



Perceptions and Experiences of Senior Medical Students on AI-assisted English doctor-patient Communication: A Qualitative Study Using Open-Ended Questionnaires

Junzhao Gu¹, Qiuheng Huang^{2,*}

¹Department of Radiation Oncology, People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

²Faculty of Chinese Medicine Science, Guangxi University of Chinese Medicine, Nanning, China

*The corresponding author's E-mail address: 774661672@qq.com

Abstract. Objective: To explore the learning perceptions, experiences, and needs of senior medical students when using an artificial intelligence (AI) assistant for English speaking practice in doctor-patient communication. **Methods:** An open-ended questionnaire survey was conducted in March 2026 among 31 senior medical students at a traditional Chinese medicine college in Guangxi Zhuang Autonomous Region, China, who had completed six months of AI-assisted English doctor-patient communication practice. Thematic analysis was employed to analyze the textual data. **Results:** Three major themes were identified: positive experiences, including enhanced psychological safety, increased practice autonomy, and instant feedback with effective pronunciation correction; major limitations, including inaccurate speech recognition, lack of emotional interaction, and unnatural dialogue rhythm; and future optimization needs, including more realistic clinical scenarios, stratified difficulty levels, and randomized interactivity. **Conclusion:** The AI assistant demonstrates significant value as a supplementary tool for English doctor-patient communication practice; however, improvements are required in speech recognition fault tolerance, emotional interaction capabilities, and personalized learning pathways.

Keywords: Artificial intelligence assistant; doctor-patient communication; English speaking; medical education; qualitative study

1 Introduction

As the internationalization of China's healthcare system accelerates, the annual number of foreign patient visits continues to rise. Consequently, the ability of medical personnel to communicate effectively in English has become an essential component of clinical competence^[1-2]. For clinical practitioners, English speaking proficiency not only affects core tasks such as health assessment, diagnostic explanation, and health education but also directly relates to patient safety and satisfaction^[3]. However,

English instruction in medical colleges remains predominantly focused on reading and translation, with speaking training—especially doctor-patient dialogue practice in professional contexts—being severely insufficient^[4].

Traditional classroom-based English speaking instruction faces multiple difficulties^[5]. First, in large-class settings, the teacher-student ratio is too low for instructors to provide individualized guidance on each student's pronunciation, intonation, and grammar^[6]. Second, class time is limited, resulting in severely inadequate frequency and duration of student oral output; most students speak for no more than two minutes per class session. Third, students commonly experience psychological anxiety when speaking English in front of peers, with "fear of speaking" and "fear of making mistakes" becoming major barriers to improving oral competence. Furthermore, traditional instruction lacks authentic doctor-patient communication scenarios, making it difficult for students to transfer their language knowledge to clinical work settings^[7].

In recent years, AI-powered speech assistants have been developed to simulate real-life dialogue scenarios, provide instant pronunciation scoring and corrective feedback, support unlimited repetitive practice, and operate without temporal or spatial constraints. In the field of general English, research has confirmed the positive effects of AI-assisted speaking practice on learners' fluency, accuracy, and learning motivation^[8]. In medical education, AI assistants have begun to be explored for doctor-patient communication training; however, AI-based speaking practice tools specifically designed for English doctor-patient communication remain at an early stage^[9], with a notable lack of in-depth qualitative analysis of learners' subjective experiences.

Current research on AI speech assistants in this domain has predominantly focused on technical feasibility, paying less attention to learners' authentic perceptions. Moreover, few qualitative studies have been conducted on AI-assisted English doctor-patient communication within the Chinese medical education context. Given this gap, the present study investigates senior medical students at a traditional Chinese medicine college in Guangxi Zhuang Autonomous Region, China, using an open-ended questionnaire to systematically explore their learning perceptions, evaluations, and suggestions for improvement.

2 Methods

2.1 Study Participants

This study adopted a purposive sampling method. Senior medical students with experience using the AI assistant and representing both traditional Chinese medicine (TCM) and nursing majors were selected as participants. These two majors were chosen because TCM focuses on traditional medical English expressions, whereas nursing emphasizes clinical nursing English communication; together, they offer complementary and representative doctor-patient English application scenarios. The participants were senior medical students from a traditional Chinese medicine college in Guangxi Zhuang Autonomous Region, China. A total of 31 students were enrolled,

including 16 TCM majors and 15 nursing majors. The inclusion criteria were: (1) full-time undergraduate students in their third year or above; (2) having passed the College English Test Band 4 (CET-4); and (3) voluntarily agreeing to participate and complete the questionnaire. The exclusion criteria were: (1) continuous interruption of the practice for more than one month; and (2) incomplete questionnaire responses or content that was obviously identical or irrelevant to the questions. All students had completed at least six months of AI-assisted English doctor-patient communication speaking practice (from October 2025 to March 2026). In this sample, there were 12 males and 19 females, with a mean age of 20.3 ± 0.9 years. Regarding English proficiency, all participants had passed CET-4, and nine had additionally passed CET-6. All participants provided written informed consent.

