

The Role of Chatbots in Foreign Language Speaking Anxiety

Liang Zhang^{1*}

¹School of Foreign Languages and Cultures, Guangdong University of Finance, Guangzhou, Guangdong, 510630, China

Abstract. With the development of society, the importance of foreign language learning has become increasingly prominent. Foreign language speaking anxiety, as the main obstacle to improving oral skills, is a problem that most learners encounter. Artificial intelligence provides a solution to this pain point - chatbots. Most studies have found that chatbots, by interacting with learners in the target foreign language, can provide learners with an independent and stress-free practice space, effectively alleviating speaking anxiety. However, some studies have shown that when the technology is not up to standard and the usage frequency is insufficient, it has no effect on reducing speaking anxiety or even has side effects. This study compared and analysed these contradictory studies. It found that the impact of chatbots on learners' speaking anxiety depends on the individual situation and whether the technology design can meet the needs of specific situations. Multiple factors can affect the impact of chatbots on learners' speaking anxiety levels, such as learners' digital literacy, initial anxiety level, technology quality, task design, and usage frequency.

1 Introduction

In today's increasingly globalized world, the importance of foreign language learning has become increasingly important. However, many learners encounter a common problem during the process of learning a foreign language - Foreign Language Speaking Anxiety. This anxiety manifests as nervousness, fear of making mistakes, and worry about being judged by others. Seriously hinders the learners' enthusiasm for oral practice and their language output ability [1]. In recent years, with the rapid development of artificial intelligence technology, chatbots have emerged and become increasingly popular. Chatbots can interact with learners in the target foreign language and provide a relaxed and stress-free language practice environment for learners. Existing studies show that learners can have one-on-one conversations with chatbots without worrying about being judged or ridiculed by others. This interaction method helps learners gradually overcome speaking anxiety, build confidence, and participate more actively in oral practice [2-5]. Some studies have pointed out that there is no significant difference in reducing anxiety between learners interacting with ChatGPT and watching YouTube [6]. Susoy found that although chatbots reduced the objective anxiety level, participants still preferred human evaluators subjectively [7]. Some studies have also

* Corresponding author: syersha518@gmail.com

pointed out that technical flaws can damage the learning experience and lead to the lack of emotional support, thereby exacerbating the stress of learners [8]. However, what causes this contradictory finding is rarely analysed by current research, which is very meaningful for how to correctly and efficiently use chatbots to reduce learners' speaking anxiety.

This study is based on empirical research in the EFL context of the Middle East and Asia from 2022 to 2025. This study evaluated the practical effectiveness of chatbots in reducing Foreign Language Speaking Anxiety. The aim was to reveal the mechanism of action and the applicable scope. It provided a crucial evidence base for theory and teaching practice.

2 Theoretical framework

2.1 Concept of Foreign Language Speaking Anxiety

According to the classic framework of Horwitz et al., Foreign Language Speaking Anxiety (FLSA) is an emotional experience related to language learning [1]. It encompasses multiple aspects such as self-perception, beliefs, feelings and behaviors. Mainly manifested in oral expression scenarios. Horwitz et al. divided FLSA into three core parts: the first is communication anxiety, that is, being afraid of oral communication; The second is test anxiety. The fear of failing the oral test; The third type is negative evaluation anxiety, which is manifested as being highly sensitive to others' evaluations and even trying to avoid them as much as possible [1]. This framework has shown excellent predictive effects and stability in studies from different cultural backgrounds, providing a unified measurement and conceptual basis for this review [9].

2.2 Core theories and their applications

2.2.1 Affective Filter Hypothesis

Krashen's Affective Filter Hypothesis holds that learners' emotional factors, such as anxiety, motivation and self-confidence, form a "filter" that hinders the effective absorption of language input [10]. Multiple studies have shown that chatbots can significantly reduce the intensity of this emotional filter by providing a private and non-judgmental practice environment [11,12]. Learners are more willing to try oral expression and receive feedback in such an environment, especially in terms of reducing FLSA. This is because chatbots eliminate the threat of social evaluation in face-to-face communication, allowing learners to repeatedly try and make mistakes in a "safe" mental space without triggering a stress response [2]. However, when the Automatic Speech Recognition (ASR) of the chatbot malfunctions, learners may instead experience technical anxiety due to the frustration of communication failure, enhancing the emotional filter [8].

