EL-MAQALAH: Journal of Arabic Language Teaching and Linguistics
Vol. 7. No. 2, 171-186, 2026

Voice-Based Generative AI as a Conversational Practice Partner in Arabic Speaking: Indonesian University Students' Experiences of Affordances and Constraints

Azwir*
Universitas Islam Negeri Ar-Raniry, Banda Aceh, Indonesia

*E-mail: azwirm.ali@ar-raniry.ac.id

ABSTRACT

Arabic programmes rarely offer opportunities for interactive speaking practice, and voice-based GenAI might be able to increase these opportunities, but little is known about the experience of Arabic learners with this type of GenAI as a conversational practice partner. This qualitative case study involved 12 students from an Arabic education programme at an Islamic university in Indonesia who participated in three voice-based sessions (15-20 minutes each) with a GenAI platform and subsequently in semi-structured interviews. Reflexive thematic analysis was used to analyse the conversation transcripts and interviews, using Affordance Theory as the main focus and the Interaction Hypothesis and Communicative Language Teaching as complementary theories. After a demonstration by a lecturer, students went from being curious and doubtful of the AI to recognising it as an interlocutor. The perceived affordances were the lower pressure environment for verbalization and turn taking through the AI's generated questions, immediate feedback (and opportunity to try again), and the confidence and repeated use of phrases with peers in class. Students corrected occasional misunderstandings (constraint 1) by providing clarification and social pressure was decreased due to the lack of nonverbal cues (constraint 2), but so were constraints on authenticity. The findings were based on a designed interaction as the AI was instructed to ask questions and provide short corrections. The study is unable to claim improvements in language fluency or the correctness of the AI's input. It recommends that voice-based AI practice is more suited to be supplemented by lecturer-led and interpersonal interaction.

Keywords: *Generative AI, Conversational Practice Partner, Arabic Speaking, Mahārah Kalam, Affordance, Constraint, Qualitative Case Study*

Citation: Azwir. "Voice-Based Generative AI as a Conversational Practice Partner in Arabic Speaking: Indonesian University Students' Experiences of Affordances and Constraints." *EL-MAQALAH: Journal of Arabic Language Teaching and Linguistics* 7.2 (2026): 171-186.

DOI: https://doi.org/10.22373/6awyzb96

Lisensi
Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

INTRODUCTION

Among the four Arabic language skills, speaking (*mahārah kalām*) plays an important role because it enables learners to use their linguistic knowledge in real-time communication. This skill is more than just a learning of words and grammar; it is a skill which must be learnt to enable the learner to be able to use words and grammar to communicate meaningfully to others. Students must learn the forms of communication and use them properly if they want to become competent communicators (Canale & Swain, 1980). This skill will be developed when the students have chances to use the language productively and interact with the speaker, and are able to carry out a conversation that involves the exchange of meaning (Hui & Yunus, 2023).

The students have limited opportunities for formal Arabic discourse in the classroom, including the limited class time, large class sizes, students' varying levels of proficiency, and unequal allocation of speaking turns in the classroom, meaning some students may have little speaking practice during the lesson (Laili et al., 2022). The pedagogical dilemma is then not only the question of how to teach *mahārah kalām*, but the amount of worthwhile, engaging speaking practice that students can receive.

This activity is important and interactive, as seen in interactionist research, which shows that the interactive nature of this activity makes it important, because in conversation, the participants will receive the input, will be prompted to produce output, and will be able to correct themselves if there are problems with the previous output produced (Long, 1996; Pica, 1994; Swain, 2005). The dialogue is, therefore, not a simple repetition of individual sentences, since the other person's answer determines how the speaker should structure his/her next turn.

The opportunities for language practice have been extended beyond the classroom using digital technologies increasingly. Huang et al. (2022) conducted a systematic review of the studies that have been conducted on Chatbots for language learning and identified the benefits of using Chatbots for language learning as accessibility, personalisation, immediacy and simulation of conversations. There are some challenges that come with their use, however, including technical problems, cognitive load, and novelty effects. Generative AI (GenAI) takes these capabilities further by enabling the use of large language models that can generate response that is relevant to the context and can adapt to user input, rather than following a scripted route. Liu et al. (2024) discovered that students proactively discussed the capabilities and limitations of GenAI as they used it in their learning process as a tutor and conversation partner. Conversational practice partner (CPP) is defined as an AI system that assumes the conversational role, asks and answers questions and keeps the topic in the conversation, forcing the learners to listen and respond. This is not the same as using AI to translate texts, teach grammar or correct sentences (Belda-Medina & Calvo-Ferrer, 2022).

