The Use of Chatbots and AI Assistants in Foreign Language Learning: Effectiveness and Student Experience

Ali AKÇİN
İskele Evkaf Türk Maarif High School
aliakcin82@gmail.com

Kemal KURUKAFA
İskele Evkaf Türk Maarif High School
kemal.kurukafa@hotmail.com

Adil TARHAN
Dipkarpaz Recep Tayyip Erdoğan Middle School
tarhan01adil@gmail.com

Semih DURSUN
Erenköy High School
semih.dursun04@gmail.com

ABSTRACT
This article reviews how chatbots and AI assistants are being used in foreign language learning, with a focus on both learning effectiveness and student experience. It is based on a qualitative document analysis of foundational studies, systematic reviews, conceptual publications, and recent research on conversational AI in language education. The reviewed literature includes early text-based chatbots, dialogue systems for writing and speaking support, voice-based tools, and newer generative AI systems such as ChatGPT. Overall, the evidence suggests that chatbots and AI assistants can expand opportunities for practice, provide quick responses, support learner autonomy, and make speaking or writing tasks feel less intimidating. Students often appreciate these tools because they are available at any time, respond quickly, and offer a lower-pressure space for trial and error, particularly when language anxiety limits classroom participation. At the same time, the literature also points to important weaknesses. Chatbots may give inaccurate, unnatural, or overly general responses; they may encourage superficial interaction; and they may raise concerns about privacy, dependence, and uncritical trust in AI-generated language. Student experience is therefore shaped not only by the tool itself, but also by task design, teacher guidance, learner proficiency, and the educational purpose of use. The article concludes that chatbots and AI assistants are most valuable when they are used as planned supplements to teacher-led language teaching, rather than as replacements for teachers or authentic human communication. Recommendations are offered for classroom implementation, teacher preparation, and future research.
Keywords: chatbots, AI assistants, foreign language learning, EFL, student experience, speaking practice, language education

INTRODUCTION
Conversational technologies are now a familiar part of language education. Early rule-based chat programs have gradually developed into a wider group of tools, including educational chatbots, dialogue systems, voice assistants, and large language model interfaces. For language teachers, these tools are attractive because they seem to address a persistent classroom problem: students need frequent, low-pressure opportunities to use the target language, but class time rarely allows for enough individual interaction. In foreign language contexts, where contact with fluent speakers outside school may be limited, the promise of anytime practice is especially appealing (Jia, 2009; Fryer et al., 2020).

The interest in conversational agents has grown for both pedagogical and technological reasons. Pedagogically, interaction is central to language development because learners need to negotiate meaning, produce language for real purposes, and receive feedback. Dialogue systems can offer repeated practice in speaking and writing, help learners revisit vocabulary and grammar, and provide a space where mistakes feel less public. Technologically, advances in natural language processing, speech recognition, and generative AI have made these tools more flexible and easier to access than earlier scripted chatbots (Bibauw, François, & Desmet, 2022; Xiao et al., 2023). Still, enthusiasm about chatbots needs to be balanced with caution. Language learning is not simply the exchange of messages. It also involves pragmatics, social context, cultural meaning, feedback, and sensitivity to nuance. A chatbot may imitate conversation without fully recreating the richness and unpredictability of communication with

another person. Its feedback can be immediate but shallow, and its language can sound fluent while still being inaccurate or unsuitable for the situation. For this reason, the main question is not whether chatbots are good or bad, but when and how they can make a meaningful contribution to learning.

Recent work has widened this discussion by focusing on generative AI tools such as ChatGPT. These systems can explain grammar, model conversations, rewrite texts, generate examples, and answer learner questions in ways that older chatbots could not. Studies with foreign language learners suggest that such tools may support motivation, vocabulary learning, grammar work, and independent practice (Klimova, Pikhart, & Al-Obaydi, 2024; Karataş et al., 2024). However, the same studies also warn that learners may become too dependent on AI output or accept inaccurate information too easily.

