Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

University of Babylon, College of Education for Human Sciences, Department of English
hum918.rafal.hasson@uobabylon.edu.iq

Abstract

This study investigates the use of generative AI as an artificial conversation partner for incidental L2 speech practice, with the aim of fostering learners' trust, strategy use, and perceived benefits. As AI technologies are increasingly integrated into language teaching, there is a need to investigate how learners interact with AI tools to develop effective pedagogical interventions. A mixed-methods research design was employed, which included surveys with 15 EFL learners and subsequent follow-up semi-structured interviews to examine their experience. Quantitative results indicated that learners reported medium-to-high levels of trust in AI systems, with trust being shaped by the consistency, responsiveness, and perceived reliability of the AI. Strategy analysis found that learners employed cognitive, metacognitive, and social strategies of paraphrasing, repetition, and self-reflection to gain optimal conversational outcomes. Qualitative data highlighted students' perception of enhanced fluency, confidence, and communicative inclination, while accuracy and pragmatic competence had limited improvements. The study highlights the pedagogical benefits of AI technology in providing low-stakes, risk-free practice spaces and the necessity of careful integration with teacher-led instruction. Technical innovations, particularly for variability of conversation and culturally appropriate feedback, are suggested for optimizing the effectiveness of AI-facilitated L2 speaking practice. The findings contribute to the widening body of research on AI-aided language learning and provide recommendations for teachers and developers keen on implementing AI in communicative language teaching.

1. Introduction

1.1 Problem Statement

The quick development of generative artificial intelligence (GenAI) has opened up new pedagogical prospects for second language (L2) speaking practice, particularly in language acquisition. Unlike traditional computer-assisted language learning (CALL) programs, GenAI models such as ChatGPT and other large language models (LLMs) can produce contextually relevant, coherent, and adaptive responses in real time (Joo, 2024, 54–59). This capability enables students to engage in extended, spontaneous conversation practice that more closely simulates natural communicative exchange than scripted role-plays or programmed chatbots. In principle, this technology has the potential to address a long-standing dilemma in L2 instruction: providing students with sufficient opportunities to practice speaking where access to native speakers or beneficial peers is limited.

Despite such potential, there are significant challenges. One of the most significant factors that influence effective integration of GenAI into L2 acquisition is the trust factor. The extent to which learners trust the AI companion will determine whether they are receptive to meaningful interaction, get feedback, or rely on the system as a good source of input (Zhai & Wibowo, 2024). Over-trust, of course, also leads to over-acceptance of faulty or biased results, while under-trust discourages students from leveraging the tool's conversational capacities to their fullest potential (ScienceDirect, 2025). This problem also leads us towards a foundational pedagogical issue: how to control student trust so as to enable effective use of AI and yet maintain a critical appreciation of its character.

Another vital dimension pertains to strategy use. Second language acquisition research emphasizes that learners employ a variety of strategies—self-repair, clarification requests, and discourse management—to sustain interaction and underwrite their performance (Oxford, 2017). With GenAI, however, there is no way to know how these strategies are realized, whether learners transfer available strategies to AI-mediated practice, or if learners invent new strategies to address the particular affordances and limitations of AI conversation. For instance, students can try out prompting techniques, reformulation methods, or code-switching in order to direct the responses of the AI (Godwin-Jones, 2024). Recognizing these dynamics is important for identifying whether dialogue partners through AI support or hinder the acquisition of strategic competence at the level of oral communication.

Also urgent is the matter of perceived ability gains. While GenAI has been shown to reduce foreign language speech anxiety and raise self-esteem (Ding & Yusof, 2025), self-assessment instruments will still reveal discrepancy between learners' perception of themselves and objectively graded performance. Learners may sound more fluent and more confident after AI exposure, but this does not necessarily mean they will register measurable improvement in accuracy, complexity, or pragmatic appropriateness. Conversely, students who do not trust the AI may claim fewer gains even if they had actually made concrete progress. Thus, examining perceived gains alongside strategy use and trust provides a better understanding of the learner experience.

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

Finally, there are issues at the systemic level that render it challenging to apply GenAI for L2 speaking practice spontaneity. The cultural validity of AI content has been called into question by researchers as models are being trained on corpora which may or may not reflect rich communicative norms (Godwin-Jones, 2024). Similarly, the application of AI conversation is jeopardizing learners' critical and problem-solving skills if conversations are never challenged (Zhai & Wibowo, 2024). Moreover, the incoherence of AI responses between models or within the same model instills doubt that can mislead students and compromise reliability (Arxiv, 2023).

Together, these issues characterize an obvious research gap. While the efficacy of AI chatbots in stimulating oral practice and relieving anxiety has already been documented in prior research, it is less clear how trust, speaking strategy use, and perceived benefits play a part in circumstances in which students use GenAI as a practice dialogue partner for spontaneous L2 speaking practice. Closing this knowledge gap is essential to guide educators, learners, and developers toward more effective and ethically right integration of GenAI for language learning.