2.2.2 Willingness to Communicate

The Willingness to Communicate (WTC) theory proposed by MacIntyre et al. regards anxiety as a direct factor hindering oral behavior [13]. This theoretical model is divided into three layers: emotion, cognition and behavior, specifically anxiety→communication confidence→willingness to communicate→actual participation. This study regards the WTC theory as the key mediating mechanism connecting anxiety state and verbal behavior. The qualitative research by Ding & Yusof found that learners will experience a low level of anxiety when interacting with chatbots [2]. The low level of anxiety can be transferred to the real classroom

setting by enhancing self-efficacy. Therefore, encourage them to demonstrate a greater willingness to communicate in future interpersonal conversations. Nadeem et al. further found that the immediate feedback of chatbots and the ability of learners to control the learning pace by themselves enhanced self-efficacy, indirectly reduced anxiety, and improved the willingness to communicate [5].

3 Empirical research review

3.1 Research subjects and regional distribution

In terms of regional distribution, the Asian region dominates [2-8]. The research subjects mainly include different groups such as middle school students, college students and non-native language learners. Cover various learning scenarios such as language learning and IELTS preparation. Among them, participants in higher education account for the largest proportion, covering educational stages from 14-year-old teenagers to adult learners.

3.2 Technical tools and intervention cycles

In terms of the selection of technical tools, Hayashi & Sato built a complete dialogue system integrating ChatGPT 3.5 and Google Cloud TTS [6]. Susoy developed a custom chatbot for speaking language assessment [7]. Kooti et al. invoke the ChatGPT API and integrate adaptive task generation and real-time error correction functions [3]. Ding & Yusof adopted Mondly, while Nadeem et al. used Duolingo Max [2,5]. Çakmak used the Replika social chatbot [8]. The intervention intensity ranged from a single assessment to 12-week immersive training, with an average daily usage of 10-30 minutes and a total duration covering 2 to 12 weeks. The medium - to long-term intervention of more than 8 weeks by Kooti et al. was more likely to observe a stable anxiety relief effect [3]. The short-term study by Hayashi & Sato demonstrated that four weeks is merely a threshold for emotional adaptation rather than an effective window for skill acquisition [6].

3.3 Research method

Riaz et al.'s study uses the quantitative experimental [4]. The experimental group used a chatbot; The control group used traditional language learning methods. Ding & Yusof and Nadeem et al. adopted the mixed-methods design, embedding interviews with 13 and 20 participants respectively [2,5]. Susoy's cross-design has 48 participants experience cross-between chatbot assessment and human assessment [7]. Effectively controlled the sequential effect and individual differences. Hayashi & Sato's research uses a longitudinal online experimental design. After the intervention period ended, semi-structured interviews were conducted separately with the participants from both groups via Zoom. In Çakmak's study, a mixed-method approach was adopted, and both quantitative and qualitative data collection tools were used before and after the intervention [8]. Kooti et al. used a quasi-experimental pretest – posttest CG design [3].

3.4 Core research results

3.4.1 *The positive discovery of speaking anxiety by chatbots*

Most studies have shown that chatbots significantly alleviate speaking anxiety. Effect sizes ranging from moderate to large ($d = 0.39$ to 1.52). For instance, in Nadeem's study, the score on the Foreign Language Classroom Anxiety Scale (FLCAS) decreased from 110.4 to 93.1 ($t(29) = 7.89$, $p < .001$, $d = 1.52$) [5]. In Riaz's study, the anxiety reduction was 36.4% ($t(98) = 2.45$, $p = 0.018$) [4]. Kooti's study used Analysis of Covariance (ANCOVA) [3]. It showed that the anxiety level of the experimental group ($M = 45.63$) was significantly lower than that of the control group ($M = 54.96$, $F(1, 57) = 10.68$, $p < .001$, $\eta^2 = .15$), and the 8-week ChatGPT intervention reduced anxiety by 27.9%. In Ding & Yusof's study, the anxiety level of the experimental group decreased from 31.85 to 27.72 ($t(29) = 5.78$, $p < 0.001$, $d = 0.76$) [2]. The improvement in the high baseline anxiety group was three times that of the low baseline anxiety group. This confirms that chatbots are more effective for learners with high initial anxiety levels. Susoy's key finding was that in the chatbot evaluation environment, the anxiety level ($M = 98.48$) was significantly lower than that in the human evaluation environment ($M = 102.94$, $p = 0.01$, $d = 0.39$) [7]. The negative correlation between anxiety and performance disappeared ($r = -0.042$, $p > 0.01$). These indicate that the chatbot may have blocked the inhibitory path from anxiety to performance. Qualitative data consistently showed that learners generally reported that the chatbot created a judgment-free safe space and provided psychological security of not having to worry about making mistakes and being ridiculed. Both Riaz's study and Ding & Yusof's study quantitatively showed that the frequency of practice was negatively correlated with the reduction in anxiety ($r = -0.58$ to -0.62 , $p < 0.01$) [2,4]. The study also found that the degree of anxiety reduction for those who practiced frequently was 2.1 times that of those who practiced less frequently. This indicates that the more times a learner practices, the better the effect will be in reducing anxiety.

