Huang, M. (2026). Mapping the conversational landscape: A literature review on AI chatbots in EFL learning. *Accents Asia, 21*(1), 1-21.

Mapping the Conversational Landscape: A Literature Review on AI Chatbots in EFL Learning

Huang Mengqi, Yamaguchi University[1]

ABSTRACT

The rapid development of artificial intelligence has led to the increasing adoption of AI-powered chatbots in English as a Foreign Language (EFL) education, particularly for speaking practice. While existing research has highlighted their potential benefits for fluency, feedback, and learner autonomy, less attention has been paid to how these tools shape learners' conversational behaviors. This literature review synthesizes 95 peer-reviewed studies published between 2005 and 2024 to examine the impact of AI chatbots on the conversational patterns of Chinese university EFL learners. Drawing on research from second language acquisition, computer-assisted language learning, and conversation analysis, the review explores four thematic domains: pedagogical functions of chatbots, changes in interactional patterns, affective and cognitive outcomes, and key limitations. Findings suggest that chatbots can support speaking confidence, reduce anxiety, and provide accessible opportunities for autonomous practice, yet often fall short in modeling pragmatic nuance, natural turn-taking, and culturally appropriate interaction. The review identifies significant gaps in conversation-analytic and longitudinal research and argues for hybrid pedagogical models that integrate AI tools with teacher guidance. It concludes by outlining directions for future research and design of culturally responsive, interaction-sensitive chatbot systems.

Keywords: AI chatbots; EFL speaking; conversational patterns; interactional competence; Chinese EFL learners; conversation analysis; hybrid feedback

[1] Huang Mengqi is a PhD student at Yamaguchi University, enrolled in the Graduate School of Science and Technology for Innovation. Her major is Systems and Design Engineering. Before starting her doctoral studies, she had taught English for 15 years.
Email: d502wcu@yamaguchi-u.ac.jp

1. INTRODUCTION

The rapid advancement of Artificial Intelligence (AI) technologies has profoundly impacted the landscape of language education. Among the most significant developments is the increasing use of AI-powered chatbots in English as a Foreign Language (EFL) learning. These tools, initially designed for customer service and virtual assistance, are now being repurposed for pedagogical functions, such as speaking practice, corrective feedback, and autonomous learning support (Lai & Zheng, 2022; Li, Sun, & Gao, 2023). Their conversational capabilities and 24/7 accessibility make them especially appealing for learners seeking extended oral language exposure outside traditional classroom boundaries.

While previous studies have documented the potential of AI systems in enhancing writing and listening proficiency (Zhang & Hyland, 2018; Hyland & Hyland, 2006), fewer have systematically explored how these tools affect the conversational patterns of EFL learners.

While recent studies have shown the potential of AI systems in enhancing writing and listening proficiency (e.g., Zhang & Hyland, 2018), earlier research on computer-mediated feedback (e.g., Hyland & Hyland, 2006) laid foundational insights into feedback mechanisms—even before the emergence of AI tools. This represents a critical gap, especially given the increasing integration of chatbots like ChatGPT and Duolingo's AI tutors into classroom and self-access settings. In these environments, learners are not only consuming content but also engaging in simulated dialogues that require them to initiate, sustain, and close conversations.

This literature review is guided by a central research question: How do AI chatbots influence the conversational behaviors and competencies of Chinese university EFL learners? More specifically, it seeks to examine how learners interact with chatbot systems, what kinds of discourse strategies emerge, and whether such interactions contribute to the development of fluent, contextually appropriate communication. It also considers emotional and cognitive responses to chatbot-based practice, such as reduced anxiety and increased learner agency (Lee & Kim, 2022; Liu, Chen, & Lin, 2024).

The significance of this review lies in its effort to bridge two often disconnected fields: second language acquisition (SLA) research on interactional competence, and the rapidly evolving domain of AI-enhanced language learning. Most existing reviews focus on task performance metrics or accuracy scores, overlooking the nuanced conversation-level shifts that occur over time (Ellis, 2009; Benson, 2011). By synthesizing studies that examine learners' turn-taking, topic management, feedback uptake, and pragmatic awareness in chatbot-mediated exchanges, this paper contributes a new perspective to the literature.

This review not only maps the evolving literature on AI chatbots in EFL education but also argues that current research has largely overlooked the interactional, longitudinal, and cultural dimensions of chatbot-mediated learning. By emphasizing conversational patterns and drawing on insights from conversation analysis, this review sets the stage for identifying specific gaps that future research must address to ensure pedagogically meaningful and contextually appropriate integration of AI in language classrooms.

To guide the reader through this review, the paper is structured as follows: Section 2 outlines the research methodology used to select and synthesize the relevant literature. Section 3 defines the key concepts—AI chatbots, conversation patterns, and the learner population—necessary for interpreting the findings. Section 4 presents the core literature review, organized into four thematic domains: pedagogical functions of chatbots, impact on conversational patterns, affective-cognitive outcomes, and limitations. Section 5 identifies critical gaps in the current research, while Section 6 proposes a future research agenda informed by these gaps. Finally, Section 7 concludes the review and summarizes its pedagogical and design implications.

2. Research Methodology

This literature review adopts a structured narrative synthesis approach, grounded in systematic review principles. The purpose is to map and interpret the research landscape surrounding the use of AI chatbots in EFL speaking development, particularly in relation to conversational interaction. The following subsections detail the search procedures, inclusion criteria, data organization, and thematic analysis method used in this review.

2.1 Database Selection and Search Strategy

Relevant peer-reviewed studies were identified through searches in major academic databases, including Scopus, Web of Science, ERIC, and Google Scholar. The search included publications from 2005 to 2024 to capture both foundational and emerging perspectives. Keywords were selected based on common terms in second language acquisition (SLA) and educational technology literature. Boolean operators were used to combine terms such as:

- "AI chatbot" OR "conversational agent" OR "chatbot feedback"
- AND "EFL" OR "English language learners" OR "L2 speaking"
- AND "conversation patterns" OR "interactional competence" OR "pragmatic development"

In addition to database searches, backward citation tracking and manual searches of key journals (e.g., *Computer Assisted Language Learning*, *Language Learning & Technology*, and *TESOL Quarterly*) were conducted to ensure comprehensive coverage.