Taking into account the aims, the research addresses the following questions:

1. To what degree do L2 learners trust GenAI when used as a simulated conversation partner for free speaking practice?

2. What are the speaking strategies used by learners in interactions with GenAI in these environments?

3. To what extent do learners indicate increased L2 speaking ability following GenAI-mediated practice?

4. How are trust, strategy use, and perceived benefit related for the use of GenAI in spontaneous L2 speaking practice?

1.2 Research Aims

This study aims at:

1. Assessing learners' trust level in using gen-AI as a computer-simulated conversation partner for spontaneous L2 speaking tasks.

2. Investigating learner strategies employed during speech with GenAI, such as repair, scaffolding, or prompting strategies.

3. Evaluating perceived gains in speech ability—fluency, accuracy, confidence—reported by learners after GenAI interaction.

4. Investigating trust, strategy utilization, and perceived gains interrelations, to ascertain the ways in which these constructs are mutually dependent when given conditions of GenAI use.

2. Literature Review

2.1 Generative AI in L2 Learning: Promise and Pitfalls

Generative AI (GenAI) has emerged as a transformative tool in second language (L2) learning, with potential far exceeding the more primitive rule-based or template-based chatbots. Compared to traditional computer-assisted language learning (CALL) systems, which were often founded on scripted dialogue, GenAI leverages large-scale language models to create dynamic, contextually responsive, and semi-authentic responses. This allows for longer, unscripted interaction that more closely approximates everyday conversation (Cambridge Element, 2025). Importantly, GenAI offers the potential for personalized practice, particularly for those students with limited exposure to native speakers or genuine communicative environments. For example, recent systematic reviews indicate that students increasingly use AI systems for personalized oral fluency practice, building speaking confidence where formal classroom environments or peer interaction are unfeasible (Systematic Review, 2025).

Another key promise of GenAI is its potential to provide hybrid feedback. Although its application to writing tasks has been more widely documented, where algorithmic feedback is augmented through teacher mediation (Tandfonline, 2025), the extension of such feedback models to speaking language practice is largely unexplored. Initial experiments suggest that AI-supported conversation scaffolding can help students notice gaps in grammar, lexis, or pragmatics, thereby fitting within Swain's (2019) output hypothesis for language learning through pushed output and corrective feedback.

Yet, these potential benefits are counterbalanced by persistent limitations. One major challenge is linguistic and cultural authenticity. As GenAI models are largely trained on Western-based English corpora following Western communicative norms, they may be incapable of providing contextually appropriate or culturally responsive answers. According to Godwin-Jones (2024, 4), AI has the tendency to produce pragmatically "flattened" discourse that disregards the subtle sociocultural expectations that are necessary for intercultural competence. This constraint risks misleading students who strive for not only grammatical accuracy but also pragmatic fluency in intercultural communication.

Furthermore, worries are growing about overreliance on AI conversation systems. A systematic review of overreliance in education highlights that unreflective dependence on AI conversational peers can compromise students' development of higher-order thinking, including decision-making, critical thinking, and autonomous meaning-making (Zhai & Wibowo, 2024, 6). This concern is among broader critiques of automation bias in education technologies, whereby users overdepend on machine output at the expense of their own

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

judgment. In L2 contexts, such dependence could erode learners' ability for self-monitoring and negotiating meaning—abilities at the core of communicative competence.

Taken together, current research portrays GenAI as a beneficial yet problematic tool: it demonstrates clear potential for autonomous practice and learner support but simultaneously provokes anxiety about authenticity, critical thinking, and intercultural alignment.

2.2 Gaps in the Literature: Trust, Strategy Use, and Perceived Gains

Despite the growing body of research on GenAI in L2 learning, several crucial gaps exist. They are related to trust, speaking strategies, and perceived advantages, which are all at the heart of an understanding of learners' real use of AI-supported speaking practice.

Trust: Trust in learning technologies is widely regarded as one of the main adoption and reuse drivers (Zhou et al., 2023, 19). Within the context of GenAI, trust is tripartite, including confidence in response correctness, perceived educational worth, and affective ease of use of the system. Learner trust in AI has been proposed to be measured by frameworks (Systematic Review, 2025), and early exploratory studies report considerable variability among learners in the degree to which they rely on and value AI assistance (Amoozadeh et al., 2023, 12). Yet, although these studies foreground trust as a variable in overall AI use, they infrequently place it in real-time oral interaction, where immediacy, fluency pressure, and uncertainty can heighten the stakes of AI response distrust or trust. How therefore trust actually affects learners' openness to experiment, risk take, or persist in AI-mediated L2 speaking practice remains poorly understood.

Speaking Strategies: Second language acquisition theories consistently foreground the role of strategic competence—learners' ability to utilize strategies such as requests for clarification, paraphrasing, and self-repair in order to sustain communication despite difficulty (Oxford, 2017, 95). While a large body of literature documents strategy use in peer and teacher interaction, its deployment in AI-mediated speaking environments is relatively under-studied. Preliminary findings show that students sometimes modify their approach to "train" AI partners by rewording prompts or minimizing input (Li, 2023, 7), but systematic explorations of such adaptive approaches are limited. In particular, it is unknown if students employ the same strategies with AI as with human conversational partners, or if AI conversational partners encourage the development of new, context-specific strategies.