3.4.2 *The negative and other discovery of speaking anxiety by chatbots*

Some studies have shown that anxiety levels have risen instead of decreased or that the effect has been insignificant. Susoy's interview revealed that among 11 participants, 10 mentioned that strict evaluation criteria and unfamiliarity with the technology had become new sources of anxiety: "This chatbot cannot understand facial expressions, and its evaluation method is very rigid, which makes me even more nervous." [7]. Although the chatbot reduced the objective anxiety level, the participants still preferred human evaluators subjectively. This indicates that the chatbot may have alleviated the emotional aspects of anxiety but failed to meet the needs of interpersonal relationships. This further confirms the crucial role of social interaction as proposed in Vygotsky's sociocultural theory. According to Çakmak's report, anxiety rose from 2.95 to 3.27 ($t(88) = 10.636$, $p < 0.001$) after Turkish students used Replika for 12 weeks [8]. This indicates the phenomenon of a separation between emotions and performance. The oral score went from 69.70 to 74.30, but anxiety did not go down. 47.2% of the students expressed feelings of fear or anger. 17 students indicated that they felt embarrassed. This was mainly due to the fact that the chatbot's understanding of intentions was not accurate enough, resulting in communication difficulties. Hayashi & Sato's research showed that there is no significant difference in reducing anxiety between those who use chatbots and those who watch YouTube videos ($F(1, 29) = 0.06$, $p > 0.05$) [6]. This indicates that if a chatbot fails to engage learners actively, its effect is no better than watching videos.

3.4.3 The key role of regulating variables

Table 1. Classification of Regulating variables and Research Findings.

Variables	Research findings
Technology quality	Çakmak found that anxiety increased instead of decreased after 12 weeks due to recognition errors using Replika (3.27>2.95) [8]. Susoy independently developed the GPT robot, reducing the anxia-performance correlation from $r=-0.50$ to $r=-0.042$ [7].
Frequency of utilization	Riaz et al. found that the correlation coefficient r between frequency and the reduction in anxiety was -0.58 ($p=.002$) [4]. The effect was reduced by 60% when the frequency was less than 2 times per week. 10 minutes a day for 4 weeks only achieved marginal improvement in anxiety [6].
Initial anxiety level	Ding & Yusof's subgroup analysis showed that the anxiety reduction in the high anxiety group (> 75 th percentile) was 6.12 points [2]. It was three times that of the low anxiety group (2.14 points), with effect size $d=0.76$ (large) vs $d=0.25$ (small).
Digital literacy	Susoy reported that the correlation coefficient r between digital literacy and the performance of chatbots was 0.353 ($p<.05$) [7]. Learns with strong digital skills saw their advantages increase by 35%. Meanwhile, those with poor digital skills not only experienced technical anxiety but also additional language-related anxiety.
Cultural background	Susoy's qualitative data shows that 10/11 Turkey learners preferred the emotional understanding and non-verbal information demonstrated by the human evaluators during the evaluation process [7]. Ding & Yusof's Chinese sample places more emphasis on the efficiency of chatbots rather than emotional support [2].
Task design	Çakmak's research indicates that anxiety has intensified, as the content provided while Replika deviated from the topic ($d = -0.82$) [8]. Ding& Yusof's Mondly maximizes the reduction of anxiety through the points badge mechanism ($d=0.76$) [2].

As shown in Table 1, anxiety changes are systematically moderated by multi-level variables. It mainly includes six aspects: technical quality, frequency of utilization, initial anxiety level, digital literacy, cultural background, and task design.

