Artificial Intelligence in Language Learning and Teaching: A Review of Applications, Benefits, and Challenges[1]

Süleyman BAŞARAN[2]

Abstract

Artificial intelligence (AI) is transforming the field of language learning and teaching. Building on decades of Computer-Assisted Language Learning (CALL), contemporary AI technologies such as natural language processing, machine learning, automated speech recognition and large language models are enabling more adaptive, interactive and data-rich learning environments than were previously possible. Drawing on recent systematic reviews, bibliometric analyses and empirical studies, this article synthesizes evidence on how AI supports language learners through intelligent tutoring systems, conversational agents, automated feedback on writing and speech, adaptive learning platforms and immersive AR/VR environments. It also examines how AI assists teachers with assessment, content creation and learning analytics. Reported benefits include greater personalization, increased opportunities for low-anxiety practice, more frequent and fine-grained feedback, and gains in motivation and engagement. At the same time, the review highlights major challenges and ethical concerns around algorithmic bias, data privacy, digital inequality, over-reliance on automation, teacher readiness and the limits of AI's contextual understanding. The article argues that AI is most productive when framed as an augmenting partner rather than a replacement for human educators and concludes with recommendations for responsible integration and future research.

Keywords: Artificial intelligence in language education, Computer-Assisted Language Learning (CALL), Intelligent tutoring and adaptive practice, Conversational agents and chatbots, Automated feedback and learning analytics.

[1]Presented orally at ICONTE Conference 2025.

[2]Prof. Dr., Fen Edebiyat Fakültesi - Batı Dilleri ve Edebiyatı Bölümü Batman Üniversitesi, https://orcid.org/0000-0002-3262-098X , suleyman.basaran@batman.edu.tr

Introduction

The rapid development of artificial intelligence is reshaping many dimensions of education, and language learning and teaching are no exception. AI now underpins adaptive language-learning apps, chatbots that simulate conversation partners, systems that automatically evaluate and comment on student writing, and immersive virtual environments where learners can interact in the target language. These developments do not emerge in a vacuum: they build on decades of experimentation with CALL, mobile learning and data-driven instruction. What is new is the scale, sophistication and accessibility of AI-driven tools, which are increasingly integrated into everyday devices and platforms.

Large-scale reviews and bibliometric studies point to a sharp increase in research on AI and language education over the last decade, with a particularly strong surge since around 2018 (Huang et al., 2023; Ivanova et al., 2024; Liang et al., 2023; Zhu & Wang, 2025). This growth is visible not only in the number of publications, but also in the diversity of topics, ranging from intelligent tutoring and adaptive vocabulary learning to generative AI for writing, chatbots for conversation practice, speech analytics for pronunciation, and VR-based environments for immersive learning. The field is becoming simultaneously broader and more specialized, with dedicated issues of journals and new conferences focusing explicitly on AI in education.

Several trends emerge from this growing body of work. First, AI is increasingly integrated into learner-facing tools that provide adaptive practice, conversational interaction and automated feedback. Learners encounter AI not as a separate module but as an embedded layer in apps and platforms they already use. Second, AI is being incorporated into teacher-facing tools that support assessment, content generation and the analysis of learner data. Teachers are asked, sometimes implicitly, to make sense of dashboards, automated scores and machine-generated suggestions. Third, there is a widening conversation about the ethical, pedagogical and policy implications of AI in education, including issues of bias, privacy, equity and the future role of human teachers (UNESCO, 2021; Hockly, 2023). Questions about what AI can do are increasingly joined by questions about what it should do, and under what conditions.

Rather than treating AI as a monolithic entity, it is helpful to consider specific families of applications and their pedagogical benefits. A grammar-checking tool, an automated writing evaluator, a VR environment and a generative chatbot are all examples of AI in action, but they differ significantly in what they offer, what data they collect and how they reshape classroom practice. This article therefore focuses on five overlapping areas: intelligent tutoring and adaptive practice; conversational AI agents and chatbots; automated feedback on writing and speaking; AI-enabled immersive environments; and AI tools that support teachers. For each area, it summarizes relevant findings, discusses benefits and limitations, and situates them within broader debates on ethics and equity.