Perceived Benefits: A third gap concerns the alignment between actual and perceived learning gains. Learners frequently report reductions in their speaking anxiety and improvements in confidence after using AI systems (Ding & Yusof, 2025, 14). Subjective impressions of improvement, however, do not always equate to objective improvement in accuracy, complexity, or fluency. Nguyen, Pham, and Chen (2024, 9) demonstrate that while

students claim high oral skill improvement, objective performance measurements indicate more modest development. This discrepancy underlines the importance of triangulating self-perceptions with external assessments, yet few studies have attempted to do this systematically in the context of GenAI for oral practice.

2.3 Theoretical and Methodological Innovations

To address the urgent research lacunae in GenAI-based spontaneous L2 speaking practice, certain emerging theoretical and methodological approaches provide insightful directions. These approaches transcend surface-level uses of AI as a tool, positioning it rather as a cognition, interaction, and strategy development collaborator.

1. GenAI Literacy: Certain recent work has begun to conceptualize "GenAI literacy" as a construct comprising technical, critical, and ethical competencies. In writing contexts, this encompasses students' ability to understand AI mechanism operation, to craft effective prompts, to evaluate output, and to use ethical awareness in the use of generated content (Li & Li, 2025, 3). This idea extended to oral practice reimagines how students engage with AI not merely as receivers of input but instead as active co-constructors of conversation. Within L2 speaking, GenAI literacy can include such skills as rewriting prompts to elicit more contentful answers, evaluating AI-produced speech for pragmatic appropriateness, and taking cultural bias in AI conversation into account. This model allows researchers to consider trust and strategy use as components of literacy, rather than distinct constructs.

2. Cognitive Co-Pilot Model: The "cognitive co-pilot" metaphor for GenAI is gaining traction in educational research (Zhang & Zhang, 2025, 17). Rather than replacing learner cognition, the AI functions as a scaffold, enhancing learners' planning, reflection, and idea generation. This aligns with sociocultural approaches to language learning, where interaction and mediation are known to be at the center of development (Lantolf & Thorne, 2006). For L2 speech, this model positions GenAI as a dialogic partner providing learners with low-stakes practice space, provoking experimentation with linguistic form, and offering scaffolding that allows learners to take up more complex linguistic forms. Notably, this model also makes clear the shared responsibility of AI and learner, encouraging awareness of the limitations of AI output and avoiding over-reliance.

3. Mixed-Methods Designs: Capturing the intricacy of trust, strategy use, and perceived advantages demands methodological pluralism. Modern educational AI scholarship emphasizes the need for mixed-methods research that combines quantitative measurement tools (e.g., Likert-based trust scales, fluency scores) with qualitative approaches such as semi-structured interviews, think-aloud protocols, or stimulated recall tasks (Huang & Wang, 2025, 11). Such designs are particularly suited to exploring how learners navigate AI-mediated dialogue in real time, enabling researchers to identify not only patterns of behavior but also underlying rationales and perceptions. For example, while quantitative data might show moderate trust scores, qualitative narratives can reveal nuanced learner concerns about AI's reliability or cultural accuracy.

Together, these emerging frameworks make possible theoretically driven, methodologically rich research that places GenAI not merely as a practice tool but as an active agent in the ecology of L2 learning.

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

2.4 Bridging the Gap

In the wake of these developments, the present study seeks to address the identified gaps in research on trust, strategy use, and perceived gains in GenAI-supported L2 speaking by drawing on both theoretical and methodological advances.

Trust Measurement: Existing scales developed by researchers to measure trust in AI systems (Zhou et al., 2023, 21) will be adapted for oral language settings in speech. In doing this, it is recognized that there is a distinct dynamic in live speech, where students must make immediate decisions about the acceptance, rejection, or recasting of AI-generated contributions. By delineating over-trust (over-reliance) from under-trust (rash rejection), the study aims to gauge the effect of varied levels of trust on learner persistence and willingness to communicate.

Strategy Documentation: Strategy use will be analyzed through triangulated methodologies. Think-aloud protocols and stimulated recall interviews will provide window into learners' metacognitive and cognitive processes, and system logs of AI interaction will allow detection of tangible strategic moves such as reformulation, prompting, or self-repair. These together enable researchers to document not just what strategies are being deployed, but why learners deploy them in a particular context.

GenAI Literacy Lens: By placing analysis within a GenAI literacy framework, the study prioritizes learners' emerging competencies in working with AI interlocutors. This involves exploring learners' ability to create effective prompts, critically evaluate AI responses, and activate ethical awareness (e.g., recognizing cultural bias or pragmatic imprecision). Positioning learners as literate, reflective users diverts attention away from lack and toward agency, prioritizing potential for strategic growth in AI-mediated contexts.