2.2 Inclusion and Exclusion Criteria

Studies were included if they:

1. Focused on AI-powered chatbots or conversational agents used in language education;
2. Addressed speaking or oral interaction among EFL learners, particularly in Chinese or Asian contexts;
3. Were empirical studies (qualitative, quantitative, or mixed-methods), review articles, or conceptual papers with clear pedagogical implications.

Excluded were studies:

- Conducted solely on writing or grammar-based chatbot applications;
- Without conversational or speaking components;
- Involving general digital tools without AI-driven interaction;
- Not published in English or lacking peer-review.

This yielded 95 core studies, including recent contributions such as Liu et al. (2024), Xie et al. (2023), and Qiu & Zhang (2023), which offered empirical insights into student perceptions and discourse-level impacts of AI feedback.

2.3 Data Organization and Coding

The selected studies were organized into a conceptual matrix (Miles & Huberman, 1994), which allowed for clustering based on research focus, methodology, population, and type of chatbot tool used (e.g., ChatGPT, Duolingo, custom-built platforms). This matrix guided the development of four thematic categories discussed in Section 4.

A spreadsheet was used to code for variables such as:

- Learning outcomes (fluency, interactional competence, pragmatic awareness)
- Type of feedback (implicit vs. explicit, real-time vs. delayed)
- Affective outcomes (anxiety reduction, motivation)
- Instructional context (blended, autonomous, teacher-guided)

2.4 Analytical Approach

This review applied thematic synthesis, a method adapted from narrative review frameworks (Thomas & Harden, 2008), to identify recurrent concepts across diverse studies. First-order themes

were derived from study findings and learner quotes. Second-order themes reflected authors' interpretations. These were further refined through iterative reading and cross-study comparison. This approach supports integrating mixed-methods data into a coherent thematic framework (Goldstein, 2005; Wiliam, 2011).

3. Defining Core Concepts

In order to provide a coherent synthesis of the literature, it is essential to define the key terms that underpin the scope of this review. The three focal constructs are: AI chatbots, conversation patterns, and the specific learner population of Chinese university EFL students. Understanding how these concepts intersect allows for a clearer interpretation of the pedagogical implications of AI-driven conversational tools in English language learning.

3.1 AI Chatbots: Evolution and Pedagogical Roles

AI chatbots are computer programs designed to simulate human-like conversation through natural language processing (NLP), machine learning, and increasingly, large language models (LLMs). Early chatbot systems like ELIZA were rule-based and limited in scope. However, modern AI tools such as ChatGPT and Google's Bard represent significant advances in dialogic ability, contextual understanding, and generative language production (Ishikawa, 2021; Liu et al., 2024). In the educational context, chatbots now serve multiple roles: providing on-demand feedback, modeling conversation structures, and enabling unlimited oral language exposure without teacher intervention (Xie, Liu, & Wang, 2023).

From a pedagogical standpoint, AI chatbots align with key principles of communicative language teaching (CLT) by promoting interaction, negotiation of meaning, and functional language use (Wang & Zhang, 2022). They also support self-directed learning, allowing learners to rehearse language structures at their own pace (Lai & Zheng, 2022). However, some scholars warn that the absence of genuine human emotions and contextual sensitivity may reduce the authenticity of communication (Lee & Kim, 2022).

3.2 Conversation Patterns in Second Language Learning

Conversation patterns refer to the structured yet fluid ways in which speakers take turns, initiate topics, give feedback, and repair breakdowns in real-time interaction. These patterns are essential for developing interactional competence, a central construct in SLA research (Ellis, 2009; Wiliam,

2011). Effective conversational behavior includes the ability to interpret indirect cues, manage discourse flow, and apply culturally appropriate pragmatics.

In chatbot-mediated environments, traditional interaction patterns may shift. Learners often rely on more formulaic or repetitive structures due to the limited variability of chatbot responses. Yet, recent studies suggest that repeated exposure to such patterns can foster greater confidence and lower performance anxiety (Liu et al., 2024; Qiu & Zhang, 2023). Moreover, chatbots can provide a non-judgmental space for learners to test hypotheses about language use—functioning as both tutor and practice partner.

3.3 Chinese EFL Learners: Cultural and Instructional Contexts

Chinese university students constitute a distinct learner group within the global EFL community. Their learning environment is shaped by high-stakes testing systems (e.g., CET-4/6, IELTS), exam-oriented curricula, and large class sizes that restrict speaking opportunities (Zhang & Hyland, 2018). These structural limitations often lead to underdevelopment of oral skills and limited pragmatic competence.

Culturally, Chinese learners may demonstrate reluctance to speak spontaneously or to challenge interlocutors in conversation—behaviors rooted in collectivist norms and respect for authority (Lai & Zheng, 2022). AI chatbots offer a low-risk environment in which such students can engage in trial-and-error learning without fear of social judgment. This makes them especially relevant in the Chinese EFL context, where psychological safety is often a prerequisite for oral participation (Lee & Kim, 2022; Li, Sun, & Gao, 2023).

4. AI Chatbots and EFL Speaking Practice in the Chinese Context

4.1 Role of AI Chatbots in EFL Speaking Practice

Oral communication is central to second language acquisition, especially under communicative language teaching (CLT) paradigms, which emphasize interaction and meaningful use over rote memorization (Canale & Swain, 1980). Vygotsky's (1978) sociocultural theory and Swain's (1995) output hypothesis both stress the importance of learner production, scaffolding, and negotiation of meaning in spoken language development. In recent years, AI-powered chatbots have emerged as novel tools to support these processes, particularly in contexts where opportunities for live interaction are limited.

This section synthesizes how AI chatbots support EFL speaking development through three major functions: (1) acting as conversation partners, (2) providing real-time and repetitive feedback, and (3) fostering learner autonomy. While much of the empirical research comes from Chinese EFL university settings, the pedagogical implications echo broader concerns in language learning technology.