The empirical data on GenAI-supported speaking practice are promising, but a positive experience should not be equated with improved performance. Wang et al. (2024) investigated text-based and voice-based GenAI chatbots in EFL speaking and identified changes in willingness to communicate, speaking anxiety, and self-perceived communicative competence for some participants, but no differences in speaking performance. From speaking, positive results were found by Tai and Chen (2024) and participants liked the chatbot's capacity to sustain a conversation, respond correctly, offer various subjects, and speak in a human-like manner. Although students do not consider conversational agents as an alternative to

communicating with a human, they seem to have positive attitudes about them (Belda-Medina & Calvo-Ferrer, 2022). There were two aspects of this experience revealed in the qualitative evidence in a study of Character.AI valued the flexibility and opportunities for practice but had difficulties with context understanding and speech recognition (Napitupulu & Dalimunte, 2025), and intelligent personal assistants have been shown to have difficulty understanding L2 speech (Moussalli & Cardoso, 2020).

The findings show that conversational AI experiences for learners should be explored through more than just in speaking. While some research has found benefits for speaking proficiency, willingness to communicate and/or anxiety reduction, some studies have recognized weaknesses in understanding context and/or recognizing speech. These results indicate various aspects of AI-supported speaking practice: the experiences of learners, the opportunities offered by the technology during speaking interactions, and the challenges presented by the technology when learners try to interact with it. These dimensions are worth of deeper reflection, especially given the use of AI not just as a language-learning tool, but as an active conversational partner.

While there is increasing interest in the use of GenAI in language learning, most of the empirical evidence related to the usage of AI for speaking practice has come from English as a Foreign Language (EFL) setting. Studies on GenAI in the field of Arabic language education are also undergoing, but they are still limited. Several issues have been cited with using AI for Arabic, such as language characteristics, the scarcity of digital resources, the accuracy of the output created by AI and moral and cultural issues (Moustafa et al., 2026). Some initial research has started to investigate whether ChatGPT can be used to practice the Arabic language (Kholid & Majeed, 2025; Setiawati et al., 2025; Fahmiyah et al., 2026). In addition, the voice-based interaction brings new questions that should be addressed with a learner focus, such as the question of the use of the Modern Standard Arabic (MSA) or spoken varieties, the accuracy of speech recognition, and the reliability of AI-generated feedback on learners' errors in spoken production. These problems are especially relevant as the quality of interaction that is mediated by AI is strongly dependent on the ability of the system to recognize and appropriately respond to learners' speech.

Thus, the research gap does not stem from the lack of research on AI and speaking, but from the limited knowledge of Arabic learners' experiences and interaction with GenAI as a conversational practice partner. Specifically, there is limited information about the speaking opportunities the learners have and the challenges learners face during conversations, as well as how these challenges are addressed. The existing research have mostly focused on technology acceptance, speaking performance, willingness to speak, or speaking anxiety (Belda-Medina & Calvo-Ferrer, 2022; Tai & Chen, 2024; Wang et al., 2024). These studies offer valuable insights, but they are not comprehensive of the learner's experience of and management of the interaction. To fill this gap, the current study relies on two types of evidence: semi-structured interviews and transcripts of students' interactions with the AI system. The transcripts of the interviews give insight into students' perceptions and experiences, and in the case of the interaction transcripts, evidence of how conversations happen in practice. Both are examined in this study, which explores the use of voice-based GenAI in the context of Arabic speaking practice in higher education that is still largely untouched by research within the context of Indonesia Islamic higher education.

Mostly the study is guided by Affordance Theory, which explores the value of a technology in terms of what learners can do with it in context, rather than what it can do for them (Chapelle, 2001; van Lier, 2000). Two other perspectives support it: The Interaction Hypothesis (Long, 1996) on how input and output, feedback, and clarification are involved in conversation; and Communicative Language Teaching (Canale & Swain, 1980; Savignon, 2002) that provides the pedagogical foundation for meaningful language use in speaking. The study is guided by three voice-based interaction sessions with a GenAI system with 12 students in one Islamic higher education institution (IHEI) in Indonesia and focuses on the following research questions: RQ1. What effect does voice-based generative AI have on Indonesian university students' experience as a conversational practice partner in Arabic speaking (*mahārah kalām*) practice? RQ2. What communication opportunities for speaking do students see and utilize in conversations in the context of using AI? RQ3. What are some of the limitations in the conversation that students face in the world of AI, and what are their reactions to these limitations? This research aims to offer empirical evidence of students' experiences with the use of voice-based AI to practice their speaking of Arabic, and to apply Affordance Theory to explain the opportunities and limitations, through the use of voice-based AI, in this process, while emphasizing the significance of its use as an addition to, rather than a replacement for, classroom teaching and interaction with human conversation partners.