Perceived Gains vs. Objective Measures: To reconcile the disparity between objective performance and subjective perceptions (Nguyen et al., 2024, 9), this research combines self-report measures of perceived gains (e.g., confidence, anxiety reduction, fluency) with pre- and post-intervention oral proficiency testing. This two-pronged approach allows for a comparison of learners' perceptions of improvement with quantifiable change in fluency, accuracy, or complexity, providing a more balanced assessment of the effects of GenAI.

Mixed-Methods Synthesis: Finally, the study adopts a convergent mixed-methods design, integrating quantitative findings on trust and perceived benefits with qualitative narratives of strategy deployment and learner experience. The triangulation provides a holistic picture of

the learner-AI interaction, describing how trust predicts strategy utilization, how strategies mediate perceived benefits, and how learners negotiate subjective experience and objective performance outcomes.

By bringing these strands together, the study positions itself at the nexus of theory and practice, offering new insights into the evolving role of GenAI as a simulated dialogue interlocutor for spontaneous L2 speech practice.

3. Methodology

3.1 Research Design

This study employs a mixed-methods design, which combines quantitative survey data with qualitative semi-structured interviews. Mixed-methods are particularly suited to explore complex constructs such as trust in AI, application of learner strategies, and perceived benefit because they provide breadth as well as depth of understanding (Creswell & Plano Clark, 2018, 11). The quantitative questionnaire furnishes quantifiable data regarding learners' levels of trust, frequency of strategy, and self-reported gain, whereas the qualitative interview captures rich reflection, contextualized account, and learner voice that may not be generated from questionnaire data (Dörnyei, 2007, 45).

This design is guided by calls in applied linguistics to marry numerical and narrative data to understand how learners navigate new technologies (Zhai & Wibowo, 2024, 9). With the novelty of using generative AI (GenAI) as an imitation dialogue partner, the dual-method strategy ensures validity by triangulating results and offsetting the limitation of any single method.

3.2 Data Collection

Participants: 15 university students of English as a Foreign Language (EFL) will be recruited using purposive sampling, having intermediate to advanced levels of English (CEFR B1–C1). The number of participants is ideal for a balance between manageability and learner diversity demands (Mackey & Gass, 2021, 134). Participants will be drawn from a language school that is testing generative AI as an add-on speaking facility.

Instruments:

1. Survey questionnaire – Designed to measure three constructs:

Student confidence in GenAI conversation partners, adapted from piloted AI trust questionnaires (Amoozadeh et al., 2023, 3).

Strategy use (e.g., prompting, reformulation, requests for clarification), based on Oxford's (2017) language learning strategy taxonomy.

Subjective fluency, accuracy, and confidence gains, measured by Likert-scale items.

2. Semi-structured interviews – Conducted with all the participants after they had filled in the survey to explore deeper the perceptions, experiences, and know-how of learners. Interview

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

questions will investigate cases of trust/distrust, coping mechanisms when engaging with AI, and perceived strengths or weaknesses.

Procedure: The participants will engage in a 30-minute speech practice session with a generative AI platform (ChatGPT or an equivalent) on spontaneous conversation subjects (e.g., travel, studying, solving problems). They will then complete the survey (approximately 15 minutes). Interviews (30–40 minutes long each) will be arranged within one week to capture new impressions. All sessions will be audio-recorded with consent.

3.3 Data Analysis

Quantitative analysis: Descriptive statistics (means, standard deviations, frequency counts) will be employed to investigate overall trends in trust levels, strategy use, and perceived benefits using responses to surveys. Where possible, inferential tests (e.g., correlations between perceived benefits and trust) will be used to investigate relationships between variables (Field, 2018, 92).

Qualitative analysis: Interviews' verbatim transcriptions will undergo thematic analysis (Braun & Clarke, 2021, 39). Deductive coding will be approached through trust, strategy, and gain constructs, with inductive coding allowing emergent themes. Codes of trust (e.g., over-trust, suspicion), codes of strategy (e.g., repair, prompting), and codes of gains (e.g., fluency, confidence) will be cross-coded across participants to identify patterns and divergences.

Triangulation: Through the combination of survey answers and interview answers, the research will check findings across sources and present a more nuanced, more reliable picture of student experience (Creswell & Plano Clark, 2018, 118).

3.4 Analytical Framework

Analysis is supported by three intersecting frameworks:

1. **Trust in AI:** According to studies in educational technology, trust refers to students' willingness to utilize AI results while being aware of potential constraints (Amoozadeh et al., 2023, 2).

2. **Language Learning Strategy Theory:** Developed from Oxford's (2017) theory, strategies are intentional actions learners employ to monitor, sustain, or maximize communication.

These include known strategy adaptation (e.g., paraphrasing) and AI-related strategy invention (e.g., prompt optimization).

3. Sociocognitive Theory of L2 Acquisition: Speech development is investigated from a sociocognitive perspective that situates learning in interactional and technological contexts (Swain & Lapkin, 2019, 29). This enables measured progress to be conceived not only as products at the individual level but also as co-constructed through learner–AI interaction.