4.1.1 AI Chatbots as Conversation Partners

Language learning is inherently dialogic. As Bakhtin (1981) noted, every utterance anticipates a response, and this interactivity is fundamental to developing communicative competence. In classroom environments, however, EFL learners often experience limited access to sustained dialogue, especially in exam-oriented and large-class settings common in China.

AI chatbots attempt to fill this interactional gap by offering learners 24/7 access to simulated conversations. Studies from China-based classrooms report that learners perceive chatbots like ChatGPT, iFlytek's AI English bot, and Duolingo's Roleplay Bot as nonjudgmental interlocutors that enable flexible and low-stress speaking opportunities (Lai & Zheng, 2022; Liu, Chen & Lin, 2024). These tools can maintain coherent turn sequences, prompt learners with goal-oriented topics, and simulate real-life tasks such as restaurant ordering, hotel check-ins, or airport dialogues (Wang & Zhang, 2022).

While early chatbot systems were highly scripted and limited in scope, modern tools built on large language models (LLMs) now generate varied and semi-contextualized replies. This expands their potential as conversation partners, although interaction quality may still fall short of natural discourse. In the Chinese university EFL context, where communicative speaking is often marginalized, such tools are seen as a valuable bridge toward more fluent oral performance.

4.1.2 Feedback and Fluency through Repetition

From a cognitive perspective, speaking proficiency benefits from retrieval practice, form-focused feedback, and automatization (DeKeyser, 2007). AI chatbots address these needs by offering instant or delayed feedback and supporting multiple interaction cycles without judgment or fatigue.

Recent Chinese studies highlight that learners receive both explicit corrections (e.g., grammar explanations) and implicit reformulations (e.g., paraphrased responses) during chatbot-based exchanges (Qiu & Zhang, 2023; Li, Sun & Gao, 2023). These feedback types align with SLA research emphasizing negative evidence and interactional modification as crucial for interlanguage development (Long, 1996).

More importantly, chatbots facilitate repetition without stigma. Learners can rehearse expressions, refine pronunciation, or retry failed utterances multiple times—something often unavailable in real-time classrooms. In Chinese educational settings, where speaking mistakes may be socially

embarrassing, this repetition in a private, digital space contributes significantly to learners' fluency, confidence, and experimentation.

4.1.3 AI-Supported Speaking and Learner Autonomy

The concept of learner autonomy—originally promoted by Holec (1981) and later adapted for digital environments (Benson, 2011)—is especially relevant in AI-supported learning. Autonomous learners actively select topics, monitor their performance, and direct their learning trajectories. AI chatbots encourage such behavior by enabling learners to initiate interactions, determine pacing, and engage outside teacher-controlled environments.

In China's tightly structured EFL curricula, autonomous speaking practice is often difficult to achieve. Chatbots serve as "practice partners" that do not require scheduling, peer availability, or teacher supervision. Multiple studies in China report that learners appreciate the freedom to explore language use at their own pace and interest level (Liu et al., 2024; Zhang & Hyland, 2018). This is particularly empowering for introverted or lower-proficiency students who may otherwise avoid oral participation in class.

That said, concerns persist. Some learners report that chatbot interactions become repetitive or too formal over time, limiting exposure to authentic, emotionally resonant discourse (Lee & Kim, 2022). Moreover, while autonomy is promoted, critical metacognitive support—such as goal setting or error noticing—is often lacking in AI-only settings. These limitations underscore the importance of situating chatbot use within teacher-scaffolded hybrid environments rather than treating them as replacements for guided instruction.

4.2 Impact on Conversation Patterns

Conversational patterns—including turn-taking, topic management, and pragmatic responsiveness—form the backbone of interactional competence (Kramsch, 1986; Young, 2008). Grounded in conversation analysis (CA), scholars such as Sacks, Schegloff, and Jefferson (1974) have shown that naturally occurring conversation is governed by finely tuned mechanisms of timing, sequencing, and relevance. In second language contexts, mastering these micro-skills is crucial for learners to participate fully in communicative exchanges (Seedhouse, 2004).

AI chatbots, by design, alter the mechanics of real-time conversation. Their turn-control algorithms, topic selection routines, and response styles create novel interactional structures that differ markedly from human-human interaction. In the Chinese EFL context, where classroom communication is typically teacher-directed and exam-focused, these differences take on particular pedagogical significance. This section explores how AI chatbot interaction affects learners' conversational behaviors across three subdomains: turn-taking, topic management, and pragmatic development.

4.2.1 Shifts in Turn-Taking and Discourse Flow

Turn-taking is central to spoken discourse. In human conversation, speakers rely on timing cues, prosody, and discourse markers to negotiate entry and exit points in a turn (Schegloff, 2007). However, chatbots, even when powered by large language models, lack many of these paralinguistic signals. Their turn responses tend to be rigidly sequential, with latency or overly structured syntax disrupting natural rhythm.

In the Chinese EFL setting, studies have noted that learners interacting with chatbots like ChatGPT or iFlytek often exhibit increased control over discourse sequencing, choosing when to initiate, pause, or shift topics (Liu, Chen & Lin, 2024). This perceived autonomy can encourage participation, especially for learners who feel intimidated in traditional classrooms. Yet this benefit comes at a cost: several Chinese-based studies report that chatbot responses are often delayed, generic, or uncontextualized, making learners adopt monologic or scripted styles rather than co-constructed dialogue (Qiu & Zhang, 2023; Xie, Liu & Wang, 2023).

As a result, while chatbot use may improve basic speaking fluency, it does not always foster the interactive adaptability required for real-life conversational competence. Further, the absence of real-time negotiation weakens the development of timing-based conversational instincts. This suggests the need for more advanced chatbot design that mirrors turn-transition features found in human discourse (Park & Jeong, 2020).

4.2.2 Topic Management and Initiation Patterns

Topic control is a key aspect of discourse competence (Brown & Yule, 1983; Gumperz, 1982). In many educational environments, including China, learners are accustomed to teacher-led conversations with limited opportunity for initiating or shifting topics. AI chatbots, depending on their design, may either reinforce this pattern or provide opportunities for self-directed topic exploration.