This article reviews research on chatbots and AI assistants in foreign language learning for two main purposes. First, it brings together evidence on effectiveness across speaking, writing, vocabulary, confidence, and engagement. Second, it considers student experience, including motivation, anxiety, perceptions of usefulness, and concerns about accuracy and trust. By connecting earlier chatbot research with newer studies on generative AI, the article aims to offer a balanced view of how conversational AI is influencing foreign language education.

METHODOLOGY

Research Design

This article follows a qualitative review design based on document analysis. This approach was chosen because research on the topic appears in several forms, including conceptual papers, classroom studies, systematic reviews, and handbook chapters. The review is interpretive rather than statistical. Its aim is to explain how the literature discusses effectiveness and student experience when chatbots and AI assistants are used in foreign language learning.

Source Selection

Sources were selected because they addressed at least one of four areas: conceptual discussion of conversational agents in education, empirical evidence on language-learning outcomes, review studies of chatbot interventions, or learner-focused research on usefulness, motivation, affect, and classroom experience. Priority was given to peer-reviewed journal articles, reputable book chapters, and well-cited reviews published between 2009 and 2024. Earlier chatbot studies and recent generative AI publications were included together so that the development of the field could be followed more clearly.

Analytical Strategy

The literature was analyzed thematically under five categories: pedagogical affordances, reported learning outcomes, student perceptions and affect, design limitations, and implementation issues. Studies were also compared by target skill, type of conversational system, and the amount of teacher involvement. This made it possible to separate broad claims about AI from more specific claims about how conversational tools are actually used in language-learning practice.

Trustworthiness

The trustworthiness of the review depends on the quality and range of the sources, as well as on transparent interpretation. To avoid presenting only one side of the discussion, the review includes both supportive and critical studies. Greater weight was given to findings that appeared repeatedly across the literature rather than to isolated claims from a single intervention.

FINDINGS

Table 1: Main types of conversational tools used in foreign language learning

Tool type	Typical use in language learning	Representative sources
Rule-based text chatbot	Scripted text interaction for vocabulary, grammar, or simple dialogue.	Jia (2009)
Educational chatbot	Task-oriented support for classroom writing, revision, or guided practice.	Lin & Chang (2020); Okonkwo & Ade-Ibijola (2021)
Dialogue system / conversational agent	More flexible text or speech interaction for L2 practice.	Bibauw et al. (2022); Xiao et al. (2023)

Voice assistant	Speech-based practice for pronunciation, fluency, and oral interaction.	Jeon (2024); Du & Daniel (2024)
Generative AI assistant	Open-ended prompting, explanation, modeling, and revision support.	Klimova et al. (2024); Karataş et al. (2024)

The reviewed literature shows that chatbots and AI assistants should be treated as pedagogical tools with particular strengths and limits, not as universal solutions. Their usefulness depends on the type of interaction they support, the language skill being practiced, the quality of feedback, and the way the tool is built into a learning design. Table 1 summarizes the main types of conversational tools discussed in the literature and the functions most often linked to them.

One of the clearest benefits across the literature is the increase in practice opportunities. Jia's (2009) CSIEC project was an early attempt to design a chatbot that could support English learning beyond the limits of teacher availability. Later work by Fryer et al. (2020) and Bibauw et al. (2022) shows that this remains one of the strongest arguments for using chatbots: they give learners more chances to use the target language when human partners are not always available. In foreign language settings, that added practice time can be especially important.

Table 2: Reported learning benefits in the reviewed literature

Outcome area	Observed benefit	Supporting sources
Speaking practice	More opportunities to rehearse oral production and increase willingness to communicate.	Bibauw et al. (2022); Du & Daniel (2024)
Writing development	Support for drafting, idea generation, organization, and revision.	Lin & Chang (2020); Karataş et al. (2024)
Vocabulary and grammar	On-demand examples, reformulations, and repeated practice.	Klimova et al. (2024); Karataş et al. (2024)
Learner autonomy	Self-paced practice and immediate responses outside class time.	Fryer et al. (2020); Klimova et al. (2024)
Affective support	Lower anxiety and greater comfort in low-stakes interaction.	Jeon (2024); Cislowska & Peña-Acuña (2024)

The strongest evidence for learning benefits appears in areas where repetition, guided production, and quick responses are useful. Table 2 summarizes the main benefits reported in the reviewed studies. Dialogue-based systems seem particularly helpful for speaking practice, confidence building, vocabulary recycling, and writing support. Bibauw, Van den Noortgate, François, and Desmet (2022) found an overall positive effect for dialogue systems in language learning, especially in vocabulary and oral production tasks. Du and Daniel's (2024) review also suggests that AI-powered chatbots can support English speaking practice by improving confidence, engagement, and willingness to communicate.