The discussion is selective rather than exhaustive, but it draws on recent systematic reviews, empirical studies and bibliometric analyses designed to map the field and identify research gaps (Huang et al., 2022; Huang et al., 2023; Liang et al., 2023; Ivanova et al., 2024; Zhu & Wang, 2025). The goal is to offer a balanced overview that can help learners, teachers and curriculum designers engage with AI critically yet constructively, treating it as a powerful but imperfect partner in language education.

AI and the Language Learner

Intelligent Tutoring Systems and Adaptive Practice

One of the longest-standing promises of AI in education is the idea of the intelligent tutoring system which is software that approximates aspects of one-to-one tutoring by modelling domain knowledge, tracking a learner's progress and adapting instructional decisions accordingly. In language education, this vision has evolved into a variety of adaptive learning systems that adjust task difficulty, sequence and feedback based on learner performance (Huang et al., 2023; Liang et al., 2023). The underlying idea is that learners rarely benefit from a "one-size-fits-all" curriculum; instead, they profit

when the system can diagnose what they know, what they struggle with and how quickly they are progressing, then feed this information back into the design of subsequent activities.

A recent systematic review of adaptive learning technologies in education more broadly found that such systems can improve learning outcomes and efficiency when they are grounded in sound pedagogical principles and when learners remain actively engaged rather than passively following algorithmic suggestions (Martin et al., 2020). Martin et al. (2020) emphasize that adaptivity is not automatically beneficial. Therefore, systems must be calibrated so that tasks remain meaningful and challenging, and learners must understand why particular activities appear and what they are expected to do with the feedback. When learners perceive the system as opaque or arbitrary, the motivational benefits of personalization can quickly erode.

In language learning, adaptive platforms embedded in commercial apps or institutional systems use statistical models and machine learning to predict which vocabulary items are at risk of being forgotten, which grammatical structures cause persistent difficulty, and which tasks are likely to be "just right" in terms of challenge (Liang et al., 2023; Zhu & Wang, 2025). A learner who repeatedly struggles with a particular tense, for instance, might see that structure recycled across different topics and modalities, while items that have been mastered receive less attention. Some systems also adjust the modality of presentation, shifting between reading, listening or cloze activities depending on performance.

The benefits of such systems, as reported in the literature, include more personalized practice trajectories, a better alignment between task difficulty and learner readiness, and increased engagement when adaptive features are made visible and understandable to learners (Martin et al., 2020; Huang et al., 2023). Learners often report that they appreciate seeing their "progress bars" move, unlocking new levels or receiving targeted challenges. At the same time, there are important concerns. If the repertoire of tasks is narrow or decontextualized, adaptivity alone cannot guarantee meaningful language use. There is a risk that learners may become highly efficient at completing micro-tasks (gap-fills, short translation items) without sufficient opportunities for extended discourse, creativity or real communication (Hockly, 2023). Moreover, if adaptivity is opaque, learners may feel that the machine "knows best" and surrender control over their own learning goals.

Effective use of intelligent systems therefore requires thoughtful integration with activities that foster interpretive, productive and interactional skills beyond what the algorithm optimizes. Teachers can, for example, treat adaptive tasks as preparation for richer communicative work in class, or invite learners to reflect on the patterns that the system identifies and to set personal goals accordingly. In this way, adaptivity becomes a resource for metacognition rather than a hidden mechanism.

Conversational AI Agents and Chatbots

Another prominent strand of AI in language education involves conversational agents; that is chatbots that interact with learners via text and/or speech. These systems range from simple rule-based bots that follow pre-scripted flows to sophisticated neural models capable of generating contextually varied responses. A systematic review of chatbot-supported language learning found that such systems have been deployed for vocabulary practice, grammar reinforcement, written conversation and, less frequently, oral interaction (Huang et al., 2022). Across studies, learners often report that chatbots provide a low-anxiety environment, are available at any time and offer more opportunities to use the target language than they might otherwise have.