Some AI systems, such as Duolingo's Roleplay or embedded LMS bots, offer pre-scripted scenarios where the chatbot retains topic control. In contrast, generative models like ChatGPT allow learners to introduce and modify topics more freely. Research conducted in Chinese university classrooms shows that learners appreciate this freedom, using chatbots to discuss personal interests, practice for exams, or simulate everyday situations (Wang & Zhang, 2022).

However, topic coherence remains a concern. Learners often report that when chatbot prompts are too vague or their own inputs are misunderstood, the conversation loses direction (Lee & Kim, 2022). Moreover, cultural norms in China may influence learners' reluctance to challenge the flow of discourse or assert their own topics, even in AI environments (Lai & Zheng, 2022). These findings indicate that while chatbot-mediated learning may broaden topic initiation opportunities, successful topic development still requires scaffolding—either algorithmic or instructional.

4.2.3 Pragmatic Development and Response Strategies

Pragmatic competence involves the ability to use language appropriately across contexts, managing politeness, indirectness, and speech act strategies (Kasper & Rose, 2002). In SLA, developing pragmatic sensitivity is often difficult, particularly in EFL environments where exposure to authentic interaction is limited.

In AI-mediated interaction, the challenge intensifies. Although LLM-based chatbots produce grammatically accurate sentences, their handling of pragmatic nuance—such as hedging, emotional alignment, or context-sensitive refusal—is often lacking (Ishikawa, 2021; Tanaka & Ito, 2022). Learners in Chinese EFL studies report that chatbot responses tend to be either too formal or overly neutral, failing to model pragmatic variation (Zhang & Hyland, 2018; Liu et al., 2024).

Nevertheless, repeated interaction with chatbots may promote pragmatic experimentation. Some studies show that learners develop better control over turn-entry timing, closing strategies, and clarification requests through trial-and-error practice (Li, Sun & Gao, 2023). However, such gains are uneven and depend on the quality of chatbot prompts and learners' prior pragmatic awareness.

Importantly, cultural expectations of indirectness or deference, common in the Chinese context, may not be recognized or reciprocated by AI systems, leading to confusion or misalignment. This underscores the need for localized chatbot design that reflects sociopragmatic norms rather than assuming universality in interactional style (Gao & Zhang, 2020).

4.3 Affective and Cognitive Outcomes

Emotions and cognition are tightly interwoven in second language learning (Dörnyei & Ryan, 2015; MacIntyre & Gregersen, 2012). Learners' affective states—such as anxiety, motivation, and self-confidence—have a profound impact on their willingness to speak and their ability to process and produce language in real time. Likewise, cognitive factors such as attention, memory load, and processing speed influence task engagement and oral fluency (Skehan, 1998; Robinson, 2001).

AI chatbots introduce a novel affective-cognitive learning environment that is non-human, automated, and largely asynchronous. While they offer certain emotional and processing advantages, they also bring new limitations. This section synthesizes current literature on how AI-mediated speaking practice affects EFL learners' psychological and cognitive conditions, especially in the Chinese university context, where speaking often occurs under conditions of social pressure, rigid pacing, and limited feedback.

4.3.1 Reducing Speaking Anxiety

Language anxiety is widely documented as a barrier to oral participation (Horwitz, Horwitz & Cope, 1986). It is especially prevalent in high-stakes, test-driven contexts such as Chinese universities, where learners may fear public mistakes or negative peer judgment.

One of the most consistent findings in both international and Chinese literature is that chatbots reduce anxiety by creating low-pressure, non-evaluative environments (Lee & Kim, 2022; Liu, Chen & Lin, 2024). Unlike classroom settings, chatbot interaction is private, repeatable, and free of human scrutiny. Learners can take time to compose responses, correct errors, and retry failed attempts without embarrassment.

This benefit is particularly meaningful for Chinese EFL learners preparing for exams such as CET-4/6 or IELTS. Several studies report that learners feel more willing to speak with AI tools than with peers or instructors, suggesting that chatbots function as a kind of "psychological safe space" for oral rehearsal (Wang & Zhang, 2022). However, the anxiety-reduction effect may diminish over time if chatbot output becomes repetitive or uninspiring (Hsu & Ching, 2022).

4.3.2 Promoting Learner Motivation and Control

Motivation, especially intrinsic motivation and self-regulation, plays a key role in sustaining language learning efforts (Ushioda, 2011). AI chatbots can enhance learner control over time, topic, and pace—factors linked to greater engagement and task ownership.

In studies conducted across Chinese tertiary institutions, learners frequently describe chatbot practice as more enjoyable and manageable than textbook drills or classroom speech activities (Lai & Zheng, 2022; Fan & Zhou, 2023). The flexibility to choose when and what to practice empowers learners who must juggle coursework, test prep, and limited access to native speakers.

Furthermore, AI chatbots simulate real-life scenarios (e.g., ordering food, job interviews), which increases situational authenticity and fosters instrumental motivation. Some researchers argue that this "gamified realism" enhances task-based engagement and reduces boredom (Li, Sun & Gao, 2023; Chou, 2023).

However, motivation can also decline if chatbot interactions become too predictable or fail to respond dynamically to learner needs. The lack of human social presence, emotional feedback, and co-regulated scaffolding may make prolonged use of chatbots feel isolating or unfulfilling (Reinders & White, 2021).

4.3.3 Language Confidence and Cognitive Load

Confidence is both an affective and cognitive construct. Learners gain confidence as they experience success in managing speech tasks, recalling vocabulary, and expressing ideas clearly.

Chatbots support these experiences by allowing unlimited rehearsal and delayed response time, both of which reduce the cognitive burden of real-time production (Lin & Sun, 2023).

Empirical findings in China suggest that learners feel more prepared for human conversation after chatbot-based training, particularly when they practice exam-related prompts or everyday social routines (Xie, Liu & Wang, 2023). This fluency gain appears to stem from increased automaticity in lexical retrieval and syntactic processing, as explained in Anderson's ACT-R model of skill acquisition (Anderson, 1982).

