisted speaking practice, the affective mechanism, namely speaking anxiety, and the context, namely perceived interaction quality.

2.1 ChatGPT-Assisted English-Speaking Practice and English-Speaking Proficiency Development

The first of the relationships in this research is related to the direct effect of ChatGPT assisted speaking practice in English speaking proficiency development. A growing body of research suggests that conversational AI can contribute to oral language growth by increasing the amount of practice learners undertake, the speed with which they receive feedback and the number of times they can revise an answer before formal assessment. In their review of AI-powered chatbots for EFL speaking, Du and Daniel (2024) reported that chatbot-supported practice often led to improvements in confidence, engagement and speaking outcomes, especially when learners used the tools consistently over a period of time. Similar conclusions can be found in wider reviews of ChatGPT for L2 learning and empirical research on AI-supported EFL teaching (Yang & Li, 2024; Kundu & Bej, 2025; Wu, 2024).

Recent work from the Chinese context is also moving in the same direction. Wang and Wen (2025) found that the use of AI tools helped to improve speaking skills as well as reducing speaking anxiety. Xun et al (2025) reported that AI supported English public speaking fostered gains in speech contents and language quality when supported by self-regulated learning processes. D. Zhang et al. (2024) found some AI tools including chatbots and intelligent tutoring systems could induce more willingness to communicate that is very related to actual oral participation and practice volume. Even for the studies that discuss informal learning mediated by AI, rather than focusing exclusively on speaking, the results indicate that the learners benefit from autonomy, repetition, and feedback-rich environments, all of which are key to speaking development (Liu et al., 2024). Accordingly, the study expects more frequent ChatGPT-assisted oral practice to be positively associated with self-reported speaking development.

H1. ChatGPT-assisted English-speaking is positively associated with English speaking proficiency development.

2.2 ChatGPT-Assisted English-Speaking Practice and Speaking Anxiety

The second relationship is related to whether or not ChatGPT assisted speaking practice can reduce speaking anxiety. This has become one of the most discussed promises of supporting speaking through AI because many EFL learners avoid oral participation, not because they can't understand anything they need to say but because they are afraid of evaluation, embarrassment or communicative failure. In the Chinese university setting, that concern is still significant. Chen (2024) found a strong link between speaking anxiety, expectancy-value beliefs and oral proficiency in Chinese undergraduate students. Gao and Tay (2023) further revealed that Chinese university EFL learners use vivid and deeply embodied way of describing public speaking anxiety, and this is suggestive of the emotional saliency of oral performance.

AI-assisted practice of speaking may be able to take some of this pressure away. C. Zhang et al. (2024), a quasi-experimental study performed on Chinese university students, showed the use of an AI speaking assistant raised the enjoyment, lessened anxiety, and boosted the willingness to communicate. Wang and Wen (2025) showed a similar trend that AI tools offered learners a more relaxed and less judgmental environment that was linked to reductions in speaking anxiety and increases in motivation. Ebadi et al. (2025) also revealed the impact of AI in speaking practice by showing that AI-powered speaking tasks increased the speaking performance and reduced the anxiety, indicating that structured interaction with AI can positively redefine the emotional experience of speaking practice. In a different line of work D. Zhang et al. (2024) demonstrated the effectiveness of AI-mediated speaking environments to transition learners from hesitation to participation by raising their willingness to communicate.

The affective-filter perspective gives a theoretical reason for this relationship to exist. When the emotional cost of speaking decreases,

learners are more willing to make longer responses, lexical risks and persist in talking after making mistakes. ChatGPT-assisted practice, by transferring oral rehearsal from a public to more private environment, should therefore be effective in attenuating anxiety. Because the results of the recent literature consistently point in this direction the following hypothesis is propounded.

H2. ChatGPT-assisted English-speaking practice is negatively associated with speaking anxiety.

2.3 Speaking Anxiety and English-Speaking Proficiency Development

The third relationship deals with the effect of speaking anxiety on English speaking proficiency development. This path is well underpinned by applied linguistics and educational psychology. Learners with high tension when speaking often stop more, pause for a longer time, avoid difficult vocabulary, and invest their cognitive resources in self-protection instead of meaning making. As a result of this, anxiety can negatively affect the amount of oral output and its quality. Recent Chinese EFL studies do not overturn this pattern. Chen (2024) found that speaking anxiety is meaningfully associated with spoken English proficiency of Chinese undergraduates. Shang and Ma (2024) reported that classroom anxiety is negatively related to English achievement and that self-efficacy has an important role in the way such emotional states influence performance.

Research on English public speaking reinforced this explanation. Zhang and Ardasheva (2019) revealed that self-efficacy in the public-speaking domain of the English language is developed from mastery experience, persuasion, and others that are easily attenuated when learners are repeatedly exposed to anxiety involving speaking episodes. Zhang et al (2020) further discovered that self-efficacy is closely linked to English public speaking performance, implying that anxiety and confidence are collaborative in determining actual oral results. Zheng et al. (2023) also reported that a video-based formative practice may reduce public speaking anxiety and aid in better performance, which indirectly supports the view that reduced anxiety contributes to better oral development conditions. Jin and Qin (2024), through their work on the topic of English-speaking enjoyment, further suggests that affective experience is of great importance in the area of spoken English, where, as the feelings of enjoyment increase, the willingness to persist with a difficult speaking task also increases in learners.

Research on English public speaking similarly shows that anxiety, confidence and self-efficacy are closely related to oral performance. Theoretically, the relationship can be explained in terms of the cognitive cost of negative emotion in real-time speaking. Speaking is a demanding task which involves lexical retrieval, grammar selection, discourse planning and listener awareness under time pressure. Anxiety seeks to compete with those same attentional resources. When it becomes stronger, the efficiency of the learner's oral system decreases. On that basis, there is support for the expectation that speaking anxiety will have a negative effect on proficiency development in the literature.

H3. Speaking anxiety is negatively associated with self-reported English-speaking proficiency development.

2.4 Speaking Anxiety as a Mediator between ChatGPT-Assisted Practice and Proficiency Development

A direct effect of ChatGPT assisted speaking practice on proficiency development is plausible, but the literature also suggests that an important part of that effect may travel through speaking anxiety. In other words, ChatGPT practice may associated with stronger self-reported proficiency development in part because it provides a lesser-pressure environment in which students are able to speak more often and with less fear. This logic relates the above two hypotheses into a mediating process. AI speaking tools seem to have an effect not only in terms of increasing practice volume but also by varying the emotional condition under which it is practiced (C. Zhang et al., 2024; Wang & Wen, 2025; Ebadi et al., 2025).