For shy or anxious learners, the fact that a chatbot does not "judge" them in the same way as a human interlocutor can be particularly valuable. They may feel freer to experiment with new structures, make mistakes and ask questions that they might not pose in front of peers. Some chatbots are designed around specific characters or scenarios (for example, a travel assistant, a restaurant server or a friendly tutor), providing a degree of role-play that can enrich the learning experience (Huang et al., 2022). Others are embedded within learning management systems and are tasked with answering questions about assignments or explaining grammar points.

However, empirical findings are mixed. Some studies report improvement in vocabulary, grammar accuracy and writing quality, especially when chatbots are tightly integrated with course content and tasks (Huang et al., 2022; Huang et al., 2023). For instance, learners may write short messages to a chatbot and receive immediate corrections or suggestions, leading to more cycles of drafting and revising than in traditional homework. Other studies point to shortcomings when chatbots misunderstand meaning, provide canned responses or fail to sustain coherent multi-turn conversation. Early rule-based systems were particularly limited in this respect, as they could only "understand" inputs that matched predefined patterns; yet even modern neural chatbots sometimes produce replies that are off-topic or shallow.

Recent work on generative AI, and ChatGPT in particular, illustrates both the potential and the pitfalls of such tools in language classrooms. Case studies and conceptual articles suggest that ChatGPT can generate rich input, provide explanations at different levels of complexity, act as a conversation partner and support creative writing tasks, sometimes leading to increased engagement and perceived learning (Ajmal et al., 2025; Hockly, 2023; Kohnke et al., 2023a). Ajmal et al. (2025), for example, describe how university students used ChatGPT to brainstorm ideas, receive suggestions on phrasing and obtain immediate feedback on drafts, and they report positive changes in students' self-confidence as writers.

At the same time, instructors report concerns about hallucinated facts, superficial or formulaic output and academic integrity when students submit AI-generated text as their own (Hockly, 2023; Kohnke et al., 2023b). Kohnke et al. (2023b) note that many instructors feel unprepared to detect AI-generated work or to design tasks that discourage mindless copying. This has prompted calls for explicit AI literacy; teaching learners how to question, verify and appropriately integrate AI-generated language rather than treating it as authoritative (Hockly, 2023; Huang et al., 2023). In practice, this might mean asking students to annotate AI-generated suggestions, explain why they accept or reject them, or use AI outputs as starting points for discussion rather than as final products.

Automated Feedback on Writing

Automated writing evaluation (AWE) systems represent one of the most mature applications of AI in language assessment. They use natural language processing and machine learning to score written texts and to highlight errors or features related to grammar, vocabulary, mechanics and aspects of organization. Over time, these systems have evolved from relatively simple grammar checkers to complex platforms that assign holistic or analytic scores and provide comments that resemble those of human teachers.

A recent systematic review of AWE use in second language writing concludes that these systems can provide timely, consistent feedback and encourage more frequent revision, particularly when learners are trained in how to interpret and respond to the feedback (Sari & Han, 2024). In many contexts, teachers struggle to provide detailed comments on multiple drafts due to time constraints. AWE systems can bridge part of this gap by offering immediate responses that students can act on before submitting work to the teacher.

Experimental and quasi-experimental studies support the claim that AWE can contribute to improvements in local accuracy (e.g., grammar and spelling) and, in some cases, to global writing quality over time (Wei et al., 2023; Yıldız & Gonen, 2024). For example, Yıldız and Gonen (2024) report that students using an AWE system in an online writing course not only reduced mechanical errors but also engaged more actively in revision, submitting more drafts and spending more time reflecting on feedback. Wei et al. (2023) found that integrating AWE into a writing course led to measurable gains compared to a control group receiving only teacher feedback, suggesting that the combination of human and machine commentary can be synergistic.