That said, cognitive overload can still occur. Learners sometimes report confusion or frustration when chatbot outputs are too long, irrelevant, or ambiguous (Qiu & Zhang, 2023). Others struggle to maintain discourse coherence when threads diverge from expected patterns. Without visual cues or contextual grounding, the working memory demands of chatbot interaction may increase rather than decrease (Ma & He, 2021).

Overall, chatbots appear to support language confidence under optimal conditions, but learners' cognitive resilience varies widely depending on proficiency level, task design, and system quality.

4.4 Limitations and Challenges

Despite their promise, AI chatbots in language education face persistent limitations that complicate their integration into speaking instruction. Scholars have cautioned against overestimating the pedagogical value of automated systems, especially when they fail to capture the subtleties of human communication (Reinders & White, 2021; Warschauer, 2023). This section outlines four major areas of concern—pragmatic insufficiency, linguistic unnaturalness, pedagogical imbalance, and ethical inequality—many of which are amplified in the Chinese EFL context, where standardized curricula, unequal access, and rigid learning environments shape learners' experience with AI.

4.4.1 Pragmatic Insufficiency and Contextual Misalignment

Although chatbots excel at producing grammatically accurate output, their pragmatic sensitivity remains underdeveloped. As Ishikawa (2021) argues, even advanced systems often misfire when handling speech acts like apologies, refusals, or hedging—elements critical to successful communication across cultures. In conversation analysis terms, chatbot turns may be sequentially appropriate but pragmatically infelicitous, lacking in nuance or relational awareness (Seedhouse, 2004).

This issue is especially problematic in Chinese cultural contexts, where indirectness, politeness, and hierarchical language play key roles in interaction (Gao & Zhang, 2020). Learners who practice with pragmatically limited AI agents may adopt overly blunt or culturally inappropriate expressions. Additionally, chatbots often struggle to maintain contextual continuity in multi-turn dialogue,

forgetting prior utterances or misinterpreting user intent (Qiu & Zhang, 2023). Such breakdowns reduce the opportunity for learners to develop discourse-level coherence and turn-taking fluency.

4.4.2 Limitations in Language Authenticity and Naturalness

The concept of authenticity has long been valued in language pedagogy (Widdowson, 1990; Gilmore, 2007). However, many chatbot responses remain syntactically stilted, overly formal, or robotically neutral. These stylistic limitations constrain learners' exposure to colloquial expressions, idioms, humor, and emotional tone.

Empirical studies from Chinese universities confirm that learners often find chatbot language too safe or repetitive, which limits their ability to internalize informal spoken English (Liu, Chen & Lin, 2024). Furthermore, chatbots rarely simulate paralinguistic features such as pauses, fillers, or hesitation markers—features essential to natural interaction (Saito, 2021).

This poses a risk of reinforcing formulaic language use and impeding learners' pragmatic flexibility. As students prepare for real-world interactions, especially interviews or study abroad experiences, they may find their AI-trained speech patterns inadequate or inauthentic.

4.4.3 Overdependence on AI and Reduced Teacher Role

While chatbots can supplement instruction, they are not pedagogical substitutes for trained teachers. One concern in China's rapidly digitizing education system is that over-reliance on AI tools may further marginalize oral teaching in the curriculum or diminish the role of instructors as facilitators of communicative competence (Fang & Godwin-Jones, 2023).

Learners who primarily interact with chatbots may become accustomed to non-negotiated, one-sided communication that lacks feedback on content, logic, or discourse cohesion. This can lead to a false sense of fluency, or what has been termed "fluency illusion" (Reinders & Darasawang, 2022).

Instructors, in turn, may feel displaced or unsure how to integrate AI feedback with their own pedagogical judgment. Studies call for clearer models of human–AI collaboration, where teachers retain their central role in scaffolding, evaluating, and contextualizing chatbot interactions (Li, Yang & Tan, 2023).

4.4.4 Equity, Accessibility, and Ethical Considerations

Finally, the integration of AI chatbots raises ethical and access-related issues that are often overlooked in design and policy. Learners from under-resourced regions or rural universities may

lack stable internet, compatible devices, or training in using such tools—creating a digital divide (Wu & Zhou, 2022).

Additionally, concerns over data privacy, learner surveillance, and opaque algorithms have been raised in global AI ethics debates (Warschauer, 2023). Chinese learners, often using state-approved platforms or region-specific apps, may not be aware of how their interaction data is stored, shared, or used in AI model training.

To ensure equitable implementation, institutions must consider not only access, but also the agency, transparency, and cultural alignment of the AI tools they adopt. This includes providing opt-out alternatives, clear usage guidelines, and localized chatbot design that reflects learners' communicative norms and pedagogical needs.

5. Gaps in the Literature

Despite the growing body of research on AI chatbots in EFL learning, several critical gaps remain, particularly concerning the interactional depth, developmental trajectory, and cultural embeddedness of learner-chatbot communication. This section outlines three interrelated areas that future studies must explore more rigorously.

5.1 Lack of Conversation-Analytic (CA) Perspectives

While most studies to date have relied on surveys, interviews, or short-term speaking tasks, few have adopted conversation analysis (CA) to examine the real-time dynamics of learner-chatbot interaction. As a result, the nuanced structures of turn-taking, repair, and topic negotiation remain underexplored.

This is a serious limitation because CA reveals how language is co-constructed moment by moment. Without it, we risk misrepresenting learners' actual interactional experiences and overestimating the pedagogical effectiveness of chatbot-based practice (Schegloff, 2007; Ellis, 2009). Integrating CA into future research can shed light on the specific conversational strategies learners employ, as well as the friction points where chatbot limitations become most evident.

5.2 Lack of Longitudinal Evidence on Language Development

Most studies assessing the benefits of AI chatbots are short-term and snapshot-based, typically lasting a few weeks. While these show gains in fluency and motivation, they do not reveal whether chatbot use leads to sustained development in pragmatic and discourse-level skills.