Literature Review

Speaking Practice as Communicative and Interactional Activity

The communicative language teaching (CLT) approach focuses primarily on the language as a medium of communication, and thus places less emphasis on the presentation of forms and more on the activity itself in the classroom (Savignon, 2002). In this tradition, communicative competence is considered to involve a grammatical, sociolinguistic, and strategic component (Canale & Swain, 1980); that is, speaking development is dependent on the learner's ability to attempt messages, adjust to the interlocutor and maintain the exchange (Hui & Yunus, 2023). Similarly, development of speech skills has been linked to the level of communicative interaction and practice (Attruk, 2025) in Arabic. From this viewpoint, mistakes are not considered failures but rather part of the process of learning to communicate, as learners will have ample opportunities to attempt again. The interactive nature of such practice is why it is important according to the Interaction Hypothesis. Interaction provides input, the need for output, and when there is a breakdown in communication, there is also a need to negotiate meaning through repetition, clarification and confirmation (Long, 1996; Pica, 1994). The Output Hypothesis also suggests that creating language can focus the learners' attention on the discrepancy between the learner's intended message and his or her actual ability to present the message (Swain, 2005). These accounts together help to differentiate between interaction and the simple exchange of speech: during conversation, the interlocutor's turn requires that the learner say something else. These accounts have been formulated in the context of human interlocutors, and this remains an open question as to whether it can be extended to AI. They are not a model of acquisition to be tested with the data in this study; rather they are used as sensitising concepts to interpret transcripts (e.g., turn taking, clarification sequences, and feedback followed by L2 repetition).

Affordance Theory and Technology-Mediated Practice

The study's principal analytic perspective is the Affordance Theory. According to van Lier (2000), in the field of language learning, an affordance is one of the possibilities for action that arises between a learner and the environment when it is perceived and utilized by the learner. When applied to technology, this implies that the features of a tool are not learning opportunities themselves, but the relationship between the features of a tool and the opportunities identified and taken up by learners is what counts (Chapelle, 2001). There are two implications that guide the analysis. First, affordances are not "declared": they have to be discovered by the learner, implying that orientation and demonstration might be important. Second, the same characteristic can be a chance or hindrance for the learner and the context (Jeon, 2022). This is supported by studies of the use of GenAI, which report that when learners use GenAI they negotiate benefits and limitations (Liu et al., 2024). An affordance is a possibility for speaking practice which students describe recognising and using, while a constraint is a characteristic of the AI-mediated practice that limited, disrupted or reduced the authenticity of the practice. The two aren't mutually exclusive, but rather viewed as interconnected.

GenAI as a Conversational Practice Partner

Chatbot studies reveal that conversational technologies have the potential of being interlocutors and simulators, while also highlighting challenges like technical issues, novelty effects, and cognitive loads (Huang et al., 2022). Studies that have explored AI as a conversation partner in speaking classes have included Yang et al. (2022), in which students spoke to an AI partner on daily routines, and Tai & Chen (2024), which examined its application in individual and paired speaking tasks, as well as AI speaking partners (Wan & Moorhouse, 2025). Conversational agents have been found to be perceived well by learners and not considered as a replacement for humans in the role of an interlocutor (Belda-Medina & Calvo-Ferrer, 2022). With the development of AI-supported speaking practice, voice interaction is becoming increasingly popular (Aldamen et al., 2026; Wang et al., 2024). Concurrently, there are warnings that the results of conversational AI systems, methods, and outcomes are highly context-dependent, and therefore cannot be generalized across systems (Lai & Lee, 2024; Li et al., 2025; Soriano-Gonzalez & Belda-Medina, 2026). Students appreciated flexibility and opportunities to practise in qualitative work reports (Napitupulu & Dalimunte, 2025); and previous work showed that intelligent personal assistants were found to be challenged with the understanding of L2 speakers (Moussalli & Cardoso, 2020). Automated feedback for speaking has also been investigated (Zou et al., 2023); however, it cannot be guaranteed to be good.

METHOD

Research Design

This study used a qualitative case study design (Stake, 1995; Yin, 2018) to examine students' experiences of voice-based GenAI as a conversational practice partner in *mahārah kalām* learning. A case study was appropriate because the aim was an in-depth understanding of experience within a bounded setting, not an assessment of effectiveness through pre- and post-tests. The case was a three-session AI-mediated speaking activity in an Arabic education

programme at one Islamic higher education institution (PTKIN) in Indonesia. The sessions were embedded in the Arabic Speaking course in the third semester. Twelve students were recruited through purposive sampling. The inclusion criteria were enrolment in the Arabic education programme, prior experience of learning *mahārah kalām*, and willingness to complete the speaking sessions and the follow-up interview.

A single GenAI platform with voice-conversation capability was used to ensure that all students could engage in similar environments. Prior to the first session, the instructor demonstrated how to have a conversation with the AI. Each student completed 3 sessions, each of about 15-20 minutes, in Modern Standard Arabic: one on personal topics, e.g., self-introduction, family, campus life, and daily activities; one on situational role-plays, e.g., in a restaurant, shop, on campus, while traveling, and in various service role-plays; and one on expressing and maintaining a stance about an opinion on a familiar topic. The same master prompt was used in all sessions.