Nevertheless, researchers consistently emphasize that AWE is more reliable for surface-level issues than for higher-order concerns such as argumentation, coherence, voice or audience awareness (Wei et al., 2023; Sari & Han, 2024). Automated scores can be misleading when used for high-stakes evaluation, and some learners either ignore the feedback or over-rely on it without reflecting on

underlying rules. There is also the danger that learners may treat AI feedback as infallible, even when it misinterprets an unconventional but acceptable structure.

For the reasons above, AWE seems to be most beneficial when used as a formative tool within a broader pedagogy that includes teacher feedback, peer review and explicit discussion of writing choices. Teachers can ask students to compare AWE comments with their own judgments, explain cases of disagreement and reflect on which types of feedback they find most useful. In this way, AWE becomes a resource for metalinguistic awareness rather than an authority that dictates "correct" writing.

Automated Speech Recognition and Pronunciation Support

Progress in automated speech recognition (ASR) has made it possible for AI systems to analyze learners' oral language and provide feedback on aspects such as segmental accuracy and fluency. Modern ASR systems can transcribe speech with high accuracy under certain conditions, detect deviations from target pronunciations and compute measures such as speech rate and/or pause frequency.

Classroom-based studies have suggested that ASR-enhanced pronunciation practice can result in improvements in comprehensibility and in the accuracy of targeted sounds, especially when practice is repeated and feedback is explicit (Jiang et al., 2023). In Jiang et al.'s (2023) study, learners used an ASR-based tool that compared their production to a model and highlighted discrepancies. Over time, learners showed measurable gains in pronunciation scores, and their self-reported awareness of segmental and supra-segmental features increased. The tool also enabled learners to practice outside class hours, increasing the amount of focused speaking they engaged in.

Such tools can support autonomous, individualized practice that would be difficult for a single teacher to provide to a large class. A teacher may, for example, introduce a set of target sounds or intonation patterns and then assign ASR-based exercises for home practice, reviewing overall class patterns later and addressing common issues. Learners gain quantity of practice and immediate feedback; teachers gain insight into patterns that might otherwise remain invisible.

However, ASR systems still perform unevenly across accents, proficiency levels and speaking conditions. Misrecognition of certain L1-influenced pronunciations can lead to incorrect feedback, which may discourage learners or reinforce inaccurate beliefs about their performance (Huang et al., 2023; Jiang et al., 2023). In addition, ASR tends to work best for read-aloud tasks or controlled speech; spontaneous, highly interactive conversation remains more challenging to analyse. ASR is therefore best embedded in a pedagogical context where learners are encouraged to treat automated feedback as indicative rather than definitive, and where teachers help them interpret and, when necessary, question the system's judgments.

Immersive and Multimodal AI-Enhanced Environments

Immersive technologies; especially virtual reality (VR) and augmented reality (AR), increasingly intersect with AI to create dynamic environments where learners can interact in the target language with virtual characters, adaptive scenarios and multimodal cues. These environments can simulate real-life contexts such as a market, a train station or a workplace, allowing learners to practice language in situations that might be difficult to recreate in a classroom.

Research on VR for L2 vocabulary learning has demonstrated that such environments can support better retention than traditional, non-immersive methods, particularly when learners can manipulate objects and encounter words in rich contexts (Legault et al., 2019). In Legault et al.'s (2019) study, learners who engaged in a VR-based vocabulary task outperformed a control group on several measures of learning and reported higher enjoyment and presence. They were able to tie words to specific locations, actions and objects, making the vocabulary more memorable.

AR and VR environments designed with motivational principles in mind can also improve engagement and performance. Hung et al. (2023) applied the ARCS model (attention, relevance, confidence and satisfaction) in designing AR/VR-based English activities for elementary school

learners and found significant improvements in both motivation and learning outcomes compared to traditional instruction. The ARCS framework helped designers ensure that the tasks were visually engaging, clearly related to learners' lives, structured to support growing confidence and rewarding in terms of feedback and progression.

A Turkish-language literature review by Alyaz and Demiryay (2023) similarly concludes that VR applications hold promise for foreign language education, while also noting practical constraints such as cost, technical complexity and teacher training needs. VR headsets and powerful computers are still relatively expensive, and not all classrooms have the space or infrastructure to support their use. Teachers may also feel uncertain about how to integrate VR sessions into their curricula, how to manage students in virtual environments and how to align VR tasks with assessment.