Conceptually, the mediation goes as follows. ChatGPT-enabled practice provides a low-stakes rehearsal environment. In that setting, the learners can have a go at an answer, get reformulation, have another try and so on with the exchange without public embarrassment. As the feeling of threat disappears, the learners become more willing to speak at length and persistent in the face of errors. This increases useful practice and should increase proficiency development gradually. The mechanism therefore starts with practice, goes through emotional relief, and it ends in improved oral performance. On the basis of this, the fourth hypothesis is proposed.

H4. Speaking anxiety statistically mediates the relationship between ChatGPT-assisted English-speaking practice and self-reported English-speaking proficiency development.

2.5 Perceived Interaction Quality as a Moderator

The final relationship explores if the perceived quality of the interaction becomes a condition for the correlation between ChatGPT-assisted speaking practice and self-reported proficiency development. Since the construct is not conceptualized as a psychological state that is an outcome of practice, it is viewed as a moderator instead of a mediator. As a contextual assessment of the practice environment, it is seen to enhance or diminish the value of the same amount of practice. While two learners might spend the same amount of time rehearsing with ChatGPT, the educational impact would likely not be the same if one learner gets adaptive/coherent/task relevant responses, and the other gets generic or confusing responses.

This placement is based on the interactional approach to learning language with the assistance of AI. There is continuous turn taking in speaking practice, which involves negotiation, reformulating and making more attempts. These processes can be more continuous when perceived interaction quality is high, as learners are more likely to keep talking, ask further questions, edit their product and accept feedback. Low perceived interaction quality may make the exchange shorter and practice a mere retrieval of answers. Hence, interaction

quality is not merely a positive attitude toward ChatGPT. It is a contextual condition under which the same practice behavior may become more or less productive.

Therefore, the current model includes PIQ as moderator. If it were tested as direct antecedent or mediator, it would provide an answer to a different question, namely whether the perceived quality by itself explains the development of proficiency or the perceived quality comes about as a result of practice. The current study, however, poses the question as to whether practice is more likely to be correlated with the development of proficiency when the AI interaction is perceived as clear, responsive, supportive and relevant to speaking needs. This results in the fifth hypothesis.

H5. Perceived interaction quality moderates the relationship between ChatGPT-assisted English-speaking practice and self-reported English-speaking proficiency development.

2.6 Research Gap

Three gaps make the present study necessary. First, a lot of recent ChatGPT and AI language learning literature has studied adoption, writing assistance, overall perceptions, or broad reviews instead of a concentrated structural model of oral development in Chinese universities. Reviews by Jiang (2022), Yang and Li (2024) and Kundu and Bej (2025) confirm rapid growth in empirical work on AI language learning, but also reveal that empirical work is, at best, patchy in terms of language skills as well as theoretical mechanisms. Speaking has been receiving increasing attention, although the evidence is currently scattered among intervention studies, attitude studies and general reviews rather than integrated explanatory models.

Second, the emotional pathway is still not well developed. Several studies report on the effect of AI tools on the reduction of anxiety and improvement in speaking/ communication related outcomes (C. Zhang et al., 2024; Wang & Wen, 2025; Ebadi et al., 2025), but fewer studies test the mediating link between AI supported practice and proficiency development (i.e., anxiety). Without that mediating analysis, it is hard to tell if the value of ChatGPT is really more practice, less anxiety, better feedback or a little of all three. For the university learners in China, this is important because speaking anxiety is not peripheral but one of the central barriers to oral English development (Chen, 2024; Shang & Ma, 2024).

Third, the field has paid too little attention to the experience of the learners themselves of the interaction. Studies tend to look at whether students use ChatGPT or whether they accept it or whether there is an increase in performance following an intervention, but fewer explain when those gains will be stronger or weaker. Perceived interaction quality helps address that issue by asking whether the AI is responsive, understandable, and useful in actual speaking practice. Recent studies on conversational agent acceptance and ChatGPT acceptance have implied that these learner judgments affect sustained engagement and learning value (Belda-Medina & Calvo-Ferrer, 2022; Alrajhi, 2024; Tram et al., 2024), but moderation tests like these are still limited in speaking research of Chinese universities.

The current manuscript addresses these gaps by constructing one model linking ChatGPT-aided English-speaking practice, speaking anxiety, English speaking proficiency development and perceived interaction quality in the Chinese university context. This design brings together the interaction-based learning, affective explanation, and contextual variation in one framework.

A further gap relates to that of measurement and contextual specificity. Existing research often reports either short-term results of interventions or general student perceptions, but fewer studies present findings of that research within a measurement model, which incorporates reliability, validity, indirect effects, moderation, and predictive relevance in one design. To universities and journal reviewers, that's important because practical recommendations are more compelling when the emotional and instructional pathways are tested together as contrasted to separately-discussed. Chinese universities also provide a unique context in which a concern for face, orientation to exams, unequal access to opportunities for speaking, and fast adoption of digital methods may influence the way students respond to AI practice. A model which is sensitive to these contextual conditions can provide more helpful guidance than generalized claims about AI in language learning.

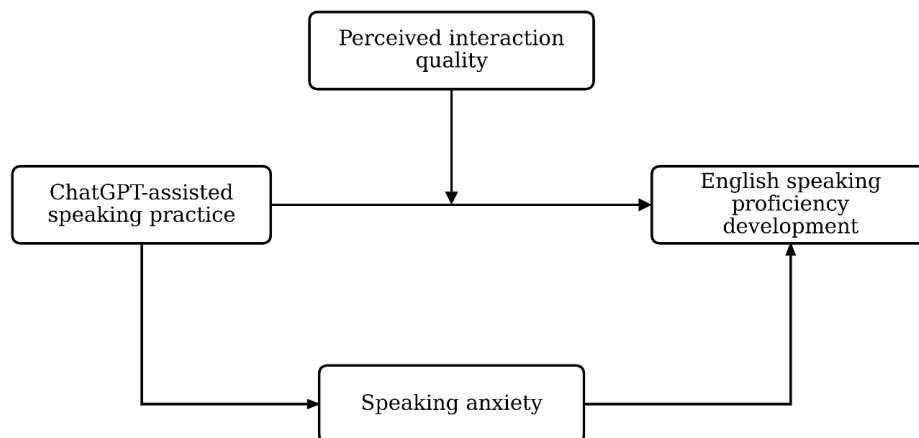


Figure 1. Theoretical Model

3. Methodology

This paper reports the quantitative phase of a sequential explanatory mixed-method design focused on Chinese university EFL learners. The target population of the study is undergraduate students taking compulsory or elective courses of English at Chinese universities. For the analysis, 412 student cases were collected as the final dataset, which is then applied for structural equation modeling in higher education language research. The sampling frame is based on six public universities located in eastern, central and western China. A demographic profile for the eventual field study is as follows: 55.3% female and 44.7% male; 27.2% first-year students, 30.1% second-year students, 25.5% third year students and 17.2% fourth year students; 38.8% English-related majors and 61.2% non-English majors. All respondents had used ChatGPT or a similar large language model that is available in Chinese for English speaking practice at least once a semester.