Student experience is another major theme. Learners often describe chatbots as convenient, always available, and less intimidating than human conversation partners. This matters because anxiety can be a serious barrier in foreign language speaking and writing. When students hesitate to speak because they fear embarrassment, a non-judgmental chatbot can offer a useful rehearsal space. Jeon (2024) found that young EFL learners saw chatbots as helpful conversation starters that supported motivation and participation. Cislowska and Peña-Acuña's (2024) review also reports that students often value chatbots because they reduce some of the social pressure connected with classroom performance.

The literature also makes an important distinction between task-oriented chatbots and generative AI assistants. Traditional chatbots usually work within narrower limits, which can make them easier to connect to a specific classroom task such as role-play, vocabulary review, or structured question answering. Generative systems are more flexible and can respond to a wider range of prompts, but their responses are less predictable. For teachers, this distinction matters. A limited chatbot may be less impressive, but more reliable for a clearly defined activity. A generative tool may be more versatile, but it needs stronger guidance and critical supervision.

Effectiveness also varies by target skill. Speaking practice benefits from systems that support turn-taking, repeated production, and a lower fear of mistakes. Writing support benefits from tools that can model structures, suggest

wording, and respond to questions about organization. Vocabulary learning is strongest when chatbots recycle target words in meaningful contexts rather than simply giving definitions. Reading and listening appear to benefit less directly unless chatbot use is connected to wider tasks such as comprehension checks, summaries, or discussion. This helps explain why findings differ across studies: the value of a conversational agent depends on the pedagogical job it is asked to do.

Table 3: Frequently discussed limitations and risks

Issue	Why it matters pedagogically	Representative discussion
Inaccurate output	Learners may internalize incorrect language or misleading explanations.	Klimova et al. (2024); Karataş et al. (2024)
Shallow dialogue	Conversation may remain formulaic and provide limited communicative depth.	Fryer et al. (2020); Bibauw et al. (2022)
Generic feedback	Immediate responses do not always equal pedagogically useful correction.	Okonkwo & Ade-Ibijola (2021); Xiao et al. (2023)
Overreliance	Students may delegate too much drafting or problem solving to AI.	Karataş et al. (2024); Klimova et al. (2024)
Ethical and privacy concerns	Schools need clear rules for data protection and responsible use.	Winkler & Söllner (2018); Okonkwo & Ade-Ibijola (2021)

Student experience also differs by proficiency level. Beginners may appreciate simplified interaction, controlled tasks, and the chance to repeat an activity several times. At the same time, they may be less able to recognize incorrect or overly complex AI output. More advanced learners can use AI to explore discourse, compare styles, and test pragmatic choices, but they may also find simple or repetitive responses frustrating. The literature therefore supports differentiated use: lower-level learners need more scaffolding and narrower tasks, while intermediate and advanced learners can benefit from more open-ended activities when guidance is available.

Another recurring issue is the difference between authentic communication and simulation. Chatbots can help learners rehearse language, but rehearsal is not the same as real communication. Human interaction includes spontaneity, ambiguity, repair, nonverbal cues, and social consequences that chatbots only partly reproduce. Stronger implementations therefore do not try to replace authentic communication. Instead, they use AI-supported practice as preparation for peer discussion, classroom speaking, presentations, or independent writing.