Two data sources were used. First, all conversations were audio-recorded and transcribed, giving 36 sessions (12 students × 3), or roughly 9–12 hours of interaction. Second, semi-structured interviews of 20–30 minutes were conducted after the speaking sessions. They explored students' experiences, the aspects they found helpful, and the limitations they encountered, and open questions allowed participants to raise issues the researcher had not anticipated. Interviews were conducted in Indonesian language

Data Analysis

Reflexive thematic analysis (Braun & Clarke, 2006; 2019; 2021) was used to analyse the data, in which themes are seen as patterns of meaning that are actively created by the researcher as they engage with the data. The analysis was carried out in six stages, namely: getting to know the transcripts and interviews; systematic coding; development of initial themes; checking themes against the coded data and the complete data set; clarifying and naming the themes; writing the analysis. Both sets of data were converted into codes. This was mostly an inductive process of coding and affordance and constraint were not first introduced as a coding frame but rather appeared as sensitising concepts during the development of the themes and formed the basis for the research questions, but did not dictate the themes. Transcripts were examined for turn-taking, topic continuation, clarification sequences, and feedback episodes, and interview accounts were compared with the corresponding transcripts to contextualise how students interpreted their interactions, without treating either source as more valid. Participant counts per theme are reported to show how widely a pattern appeared in this sample, not to claim prevalence.

Trustworthiness and Reflexivity

In terms of reflexivity, the four criteria of trustworthiness (credibility, transferability, dependability, and confirmability) proposed by Lincoln & Guba (1985) were employed. The triangulation of interview accounts with conversation transcripts and multiple returnings to the data in the process of transcription, coding and theme development lent credibility. Transferability was supported by a detailed description of the context, participants, platform, prompt, and procedures. A consistent procedure across sessions and interviews and the use of

excerpts for interpretation provided support for dependability and confirmability. Reflexive thematic analysis is used and the researcher's role is recognised.

Ethical Considerations

The purpose and procedures were explained to the participants who provided informed consent. There were no academic consequences to withdrawal and participation was voluntary. The participants were asked not to provide any private or sensitive personal data into the AI platform and identities were anonymised using codes. The recordings, transcripts and interview data were kept securely and solely for research.

Results and Discussion

From Doubt to Recognising the AI as a Partner

P1, P3, and P5 described a starting point at which they could not imagine how speaking practice with AI would work, because they associated speaking practice with other people. Interest appeared after the lecturer demonstrated a simulated conversation, but it was accompanied by doubt about whether the AI would respond in Arabic at all.

> "Initially, I could not imagine how speaking practice with AI could be done as I was only familiar with practicing with humans. However, after the lecturer gave an example or demonstration in class, and with some effort at listening, I became curious. Could AI actually speak Arabic with me? What if it just remained silent? Yes, at first, I was still hesitant."

This pattern shows that affordances are not immediately visible. A system capable of conversation does not become a speaking partner for learners until they can see how it can be used as one (van Lier, 2000; Chapelle, 2001), and the lecturer's demonstration appears to have helped make that use visible. The finding is consistent with work showing that learners perceive chatbot affordances differently (Jeon, 2022) and that understanding a technology's usefulness shapes experiences of AI-assisted language learning (An et al., 2023; Liu & Ma, 2023).

Two cautions apply. First, the doubt reported here concerned whether the AI would respond, not whether its responses would be correct, so it should not be read as critical evaluation of accuracy. Second, three accounts cannot show that the demonstration caused the change. They suggest that orientation deserves attention when AI is introduced as a speaking partner.

A Lower Pressure Space for Attempting Speech

Six participants (P2, P4, P7, and P9, with similar accounts from P1 and P12) described speaking with AI as less pressured than speaking with people. They reported that fear of making mistakes diminished, that they could speak without being sure their sentences were perfect, and that they could ask the AI to repeat what they had not understood.

> "It actually feels good when the AI responds. It even answers and encourages me to ask questions and speak again. The fear of making mistakes that I usually experience when communicating with humans seems to disappear. After that, I can just speak freely, sometimes even saying things randomly that are unrelated to the topic being discussed."

> "I can also ask it to repeat its response if there is something I do not understand."

This theme should be broken down into two attributes. The first is affective, and the second is instrumental. Students also saw a decrease in social pressure as expected, in line with previous research that found a correlation between GenAI-supported speaking and willingness to communicate, as well as speaking anxiety (Wang et al., 2024). Since the study did not measure anxiety, this theme refers to a perceived reduction in social pressure, rather than a decrease in actual anxiety, and the impact of digital environments on emotions has been found to vary between learners and contexts (Soyoof et al., 2023; Zadorozhnyy & Lee, 2023). The second feature is agential. Students were given the opportunity to repeat as well as to be away from an audience, thereby giving them control of the interaction.

The theme also contains a tension. One account described speaking freely to the point of producing talk unrelated to the topic. Low pressure may encourage risk-taking, but it may also reduce accountability to the conversational task, and the interview data alone cannot show whether practice stayed purposeful.

Sustained Turn-Taking Through AI-Initiated Questions

P1, P6, P8, and P10 described the AI as asking many questions and giving them something to respond to, and the transcripts show this pattern. In Excerpt 1, the AI answers the student's stated intention with a related question and then follows the student's reply with another.