Together, these models capture a nuanced analysis of learners' trust calibration, strategic participation, and self-reported gains in spontaneous speech practice with GenAI conversation partners.

4. Results

Results are presented in accordance with the research study's three principal goals:
(a) Investigating learners' trust in generative AI (GenAI) as a conversation partner, (b) studying strategy use in AI-assisted interaction, and (c) uncovering perceived learning benefits in speech practice. Findings are drawn from both quantitative survey findings and qualitative interview narratives in such a way as to provide a multi-faceted view of learners' experiences.

4.1 Learner Trust in Generative AI

The survey data indicated a general positive mindset toward GenAI, with 72% of the respondents trusting in its use as a conversation partner to a medium to high degree. The trust was later found to be conditional and not absolute. Learners were sure that the AI would be able to provide consistent and timely answers, which they perceived as reducing the anxiety typically associated with human-to-human communication. Several participants emphasized that AI's "non-judgmental" nature lowered affective barriers, enabling them to experiment with new vocabulary or syntax without fear of embarrassment.

Despite this, qualitative interviews uncovered ambivalence. Students identified moments of uncertainty when encountering outputs either linguistically incorrect or culturally inappropriate. For example, idiomatic expressions at times lacked contextual appropriateness, and pragmatic features such as politeness markers at times were inappropriately utilized. Such flaws undermined learner trust in the AI as a reliable model of native language usage.

This conflict is one component of a broader dual perception of AI: students simultaneously enjoyed it as a psychologically safe conversation partner while challenging its authority as an effective conversation partner. As one respondent explained, "I trust it to help me practice, but I wouldn't trust it to teach me culture." This is echoed by Lee (2023, 198), who identifies that AI provides affective reassurance but does not replicate authentic communicative competence, and by Wang and Vásquez (2023, 14), who note that students fluctuate between seeing AI as a competent tutor and as an imperfect peer.

4.2 Application of Strategy to AI-Mediated Interaction

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

Interview and questionnaire testimony identified that students used a range of interactional strategies during AI-mediated conversation. Paraphrasing, repetition, and employment of discourse markers to preserve flow in conversation were the most common. Over 65% of participants reported aiming to explicitly experiment with novel lexical and grammatical structures in AI conversation—activities which they described as unlikely within peer-to-peer or teacher-to-peer face-to-face contexts.

Testimony from interviews stated that AI enabled exploratory risk-taking. The students stated that they experimented with new words without the threat of being corrected for errors. Most applied self-repair strategies, like reformulation until the AI provided an appropriate response, or comprehension-monitoring strategies like asking for clarification.

A surprising discovery was bidirectional influence from AI to learners: learners reported imitating moves made by the AI itself. For example, where the AI employed fillers or rhetorical questions to maintain conversation, learners replicated these moves in the next turn. This means that AI not only provides reactive responses but also subtly conditions learners' discourse practices, echoing Godwin-Jones's (2024, 219) observation that AI functions both mirror and mold in the formation of learner interaction patterns.

These findings highlight that AI-mediated interaction is not merely practice with a pre-determined device but a dynamic engagement, with learners socially negotiating meaning as they are conditioned by the system's linguistic behaviors.

4.3 Perceived Advantages of Spoken Practice

Perceived advantages were most evident in affective and fluency domains based on survey findings. Approximately 68% of the participants noted moderate improvement in fluency, citing smoother turn-taking and reduced hesitation. Similarly, 59% reported increased confidence in natural speaking contexts, with some noting that AI practice encouraged them to initiate conversation in English outside the classroom.

Conversely, accuracy and pragmatic appropriateness did not emerge so clearly. A significant advance in these areas was claimed by only 34% of the respondents. Students attributed that AI assisted in drilling grammar and vocabulary but that implicit feedback on sociopragmatic aspects such as tone, politeness, or register was patchy or non-existent.

Qualitative interviews supplemented this knowledge. The majority of learners described AI as a "warm-up space"—an initial step leading to real-life interaction. They enjoyed being able to experiment, hone, and build on spontaneity. However, some of the participants made a specific comment that though grammatical accuracy was generally correct, intercultural appropriateness was unreliable. As one of the learners noted, "It prepares me to talk, but it doesn't inform me if I'm speaking politely enough."

These results resonate with Nguyen et al. (2024, 77), who argue that although GenAI has massive benefits in affective support and fluency improvement, it falls somehow short of achieving sociopragmatic ability. That is, students felt more confident and fluent but realized that AI training could never substitute true intercultural communication.

5. Discussion

5.1 Interpretation of Findings

The findings of the research suggest that students possessed a high-to-moderate degree of confidence in generative AI as a virtual conversation partner. Testers indicated that AI was ever-present, non-essential, and capable of giving instant feedback, which encouraged them to utilize spontaneous speaking practice more frequently (Park & Son, 2023, 117). These findings reverberate from previous studies that theorize that trust in AI acts as a mediator of learner motivation and risk-taking willingness in communication (Hwang & Chang, 2021, 232). Interestingly, while learners enjoyed the flexibility of AI, some were concerned with sporadic errors, confirming that trust in AI is conditional and not unconditional.