Four latent constructs were identified that had separate measurement boundaries. ChatGPT-assisted English-speaking practice (CAESP) was operationalized as the frequency and intensity of oral practice with ChatGPT or another large language model. Its indicators emphasized the frequency and the extent of students' use of AI in oral rehearsal, prompt-based speaking, mock dialogues, practice and revision of presentation, and spoken response revision. The construct did not include perceived usefulness, satisfaction, feedback quality or confidence, as these aspects of the evaluation are captured by the perceived quality of the interaction or their own self-reported proficiency development. Speaking anxiety (SA) focused on emotional distress during speaking English; nervousness, fear of error and tension in evaluation or public speaking situations. The operationalization of English-speaking proficiency development (ESPD) was in terms of self-reported development in the current semester in terms of fluency, clarity, vocabulary use and communicative confidence. It therefore must be understood as perceived proficiency development and not as an external rating of speaking. Perceived interaction quality (PIQ) was used to collect learner judgements of the quality of the interaction with the AI, such as: timely delivery, understandability, responsiveness, supportiveness and adapting to learner's speech needs.

The questionnaire items were all organized on a five-point Likert scale of 1 = strongly disagree to 5 = strongly agree. The wording of the items was based on recent research in chatbot-aided language learning, Chinese EFL speaking anxiety, and AI-aided language learning. PLS-SEM was selected because the primary goal of the quantitative part was not to confirm a well-established theory but to use a model for prediction purposes. The model includes a direct path, mediation path and moderation path and explanatory variance and predictive relevance are assessed by R^2 , Q^2_{predict} , RMSE and MAE. PLS-SEM is suitable for the latter design because it can be applied for latent-variable modelling with complex structural relationships and is centered on the magnitude and stability of path estimates, explained variance and predictive usefulness (Hair et al., 2019). If overall model fit were a primary goal, proving a well-developed theory, then covariance-based SEM would be more appropriate. This study takes ChatGPT-assisted speaking practice in Chinese universities as an emerging research context in which prediction, path stability and exploratory theoretical development dominate.

Before the full instrument was finalized, the design was conducted with a pilot check of 32 Chinese university students for the clarity of the wording and the relevance of the items. Minor wording adjustments were done to make sure that "speaking practice" is interpreted in a consistent way as oral rehearsal instead of text chat only. The final questionnaire was distributed online via University English course groups and student learning platforms. Participation is voluntary and informed consent was being made aware before students could access the questionnaire.

Common method bias was taken into account due to the fact that all quantitative variables were measured using the same questionnaire at the same time. Several procedural safeguards were used: Voluntary participation, Avoidance of ambiguous wording, Separation of constructs in the questionnaire, and Pilot testing to reduce ambiguity. However, in a single-source survey, common method variance cannot be completely eliminated (Podsakoff et al., 2003; Kock, 2015). Therefore, the findings are considered as associations between the self-reported constructs and not as causal evidence.

The quantitative analysis is done in five steps. Descriptive checks of item distributions are done first. Secondly, Cronbach's alpha, Dijkstra-Henseler's rho, Jöreskog's rho and average variance extracted are used to check for reliability and convergent validity. Third, HTMT and Fornell-Larcker criterion are used to evaluate the discriminant validity. HTMT remains a test because it is sensitive to the separation of constructs in variance-based SEM, particularly when constructs are theoretically related and require empirical separation (Henseler et al., 2015). Fourthly, as to the direct path, the mediation path and the moderation path, it is tested in the structural model. Fifth: Q^2_{predict} , RMSE and MAE are used to measure the predictive relevance.

3.1 Questionnaire Profile

Table 1 presents the questionnaire profile. The construct structure was kept concise because the goal was to ensure adequate reliability and content coverage for each variable. To minimize construct overlap, the questionnaire separated practice behavior, emotional response, outcome perception and interaction judgment at the item level. CAESP items were presented as oral practice behavior, items of PIQ as judgements about the quality of interaction, items of SA as emotional reaction and items of ESPD as perceived changes in speaking ability. The separation was verified during pilot testing with 32 students, and words were adjusted to ensure that the reference of "speaking practice" was not just text chat but oral rehearsal.

Table 1. Questionnaire profile

Variable	Items	Scale
ChatGPT-assisted English-speaking practice	Four	(Belda-Medina & Calvo-Ferrer, 2022; Tram et al., 2024; Muniandy & Selvanathan, 2025)
Speaking anxiety	Four	(Chen, 2024; Gao & Tay, 2023)
Perceived interaction quality	Five	(Alrajhi, 2024; Tram et al., 2024; Belda-Medina & Calvo-Ferrer, 2022)
English speaking proficiency development	Four	(Wang & Wen, 2025; Xun et al., 2025; C. Zhang et al., 2024)

4. Results

Results section is the part of the paper that reports the statistical results obtained from the sample of 412 cases. As the model takes one direct predictor, one mediator, one moderator and one focal outcome into account, the findings are shown in: reliability and validity, model estimation, discriminant validity, effect overview, model fit and predictive relevance and structural path analysis.

4.1 Variables Reliability and Validity

Reliability and convergent validity of the 4 core constructs are reported in Table 2. The values are satisfactory in all measures. ChatGPT-Assisted English-Speaking Practice Cronbach's $\alpha = 0.8758$, Dijkstra-Henseler's $\rho = 0.8778$, Joreskog's $\rho = 0.8816$, AVE=0.6435 These values provide evidence of good internal consistency and acceptable convergent validity. Speaking anxiety is also a good measurement with $\alpha = 0.8692$, $\rho_A = 0.8703$, $\rho_C = 0.8742$, and AVE = 0.6277. English speaking proficiency development follows the same threshold pattern, with $\alpha = 0.8666$, $\rho_A = 0.8676$, $\rho_C = 0.8711$ and AVE = 0.6215. Perceived interaction quality has the highest overall consistency in the model and the values are $\alpha = 0.8792$, $\rho_A = 0.8807$, $\rho_C = 0.8849$, and AVE = 0.5972.