Excerpt 1

Role-play at a restaurant

Turn	Utterance	Gloss
Student	السلام عليكم، أريد أن أذهب إلى المطعم.	Peace be upon you, I want to go to the restaurant.
AI	وعليكم السلام. ماذا تريد أن تأكل في المطعم؟	And upon you peace. What do you want to eat at the restaurant?
Student	أريد أن آكل الأرز والدجاج.	I want to eat rice and chicken.
AI	هل تذهب إلى المطعم وحدك أم مع أصدقائك؟	Do you go to the restaurant alone or with your friends?

Each turn of the AI requires input, which the student provides output for, and the next AI turn is built on that output. Long (1996) and Pica (1994) show how interactionists treat this type of exchange as the basis for practice. In regard to affordance, the constant and rapid questions meant that students did not need to wait for a human partner to keep talking. The flow of the conversation was along the lines of what has been reported for chatbots: being able to flow with the conversation and providing adequate responses (Tai and Chen, 2024). Likewise, chatbots have been reported to be good conversation partners for speaking practice (Wan and Moorhouse, 2025; Yang et al, 2022).

Two qualifications are needed. First, the questioning pattern was specified in the prompt (ask one question at a time), so students experienced a designed interaction. The theme shows that this design was experienced as supporting continued speaking; it does not show that the

AI spontaneously acted as an interlocutor. Second, the excerpt consists of short, simple turns, and an AI that leads with questions positions the student mainly as a respondent. Whether students also initiated topics, elaborated, or produced extended turns is not shown.

Immediate Feedback and the Chance to Retry

P2, P5, P8, and P11 described receiving feedback during the conversation and being able to try again straight away. Excerpt 2 illustrates the sequence.

Excerpt 2

Feedback within conversation

Turn	Utterance	Gloss
Student	أنا ذهبت إلى الجامعة أمس.	I went to the university yesterday.
AI	يمكنك أن تقول: ذهبتُ إلى الجامعة أمس.	You can say: I went to the university yesterday.
Student	ذهبتُ إلى الجامعة أمس.	I went to the university yesterday.

Feedback here is embedded in the conversation and followed by the student's repetition of the reformulated utterance, a sequence resembling a reformulation followed by learner uptake. It should not be read as strong evidence of modified output in the sense described by Swain (2005). The student's first utterance is not ungrammatical: the AI's version omits the independent pronoun, which is a stylistic choice, and the student's second turn is repetition rather than a self-generated modification. What the excerpt does show is that feedback was available within the flow of talk and that the interaction invited immediate retry.

The importance of this affordance relies on the quality of feedback that data does not determine. Research integrates AI with feedback functions (Huang et al., 2022). Automated feedback on speaking is also researched (Zou et al., 2023). The pedagogical functions of conversational AI are also explored (Lai & Lee, 2024; Li et al., 2025). The accuracy of GenAI with Arabic is of concern (Moustafa et al., 2026). As such students are required to analyze the AI feedback critically and corroborate it with the lecturer.

Perceived Confidence and Reuse of Expressions

P1, P4, P6, and P9 reported feeling more confident after repeated conversations, and they described imitating expressions and styles of language they met in the AI's speech when talking to classmates.

"After chatting continuously, why do I now feel more confident? Because AI can provide many things. I can imitate its way of expressing things and apply it when speaking with my classmates."

There are two points to make here: The study did not measure any change in confidence or proficiency; therefore, the self-report here does not necessarily reflect a relationship between AI interactions and self-perceived communicative competence as reported by Wang et al. (2024). Reuse of expressions is likewise reported, not observed. As perceptions, these accounts indicate that students believed the AI was a language to practice and then bring into the real

world of language use, as evidenced in other studies that AI-based informal learning can contribute to language learning beyond the classroom (Soyoof et al., 2023; Liu et al., 2024).

Two things in this report need to be kept separate. The first concerns 'perceived confidence'. This was a self-report, and the study did not evaluate change in confidence or proficiency, so the account is consistent with, but does not show, a link between the AI interaction and self-perceived communication competence (Wang et al., 2024). Reuse of expressions is also reported and not observed. The students' responses indicated that they saw the AI as a language rehearsal tool and an agent that they could control, which aligns with the findings of Soyoof et al. (2023) and Liu et al. (2024) that AI-assisted informal learning can extend language use beyond the classroom."

Misunderstanding and Repair Through Clarification

Participants 3, 7, and 10 reported the AI sometimes responded in ways that were unrelated to the input. They said it happened infrequently and usually was because of unclear or ambiguous input. They said most of the time it could be resolved by saying it differently.

> "Occasionally, if I ask ChatGPT a question, it will answer something that may not be what I really wanted to say; I think that's because my question was not very clear in what I meant but it doesn't happen that much, just reword it again and it can understand and respond appropriately."