4 Conclusion

4.1 Summary of research findings

This review analyzes the impact of chatbots on FLSA. The results show that this impact varies depending on the context and is rather complex. More than half of the studies indicate that chatbots can significantly reduce speaking anxiety, with the effect ranging from moderate to significant ($d = 0.39$ to 1.52). The chatbot creates a non-judgmental and safe environment, providing immediate feedback. This reduces the situational anxiety of learners. However, other studies have shown that anxiety either increases or the effect is not significant. The reasons include technical issues (such as voice recognition failures, stereotyped feedback) and a lack of understanding of emotions. These conflicting results indicate that the impact of chatbots on anxiety is not an inherent characteristic of the technology itself. It is the result of the dynamic interaction of various factors.

4.2 Implications of teaching practice

When using chatbots in teaching practice, it is essential to ensure that these bots can provide emotional intelligence support and accurate diagnostic feedback. This means that chatbots not only need to accurately correct pronunciation and grammar errors, but also should offer emotional responses to help students have a better emotional experience. Provide digital literacy training for students, especially for those with weaker digital literacy skills. This will help them use chatbots better. Ensure that the intervention period is long enough. It is recommended to use chatbots for at least 20 minutes per day for more than 8 weeks to achieve the best results. Teachers should also transform from traditional teaching leaders to technical intermediaries to help students better utilize chatbots. Consider factors such as students' digital literacy and anxiety levels to implement personalized teaching.

4.3 Future research direction

Cross-cultural research can be conducted to determine the role of chatbots in different cultural contexts. Long-term tracking of learners can be carried out to more comprehensively evaluate the impact of chatbots on learners. Strengthening research on technical design can further improve the design of chatbots. Particularly in the areas of emotional computing and feedback mechanisms. Research on the teacher-chatbot collaboration model can be carried out to understand how teachers and chatbots can better cooperate to enhance teaching quality.

References

1. E. K. Horwitz, M. B. Horwitz, J. Cope. Foreign language classroom anxiety. *Mod. Lang. J.* **70**(2):125-132 (1986).
2. D. Ding, M. B. Yusof. Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: a mixed methods study. *Hum. Soc. Sci. Commun.* **12**(1):1-16 (2025).
3. M. Kooti, M. Abyavi, H. Mombeini, P. Allahdini. The Impact of AI-Integrated Language Instruction on EFL Learners' Speaking Proficiency, Speaking Anxiety, and Foreign Language Motivation. *Assess. Pract. Educ. Sci.* **3**(3):1-13 (2025).
4. S. Riaz, S. Adam, S. Mujeeb, M. A. U. I. Qureshi. Use of ai-powered chatbots in reducing speaking anxiety among english language learners: A psychological perspective. *Pak. J. Soc. Sci. Rev.* **4**(5):438-451 (2025).
5. S. Nadeem, S. Iftikhar, A. Rasheed. Addressing the fear of speaking: The role of ai in reducing speaking anxiety in esl learners. *J. Appl. Linguist. TESOL (JALT)* **8**(2):825-833 (2025).
6. K. Hayashi, T. Sato. The effectiveness of ChatGPT in enhancing English language proficiency and reducing second language anxiety (L2). In *WorldCALL2023: Official Conference Proceedings*, 201-208 (2024).
7. Z. Susoy. Reducing Anxiety and Enhancing Performance: The Impact of AI Chatbots versus Human Facilitation on EFL Speaking Assessment Outcomes. (2025).
8. F. Çakmak. Chatbot-Human Interaction and Its Effects on EFL Students' L2 Speaking Performance and Anxiety. *Novitas-ROYAL (Res. Youth Lang.)* **16**(2):113-131 (2022).
9. E. Botes, J. M. Dewaele, S. Greiff. The foreign language classroom anxiety scale and academic achievement: an overview of the prevailing literature and a meta-analysis. *J. Psychol. Lang. Learn.* **2**(1):26-56 (2020).
10. S. Krashen. *Principles and practice in second language acquisition*. (1982).

11. A. Kim, Y. Su. How implementing an AI chatbot impacts Korean as a foreign language learners' willingness to communicate in Korean. *System* **122**, 103256 (2024).
12. Y. Yuan. An empirical study of the efficacy of AI chatbots for English as a foreign language learning in primary education. *Interact. Learn. Environ.* **32**(10):6774-6789 (2024).
13. P. D. MacIntyre, R. Clément, Z. Dörnyei, K. A. Noels. Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *Mod. Lang. J.* **82**(4):545-562 (1998).