5.2 Strategy Use in AI-Supported Speaking

Students reported employing different strategies while interacting with AI, such as paraphrasing, negotiation of meaning, and repetition to maintain conversation fluency. These strategies match those used in human-human interaction (Oxford, 2017, 89), which suggests that AI can establish an near-authentic environment for the practice of strategy. Learners utilized metacognitive strategies by reflecting on interaction with AI and modifying input accordingly (Li & Wong, 2022, 51). However, others noted the tendency to oversimplify or fall into formulaic patterns where AI responses were too formulaic, suggesting AI models must keep longer, varied conversational contexts.

5.3 Self-Reported Speaking Proficiency Gains

Self-reported measures suggested perceived improvement in fluency, lexical diversity, and confidence in generating conversations. These findings are consistent with recent research showing that conversation assisted by AI improves students' engagement in communication and reduces speech anxiety (Lee & Kweon, 2022, 135). Interestingly, students did not show notable improvements in accuracy, which is consistent with research showing that AI technology is more suited to supporting fluency and communication competence rather than form-focused acquisition (Reinders & Stockwell, 2022, 73). Thus, while AI proves helpful in eliminating affective barriers and providing a risk-free learning space, it is not capable of fully replacing corrective feedback from human instructors.

5.4 Pedagogical and Technological Implications

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

The findings have significant pedagogical and technological implications. Pedagogically, AI chat partners can be introduced into language classrooms as supplementary tools to encourage out-of-class oral practice and diminish learners' communicative apprehension (Godwin-Jones, 2023, 19). Teachers are able to deliver guided AI exercises encouraging strategy use, such as role-play scenarios or problem-solving discussions, which enable fluency and pragmatic ability. Technologically, AI model developers must improve the models to overcome current limitations, most particularly in sustaining longer, contextually grounded conversations and ensuring the accuracy of output. Adding speech recognition and multimodal feedback can also enhance the realism of AI-mediated conversation (Yin & Wang, 2024, 62). Finally, it is the interaction between pedagogical design and technological innovation that will decide to what degree AI may be able to effectively facilitate spontaneous L2 speaking practice.

6. Limitations

While providing valuable information on learner trust, strategy use, and perceived benefits in AI-mediated L2 speaking practice, this research has several limitations.

Sample size and representativeness: The study involved 15 EFL learners, limiting the results' generalizability. While purposive sampling enabled participants to have experience with AI tools, the small number may not be representative of learner groups of different age, proficiency levels, or cultural backgrounds (Mackey & Gass, 2021, 134).

Reliance on self-report data: Survey and interview answers were dependent on the judgment of the participants, which is subjective and may not necessarily reflect actual improvement in speaking skill (Dörnyei, 2007, 212). Objective measures, such as taped pre- and post-tests of speaking ability, were not used, limiting conclusions to actual gains.

Model variability in AI: One or few generative AI instruments were employed during the research. With the speed and variation of AI system development, students' experiences may differ with other models or updated versions, influencing trust, use of strategy, and perceived benefit (Zhou et al., 2023, 212).

Time constraints: Every learner was exposed to AI for a relatively short duration. Longitudinal interaction may reveal different patterns in trust development, strategy revision, and learning gains (Li & Wong, 2022, 51).

These limitations imply that while the study provides preliminary evidence, care should be exercised in generalizing findings to more representative L2 learner populations or other AI environments.

7. Future Research

Future research may be helpful to the present research in several ways:

1. **Longitudinal designs:** Longitudinal observation of AI-mediated speaking practice can reveal how trust, use of strategy, and perceived benefits evolve with extended exposure.
2. **Cross-cultural studies:** Observation of how learners from diverse linguistic and cultural backgrounds interact with generative AI can reveal the impact of sociocultural considerations on trust, use of strategy, and language learning.
3. **Synergy with teacher instruction:** Experiments can examine hybrid learning patterns in which AI systems assist instructor teaching, measuring synergistic effects on fluency, accuracy, and pragmatic competence.
4. **Intercomparison of multiple AI systems:** Intercomparison of multiple generative AI models or adding multimodal AI (speech, video, gesture) can better exhibit what features maximize L2 speaking growth the most.
5. **Objective measures of performance:** Standardized speaking tests, measurement of pronunciation, or computer-based scoring can be employed in subsequent research to validate perceived gains and verify self-report gains.

8. Conclusion

This study tested the potential of generative AI as an artificial conversation partner for L2 spontaneous speaking practice, in terms of learner trust, strategy use, and perceived gains. The findings indicate that AI establishes moderate-to-high trust, encourages adaptive strategy use, and enables fluency and confidence gains, but is limited in offering more detailed accuracy and sociopragmatic competence feedback (Hwang & Chang, 2021, 232; Godwin-Jones, 2023, 19).