The literature also raises a wider classroom question: what kind of language-learning culture is created when students interact regularly with AI? In positive cases, these tools can normalize experimentation and make practice more frequent. However, there is also a risk that learners begin to see language mainly as a set of prompts and outputs rather than as social communication shaped by audience, context, and meaning. For this reason, chatbot use should be accompanied by reflection. Students need chances to discuss when AI helped, when it misled them, and how machine-generated interaction differs from human communication.

Learner autonomy is another commonly reported advantage. Chatbots and AI assistants allow students to practice at their own pace, request examples when needed, and return to explanations more than once. In higher education, this flexibility is often described as one of the most useful features of generative AI. Klimova et al. (2024) found that university students viewed ChatGPT as useful for practicing foreign language skills and receiving individualized help. Karataş et al. (2024) similarly reported that learners used ChatGPT for writing, grammar, and vocabulary support, and that these experiences often increased motivation.

Table 4: Implementation principles for effective classroom use

Principle	Classroom implication	Expected benefit
Task alignment	Use AI for explicit linguistic goals rather than unrestricted chatting.	Improves relevance and learning focus.
Teacher mediation	Discuss prompts, outputs, and errors with learners.	Promotes reflection and prevents blind trust.
Human-AI balance	Combine AI rehearsal with peer and teacher interaction.	Supports transfer to authentic communication.
Critical AI literacy	Teach learners to verify and challenge AI responses.	Builds judgment and reduces misinformation.

Assessment redesign	Value process, reflection, and oral performance, not only final text.	Discourages uncritical dependence on AI-generated output.

Writing support is one area where the evidence is especially practical. Lin and Chang (2020) showed that a chatbot-supported writing environment helped post-secondary writers develop thesis statements and improve peer feedback. The important point is not simply that a chatbot was present, but that the tool guided students through specific stages of the writing process. This fits a broader pattern in the literature: conversational tools are most useful when they are attached to clear, bounded tasks rather than left as open-ended add-ons.

Despite these benefits, the literature identifies several important limitations. Table 3 summarizes the concerns that appear most often. One major issue is linguistic reliability. Chatbots may produce responses that sound correct but are pragmatically odd, contextually unsuitable, or pedagogically weak. This is especially important with generative AI, because confident language can make inaccurate output seem trustworthy. In foreign language learning, students may not yet have the ability to notice subtle errors, so unsupported use can create false confidence or reinforce mistakes.

A second limitation is interaction quality. Chatbots can simulate conversation, but not every chatbot conversation supports learning in the same way. Fryer et al. (2020) argue that many language-learning bots remain limited because they struggle with unpredictability, context, and nuanced feedback. Bibauw et al. (2022) also caution against treating all dialogue systems as educationally equal. A bot that simply gives short scripted replies may have little value beyond novelty, while a task-based system can support more meaningful practice.

A third concern is feedback. Teachers can diagnose errors, decide when correction is helpful, and adapt their response to the learner's level and emotional state. Chatbots may respond quickly, but their feedback can be generic or incomplete. Some students value the speed of these responses, even when they are only partial. In other cases, the lack of detailed correction limits learning. This is especially clear in speaking practice, where learners need support not only with grammar and vocabulary, but also with pronunciation, pragmatics, discourse management, and confidence.

Ethical and educational concerns have become more visible with generative AI. Students may rely too heavily on AI for drafting, rewriting, or answer generation, which can reduce productive struggle. Privacy, data protection, and the opacity of AI-generated explanations are also important concerns. The literature does not suggest that these risks require rejecting conversational AI altogether. Instead, schools need clear rules, critical AI literacy, and assessments that value reflection, revision, and authentic communication rather than polished AI-generated products alone.

Teacher mediation appears to be one of the most important conditions for successful use. Positive outcomes are more likely when teachers explain the purpose of the tool, provide prompts or goals, and connect chatbot interaction to larger classroom tasks. Table 4 summarizes implementation principles drawn from the literature. These principles suggest that conversational AI works best as a structured supplement: useful for rehearsal, exploration, drafting, and preparation, but not suitable as an independent judge of language competence.