Taken together, the results indicate that the measurement structure is reliable enough for use in structural analysis. All alpha values are greater than the traditional value of 0.70 and all AVE values are greater than 0.50, indicating that these constructs all capture more than half of the variance in their respective indicators. This is important for the present study because the model is conceptually reliant on but empirically distinct constructs. Without good measurement, it would be very hard to distinguish the influence part of ChatGPT assisted practice from the emotional pathway of anxiety or the contextual contribution of perceived interaction quality.

Table 2. Variables reliability and validity

Construct	Dijkstra-Henseler's rho (ρ_A)	Jöreskog's rho (ρ_C)	Cronbach's alpha (α)	AVE
ChatGPT-assisted speaking practice	0.8778	0.8816	0.8758	0.6435
Speaking anxiety	0.8703	0.8742	0.8692	0.6277
English speaking proficiency development	0.8676	0.8711	0.8666	0.6215
Perceived interaction quality	0.8807	0.8849	0.8792	0.5972

4.2 Estimated Model

The estimated structural model is summarized in Figure 2. The results align with the hypothesized directions as ChatGPT-assisted speaking practice is negatively associated with speaking anxiety and positively associated with self-reported proficiency development. The interaction term of ChatGPT-assisted practice and perceived interaction quality is positive, while the speaking anxiety is negative with self-reported proficiency development. This figure reflects, therefore, an association-based view in which the three factors work together: practice behaviour, emotional experience and perceived quality of the interaction.

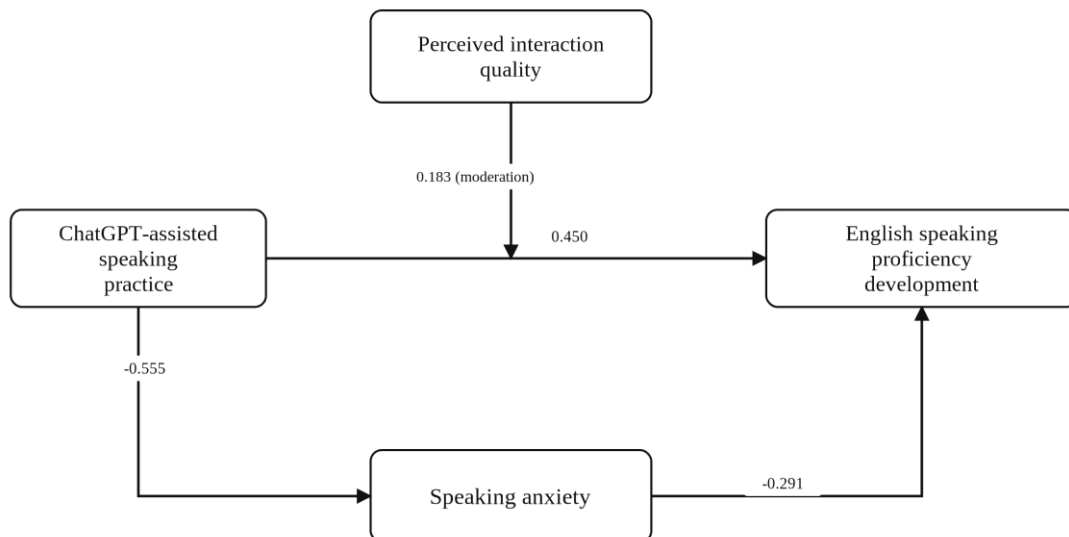


Figure 2. Estimated Model

4.3 Measurement Items Fitness Statistics

Table 3 presents the standardized loadings for all of the measurement items. The four indicators of English-speaking practice with ChatGPT score between 0.7199 and 0.8817, indicating that each of them contributes meaningfully to the construct. The speaking anxiety items have item-construct alignment scores that range from 0.7210 to 0.8593, again indicating strong item-construct alignment. English speaking proficiency development indicators are highly varied from 0.7466 to 0.8232 while the perceived interaction quality indicators also range from 0.7197 to 0.8547.

These item level values are important as the model focuses on recent, closely related constructs. ChatGPT assisted practice and perceived quality of interaction could certainly overlap conceptually if the indicators were weak or not well targeted. From the loadings, we know that the current structure does not have that problem. Each item cluster makes good sense and no loading is so low that it threatens the interpretation of the construct. This provides additional support to the use of questionnaire in future field data collection in Chinese universities.

Table 3. Measurement Items Fitness Statistics

Indicator	ChatGPT-assisted speaking practice	Speaking anxiety	English speaking proficiency development	Perceived interaction quality
CAESP1	0.8243			
CAESP2	0.8817			
CAESP3	0.7739			
CAESP4	0.7199			
SA1		0.8331		
SA2		0.8593		
SA3		0.7474		
SA4		0.7210		
ESPD1			0.8232	
ESPD2			0.8102	
ESPD3			0.7711	
ESPD4			0.7466	
PIQ1				0.8547
PIQ2				0.8078
PIQ3				0.7414
PIQ4				0.7197
PIQ5				0.7316

4.4 Discriminant Validity

Discriminant validity was tested in two complementary ways. The HTMT matrix is reported in Table 4. All the values are below the conservative threshold value 0.85. The highest HTMT value occurs between ChatGPT assisted English speaking practice and English-speaking proficiency development (0.7260), which is expected because the theoretical relationships exist between those constructs. The HTMT values between speaking anxiety and the positive constructs are moderate to low, so speaking anxiety is related but indistinguishable from these other variables. The lowest HTMT value is between speaking anxiety and perceived interaction quality (0.1467), which indicates very clear separation between the emotional construct and the contextual quality construct.

Table 5 augments the HTMT results using the Fornell-Larcker criterion. The square root of AVE on each diagonal is greater, in absolute terms, than the corresponding off diagonal correlations. This corroborates the conclusion that each construct is more similar in terms of variance to its own indicators than to other latent variables in the model. Together, the two checks ensure that the measurement model is sufficiently different for structural testing.