The data suggest breakdowns did not end the conversation, and in some cases led to learners initiating clarification and/or a new AI response (Pica, 1994; Long, 1996). The data do not show the AI frequently errs; they come from 3 participants and describe infrequent occurrences. The findings also point to a risk. Participants located the cause of breakdowns in their own unclear language. In voice-based practice, however, a breakdown may stem from how the system recognises learner speech rather than from the learner's ambiguity, and L2 speech is known to be difficult for such systems (Moussalli & Cardoso, 2020; Napitupulu & Dalimunte, 2025). If learners attribute system failures to themselves, they may over-correct, and lose confidence in their speech. The affordance of repair thus depends on learners' ability to manage breakdowns, and hence, there is a case to be made to teach clarification strategies explicitly.

The Absence of Nonverbal Cues as Constraint and Affordance (RQ2 and RQ3)

P2, P6, and P12 described the lack of a face or visible expression in AI interaction, and they did not experience it as a deficit. They found it comfortable because they did not have to fear being laughed at or seeing negative reactions from peers when they made mistakes.

> "Sometimes it's better to communicate with the AI than with other people because I don't have to worry about being laughed at or judged because they can't see my face or react, I can speak freely and try to express my ideas without feeling embarrassed."

This is thus a constraint and an affordance. In relation to human communication, the lack in facial expression, gesture and social reaction reduces the authenticity of the interaction as communicative competence involves sociolinguistic and strategic aspects which occur in real social contexts (Canale & Swain, 1980). The same absence reduces the pressure to attempt language for those who feel it when they are in front of others. An affordance perspective is

the interdependence predicted by the value of a characteristic depending upon the relationship between the learner and environment (van Lier, 2000; Jeon, 2022).

Interdependent Affordances and Constraints, and the Role of Human Mediation

The themes collectively illustrate AImediated practice as a context where affordances a nd constraints interact. There was a lower social pressure setting in which students could speak, be asked follow-up questions, be given reformulations, be able to try again, and clarify. In the same setting, nonverbal feedback was not provided, regular breakdowns were observed, and feedback from the AI was dependent on the quality of the responses. The original draft proposed a progression from uncertainty to confidence but the data are cross-sectional accounts from different participants and do not substantiate the progression.

Three points deserve emphasis. First, human mediation shaped the experience. The lecturer's demonstration made the partner role visible (Theme 1), and students' interpretation of AI feedback and communication breakdowns required judgement that lecturers could support. Second, much of what students valued was designed in. Question-asking and feedback were specified in the prompt, so the pedagogical design of the interaction, and not only the technology, shapes the affordances. Third, nothing in the data suggests that AI is better than human interaction. Students' accounts of missing nonverbal cues and of carrying expressions over to classmates position AI as a complement.

The sessions were conducted in MSA, so the findings concern practice in the formal variety taught in programmes. Everyday situations such as ordering in a restaurant are usually conducted in spoken varieties, which limits the authenticity of the role-plays. The recent study on the research of GenAI in Arabic learning suggests that accuracy is the focus of GenAI in Arabic learning, the linguistic features that the Arabic learners need to embrace, and the requirement for human monitoring (Moustafa et al., 2026); meanwhile, the study conducted on ChatGPT in Arabic speaking shows that there are more people who want to use this technology in Arabic learning (Kholid & Majeed, 2025; Setiawati et al., 2025; Fahmiyah et al., 2026).

The findings in the present study broadly align with the previous studies on communicative language learning and AI-based speaking practice, but provide some specific implications. The researchers Hui and Yunus (2023) have shown that meaningful interaction is a crucial element in teaching speaking using Communicative Language Teaching (CLT). This outlook clearly shines through in the current results: Students appreciated ChatGPT as a tool for enriching their vocabulary or providing grammatical explanations, but also as a tool to engage in conversations and keep them going. Similarly, Huang et al. (2022) mentioned the possibility of using chatbots as interlocutors, which are able to respond in a timely manner, engage in personalized interaction, and allow for repeated language practice. Belda-Medina & Calvo-Ferrer (2022) also investigated the possibility of using conversational chatbots as language-learning companions. The current study is in line with these perspectives, as it reveals how students utilized ChatGPT to start and respond to subsequent questions and to sustain conversations in Arabic.

The results also align with Yang et al. (2022) and Liu et al. (2024) on the aspect of using AI as a conversation partner, other than the traditional class interaction. Both Yang et al. and Liu et al. proved that learners engaged in conversational interactions when using chatbots for tasks and also used generative AI as a tutor and conversation partner in informal language

learning. The present study contributes to these findings by investigating the nature of these interactions and how they are perceived by the learners in this case, the Indonesian university students studying Arabic, specifically dialogue continuity, feedback and the opportunities to express ideas. Moreover, the results corroborate those of Tai and Chen (2024), whose research indicates that using a generative AI chatbot positively affects EFL speaking performance and learners' affective responses, such as speaking confidence and decreased speaking anxiety. Wang et al. (2024) also concluded that conversational AI could help reduce speaking anxiety and increase willingness to communicate, but that improvements in affective outcomes did not necessarily lead to statistically significant improvements in speaking outcomes. This distinction is significant for the present study as it reflects students' subjective experiences of the experience, and not their objective improvements in their speaking skills.