2.2 Teaching Method and AI Assistant Configuration

Teachers configured AI practice tasks on the Xuexitong (Learning Pass) platform, which included doctor-patient communication scenarios (e.g., admission assessment, pain inquiry, treatment explanation, and medication guidance). The AI agent played the role of a patient, while students acted as nurses (nursing majors) or doctors (traditional Chinese medicine majors) according to their discipline. The system preset dialogue opening statements, scenario settings, role requirements, and scoring criteria (pronunciation accuracy, grammatical correctness, speech naturalness, and use of medical terminology). Students engaged in immersive English speaking conversations with the AI for at least two sessions per week, each lasting 5-10 minutes, over a continuous period of six months.

2.3 Data Collection and Analysis

A self-developed open-ended questionnaire was used. Based on a literature review and content validity assessment by two medical English professors and one nursing education expert, five core questions were finalized: overall learning experience and differences from traditional classroom practice; perceived help and limitations of the AI in doctor-patient communication, terminology pronunciation, and English interaction; functions or aspects that were satisfactory or inconvenient and unrealistic; willingness to continue using the AI model and the reasons; and suggestions for adjusting the content, difficulty level, and interaction mode to better fit clinical learning needs. The questionnaire also collected general information, including age, sex, major, year of study, and English proficiency. On March 31, 2026, the questionnaires were distributed on-site, with a 30-minute time limit for completion, and were collected immediately afterward. A total of 31 questionnaires were distributed, and 31 valid questionnaires were returned (100% response rate). Thematic analysis was employed to analyze the textual data^[10]. Two researchers jointly determined 210 coding units and then coded the data independently. The initial coding agreement between the two researchers was 180 out of 210 units, yielding a percentage agreement of 85.7% (180/210). Discrepancies in coding were resolved

through discussion and consensus. All coding and consistency comparisons were performed using NVivo 12 software.

3 Results

Analysis of the 31 feedback responses yielded three core themes: positive experiences, major limitations and challenges, and future optimization needs. The themes are detailed below.

3.1 Theme 1: Positive Experiences

3.1.1 Enhanced Psychological Safety and Increased Willingness to Speak.

Several students reported that the AI environment reduced the psychological pressure of speaking practice. Student 6 (hereafter S6) stated, “Using AI for speaking practice makes me feel much more relaxed; I’m not embarrassed when I make mistakes.” S14 noted, “The AI allows one-on-one practice anytime, removes the pressure of speaking in front of others, and gives instant error correction.”

3.1.2 Greater Practice Autonomy and Frequency.

The AI assistant was not constrained by time or location, enabling students to practice whenever and wherever they wished. S13 said, “I can go back and review my previous practices and then correct my mistakes.” S4 wrote, “My practice frequency and speaking time have both increased significantly, and my nursing English speaking proficiency has improved noticeably.”

3.1.3 Instant Feedback and Effective Pronunciation Correction.

Many students affirmed the value of the AI’s instant pronunciation scoring and error correction. S3 noted, “The real-time pronunciation scoring and error correction quickly identify my weak points in pronunciation.” S20 added, “The correction is accurate, and it rapidly improves my speaking shortcomings.”

3.2 Theme 2: Major Limitations and Challenges

3.2.1 Inaccurate Speech Recognition Leading to Frustration.

Several students reported problems with speech recognition. S1 frankly stated, “Sometimes the AI can’t understand what I’m saying; it’s less flexible than a real person.” S9 said, “Because my own pronunciation is not very good, the phone often recognizes words incorrectly, causing confusion in the conversation.”

3.2.2 Severe lack of Emotional Interaction and Dynamic Scenarios.

Many students (e.g., S7, S18) pointed out that the AI was unable to simulate real patients’ emotions, tone changes, or unexpected reactions. S7 wrote, “It can’t recreate real patients’ emotions or sudden situations; the scenarios are too stereotyped.” S18

emphasized, “The AI’s responses are rigid and formulaic; it can’t simulate the many unexpected situations that occur in a hospital.”