Table 4. Discriminant Validity: Heterotrait-Monotrait Ratio of Correlations (HTMT)

Construct	1	2	3	4
ChatGPT-assisted speaking practice				
Speaking anxiety	0.6355			
English speaking proficiency development	0.7260	0.6480		
Perceived interaction quality	0.3470	0.1467	0.2705	

Note: 1 = ChatGPT-assisted speaking practice; 2 = Speaking anxiety; 3 = English speaking proficiency development; 4 = Perceived interaction quality.

Table 5. Discriminant Validity: Fornell-Larcker Criterion

Construct	1	2	3	4
ChatGPT-assisted speaking practice	0.8022			
Speaking anxiety	-0.5554	0.7923		
English speaking proficiency development	0.6317	-0.5637	0.7883	
Perceived interaction quality	0.3054	-0.1293	0.2357	0.7728

Note: Diagonal entries are the square roots of AVE. Off-diagonal entries are latent correlations.

4.5 Variables Effects Overview

Table 6 presents a summary of the key direct, indirect, and total effects. ChatGPT-assisted English-speaking practice. Negative path coefficient of ChatGPT on Speaking Anxiety ($\beta = -0.5554$) Moderate to strong effect size ($f^2 = 0.4461$) On English speaking proficiency development, ChatGPT-assisted practice has a direct effect 0.4497 and an indirect effect 0.1615 through lowered speaking anxiety resulting in an overall effect of 0.6112. This suggests that the gain of Chat GPT practice not only comes immediately but it is channeled through an emotional mechanism. Speaking anxiety has a negative direct effect on proficiency development ($\beta = -0.2908$), and perceived interaction quality has a positive effect on proficiency development ($\beta = 0.1827$), the interaction between ChatGPT-assisted practice and perceived interaction quality. In other words, the association between practice and self-reported proficiency development is stronger when perceived interaction quality is higher.

This general pattern has a theoretical meaning. The greatest influence of all in the model is ChatGPT-assisted speaking practice, which aligns with the belief that repeated low-stakes oral production is the key driver of development. However, the power of the indirect path proves that it is impossible to neglect emotional conditions. The moderation effect, unlike the direct practice effect, though smaller in strength, suggests that the same amount of practice might have different results depending on the way students judge the quality of the interaction.

Table 6. Variables effects overview

Effect	Beta	Indirect effects	Total effect	Cohen's f^2
ChatGPT-assisted speaking practice -> Speaking anxiety	-0.5554		-0.5554	0.4461
ChatGPT-assisted speaking practice -> English speaking proficiency development	0.4497	0.1615	0.6112	0.2586
Speaking anxiety -> English speaking proficiency development	-0.2908		-0.2908	0.1165
ChatGPT-assisted speaking practice \times Perceived interaction quality -> English speaking proficiency development	0.1827		0.1827	0.0714

4.6 R-squared Statistics Model Goodness of Fit Statistics

Table 7 shows the variance explained in the endogenous constructs. The model explained the variance of 30.85% on speaking anxiety and 50.34% on English speaking proficiency development. The adjusted R^2 values are not very different from the raw R^2 values; thus, the amount of explained variance does not seem to be exaggerated by the addition of unnecessary predictors. The predictive relevance is also acceptable. Q^2 predict obtains 0.3015 for speaking anxiety and 0.4894 for proficiency development. RMSE and MAE values are within a reasonable range for a social science model that is constructed with latent constructs.

These values imply a useful amount of explanatory power for the present topic. Speaking proficiency development is influenced by many factors which are not included here, such as prior proficiency, teacher support, peer interaction, course design and access to oral English outside class. Explaining just over half the variance with four focal constructs is therefore significant. The equation of anxiety also has meanings because the results of emotions are often influenced by many personal and contextual conditions. The current model still accounts for almost one third of that variance from ChatGPT-assisted speaking practice alone.

Table 7. R-square statistics Model Goodness of Fit Statistics

Construct	Coefficient of determination (R^2)	Adjusted R^2	Q^2 predict	RMSE	MAE
Speaking anxiety	0.3085	0.3068	0.3015	0.8358	0.6728
English speaking proficiency development	0.5034	0.4985	0.4894	0.7146	0.5635

4.7 Structural Model for Path Analysis

The structural model for path analysis is given in Figure 3. The diagram illustrates one exogenous predictor, one mediator, one moderator, and one focal outcome. The signs of the coefficients are in the theorized direction. ChatGPT-assisted English-speaking practice is statistically linked to lower speaking anxiety and is directly associated with stronger self-reported proficiency development. Speaking anxiety decreases proficiency development. Perceived interaction quality builds the practice-to-proficiency path. The above figure therefore offers a coherent perspective of functioning of AI supported oral practice in Chinese university English learning.

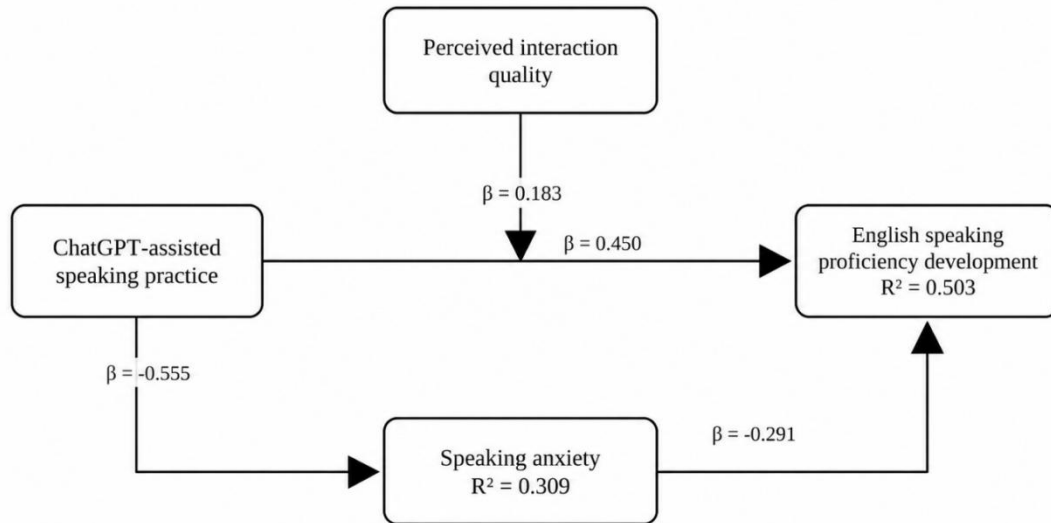


Figure 3. Structural Model for Path Analysis

4.8 Path Analysis

The bootstrapped path analysis is presented in Table 8. H1 is supported: ChatGPT-assisted English-speaking practice is positively associated with self-reported English-speaking proficiency development ($\beta = 0.450, t = 10.486, p < 0.001$). H2 is supported: ChatGPT-assisted English-speaking practice is negatively associated with speaking anxiety ($\beta = -0.555, t = 14.295, p < 0.001$). H3 is also supported: speaking anxiety is negatively associated with self-reported proficiency development ($\beta = -0.291, t = 7.105, p < 0.001$).