Attruk (2025) presents a wider discussion on challenges in developing language skills in speaking in the context of Arabic language education in Indonesia, and Moustafa et al. (2026) discuss the opportunities and limitations of the generative artificial intelligence in Arabic language teaching and learning. The present results in this discussion are intended to illustrate the experience of students in using ChatGPT as a practical conversational agent in the learning process of *mahārah kalām*. Specifically, the results identify two points that warrant further investigation: that when the AI-generated answers do not convey the same meaning as intended by the student, the student will need to rephrase or explain, and that the lack of facial expression in the AI response, which can be a source of concern for some students in face-to-face communication, is seen by some students as a comfort because it reduces the need to fear negative social reactions. The experiences indicate that the potential of AI-based computer-mediated speaking practice is not solely in the ability to produce conversational replies, but also in the way learners experience and interact with the opportunities and constraints of communicative interaction.

These results present five implications for practice. It is recommended to orient students with a demonstration prior to independent AI practice and having the prompt designed intentionally and reported, as it influences the interaction. Students must also learn strategies for clarifying and be able to distinguish between recognition and comprehension problems. The use of AI feedback should be considered provisional and checked by AI practice and lecturer, and the practice should be linked to tasks with human partners.

CONCLUSION

Students started out with curiosity and doubt about if the AI would respond at all. After a demonstration by the lecturer and direct interaction, students came to recognize the AI as an interlocutor. The partner role was not shown in advance and was dependent on orientation. Students recognized and used four affordances: a less-stressed environment to try speaking, turn-taking through the AI posing questions, prompt feedback with an opportunity for retries, and (as self-reported) increased confidence and use of expressions with classmates. Because affirmation and feedback as well as posing questions were all in the prompt, these affordances represent intended interaction design. Students experienced AI misunderstandings which they repaired through clarification, and the absence of non-verbal cues (reducing the authenticity of the interaction and adding social pressure). The absence of non-verbal cues is a trade-off between a constraint and affordance.

Empirically, the study documents learners' situated experiences of voice-based AI practice in Arabic, drawing on interviews and on transcripts of the conversations themselves. Conceptually, it shows affordances and constraints in AI-mediated oral practice as interdependent: absent nonverbal cues limit authenticity and lower pressure, and the usefulness of repair depends on learners' ability to diagnose a breakdown. This study indicates that AI can be pedagogically beneficial rather than detrimental to classroom teaching and other human dialogue. Learners' experience of the practice varied depending on the way that lecturers demonstrated their own use of AI, design of prompts, ways in which learners interacted with and critically engaged with AI feedback, clarification strategies and reasoning provided, as well as opportunities to relate the AI practice to interactions with others. The results do not indicate an advantage of using AI for improving speaking skills or that AI is better than using a conversation partner.

There are several limitations to these findings. The study involved only 12 students from one school and took place over three sessions. As a result, it does not show what might happen with longer use or once students become more familiar with the tool. The findings may also not apply to other schools or groups of students. In addition, the interaction was based on one prompt and one platform, so using a different prompt or platform could lead to different results. Finally, the data came from students' own reports, which may have been influenced by what they thought the researcher wanted to hear, particularly as the researcher or interviewer was also in a teaching role.

Future studies can explore experiences over longer periods using longitudinal research methods. Research can examine different institutions and levels of proficiency. Research can study different learning platforms and different prompts and compare human partner conditions. Research can use mixed methods to study experiences and evaluate speaking performances and compare them to evaluations of AI interactions done by experts. There are Arabic-specific issues such as Arabic learner language recognition and the relationship of MSA to colloquial Arabic. Research can observe whether linguistic sequences generated with the help of AI are used in conversations with classmates and take the perspective of lecturers to study to what extent practice with AI can take the place of practice with peers in Arabic class speaking.

REFERENCE

Aldamen, H., Almashour, M., Al-Deaibes, M., & Alsharefeen, R. (2026). AI as a reflective partner in EFL speaking: A qualitative study informed by cyborg psychology. Frontiers in Psychology, 17, 1786617. https://doi.org/10.3389/fpsyg.2026.1786617

An, X., Chai, C. S., Li, Y., Zhou, Y., & Yang, B. (2023). Modeling students' perceptions of artificial intelligence assisted language learning. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2023.2246519

Attruk, M. F. H. (2025). The development of speaking skills in Arabic language learning in Indonesia: A literature review. *EDUCATUM Scientific Journal of Education*, *3*(3). https://doi.org/10.59165/educatum.v3i3.193

Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, *12*(17), Article 8427. https://doi.org/10.3390/app12178427

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

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

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*(1), 1–47. https://doi.org/10.1093/applin/1.1.1

Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing and research*. Cambridge University Press.