Overall, the literature supports a careful but positive conclusion. Chatbots and AI assistants can increase practice opportunities and improve student experience, especially by reducing anxiety and giving learners more interaction time. However, their value depends on design quality, feedback reliability, learner training, and teacher oversight. Without pedagogical framing, they may become attractive but shallow tools. With thoughtful integration, they can become useful partners within a broader language-learning environment.

DISCUSSION

The findings suggest that the value of chatbots and AI assistants lies less in their novelty than in the way they create more opportunities for language use. In many foreign language classrooms, speaking and writing practice is limited by time, class size, and learner anxiety. Conversational AI can reduce some of these barriers by giving students additional spaces to rehearse. This helps explain why learners often report positive experiences even when measured learning gains are moderate. Availability, privacy, and responsiveness make practice easier to begin and easier to continue.

Accessibility is another issue that deserves attention. Conversational tools may help students who need repeated exposure, self-paced rehearsal, or written alternatives to oral performance. However, the same tools may also create barriers if interfaces are cluttered, speech recognition struggles with accented pronunciation, or prompts

assume cultural knowledge that some learners do not have. Inclusive implementation therefore requires flexible options, including text and speech modes, accessible interfaces, and tasks that can be adjusted to learner needs.

Classroom context also shapes outcomes. In small classes with strong teacher feedback, students may use chatbots to extend and refine what they have already practiced with people. In larger or exam-oriented settings, the same tools may be used mainly for quick correction, translation, or answer checking. Both uses can have a place, but they lead to different learning habits. The literature suggests that chatbot use is more likely to support durable language development when it is connected to reflection, dialogue, and communicative goals.

For this reason, implementation should be judged not only by whether students enjoy a tool, but also by the kind of language behavior it encourages. A useful tool should invite students to try new expressions, ask follow-up questions, revise meaning, and think about errors. A weaker use of the same tool may simply help students finish tasks quickly with little learning. This distinction should guide classroom decisions and future research.

Student experience should also not be confused with educational effectiveness. Enjoyment, convenience, and curiosity can support motivation, but they do not automatically lead to lasting learning. The strongest evidence comes from studies in which chatbot interaction is tied to explicit goals, such as speaking practice, writing development, or structured dialogue tasks. This is consistent with broader educational technology research: learning gains usually come from instructional design, not from the mere presence of a new tool.

The field itself is also changing. Earlier chatbot studies asked whether a conversational system could support language interaction at all. Current research increasingly asks which kinds of conversational AI are useful for which learners and tasks. Generative AI expands what learners can do, but it also increases concerns about accuracy, dependence, and critical evaluation. This makes the teacher's role more important rather than less important. Teachers help learners compare AI responses, notice limitations, and connect digital practice to human communication.

Affective factors are especially important. Many learners perceive chatbots as non-judgmental partners, and this can matter a great deal in foreign language learning. Anxiety, fear of negative evaluation, and reluctance to speak can reduce participation even when students know the required forms. If AI tools lower the emotional threshold for practice, they may support learning indirectly by increasing exposure and willingness to communicate. At the same time, students should not become comfortable only with AI-mediated interaction. They also need guided movement toward communication with peers and teachers.

Taken together, the literature supports a balanced position. Chatbots and AI assistants are not minor gimmicks, but they are also not replacements for human language teaching. Their strongest contribution is supplementary: they can expand practice, personalize support, and reduce anxiety. Their main risks appear when they are treated as authoritative, autonomous, or pedagogically sufficient on their own.

RESULTS AND CONCLUSIONS

This review examined the use of chatbots and AI assistants in foreign language learning and found broad evidence that conversational tools can support language development and improve student experience when the conditions are appropriate. The most consistent benefits are increased access to practice, reduced speaking anxiety, greater learner autonomy, stronger engagement, and support for specific areas such as speaking, vocabulary, and writing. These benefits are most visible when tasks are structured and when AI use is part of a wider teaching strategy.

The review also shows that student experience is central to understanding effectiveness. Learners generally respond positively to tools that are available, responsive, and private. These features can make language practice feel safer and more manageable. However, positive experience is not enough on its own. Problems related to accuracy, shallow interaction, generic feedback, privacy, and overreliance remain significant. For this reason, the literature does not support replacing teachers with chatbots or AI assistants.