H4 is supported by a significant indirect path from ChatGPT-assisted speaking practice to self-reported proficiency development through speaking anxiety ($\beta = 0.162, t = 6.311, p < 0.001$). This indicates that lower speaking anxiety is statistically involved in the association between practice and perceived proficiency development. H5 is also supported. The interaction term between ChatGPT-assisted practice and perceived interaction quality is positively associated with self-reported proficiency development ($\beta = 0.183, t = 4.744, p < 0.001$). In practical terms, the positive practice to proficiency association is stronger when learners judge the AI interaction as clear, responsive and useful.

Overall, the path analysis is consistent with the proposed conceptual model. It does not establish causality, but it shows that behavioral practice, anxiety and perceived interaction quality are significantly related in the expected directions. The findings therefore support an association-based interpretation of ChatGPT-assisted oral practice in Chinese university EFL learning.

Table 8. Path Analysis

	Original Sample	STDEV	T Statistics	P Values
ChatGPT-assisted English-speaking is positively associated with English speaking proficiency development.	0.450	0.0429	10.486	0.000
ChatGPT-assisted English-speaking practice is negatively associated with speaking anxiety.	-0.555	0.0389	14.295	0.000
Speaking anxiety is negatively associated with self-reported English-speaking proficiency development.	-0.291	0.0409	7.105	0.000
Speaking anxiety statistically mediates the relationship between ChatGPT-assisted English-speaking practice and self-reported English-speaking proficiency development.	0.162	0.0256	6.311	0.000
Perceived interaction quality moderates the relationship between ChatGPT-assisted English-speaking practice and self-reported English-speaking proficiency development.	0.183	0.0385	4.744	0.000

5. Discussion

The results reveal some support for the association-based explanation of ChatGPT-assisted speaking practice in Chinese universities. There are positive direct and indirect links between ChatGPT-assisted English-speaking practice and English-speaking proficiency development as measured by self-efficacy, while the quality of perceived interaction is a moderator between the practice of ChatGPT and English-speaking proficiency development. The findings must not be interpreted as evidence of proficiency growth or reduction in anxiety as a result of ChatGPT. Instead, they suggest that the more often and intensively students engage with ChatGPT through oral tasks, the less anxiety they experience in their oral communication and the more they feel they have improved, particularly if they find the interaction with the AI clear and responsive, and if the tasks are relevant..

H1 proposed that ChatGPT facilitated English speaking practice would positively affect proficiency development. The results support this proposition. This finding is consistent with the recent research that AI-supported conversation environments allow for more opportunities of output, repeating, and revision (Du & Daniel, 2024; Yang & Li, 2024; Wang & Wen, 2025; Xun et al., 2025). It is also consistent with

the case-based evidence that ChatGPT can serve as an effective conversational partner in a flipped or self-directed learning setting (Muniandy & Selvanathan, 2025). In the Chinese university environment, where available time for formal speaking may be limited, the value of AI may be in increasing the frequency of practice rather than as a substitute for classroom teaching. Repeated practice with responsive prompts seems to produce cumulative growth in fluency, organization and communicative confidence.

H2 proposed that ChatGPT-assisted speaking practice would decrease speaking anxiety and the results point strongly in that direction. This is consistent with the recent Chinese evidence that AI speaking assistants as well as tools related to them can reduce foreign language anxiety while increasing enjoyment and willingness to communicate (C. Zhang et al., 2024; Wang & Wen, 2025). The result is also consistent with the larger body of review evidence summarized by Zhang and Liu (2025) and Torres and Kahveci (2025) who say that AI-assisted language learning often creates a more positive emotional climate than traditional high pressure speaking environments. For Chinese university learners, we see this is important because oral performance is often entangled with visibility in public, and fear of embarrassment. ChatGPT-assisted practice appears to provide an intermediate space between silence and showing off in front of the class, which gives students an opportunity to speak without the full burden of social evaluation.

H3 stated that speaking anxiety would have negative effects on proficiency development and the expectation is confirmed as well. This result turns out to be very consistent with studies of Chinese undergraduates and English public speaking, where anxiety is also a consistent predictor of weaker performance and lower confidence (Chen, 2024; Gao & Tay, 2023; Shang & Ma, 2024; Zhang et al., 2020). The present model contributes to that literature by situating the concept of anxiety in an AI-assisted learning model. Rather than considering speaking anxiety an exclusively class-based phenomenon, the results appear to indicate that speaking anxiety continues to be a significant factor even when learners transfer to a technology-supported speaking practice. What changes is the possibility that AI-supported environments is statistically linked to lowering of anxiety, rather than the importance of anxiety.

H4 suggested that speaking anxiety would mediate the link between ChatGPT-assisted speaking practice and proficiency development. The strong indirect influence confirms that emotional change is one way that ChatGPT practice contributes to oral growth. This result is especially important because it makes it clear why speaking with the help of AI might be effective. The advantage of this is not just that students speak more frequently, it is that they are speaking in less threatening circumstances. That combination makes it possible to make practice more productive. This interpretation is consistent with the pattern reported in AI-speaking studies where there is a co-occurrence of emotional gains and performance gains (C. Zhang et al., 2024; Wang & Wen, 2025; Ebadi et al., 2025). In practical terms, any implementation of ChatGPT for speaking at a university should, therefore, pay attention not only to how many practice tasks are put, but also how the students experience these practice tasks as being emotionally manageable.

H5 suggested that perceived interaction quality would enhance the impact of ChatGPT-assisted speaking practice on proficiency development. The positive moderation result provides support for this claim and extends existing literature in a useful way. Much research in the last several years describes the benefits of AI interaction, but fewer studies explore the role of learner judgments of the interaction. The present finding suggests that oral practice is more effective if students feel that ChatGPT responds without delay, stays on topic, provides understandable feedback and adjusts to the level of the student or the need of the task. This is in line with earlier research on conversational agents and self-learning acceptance which highlights usefulness, clarity and learner centered responsiveness (Belda-Medina and Calvo-Ferrer, 2022; Alrajhi, 2024; Tram et al., 2024). To put it all together, AI availability is not sufficient. The interaction should be worthwhile.