Fahmiyah, M., Faroji, A., & Ubaidillah. (2026). Implementation of an interactive dialogue training model based on ChatGPT to improve speaking skills. *El-Tsaqafah: Jurnal Jurusan PBA*, *25*(1), 103–114. https://doi.org/10.20414/tsaqafah.v25i1.14702

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, *38*(1), 237–257. https://doi.org/10.1111/jcal.12610

Hui, S. M., & Yunus, M. M. (2023). Revisiting communicative language teaching approach in teaching ESL speaking skills. Journal of Language Teaching and Research, 14(6), 1515–1523. https://doi.org/10.17507/jltr.1406.09

Jeon, J. (2024). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning*, *37*(1–2), 1–26. https://doi.org/10.1080/09588221.2021.2021241

Kholid, A. A., & Majeed, M. Y. (2025). Enhancing Arabic speaking skills via ChatGPT: An exploratory study at Nahdlatul Ulama Sunan Giri University. *al-Ittijah: Jurnal Keilmuan dan Kependidikan Bahasa Arab*, *17*(2), 148–162. https://doi.org/10.32678/alittijah.v17i2.12206

Lai, W., & Lee, J. S. (2024). A systematic review of conversational AI tools in ELT: Publication trends, tools, research methods, learning outcomes, and antecedents. Computers and Education: Artificial Intelligence, 7, 100291. https://doi.org/10.1016/j.caeai.2024.100291

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Liu, G. L., Darvin, R., & Ma, C. (2025). Exploring AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation of Chinese EFL learners' AI adoption

and experiences. *Computer Assisted Language Learning*, *38*(7), 1632–1660. https://doi.org/10.1080/09588221.2024.2310288

Liu, G. L., & Ma, C. (2024). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, *18*(2), 125–138. https://doi.org/10.1080/17501229.2023.2240316

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning*, *33*(8), 865–890. https://doi.org/10.1080/09588221.2019.1595664

Moustafa, A. H. A., Al-Hamad, M. F., Qureshi, M. A., & Qadir, J. (2026). Generative AI for learning and teaching Arabic: Opportunities, challenges, and ethical considerations. *IEEE Access*, *14*. https://doi.org/10.1109/ACCESS.2026.3651864

Napitupulu, M. F., & Dalimunte, A. A. (2025). A study of students' perception of Character AI in practicing English speaking fluency. *CELTIC: A Journal of Culture, English Language Teaching & Literature*, *12*(1), 384–404. https://doi.org/10.22219/celtic.v12i1.40721

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, *44*(3), 493–527. https://doi.org/10.1111/j.1467-1770.1994.tb01115.x

Savignon, S. J. (2002). *Interpreting communicative language teaching: Contexts and concerns in teacher education*. Yale University Press.

Setiawati, A. G., Hanifah, U., & Alfisyah. (2025). The use of Chat GPT platform as a learning medium to improve maharah kalam. *Jurnal Al Bayan: Jurnal Jurusan Pendidikan Bahasa Arab*, *17*(2), 544–564. https://doi.org/10.24042/jmgf6711

Soriano-Gonzalez, L., & Belda-Medina, J. (2026). Chatbots for oral proficiency in English language learning: A systematic review. Telematics and Informatics Reports, 23, 100346. https://doi.org/10.1016/j.teler.2026.100346

Soyoof, A., Reynolds, B. L., Vazquez-Calvo, B., & McLay, K. (2023). Informal digital learning of English (IDLE): A scoping review of what has been done and a look towards what is to come. *Computer Assisted Language Learning*, *36*(4), 608–640. https://doi.org/10.1080/09588221.2021.1936562

Stake, R. E. (1995). *The art of case study research*. SAGE Publications.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Lawrence Erlbaum Associates.

Tai, T.-Y., & Chen, H. H.-J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers & Education*, *220*, Article 105112. https://doi.org/10.1016/j.compedu.2024.105112

van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 245–259). Oxford University Press.

Wan, Y., & Moorhouse, B. L. (2025). Using Call Annie as a generative artificial intelligence speaking partner for language learners. *RELC Journal*, *56*(2), https://doi.org/10.1177/00336882231224813

Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: An analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System*, *127*, Article 103533. https://doi.org/10.1016/j.system.2024.103533

Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL*, *34*(3), 327–343. https://doi.org/10.1017/S0958344022000039

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.

Zadorozhnyy, A., & Lee, J. S. (2023). Informal digital learning of English and willingness to communicate in a second language: Self-efficacy beliefs as a mediator. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2023.2215279

Zou, B., Du, Y., Wang, Z., Chen, J., & Zhang, W. (2023). An Investigation Into Artificial Intelligence Speech Evaluation Programs With Automatic Feedback for Developing EFL Learners' Speaking Skills. *Sage Open*, *13*(3). https://doi.org/10.1177/21582440231193818