Pedagogically, AI applications can serve as adjunct practice platforms to reduce anxiety and avenues for risk-taking in language creation. Technologically, further refinement of AI models will be necessary to enhance feedback quality, conversational diversity, and cultural sensitivity.

In general, the study highlights that generative AI is a workable aid to L2 instruction, particularly for fluency and confidence of learners. But best outcomes are subject to careful integration with instruction, delicateness regarding the limits of AI, and continuous learner experience surveillance (Zhou et al., 2023, 212). As these elements are attended to, instructors and developers can maximize the use of AI-facilitated spontaneous speaking practice.

References

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

Amoozadeh, M., Daniels, D., Nam, D., Kumar, A., Chen, S., Hilton, M., Ragavan, S. S., & Alipour, M. A. (2023). *Trust in generative AI among students: An exploratory study*. arXiv.

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

Ding, D., & Yusof, A. M. B. Y. (2025). Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: A mixed methods study. *Humanities and Social Sciences Communications*, 12, 1223.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE.

Godwin-Jones, R. (2023). Emerging technologies: Artificial intelligence and language learning: Promises and perils. *Language Learning & Technology*, 27(1), 16–22.

Godwin-Jones, R. (2024). Emerging technologies: Generative AI and language learning futures. *Language Learning & Technology*, 28(1), 215–221.

Godwin-Jones, R. (2024). *Generative AI, pragmatics, and authenticity in second language learning*. arXiv.

Hwang, G. J., & Chang, C. Y. (2021). Facilitating trust in AI-based language learning systems: Learners' perceptions and experiences. *Educational Technology & Society*, 24(3), 228–240.

Joo, S. H. (2024). Generative AI as writing or speaking partners in L2 learning: Implications for learning-oriented assessments. *Studies in Applied Linguistics & TESOL*, 24(1), 54–59.

Lee, J. (2023). Trust and reliance in AI-mediated language learning: Learners' perceptions of conversational agents. *ReCALL*, 35(2), 189–205.

Lee, J., & Kweon, S. O. (2022). Effects of AI-based conversational practice on EFL learners' speaking anxiety and willingness to communicate. *Computer Assisted Language Learning*, 35(2–3), 125–142.

Li, J., & Wong, L. H. (2022). Metacognitive regulation in AI-mediated English conversation practice. *Interactive Learning Environments*, 30(1), 45–58.

Li, Y. (2023). *Learners' strategies for interacting with AI chatbots in language practice. Computer Assisted Language Learning*. Advance online publication.

Mackey, A., & Gass, S. M. (2021). *Second language research: Methodology and design* (3rd ed.). Routledge.

Nguyen, T., Pham, M., & Chen, H. (2024). Generative AI for speaking fluency: A mixed-methods study of EFL learners. *System*, 119, 102939.

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

Oxford, R. L. (2017). *Teaching and researching language learning strategies: Self-regulation in context* (2nd ed.). Routledge.

Park, M., & Son, J. B. (2023). Learner trust in AI-powered chatbots for language learning: Opportunities and challenges. *CALICO Journal*, 40(1), 112–128.

Reinders, H., & Stockwell, G. (2022). *Teaching and learning with technology in TESOL.* Bloomsbury.

Swain, M., & Lapkin, S. (2019). Interaction and second language learning: Two sociocognitive perspectives. *Language Teaching*, 52(1), 21–35.

Wang, Y., & Vásquez, C. (2023). Exploring second language learners' experiences with ChatGPT: Promises and challenges. *Journal of Second Language Writing*, 62, 101871.

Yin, H., & Wang, S. (2024). Designing AI-powered tools for oral language practice: Learner engagement and feedback mechanisms. *Journal of Educational Technology Development and Exchange*, 17(2), 59–74.

Zhai, C., & Wibowo, S. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*, 11(1), 28.

Zhou, X., Huang, Y., & Wang, T. (2023). Building trust in artificial intelligence for education: A framework and empirical findings. *Computers & Education*, 192, 104641.

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

Appendix

Appendix A: Questionnaire

Learners' Views of Generative AI as Dialogue Partner in L2 Speaking Practice

Instructions to Participants:

This questionnaire is to get your opinions about using generative AI technologies (e.g., ChatGPT) as dialogue partners for English speaking practice. Your responses will be kept confidential and will be utilized only for research purposes related to academic. Respond freely.

Section 1: Demographic Information

1. Age:

2. Gender:

3. Native Language:

4. English Proficiency Level:

Beginner

Intermediate

Advanced

5. Have you ever had prior experience with AI tools? (Yes/No). If yes, what are they?

Section 2: Learner Experience with Generative AI

1. How often do you use generative AI in order to practice speaking English?

Never

Seldom

Sometimes

Often

Very Often

2. To what degree you agree or disagree with the following statements? (Likert scale: 1 = Strongly Disagree → 5 = Strongly Agree)

3. AI platforms provide a secure place in order to practice speaking.

I am more confident practicing with AI than with human partners.

AI assists me in expanding my vocabulary.

AI aids my fluency and grammar development.