These findings all point to an interaction-affect-context explanation of an AI-assisted speaking development. The predictor is actual ChatGPT speaking practice by the learner, the mechanism is linked with lowering of speaking anxiety and the contextual amplifier is perceived quality of interaction. This integrated view is valuable because of the fact that current debates on AI in language education tend to oscillate between celebration and skepticism at a very fast pace. The present model can be a more specific conclusion: ChatGPT assisted speaking practice can be educationally productive in Chinese universities when it expands practice, lowers emotional barriers, and provides interactions perceived by learners as genuinely useful.

The study has a number of implications for theory, pedagogy and future research. Theoretically, it brings together three threads of recent scholarship which are far too often treated in isolation: interaction-based explanations of AI assisted language practice, affective explanations of oral performance and learner-centered judgments of AI quality. By putting these strands in one structural model, the manuscript suggests that AI-supported speaking development is not just a case of access to technology. It is a question of altering the conditions of practice and learner emotion in light of changes in technology. This adds to the current advances in AI-assisted language learning research to shift from descriptive accounts of tool use to descriptive models of learning processes (Crompton et al., 2024; Wu, 2024; Kundu and Bej, 2025).

Pedagogically, the results indicate that Chinese university teachers should not simply treat ChatGPT as a tool for helping writing or generating ideas. When carefully guided, it may also be used as a low-stakes oral rehearsal partner. Teachers can create mini tasks for role play, interviews, presentations, impromptu speech and vocabulary building dialogues that students can complete with ChatGPT prior to class. Such tasks may be particularly helpful for the quieter student who abstains from participating in the oral discussions. However, the moderation result also shows that task design is important. Teachers must equip students with prompt templates and feedback routines to make the AI interaction seem better in quality. Without that support students can fall into shallow or repetitive exchanges.

Institutionally, there may be ways for language centers and English departments in Chinese universities to incorporate ChatGPT-enabled

speaking practice into blended learning systems. This does not require that teachers be replaced. On the contrary, teacher guidance is still important in supporting students in judging AI responses, recognizing inaccurate feedback, transferring AI-based rehearsal to actual classroom communication. The likely value of ChatGPT is not in supplanting human judgment or classroom interaction, but in increasing the amount of practice time and emotional safety. Departments can therefore utilize ChatGPT as an additional oral practice environment.

There are also ethical and administrative implications. Universities that encourage speaking work with ChatGPT should have clear guidelines about privacy, about what is appropriate to disclose, and the limitations of AI feedback. Students require to know that conversational AI has the ability to offer support in the form of rehearsal, yet it may not be inaccurate, or it may give generic and flat replies with a cultural context. Teachers are therefore still at the centre of defining criteria for speaking, monitoring transfer from AI practice to classroom performance and guiding students in knowing what constitutes useful feedback and what constitutes weak feedback. Institutional support in the form of short training sessions, model prompts, and departmental policy can reduce the confusion and help the students use AI for speaking practice in a disciplined way.

6. Limitations and Future Research

There are some caveats to be noted. First, the quantitative data are cross-sectional and based on self-report. The results thus reflect statistically significant relationships and not cause and effect changes. More longitudinal and experimental studies should be conducted to see if the use of ChatGPT for speaking practice leads to a lasting decrease in anxiety or to an improvement in the students' speaking as rated by others. Second, the progress in English speaking was assessed in terms of students' subjective perceptions about the improvement of their speaking fluency, comprehensibility, vocabulary and confidence in speaking and communication during the semester. This is significant with regard to understanding the learner's experience, but should not be seen as a synonym for objective speaking proficiency. Teacher ratings, standardized oral tasks, speech recordings, pronunciation measurements and fluency measurements should be included in future studies.

Third, the measures of practice, anxiety, interaction quality and proficiency development were filled out by the same respondents at one time point, which may have resulted in common method variance in the observed associations. The study attempted to mitigate this risk by separating the constructs, using the pilot checking and wording of the neutral items, but cannot be taken away completely. In future studies multiple data sources and time separated measurement should be incorporated to further minimize method bias. Fourth, the study concentrates on the Chinese University students who had applied ChatGPT or any other large language model to speaking exercise. The results may not be directly applicable to other levels of education, institutional contexts and AI systems. Future research should involve the use of different AI platforms; different types of speaking tasks and teacher guided implementation models.

7. Conclusion

This manuscript explored the connection between ChatGPT-assisted English-speaking practice, speaking anxiety and self-reported English-speaking proficiency development in Chinese universities, and the interaction quality was proposed as a moderator. All five hypotheses were supported at the associational level. Those who mentioned using ChatGPT to practice speaking more often and deeply also indicated that they had less fear of speaking and were more convinced that their speaking level had improved. Speaking anxiety was a statistically intervening variable between the practice to proficiency relationship, and perceived interaction quality was a supporting variable. The findings indicate that AI-oral practice is most likely to be beneficial for learners when used frequently, when they feel emotionally supported by the practice, and when they feel they have a clear, responsive and supportive experience with the AI. The study provides a narrow model of AI-supported speaking practice for students in higher education institutions in the context of Chinese language education, but it is important to note that further longitudinal and performance studies are required to confirm the causal and objective speaking outcomes.

Declarations

Use of AI-Assisted Tools

During revision, the authors used ChatGPT for language editing, text structuring and clarity checking. The tool was not used to generate data, fabricate or select references, conduct statistical analysis, interpret findings or determine authorship. All AI-assisted outputs were reviewed, edited and verified by the authors, who remain fully responsible for the accuracy, integrity and originality of the manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisors for their invaluable guidance, patience, and continuous encouragement throughout my doctoral journey. Their insightful advice and academic support greatly contributed to the completion of this study.

I would also like to extend my sincere appreciation to all the students who participated in this research. Without their cooperation and support, this study would not have been possible.

My heartfelt thanks also go to Universiti Kebangsaan Malaysia for providing the academic environment and research resources necessary for this study.

Finally, I would like to express my deepest gratitude to my family and friends for their endless love, understanding, and encouragement throughout this challenging journey.

Author contributions

CUI QINGQING was responsible for the study design, literature review, data collection, data analysis, and manuscript drafting. Prof. Supyan supervised the research, provided methodological guidance, and revised the manuscript. All authors read and approved the final manuscript.