AI plays several roles in such environments: controlling non-player characters that respond to learner input, adapting difficulty in real time, and capturing data on learner performance for later analysis (Huang et al., 2023; Liang et al., 2023). For example, AI-driven characters can adjust their speech rate or vocabulary based on learner responses, or the system can track which objects and prompts learners interact with and adjust subsequent tasks accordingly. Immersive AI-enhanced systems are still relatively resource-intensive, but they illustrate what "thick" digital environments for language use might look like in the near future, especially as hardware becomes more affordable.

AI as a Partner for Language Teachers

Although many AI applications target learners directly, teachers are increasingly using AI as a tool to support their own work. This includes assistance with assessment and feedback, content creation, differentiation and data-informed decision-making. The teacher's role is shifting not only in what they do, but also in how they make sense of the tools that learners bring into the classroom.

One of the most attractive uses of AI is related to assessment and feedback. AWE systems can help teachers manage large volumes of writing by flagging local errors, suggesting corrections and providing preliminary analytic scores (Sari & Han, 2024; Yıldız & Gonen, 2024). When teachers retain final responsibility for grading and use automated feedback to supplement their own, studies suggest that overall feedback can become more timely and that teacher time can be redirected toward more complex aspects of writing such as content and organization (Wei et al., 2023). Rather than correcting every article or tense error, teachers can focus on how students structure arguments, integrate sources or develop ideas.

Similarly, ASR-based analytics can provide information on learner fluency and segmental accuracy that teachers can combine with their holistic assessment of communicative effectiveness (Jiang et al., 2023). A teacher might, for instance, use ASR summaries to see which sounds or stress patterns are most problematic across a class, then design targeted pronunciation activities. Used in this way, AI extends rather than replaces teachers' diagnostic powers.

Teachers are also beginning to use AI-powered teaching assistants and chatbots to handle routine queries, model example dialogues or provide extra practice materials. Reviews and case reports indicate that chatbots integrated into language courses can answer common questions about assignments, explain grammar points and offer low-stakes conversation opportunities, thus reducing repetitive workload for instructors (Huang et al., 2022; Huang et al., 2023). Instructors who teach large classes or online courses may particularly appreciate a chatbot that can respond to frequently asked questions outside office hours.

Generative AI tools such as ChatGPT are being explored as content-creation partners. Instructors report using them to generate reading passages at different proficiency levels, create vocabulary exercises, propose role-play scenarios or suggest alternative phrasings and explanations (Hockly, 2023; Kohnke et al., 2023a; Ajmal et al., 2025). A teacher might, for example, ask ChatGPT to write three versions of a story; for example, at A2, B1 and B2 level on the same topic, then edit these drafts to match curricular goals. Others may use AI to draft feedback comments that they then refine, particularly when working with large numbers of students.

Surveys of university language instructors suggest both enthusiasm for these possibilities and anxiety about errors, cultural inappropriateness and the time required to verify and adapt AI outputs (Kohnke et al., 2023b). Many instructors worry that they might inadvertently pass on AI-generated misconceptions or insensitivities if they rely too heavily on unedited content. The emerging consensus is that generative AI is most useful in early stages of design, for instance as a way to brainstorm and generate draft materials, while human teachers are responsible for editing, contextualization and alignment with pedagogical objectives (Hockly, 2023; Kohnke et al., 2023a, 2023b).

Finally, AI underpins many learning analytics dashboards that aggregate data from learners' interactions with digital platforms. Syntheses of AI in language education note that such analytics can help teachers identify learners who are struggling, detect patterns of errors and monitor engagement (Huang et al., 2023; Liang et al., 2023). For example, a dashboard might show that a subset of students consistently underperforms on listening tasks or abandons speaking exercises early, prompting the teacher to investigate possible causes and offer support. Yet the usefulness of these tools depends heavily on teachers' data literacy and on institutional policies governing data collection and interpretation. If analytics are treated as neutral or self-explanatory, there is a danger of over-simplifying complex learning processes or inadvertently stigmatizing certain groups of students.