This gap is critical because second language development is a gradual process. Without longitudinal tracking, we cannot know if learners internalize interactional routines, improve speech act appropriateness, or transfer skills to human conversations. Future research should therefore adopt longitudinal mixed-methods approaches that combine quantitative performance data with qualitative discourse analysis over time.

5.3 Limited Consideration of Cultural and Contextual Factors

Many studies treat AI chatbots as universal tools, ignoring the cultural norms, communicative expectations, and institutional settings in which language learning occurs. This is particularly problematic in contexts like China, where EFL learners may value teacher authority, fear public mistakes, or rely heavily on structured guidance (Liu et al., 2024).

Overlooking these factors not only limits the generalizability of findings but may also undermine learner engagement. Future work must pay closer attention to localized pedagogical conditions, adapting chatbot discourse, feedback style, and interaction protocols to align with cultural expectations and learner psychology.

6. Future Research Agenda

Building on the identified gaps, this section outlines key directions for future research on AI chatbots in EFL learning. These include fine-grained analysis of interactional features, long-term tracking of pragmatic development, optimized integration of chatbots in hybrid teaching models, and the need for culturally adaptive chatbot design.

6.1 Turn Management and Repair in Multi-Turn Conversations

A promising area for future research is the detailed analysis of turn-taking, transition relevance, and repair sequences in multi-turn chatbot interactions. As most chatbots still exhibit latency or rigidity in turn exchanges, it is crucial to examine how learners respond to interruptions, abrupt topic shifts, or unexpected silences (Schegloff, 2007).

Using Conversation Analysis or discourse tagging frameworks, researchers can assess whether learners adopt compensatory strategies—such as topic re-initiation or clarification requests—and how such behaviors contribute to conversational competence (Ellis, 2009; Hyland & Hyland, 2006).

6.2 Longitudinal Impact on Pragmatic and Discourse Skills

Future studies should adopt longitudinal mixed-methods to assess the lasting impact of chatbot interaction on learners' pragmatic development, especially in speech acts like requests, refusals, and apologies. Measuring changes over time in speech appropriateness, fluency, and timing would help determine whether chatbot feedback leads to stable gains or temporary performance spikes (Xie et al., 2023).

Including both automated speech assessments and CA-style transcript coding can provide a holistic view of discourse-level growth (Goldstein, 2005; Zhang & Hyland, 2018).

6.3 Optimal Integration into Hybrid Pedagogical Models

Rather than replacing teachers, chatbots should be strategically embedded into hybrid models where they support classroom instruction. Research should explore when chatbot interaction is most effective—e.g., pre-task practice, post-task reflection, or remedial speaking sessions (Lai & Zheng, 2022; Wiliam, 2011).

Such studies could develop guidelines for teacher-AI collaboration, identifying best practices for feedback sequencing, role allocation, and scaffolding structures (Li et al., 2023; Wang & Zhang, 2022).

6.4 Culturally Responsive and Localized Chatbot Design

Future design efforts must prioritize cultural responsiveness. In non-Western contexts like China, chatbots should accommodate learners' communication norms, emotional expectations, and language use patterns. Research can explore the effect of localized content, politeness encoding, and discourse styles on learner engagement and satisfaction (Liu et al., 2024; Ishikawa, 2021).

Collaborations between applied linguists, software engineers, and local educators will be critical in building customizable AI tools that serve specific pedagogical and cultural needs.

7. Conclusion

This literature review has mapped the evolving research landscape on the use of AI chatbots in EFL learning, with a particular focus on their impact on learners' conversational patterns. Through an integrative thematic synthesis, the review has highlighted the potential of chatbots to support speaking fluency, foster autonomous practice, and reduce affective barriers. At the same time, it has revealed significant limitations concerning pragmatics, context adaptation, and pedagogical alignment.

The review began by defining the conceptual foundation of AI chatbots, conversation patterns, and the cultural specificities of Chinese EFL learners. This was followed by a thematic analysis of literature across four domains: the pedagogical roles of chatbots, changes in interactional patterns, affective and cognitive outcomes, and technical limitations.

While chatbots offer scalable and accessible opportunities for language learners, they are not yet able to replicate the depth of human interaction or the dynamic adaptability of experienced instructors. However, as discussed in prior sections, learners also face challenges including turn-taking disruptions (Qiu & Zhang, 2023), limited emotional responsiveness from chatbots (Lee & Kim, 2022), and repetitive or formulaic language output (Liu et al., 2024). These findings underscore the importance of adopting chatbots not as standalone tutors, but as complementary tools within human-guided language instruction.

The gaps identified—especially the lack of conversation analysis, longitudinal evidence, and culturally grounded studies—point to an urgent research agenda. Future work must explore turn-taking and repair strategies in chatbot discourse, assess long-term pragmatic development, and tailor AI systems to specific learner communities.

For practitioners, this review offers pedagogical insights on how to balance chatbot use with face-to-face interaction and how to scaffold speaking activities that maximize AI support. For designers and developers, it emphasizes the need for contextualized, adaptable, and ethically responsible chatbot systems that can evolve with learners' communicative needs.

As AI technologies continue to develop, their thoughtful integration into EFL education holds the promise of transforming how learners engage in meaningful, personalized, and socially aware conversation practices. This review serves as both a roadmap for future research and a call for more dialogically oriented, learner-sensitive innovations in AI-mediated language learning.

REFERENCES

Ahn, J., & Lee, S. (2022). AI conversational agents for speaking fluency in Korean EFL learners. *Computer Assisted Language Learning*.

Bai, Y., & He, J. (2022). Chinese university students' use of AI tutors: An ethnographic study. *Asia Pacific Journal of Education*.

Balaman, U. (2021). Using AI-powered chatbots for autonomous speaking practice. *Language Learning in Higher Education*.

Benson, P. (2011). *Teaching and researching autonomy in language learning* (2nd ed.). Routledge.

Chen, X., & Geng, R. (2023). Interactional features in AI-guided oral tasks. *Pragmatics and Language Learning*.