3.2.3 Unrealistic Dialogue Rhythm and Structure.

Some students reported that the AI’s dialogue pace was too fast or that its sentence segmentation was unnatural. S2 noted, “The dialogue is too compact, leaving insufficient time to think, and the AI’s sentence breaks are problematic.” S8 mentioned, “The speaking speed is too fast.”

3.3 Theme 3: Future Optimization Needs

3.3.1 More Realistic Clinical Scenarios and Specialized Dialogues.

Many students hoped that the AI would cover a richer set of scenarios closer to actual clinical practice. S5 suggested, “Add high-frequency clinical scenarios such as pain assessment, vital signs communication, pre-operative education, emergency communication, and privacy disclosure.” S17 proposed, “Add more real clinical dialogues like oxygen therapy guidance, post-operative education, chronic disease communication, and foreign-related emergency care.”

3.3.2 Stratified Difficulty Levels to Suit Different Proficiency Levels.

Several students recommended graded learning. S3 hoped for “three gradients: basic, specialized, and advanced international.” S11 also mentioned, “Divide into basic entry (daily consultation), intermediate (specialty communication), and advanced (complex condition explanation and emergency dialogue).”

3.3.3 Enhanced Random Interactivity and Role-Play Functions.

Many students wanted the AI to simulate more realistic and unpredictable conversations. S22 wrote, “Add random Q&A, emergency scenario dialogues, and patient emotional comfort conversations to simulate real-life interactions.” S16 suggested, “Don’t stick to fixed question-and-answer; use random questions to train impromptu English response skills.”

4 Discussion

4.1 The Unique Value and Role of AI Assistants as a Tool for English Doctor-Patient Communication Practice

The results of this study indicate that the AI assistant offers significant advantages in reducing speaking practice anxiety, increasing practice frequency, and providing instant pronunciation correction, which is consistent with previous research^[11]. As illustrated by the student quotations in the results, the AI provides a “no-cost error” space that encourages students to attempt complex sentence structures and professional expressions. Therefore, the AI should not be viewed as a substitute for

teachers but rather as an “after-class practice partner” or a “basic training platform” that complements classroom scenario simulations and teacher feedback. The themes of “enhanced psychological safety” and “increased practice autonomy” identified in the results support this conclusion.

4.2 The Contradiction Between Technical Bottlenecks and the Specificity of Medical Language

The insufficient accuracy of AI speech recognition in medical settings stems not only from technical limitations but also from the specific characteristics of medical English. Medical terminology is largely derived from Latin and Greek, with complex and easily confusable pronunciations (e.g., “ileum” vs. “ilium”). Moreover, non-standard pronunciation, accents, and variations in speaking speed can all affect recognition performance. Current mainstream automatic speech recognition (ASR) systems are mostly trained on general corpora and therefore have limited capacity to recognize medical-specific speech^[12]. The “recognition errors” and “frustration” reported by multiple students in the results confirm this bottleneck. Future efforts should develop medical-domain speech recognition models, appropriately increase tolerance for learners’ pronunciation deviations, or provide more intelligent “approximate matching” feedback.

4.3 Lack of Emotional Interaction: From “Language Training” to “Communication Training”

Doctor-patient communication is inherently a highly contextualized and emotionally embedded interactive process. The fact that many students in this study reported that the AI “cannot simulate patient emotions” and “lacks the sense of real-life interaction” exposes the current AI’s shortfall in “communication training” rather than merely “language training.” Some scholars have noted that AI dialogue systems that remain at the level of simple information exchange have difficulty promoting the development of students’ empathy, emotion recognition, and coping strategies^[13]. Future AI assistants could attempt to incorporate multimodal interaction (e.g., voice emotion recognition, changes in virtual avatar expressions) or generative models based on authentic doctor-patient dialogue corpora to enhance the emotional and contextual authenticity of conversations.

4.4 Personalization of Learning Pathways: From “Uniform Content” to “Adaptive Practice”

Students’ suggestions regarding “stratified difficulty,” “scenario classification,” and “randomized interactivity” essentially reflect a demand for personalized learning pathways. The development of medical English competence is not linear but rather a multidimensional skill set grounded in work-related scenarios. AI systems could dynamically recommend practice content based on students’ historical performance (e.g., types of pronunciation errors, points of hesitation in dialogues, grammatical

weaknesses), thereby realizing a closed loop from competency profiling to weakness localization and then to scenario recommendation^[14]. This type of adaptive mechanism has already been initially implemented in general language learning platforms (e.g., Duolingo, ELSA) and is worth emulating for medical English AI systems.