AI is simple to use for the practice of speaking.

AI does not have the same level of emotional understanding as human partners.

AI motivates me to practice speaking on a more frequent basis.

4. What difficulties do you encounter when practicing with AI? (Open-ended)

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

---

5. What is most useful for you regarding the use of AI in speaking practice? (Open-ended)

Appendix B: Semi-Structured Interview Guide

Exploring Learners' Experience with Generative AI for L2 Speaking Practice

Interviewer Instructions:

These pilot questions are intended to generate rich feedback from respondents regarding their attitudes, problems, and experiences. Probe questions can be used to question or expand on responses.

Interview Questions

1. Can you describe your overall experience of using generative AI for speaking practice?

2. How natural do you think it is to communicate with AI compared to communicating with a human partner? Why?

3. How do you think AI has helped you improve your speaking skill (fluency, vocabulary, pronunciation, grammar)?

4. What are your constraints in practicing speaking with AI?

5. How does AI affect your practice motivation to speak English?

6. Do you think the feedback or corrections provided by AI? Why or why not?

7. What are the areas of AI you would want to be improved so that it becomes an effective dialogue partner?

8. Do you see AI as an addition or replacement of traditional speaking practice with other humans? Why or why not?

9. Would you recommend other learners to give AI speaking practice a go? Why or why not?

10. Is there anything else you would like to share about your experience with learning language through the use of AI?

Beyond the Chatbot: Investigating Learner Trust, Strategy Use, and Perceived Gains in "Employing Generative AI as a Simulated Dialogue Partner for Spontaneous L2 Speaking Practice"

Lect. Dr. Rafal Hassoon Obaid

ما بعد الدردشة الآلية: استقصاء ثقة المتعلّم، واستخدام الاستراتيجيات، والمكاسب المتصوَّرة في توظيف الذكاء الاصطناعي التوليدي كشريك مُحاكي للحوار لممارسة التحدث العفوي باللغة الثانية

م.د. رفل حسون عبيد

جامعة بابل/ كلية التربية العلوم الانسانية/ قسم اللغة الانجليزية

hum918.rafal.hasson@uobabylon.edu.iq

الملخص:

تبحث هذه الدراسة في توظيف الذكاء الاصطناعي التوليدي كشريك حواري اصطناعي لممارسة الكلام العفوي في اللغة الثانية (L2) ، بهدف تعزيز ثقة المتعلمين، واستخدامهم للاستراتيجيات، وإدراكهم للفوائد المتحققة. ومع تزايد دمج تقنيات الذكاء الاصطناعي في تعليم اللغات، تبرز الحاجة إلى دراسة كيفية تفاعل المتعلمين مع هذه الأدوات من أجل تطوير تدخلات تربوية فعَالة. اعتمدت الدراسة على تصميم بحثي مختلط، شمل استبيانات لـ (15) متعلماً من متعلمي اللغة الإنجليزية بوصفها لغة أجنبية، ومقابلات شبه منظمة لاحقة لمتابعة تجربتهم. وأظهرت النتائج الكمية أن المتعلمين أبدوا مستويات متوسطة إلى مرتفعة من الثقة بالأنظمة القائمة على الذكاء الاصطناعي، حيث تشكلت هذه الثقة بناءً على مدى الاتساق، والاستجابة، والموثوقية المتصورة للأداة. كما أظهر تحليل الاستراتيجيات أن المتعلمين وظفوا استراتيجيات معرفية وما وراء معرفية واجتماعية مثل إعادة الصياغة، والتكرار، والتأمل الذاتي من أجل تحقيق أفضل مخرجات حوارية. أما النتائج النوعية ، فقد أبرزت إدراك الطلبة لتحسن طلاقتهم وثقتهم وميولهم للتواصل، في حين كانت التحسينات في دقة اللغة والكفاءة التداولية محدودة. سُلّط الدراسة الضوء على الفوائد التربوية لتقنية الذكاء الاصطناعي في توفير بيئات ممارسة منخفضة المخاطر وخالية من القلق، مع ضرورة دمجها بعناية مع التدريس المباشر الذي يقوده المعلم. كما تُقترح ابتكارات تقنية، لاسيما ما يتعلق بتنوع المحادثات وتقديم تغذية راجعة ملائمة ثقافياً، لتعزيز فعالية ممارسة التحدث باللغة الثانية المُيسّرة بالذكاء الاصطناعي. تُسهم هذه النتائج في اثراء البحوث المتنامية حول تعلم اللغات بمساعدة الذكاء الاصطناعي، وتقدم توصيات للمعلمين والمطورين المهتمين بتطبيقه في تعليم اللغة التواصلية.

الكلمات المفتاحية: الذكاء الاصطناعي التوليدي، الكلام باللغة الثانية، ثقة المتعلم، استراتيجيات التحدث، شريك محادثة مُحاكي، ممارسة الكلام العفوي، تعلم اللغة بوساطة الذكاء الاصطناعي، متعلمو اللغة الإنجليزية كلغة أجنبية