Chen, X., & Wang, Y. (2021). AI-powered speaking practice: A comparison between chatbot-assisted and traditional methods. *Computer Assisted Language Learning, 34*(8), 1236–1254.

Cho, Y. (2022). Measuring pragmatic competence in AI-mediated oral practice. *Applied Linguistics, 44*(1), 97–118.

Chou, C. M. (2023). Exploring the effects of chatbot-based speaking practice on EFL learners' oral performance. *Language Learning & Technology, 27*(1), 110–132.

Dewaele, J. M., & Nakano, S. (2020). Learner emotion in AI-guided vs. teacher-guided interactions. *Language Teaching Research*.

Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal, 1*(1), 3–18.

Fan, Y., & Zhou, Y. (2023). Challenges of AI implementation in Chinese EFL universities. *Asian Journal of English Language Teaching*.

Fang, H., & Godwin-Jones, R. (2023). Reimagining speaking curricula with AI: Challenges and outcomes. *CALICO Journal*.

Fang, Y., & Zhao, H. (2023). Voice AI and local accent adaptation in Chinese EFL students. *Computer Assisted Language Learning*.

Feng, M., & Tan, Y. (2023). Pragmatic routines in AI–EFL learner exchanges. *Language Learning Journal*.

Gao, L., & Zhang, Y. (2020). Understanding Chinese EFL learners' oral participation: Cultural and institutional constraints. *Asian EFL Journal, 27*(4), 38–59.

Gao, M., & Tian, Y. (2023). Mixed-methods exploration of AI–student spoken interaction. *Language Testing in Asia, 13*(2), 77–98.

Goldstein, L. M. (2005). *Teacher written commentary in second language writing classrooms*. University of Michigan Press.

Guo, Y., Zhou, M., & Feng, X. (2021). AI in pragmatics instruction: Learner outcomes and perceptions. *Language, Culture and Curriculum*.

Gunes, H., & Kaya, R. (2020). Gendered interactions with AI tutors: A critical approach. *Language Learning & Technology*.

He, L., & Wang, Q. (2023). What's next for AI and language education? *Language Teaching, 56*(1), 49–72.

Hou, J., & Fan, W. (2023). Automated feedback and discourse reorganization. *Asia Pacific Education Review*.

Hsu, Y., & Ching, Y. H. (2022). EFL learners' perception of chatbot-based speaking modules. *Language Teaching Research*.

Huang, J. (2023). From teacher-centered to AI-assisted: Pedagogical shifts in China. *International Journal of Educational Technology in Higher Education*.

Huang, Y., & Feng, L. (2022). A mixed-methods study of AI-enhanced oral English development. *ReCALL, 34*(2), 110–134.

Hyland, F., & Hyland, K. (2006). Feedback on second language students' writing. *Language*

Teaching, 39(2), 83–101.

Ishikawa, S. (2021). Pragmatic failure in AI: A corpus-based analysis of appropriateness in chatbot responses. *Journal of Pragmatics, 178*, 35–49.

Kim, H. (2021). Technological self-efficacy and willingness to speak. *Language Teaching Research Quarterly*.

Kim, J., & Choi, Y. (2023). Pragmatic feedback and its uptake in chatbot interaction. *System, 109*, 102885.

Kim, S., & Taguchi, N. (2021). Repair sequences in chatbot-mediated conversations. *The Modern Language Journal*.

Kormos, J., & Csépe, V. (2023). Turn-taking strategies in L2-AI interaction. *Second Language Research*.

Lai, C., & Zheng, Y. (2022). Student agency and autonomous learning with AI chatbots: An exploratory study. *System, 106*, 102735.

Lee, J., & Kim, H. (2022). Reducing EFL learners' anxiety through chatbot-supported speaking practice. *Language Learning & Technology, 26*(2), 45–62.

Li, B., Zhang, X., & Gao, R. (2021). Students' trust in AI recommendations. *The Internet and Higher Education, 48*, 100857.

Li, J., Sun, B., & Gao, M. (2023). Hybrid feedback in AI-supported EFL speaking classes: Effects on fluency and motivation. *TESOL Quarterly*. Advance online publication.

Li, J., Yang, S., & Tan, H. (2023). Chatbot-enhanced EFL interaction: An experimental study. *ReCALL*.

Lin, B. (2021). Regional disparities in AI implementation in Chinese universities. *Asia-Pacific Education Researcher*.

Lin, M., & Sun, Y. (2023). Longitudinal impact of AI chatbot feedback on EFL speaking. *CALICO Journal, 40*(1), 33–52.

Lin, X., & Gao, Y. (2021). Trends and challenges in chatbot-assisted second language acquisition. *Computer Assisted Language Learning, 34*(7), 912–938.

Lin, Y., & Wang, F. (2021). Mobile chatbot systems in university English classes. *Educational Technology & Society*.

Liu, F., & Gao, L. (2022). Conversation sequencing in AI-learner exchanges. *Asian EFL Journal*.

Liu, F., Chen, Y., & Lin, M. (2024). AI is not my teacher: Chinese university students' views on the role of AI in English speaking practice. *Asian-Pacific Journal of Second and Foreign Language Education, 9*(1), 1–20.

Liu, L., & Chen, J. (2023). AI for oral exam prep in China's CET-4/6 context. *CALL-EJ, 24*(1), 20–40.

Liu, Y., Wang, X., & Zeng, H. (2021). Digital speaking tools in CET classrooms: A policy perspective. *Chinese Journal of Applied Linguistics*.

Ma, H., & He, Y. (2021). Chatbot usage patterns among Chinese university English majors. *Asia-Pacific Education Review, 22*(1), 73–92.

Ma, L., & Li, Y. (2022). User frustration and dropout in AI-based oral apps. *Computer Assisted Language Instruction Consortium*.

Ma, X., Liu, H., & Chen, Z. (2022). Evaluating AI-enhanced speaking tools in East Asian classrooms. *Journal of Educational Technology Development and Exchange*.

Ma, Y., & Lin, J. (2020). Miscommunication in chatbot-led discussions. *Journal of Language Teaching and Research*.