Funding

Not applicable.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

Ethical approval for this study was obtained from the Research Ethics Committee of Universiti Kebangsaan Malaysia.

ETHICAL NO: JEP-2025-1275

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

- Alrajhi, A. S. (2024). Artificial intelligence pedagogical chatbots as L2 conversational agents. *Cogent Education*, 11(1), 2327789. <https://doi.org/10.1080/2331186X.2024.2327789>
- Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, 12(17), 8427. <https://doi.org/10.3390/app12178427>
- Chen, Z. (2024). A study of Chinese undergraduate students' English language speaking anxiety, expectancy-value beliefs and spoken English proficiency. *SAGE Open*, 14(1), 21582440231219312. <https://doi.org/10.1177/21582440231219312>
- Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*, 55(6), 2503-2529. <https://doi.org/10.1111/bjet.13460>
- Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, 6, 100230. <https://doi.org/10.1016/j.caeai.2024.100230>
- Ebadi, S., Velayati, S., Ramezanzadeh, A., & Salman, A. R. (2025). Exploring the impact of AI-powered speaking tasks on EFL learners' speaking performance and anxiety: An activity theory study. *Acta Psychologica*, 259, 105391. <https://doi.org/10.1016/j.actpsy.2025.105391>
- Gao, F., & Tay, D. Z. (2023). Metaphor use in describing English public speaking anxiety by Chinese university EFL learners. *System*, 118, 103091. <https://doi.org/10.1016/j.system.2023.103091>
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/10.1108/eb-11-2018-0203>

- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115-135. <https://doi.org/10.1007/s11747-014-0403-8>
- Jiang, R. (2022). How does artificial intelligence empower EFL teaching and learning nowadays? A review on artificial intelligence in the EFL context. *Frontiers in Psychology*, 13, 1049401. <https://doi.org/10.3389/fpsyg.2022.1049401>
- Jin, Y., & Qin, L. (2024). Examining Chinese university students' English speaking enjoyment: Scale development and validation. *System*, 124, 103382. <https://doi.org/10.1016/j.system.2024.103382>
- Klimova, B., & Ibna Seraj, P. M. (2023). The use of chatbots in university EFL settings: Research trends and pedagogical implications. *Frontiers in Psychology*, 14, 1131506. <https://doi.org/10.3389/fpsyg.2023.1131506>
- Kock, N. (2015). Common method bias in PLS-SEM. *International Journal of e-Collaboration*, 11(4), 1-10. <https://doi.org/10.4018/ijec.2015100101>
- Kundu, A., & Bej, T. (2025). Transforming EFL teaching with AI: A systematic review of empirical studies. *International Journal of Artificial Intelligence in Education*, 35, 2281-2314. <https://doi.org/10.1007/s40593-025-00470-0>
- Liu, G. L., Darwin, R., & Ma, C. (2024). Unpacking the role of motivation and enjoyment in AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation in the Chinese context. *Computers in Human Behavior*, 160, 108362. <https://doi.org/10.1016/j.chb.2024.108362>
- Muniandy, J., & Selvanathan, M. (2025). ChatGPT, a partnering tool to improve ESL learners' speaking skills: Case study in a public university, Malaysia. *Teaching Public Administration*, 43(1), 4-20. <https://doi.org/10.1177/01447394241230152>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Shang, Y., & Ma, L. (2024). Classroom anxiety, learning motivation, and English achievement of Chinese college students: The mediating role of self-efficacy. *Acta Psychologica*, 251, 104550. <https://doi.org/10.1016/j.actpsy.2024.104550>
- Torres, P. J., & Kahveci, Y. E. (2025). Effectiveness of artificial intelligence (AI) in language teaching. *Computers and Education: Artificial Intelligence*, 9, 100522. <https://doi.org/10.1016/j.caeai.2025.100522>
- Tram, N. H. M., Nguyen, T. T., & Tran, C. D. (2024). ChatGPT as a tool for self-learning English among EFL learners: A multi-methods study. *System*, 127, 103528. <https://doi.org/10.1016/j.system.2024.103528>
- Wang, X., & Wen, Y. (2025). The role of artificial intelligence tools on Chinese EFL learners' speaking skills, speaking anxiety, and motivation. *Acta Psychologica*, 261, 105957. <https://doi.org/10.1016/j.actpsy.2025.105957>
- Wu, X. Y. (2024). Artificial intelligence in L2 learning: A meta-analysis of contextual, instructional, and social-emotional moderators. *System*, 126, 103498. <https://doi.org/10.1016/j.system.2024.103498>
- Xun, H., Zhou, C., Li, Y., & Wu, Y. (2025). AI-supported English public speaking in the Chinese EFL context: Insights from self-regulated learning and positive psychology. *Learning and Motivation*, 92, 102211. <https://doi.org/10.1016/j.lmot.2025.102211>
- Yang, L., & Li, R. (2024). ChatGPT for L2 learning: Current status and implications. *System*, 124, 103351. <https://doi.org/10.1016/j.system.2024.103351>
- Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System*, 121, 103259. <https://doi.org/10.1016/j.system.2024.103259>
- Zhang, D., Wu, J. G., & Fu, Z. (2024). From shy to fly: Facilitating EFL learners' willingness to communicate with an AI chatbot and an intelligent tutoring system. *System*, 127, 103501. <https://doi.org/10.1016/j.system.2024.103501>
- Zhang, S., & Liu, X. (2025). Learner emotions in AI-assisted English as a second/foreign language learning: A systematic review of empirical studies. *Frontiers in Psychology*, 16, 1652806. <https://doi.org/10.3389/fpsyg.2025.1652806>
- Zhang, X., & Ardasheva, Y. (2019). Sources of college EFL learners' self-efficacy in the English public speaking domain. *English for Specific Purposes*, 53, 47-59. <https://doi.org/10.1016/j.esp.2018.09.004>
- Zhang, X., Ardasheva, Y., & Austin, B. W. (2020). Self-efficacy and English public speaking performance: A mixed method approach. *English for Specific Purposes*, 59, 1-16. <https://doi.org/10.1016/j.esp.2020.02.001>
- Zheng, C., Wang, L., & Chai, C. S. (2023). Self-assessment first or peer-assessment first: Effects of video-based formative practice on learners' English public speaking anxiety and performance. *Computer Assisted Language Learning*, 36(4), 806-839. <https://doi.org/10.1080/09588221.2021.1946562>