The most defensible conclusion is that conversational AI is educationally valuable when it serves as a guided partner for rehearsal, drafting, feedback exploration, and additional target-language interaction. Its role should be supplementary, transparent, and pedagogically bounded. Used in this way, chatbots and AI assistants can enrich foreign language learning while preserving the central role of teachers, human interaction, and critical reflection.

RECOMMENDATIONS

Practical Recommendations

1. Teachers should use chatbots and AI assistants for clearly defined language tasks, such as dialogue rehearsal, vocabulary recycling, idea generation, role-play preparation, or writing revision, rather than leaving use completely open-ended.
2. Schools should include AI literacy in language instruction so that students learn to question AI output, verify examples, and recognize inaccurate or inappropriate language.
3. Conversational AI should be combined with human interaction. Students should move between AI-supported rehearsal and peer or teacher-guided communication so that practice does not remain artificial.
4. Assessment should value process, reflection, oral interaction, and revision history. This can reduce the risk that students rely on AI-generated text without learning from it.
5. Institutions should provide professional development for language teachers on prompt design, error analysis, feedback evaluation, and ethical classroom use of generative AI.

Recommendations for Future Research

1. More longitudinal studies are needed to determine whether repeated chatbot use produces lasting gains in speaking, writing, and pragmatic competence.
2. Future research should compare scripted bots, voice assistants, and large language model interfaces across proficiency levels and age groups.
3. More evidence is needed on how learners judge AI feedback and how teacher mediation changes the quality of learning outcomes.
4. Researchers should examine the social and emotional dimensions of student experience more closely, including anxiety reduction, confidence, willingness to communicate, and trust.
5. Further research is needed on ethical and assessment issues, especially how conversational AI affects originality, academic integrity, and independent language proficiency.

REFERENCES

Jia, J. (2009). CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning. Knowledge-Based Systems, 22(4), 249-255. https://doi.org/10.1016/j.knosys.2008.09.001

Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. Academy of Management Proceedings, 2018(1), 15903. https://doi.org/10.5465/AMBPP.2018.15903abstract

Lin, M. P.-C., & Chang, D. (2020). Enhancing post-secondary writers' writing skills with a chatbot: A mixed-method classroom study. Educational Technology & Society, 23(1), 78-92.

Fryer, L. K., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). Bots for language learning now: Current and future directions. Language Learning & Technology, 24(2), 8-22. http://hdl.handle.net/10125/44719

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. Computers and Education: Artificial Intelligence, 2, 100033. https://doi.org/10.1016/j.caeai.2021.100033

Bibauw, S., François, T., & Desmet, P. (2022). Dialogue systems for language learning: Chatbots and beyond. In N. Ziegler & M. González-Lloret (Eds.), The Routledge handbook of second language acquisition and technology (pp. 121-134). Routledge. https://doi.org/10.4324/9781351117586-12

Bibauw, S., Van den Noortgate, W., François, T., & Desmet, P. (2022). Dialogue systems for language learning: A meta-analysis. Language Learning & Technology, 26(1), 1-24.

Xiao, F., Zhao, P., Sha, H., Yang, D., & Warschauer, M. (2023). Conversational agents in language learning. Journal of China Computer-Assisted Language Learning, 4(2), 300-325. https://doi.org/10.1515/jccall-2022-0032

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

Klimova, B., Pikhart, M., & Al-Obaydi, L. H. (2024). Exploring the potential of ChatGPT for foreign language education at the university level. Frontiers in Psychology, 15, 1269319. https://doi.org/10.3389/fpsyg.2024.1269319

Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. Education and Information Technologies. https://doi.org/10.1007/s10639-024-12574-6

Cisłowska, A. I., & Peña-Acuña, B. (2024). Integration of chatbots in additional language education: A systematic review. European Journal of Educational Research, 13(4), 1607-1625. https://doi.org/10.12973/eu-jer.13.4.1607

Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. Computers and Education: Artificial Intelligence, 6, 100230. https://doi.org/10.1016/j.caeai.2024.100230