Challenges, Ethics and Equity

While the promise of AI in language education is substantial, the literature also draws attention to significant challenges and ethical concerns. These issues are not peripheral; they shape how AI is experienced by learners and teachers and influence whether AI contributes to more just or more unequal educational systems.

Algorithmic bias is a central concern. Because AI systems learn from data, they inevitably reproduce patterns found in those data. In language education, this can lead to better performance on "standard" varieties of English or other target languages, while misrecognizing regional or minoritized accents, dialects and learner varieties (Huang et al., 2023; Jiang et al., 2023). For ASR-based pronunciation tools, this may mean that certain learners receive more frequent and accurate feedback than others; for automated scoring systems, it may mean that some groups are systematically mis-scored. If, for example, training data include few samples of a particular accent, the system may treat that accent as "incorrect" even when it is entirely intelligible.

Fairness-oriented AI research therefore emphasizes the need for diverse training data, ongoing evaluation and transparency about system limitations, especially when AI tools are used in high-stakes assessment contexts (UNESCO, 2021; Huang et al., 2023). For language educators, this means asking difficult questions about whose language varieties are represented in AI systems, how errors are defined and what happens when students' identities and voices do not fit the system's expectations. It also implies advocating for institutions and policy-makers to adopt fairness standards that are sensitive to linguistic diversity.

Data privacy and security are equally pressing. Many AI platforms log detailed traces of learner behaviour, including keystrokes, speech recordings, error patterns and engagement metrics. UNESCO's guidance on AI in education stresses that such data should be collected and processed only with informed consent, strong security safeguards and clear rules regarding ownership, retention and secondary use (UNESCO, 2021). For language educators and institutions, this implies the need for transparent communication with learners about what data are collected and why, options to opt out where feasible, and restrictions on commercial reuse of student data. When students use third-party tools outside institutional control, these protections may be weaker, raising additional questions about responsibility.

Another widely discussed issue is over-reliance. Studies of generative AI in educational contexts suggest that some learners are tempted to delegate core tasks such as writing essays or translating texts entirely to AI systems, thereby limiting opportunities for productive struggle and the development of critical thinking (Hockly, 2023; Kohnke et al., 2023b; Ajmal et al., 2025). If learners consistently ask AI to "write for them" rather than with them, their own productive skills may stagnate. Teachers, too, may risk deskilling if they allow AI to take over functions such as assessment

design or feedback without maintaining their own expertise. Rather than framing AI as a replacement for human skill, authors argue for a model of human-AI partnership, where teachers use AI to extend their capacity while retaining responsibility for pedagogical decisions (Hockly, 2023; Huang et al., 2023; Zhu & Wang, 2025).

Teacher readiness and professional development emerge as critical enabling factors. Systematic reviews consistently note that teachers' beliefs, confidence and training strongly influence whether AI is used meaningfully, superficially or not at all (Huang et al., 2023; Liang et al., 2023; Zhu & Wang, 2025). Many instructors feel underprepared to evaluate AI tools, integrate them into curricula, or discuss their limitations and ethical implications with students (Kohnke et al., 2023b). Some may be enthusiastic but lack technical skills whereas others may be sceptical or anxious about being replaced. Professional development initiatives therefore need to move beyond purely technical training to include critical examination of AI's strengths and weaknesses, collaborative exploration of AI-supported task designs and discussion of policy and ethics (UNESCO, 2021). When teachers are invited to experiment with AI in supported settings, share experiences with colleagues and reflect on their own values, they are more likely to adopt AI in ways that align with their pedagogical commitments.

Finally, the issue of equity and the digital divide runs through much of the literature. Advanced AI tools often require up-to-date devices, stable broadband and sometimes paid subscriptions, which may not be available to all learners or institutions. Without deliberate policy measures, AI-enhanced language learning risks amplifying existing inequalities; that is, students in well-resourced contexts may benefit from rich, adaptive and immersive environments, while others are limited to basic or offline options (UNESCO, 2021; Huang et al., 2023). Even within the same institution, differences in access to hardware or quiet spaces for online practice can shape who is able to take full advantage of AI-based tools.