4.5 Implications for Teaching and System Design

Based on the findings of this study, the following recommendations are proposed. Regarding content, an English dialogue repository covering high-frequency doctor-patient communication scenarios such as emergency care, pre- and post-operative care, and health education should be established. Regarding technology, medical-domain speech recognition models and emotion recognition modules should be introduced to improve fault tolerance. Regarding functionality, features such as slow mode, pause for thinking, role-play, and random questioning should be supported, along with the provision of optimized phrase reports. Regarding teaching, AI practice should be integrated into pre-class and post-class activities, while class time should focus on high-difficulty scenario simulations and teacher-guided instruction.

5 Conclusion

AI assistants show clear strengths in reducing speaking anxiety and enhancing pronunciation practice for English doctor-patient communication, yet face critical gaps in speech recognition, emotional interaction, and scenario realism. Students strongly support continued use and call for more realistic, stratified, and interactive designs. Future development should transform AI from a language tool into a communication training partner, advancing in anthropomorphism, clinical embedding, and personalization.

Acknowledgment

This study was funded by the Scientific Research Program of the Faculty of Chinese Medicine Science, Guangxi University of Chinese Medicine (No.2023MS024) and the University-Level Scientific Research Program (Special Project for Marxism) of the Faculty of Chinese Medicine Science, Guangxi University of Chinese Medicine (No. 2024MYM008).

References

1. Geethi, R. (2024). Linguistic Trends in Doctor-Patient Communication. *International Journal of English Language, Education and Literature Studies (IJEEL)*, 3(3), 1-5. DOI: 10.22161/ijeel.3.3.1.

2. Lampus, N.S., & Wuisan, D.S.S. (2024). Correlation between Doctor-Patient Communication with Patient Satisfaction and Loyalty. *Medical Scope Journal*, 6(2), 149-158. DOI: 10.35790/msj.v6i2.53161.
3. Peng, T., et al.(2025). A controlled study on salon training model based on Calgary-Cambridge guidelines to improve doctor-patient communication skills of general practice residents. *Chinese General Practice*, 28(1), 71-76. DOI: 10.12114/j.issn.1007-9572.2023.0729
4. Yan, H. H., & Lin, Z. H. (2022). Needs analysis and teaching reflection on five-year program medical English after new educational reform. *China Continuing Medical Education*, 14(23), 24-30. DOI:10.3969/j.issn.1674-9308.2022.23.006.
5. Phuong, H.T.B., Du, T.T., & The, N.H. (2025). Applying the File-Card Model in Teaching and Learning English Speaking Skills for English-Majored Students. *Arab World English Journal*, 16(4). DOI: 10.24093/awej/vol16no4.5.
6. Wang, Q. D., et al. (2023). Reform of medical English education and teaching from the perspective of medical humanities. *Medical Education Research and Practice*, 31(4), 485-488, 499. DOI:10.13555/j.cnki.c.m.e.2023.04.017.
7. Li, W. Y., et al. (2026). Design and practice of transition course from general English to medical English for nursing students in "top-up" programs. *Chinese Journal of Nursing Education*, 23(1), 48-54. DOI:10.3761/j.issn.1672-9234.2026.01.007.
8. Wu, L. J. (2025). An empirical study on large language models empowering English speaking teaching. *Modern Educational Technology*, 35(10), 33-41. DOI:10.3969/j.issn.1009-8097.2025.10.004.
9. Wang, Z. Z., & Xue, W. J. (2026). Construction of AI digital intelligent course for medical humanities English and exploration of blended teaching. *Medical Education Research and Practice*, 34(3), 475-481. DOI:10.13555/j.cnki.c.m.e.2026.03.019.
10. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. DOI:10.1191/1478088706qp063oa.
11. Huang, F., & Tang, Z. Q. (2026). AI-empowered English oral learning: A study on the relationship between perceived enjoyment and continuous use intention. *Journal of Xi'an International Studies University*, 34(1), 53-59. DOI:10.16362/j.cnki.cn61-1457/h.2026.01.009.
12. Li, Y. J., et al.(2021). Discussion on difficulties and solutions of intelligent speech technology in medical scenarios. *China Digital Medicine*, 16(8), 7-11. DOI:10.3969/j.issn.1673-7571.2021.08.002.
13. Jacobs, C., Johnson, H., Tan, N., et al. (2025). Application of AI Communication Training Tools in Medical Undergraduate Education: Mixed Methods Feasibility Study Within a Primary Care Context. *JMIR Medical Education*, 11. DOI: 10.2196/70766.
14. Xu, J. J. (2025). Strategies for implementing personalized foreign language education in the context of large language models. *Foreign Language Teaching and Research*, 57(1), 81-91. DOI:10.19923/j.cnki.fltr.2025.01.003.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