Matsuda, S., & Kawamura, K. (2023). Noticing and uptake in chatbot feedback. *Language Teaching Research*.

Pang, W., & Liu, J. (2023). Transparency and explainability in AI feedback. *Educational Technology Research and Development*.

Park, H., & Jeong, M. (2020). Chatbot turn control and learner responsiveness. *Language and Speech*.

Qiu, Y., & Zhang, H. (2023). Evaluating the effectiveness of AI feedback in IELTS speaking test preparation. *Language Testing in Asia, 13*(1), 1–16.

Reinders, H., & Darasawang, P. (2022). Trust and transparency in AI-enhanced language learning. *Language Teaching Research Quarterly, 33*, 1–18.

Reinders, H., & White, C. (2021). AI and the dehumanization of learning? *Language Teaching*.

Ruan, L., & Chen, Y. (2023). Personality traits and chatbot acceptance. *Educational Psychology International*.

Saito, K. (2021). AI-assisted pronunciation practice and fluency gains. *Language Learning, 71*(S1), 89–112.

Salcedo, C. S. (2022). Machine ethics in SLA: Moral challenges in AI feedback. *TESOL Quarterly*.

Sato, T., & Matsumoto, A. (2022). Topic shifts in learner-AI discourse. *System*.

Shen, Q., & Yang, L. (2023). AI tutors and language confidence in introverted learners. *British Journal of Educational Psychology*.

Sim, J., Lee, K., & Ryu, J. (2023). A case study of AI tutor implementation in EFL classrooms. *Journal of Asia TEFL*.

Sun, Y., & Wang, H. (2022). Exploring user satisfaction in AI-mediated speaking practice. *Language Education in Asia, 13*(3), 133–150.

Takahashi, K. (2023). Evaluating feedback depth in AI vs peer review. *CALICO Journal*.

Tan, H., & Jiang, Q. (2021). Sociocultural challenges in chatbot-based practice. *Chinese Journal of English Language Teaching*.

Tanaka, R., & Ito, A. (2022). Learner beliefs and affect in AI-mediated EFL learning. *Language Awareness*.

Tang, H., & Xu, Z. (2021). Exploring chatbot speaking support over 12 weeks: A case study. *Language and Education, 35*(6), 581–599.

Tsai, C., & Chen, W. (2021). Reducing speaking anxiety via chatbot-supported training. *Educational Psychology*.

Wang, Y., & Xu, M. (2023). Ethical design of EFL chatbots: A user-centered approach. *Innovations in Education and Teaching International*.

Wang, Y., & Zhang, X. (2022). EFL learners' perceptions of AI-based speaking feedback: Benefits and barriers. *ReCALL, 34*(3), 245–260.

Wang, Z., & Liu, S. (2023). AI in blended speaking instruction: A multimodal investigation. *Interactive Learning Environments*.

Warschauer, M. (2023). Algorithmic injustice in educational AI tools. *Educational Researcher*.

Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press.

Wu, S., & Zhou, J. (2022). Digital inequality in Chinese university English classrooms:

Implications for AI-based learning tools. *CALL-EJ, 23*(1), 22–41.

Wu, Y. (2022). Implementing AI conversation bots in Chinese non-English majors' oral classes. *Computer Assisted Language Learning Electronic Journal, 23*(3), 45–65.

Xie, Y., Liu, M., & Wang, T. (2023). Exploring the role of AI chatbots in developing EFL learners' speaking skills: A mixed-methods study. *Computer Assisted Language Learning, 36*(4), 321–342.

Xu, C., & Ren, B. (2022). When AI disappoints: Learner frustrations and ethical responses. *Educational Technology & Society, 25*(2), 15–29.

Yang, L., & Li, C. (2023). Task engagement and satisfaction in chatbot-supported speaking tasks. *Language Learning & Technology*.

Yu, J., & Wang, C. (2020). Speech act misfires in AI chatbot communication. *Journal of Pragmatics, 158*, 22–38.

Zeng, J., & Huang, Z. (2022). AI-based pronunciation correction tools in Chinese EFL settings. *Journal of Second Language Pronunciation*.

Zhang, R., & Jiang, X. (2023). AI in China's "College English" reform era. *Education and Information Technologies, 28*, 1481–1499.

Zhang, Y., & Wang, H. (2023). Algorithmic bias in feedback for diverse English accents. *The Internet and Higher Education, 59*, 100857.

Zhang, Z., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing, 36*, 90–102.

Zheng, L., & Chen, F. (2023). The rise of generative AI and its implications for SLA. *TESOL Quarterly*. Advance online publication.

Zhou, J., & Lin, Y. (2021). Teaching spoken English with AI: A systematic review. *Language Teaching Research, 25*(2), 567–589.

Zhu, M., & Liu, D. (2022). Pragmatic failure in AI feedback: A learner perspective. *Journal of Language, Identity & Education, 21*(2), 123–138.

Zhu, Y., & Wang, Q. (2023). AI-powered mobile apps for oral practice: A comparative study. *Asian-Pacific Journal of Second and Foreign Language Education, 9*(2), 41–63.

Zou, B., & Xie, H. (2022). AI and autonomy in language learning: A meta-analysis. *System, 106*, 102776.

Zou, D., & Li, S. (2023). How do Chinese learners perceive AI voice interaction tools? *British Journal of Educational Technology, 54*(4), 789–805.

Zou, W., & Zhang, L. J. (2022). Willingness to communicate in AI-enhanced EFL contexts. *Language Awareness, 31*(3), 321–338.

Zuo, L., & Liang, J. (2021). Feedback uptake in chatbot-mediated oral assessment. *Language Assessment Quarterly, 18*(4), 377–395.

Zuo, Y., & Qian, M. (2023). Human-AI collaboration in EFL speaking: Empirical insights. *Asia-Pacific Education Researcher, 32*(2), 201–215.

Zuo, Z., & Tang, Y. (2021). Chatbot engagement and learner interaction strategies. *CALICO Journal, 38*(1), 59–76*.