Researchers and policymakers therefore call for strategies such as institutional licensing, the development of low-bandwidth and mobile-friendly tools, and open-source alternatives to ensure that AI does not become a new axis of educational exclusion (UNESCO, 2021; Zhu & Wang, 2025). For language teachers, equity concerns may translate into concrete choices such as selecting tools that work on older phones, providing non-digital alternatives where appropriate, and designing tasks so that students are not penalized for circumstances beyond their control.

Overlaying all of these issues is the more fundamental limitation that even the most advanced AI systems do not "understand" language and culture in human ways. Large language models can produce fluent but factually incorrect or pragmatically inappropriate responses, a phenomenon widely known as hallucination (Hockly, 2023; Huang et al., 2023). They may also reproduce stereotypical or biased content if such patterns appear in their training data. Studies evaluating such models in educational settings underscore the need for learners and teachers to verify AI outputs, triangulate information and resist the temptation to treat fluent language as evidence of truth (Zhu & Wang, 2025). Developing AI literacy comprising skills for questioning, checking and situating AI-generated language is therefore as important as learning to use the tools technically.

Conclusion and Future Directions

The research surveyed in this article depicts different aspects of AI in language learning and teaching. On the one hand, AI-enhanced tutoring, chatbots, automated feedback, adaptive platforms and immersive environments can expand opportunities for practice, offer more tailored support and alleviate some routine aspects of teachers' workload. They can provide learners with additional channels for receiving input, experimenting with output and receiving feedback, often in contexts where human resources are stretched thin (Huang et al., 2023; Liang et al., 2023; Zhu & Wang, 2025). Intelligent systems can track patterns across large datasets that would be difficult for a single teacher to detect, and they can operate continuously, offering micro-interventions at moments when learners are ready.

On the other hand, the benefits of AI are dependent on careful pedagogical design, ethical governance and sustained human oversight. Intelligent systems can optimize what they are

programmed to measure, but they cannot on their own ensure the development of critical, creative and socially situated language use. Teachers remain central in framing tasks, interpreting data, mediating feedback and nurturing the interpersonal and intercultural dimensions of language learning that AI cannot replicate (UNESCO, 2021; Hockly, 2023). The risk is not that AI will suddenly replace teachers, but that it may subtly shape what counts as "successful" learning in ways that prioritize what can be easily measured and automated.

Across the literature, three broad messages recur. First, AI is best understood as a powerful amplifier rather than a substitute for human teaching. When used judiciously, it can free time for relational and interpretive work while supporting learners with additional practice and feedback. Second, context and criticality matter; tools need to be integrated into specific curricula, with explicit guidance for learners on how to use them reflectively rather than passively. Third, ethics and equity are not peripheral concerns but core design constraints; addressing bias, privacy and access is essential if AI is to contribute to more just rather than more unequal language education systems (UNESCO, 2021; Huang et al., 2023; Zhu & Wang, 2025).

Future research will need to move beyond short-term pilot studies toward more rigorous, longitudinal and comparative designs. There is a need for studies that examine how AI-supported and non-AI-supported approaches differ in their long-term effects on proficiency, motivation, learner identity and autonomy across diverse populations and languages. Large-scale systematic reviews have begun to sketch this landscape (Huang et al., 2023; Liang et al., 2023; Zhu & Wang, 2025), but more context-sensitive qualitative work is also required to understand how learners and teachers experience AI in their everyday practices.

It will also be important to explore models of human-AI collaboration in which teachers and learners co-construct tasks, feedback and assessments with AI, rather than simply consuming pre-packaged solutions. Such models might involve students critically rewriting AI-generated texts, negotiating assessment criteria in light of AI analytics or designing their own prompts and workflows for AI tools (Kohnke et al., 2023a, 2023b; Ajmal et al., 2025). Finally, research on teacher professional development, policy frameworks and institutional cultures will be crucial in shaping whether AI deepens or diminishes the human core of language education.

If approached with care, humility and critical imagination, AI can become a valuable partner in language learning and teaching; not by replacing teachers or learners, but by augmenting their capacities and opening new spaces for communication, reflection and creativity.

Disclosure Statements

Contribution rate statement of the researchers:

The author conducted the whole study.

Conflict of interest statement:

The authors declare that there is no conflict of interest.

CRediT Authorship Contribution Statement

Süleyman BAŞARAN: Conceptualization, methodology, data collection, data analysis, writing – review & editing.

References

Ajmal, A., Aslam, S., Yousaf, M. S., & Raza, M. A. (2025). Using ChatGPT to enhance ESL writing skills in Pakistani higher education. *International Journal of Social Sciences Bulletin,* 3(3), 666–673. https://ijssbulletin.com/index.php/IJSSB/article/view/535

Alyaz, Y., Demiryay, N. (2023). Yabancı dil öğrenimi ve öğretiminde sanal gerçeklik uygulamaları. *Diyalog*, 11(1), 107–127. https://doi.org/10.37583/diyalog.1312776

Hockly, N. (2023). Artificial intelligence in English language teaching: The good, the bad and the ugly. *RELC Journal*, 54(2), 445–451. https://doi.org/10.1177/00336882231168504

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

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, 26(1), 112–131. https://doi.org/10.30191/ETS.202301_26(1).0009

Hung, C.-Y., Lin, Y.-T., Yu, S.-J., & Sun, J. C.-Y. (2023). Effects of AR- and VR-based wearables in teaching English: The application of an ARCS model-based learning design to improve elementary school students' learning motivation and performance. *Journal of Computer Assisted Learning*, 39(5), 1510–1527. https://doi.org/10.1111/jcal.12814

Ivanova, M., Grosseck, G., & Holotescu, C. (2024). Unveiling insights: A bibliometric analysis of artificial intelligence in teaching. *Informatics*, 11(1), 10. https://doi.org/10.3390/informatics11010010

Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., & Wu, N. (2023). Exploring the effects of automatic speech recognition technology on oral accuracy and fluency in a flipped classroom. Journal of Computer Assisted Learning, 39(1), 125-140.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023a). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. https://doi.org/10.1177/00336882231162868

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023b). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, 5, 100156. https://doi.org/10.1016/j.caeai.2023.100156

Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages*, 4(1), 13. https://doi.org/10.3390/languages4010013

Liang, J.-C., Hwang, G.-J., Chen, M.-R. A., & Darmawansah, D. (2023). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*, 31(7), 4270–4296. https://doi.org/10.1080/10494820.2021.1958348

Martin, F., Chen, Y., Moore, R. L., & Westine, C. D. (2020). Systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018. *Educational Technology Research and Development*, 68(4), 1903-1929.

Sari, E., & Han, T. (2024). The impact of automated writing evaluation on English as a foreign language learners' writing self-efficacy, self-regulation, anxiety, and performance. Journal of Computer Assisted Learning, 40(5), 2065–2080. https://doi.org/10.1111/jcal.13004

UNESCO. (2021). *Artificial intelligence in education: Guidance for policy-makers*. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000376709

Wei, P., Wang, X., & Dong, H. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: A randomized controlled trial. Frontiers in Psychology, 14, 1249991. https://doi.org/10.3389/fpsyg.2023.1249991

Yıldız, H. & Gonen, İ. K., (2024). Automated writing evaluation system for writing in online learning environments. *Turkish Online Journal of Distance Education*, 24(3), 169–190. https://doi.org/10.17718/tojde.1169727

Zhu, L., & Wang, T. (2025). A systematic review of research on artificial intelligence in language education: Current status and future implications. *Language Learning & Technology*, 29(1), 1–26. https://www.lltjournal.org

****

Ethical Declaration and Committee Approval

In this research, the principles of scientific research and publication ethics were followed.