Global Spectrum of Research and Humanities
ISSN: 3007-9136
Vol. 3 No. 3 (2026): Autumn

Artificial Intelligence in Language Education: Human-AI Interaction, Pedagogical Design, and Ethical Challenges

Zarifa Sadigzade
Lecturer, Nakhchivan State University, Azerbaijan
Email: zarifasadig@gmail.com
ORCID: https://orcid.org/0009-0007-1179-1214
DOI: https://doi.org/10.69760/gsrh.0260303001

Abstract

The integration of artificial intelligence (AI) into language education has moved well beyond the deployment of automated grammar checkers and rudimentary machine translation tools. Today, AI reshapes the fundamental relationships among teachers, learners, language, and knowledge. This article presents a theoretically grounded analysis of how AI systems are reconfiguring language teaching, language learning, translation-mediated tasks, and academic writing in both first- and second-language educational contexts. Drawing on the literature in computer-assisted language learning (CALL), educational technology, and human-computer interaction, the article develops an original Triadic Human-AI Interaction Framework that conceptualises the dynamic and mutually constitutive roles of AI systems, teachers, and learners. The framework foregrounds human-AI collaboration as a relational and pedagogically purposeful process rather than a matter of mere tool adoption. Four thematic domains are examined in depth: (1) AI-mediated instructional design and teacher mediation; (2) learner autonomy and personalisation; (3) AI-assisted feedback and assessment; and (4) academic writing and translation-related tasks. The article further addresses a constellation of ethical challenges — privacy, algorithmic bias, transparency, academic integrity, the risk of overreliance, and teacher AI literacy — arguing that these issues are not peripheral but constitutive concerns for responsible AI integration. Implications for pedagogical theory and classroom practice are offered, together with an agenda for future research.

Keywords: *artificial intelligence; language education; human-AI interaction; pedagogical design; learner autonomy; teacher mediation; academic writing; machine translation; ethical challenges; CALL*

1. Introduction

The relationship between language and technology is neither new nor uncomplicated. From the earliest language laboratories of the mid-twentieth century to the networked multimedia environments of the digital era, each technological wave has prompted educators, linguists, and theorists to reconsider what it means to teach and learn a language. Artificial intelligence, however, represents something qualitatively different from its predecessors.

Where earlier technologies extended the reach of pedagogical intention — providing more input, faster feedback, or richer multimedia — AI systems increasingly intervene in the cognitive and communicative processes that constitute language learning itself. They evaluate, adapt, generate, translate, and converse. They operate not merely as conduits for instructional content but as active participants in the epistemic and communicative labour of language education.

The pace of this transformation has been rapid and, at times, disorienting. Large language models (LLMs), neural machine translation (NMT) systems, and conversational AI agents have entered classrooms, writing centres, and assessment regimes with a speed that has outpaced the theoretical frameworks available to make sense of them. While a growing body of empirical research documents learner attitudes, chatbot effectiveness, and writing outcomes, theoretical synthesis has lagged behind. This article seeks to close that gap by offering a coherent theoretical account of how AI is reshaping language education across four interrelated domains: classroom instruction and teacher mediation, learner autonomy and personalisation, feedback and assessment, and academic writing and translation.

The article proceeds from the conviction that neither techno-optimism nor techno-pessimism offers adequate purchase on the complexity of AI in language education. The question is not whether AI is good or bad for language learning in the abstract, but rather: under what conditions, with what pedagogical intentions, and subject to what ethical safeguards can human-AI interaction productively support language development? Answering this question requires theory — and theory adequate to the relational, ethical, and pedagogical dimensions of the phenomenon.

To that end, the article makes three principal contributions. First, it synthesises the extant literature across CALL, educational technology, and applied linguistics to construct a comprehensive Triadic Human-AI Interaction Framework. Second, it applies this framework to analyse each of the four thematic domains named above. Third, it situates the pedagogical analysis within a sustained examination of ethical challenges, arguing that responsible AI integration requires not only technical design choices but shifts in professional culture, institutional policy, and curricular philosophy.

2. Theoretical Foundations and Literature Review

2.1 From CALL to AI-Mediated Language Education

Computer-assisted language learning emerged in the 1960s as a discipline concerned with the pedagogical application of computational tools to language instruction. Its history has been characterised by iterative recalibration in response to shifting paradigms — from behaviourist drill-and-practice software, through communicative and constructivist approaches, to the networked and mobile environments of the contemporary moment (Schank & Childers, 1983). Each phase brought new questions about the role of the machine in relation to the teacher and the learner. The advent of AI-powered language tools marks what may be the most consequential phase of this evolution.

Early AI applications in language education were characterised by rule-based systems with limited flexibility. Machine translation (MT), for example, moved from rule-based to statistical to neural architectures over several decades, with each transition bringing markedly improved output quality and correspondingly altered implications for language pedagogy (He et al., 2016). The introduction of neural machine translation systems capable of producing fluent, contextually sensitive translations challenged previously settled assumptions about what MT could and could not do in educational contexts (Garcia & Pena, 2011; Clifford, Merschel,

& Munne, 2013; Lee, 2020). Similarly, the development of conversational AI — from simple pattern-matching chatbots to transformer-based dialogue systems — opened new possibilities for extended, contextualised interaction in the target language (Haristiani, 2019; Belda-Medina & Calvo-Ferrer, 2022).

Huang et al. (2023) document the trajectory of AI research in language education, identifying a progressive broadening from narrow task-specific applications toward integrated, multi-modal intelligent systems. Schmidt and Strasser (2022) call for what they term 'intelligent practice' — pedagogical approaches that harness AI capabilities while preserving the essentially human dimensions of language learning. This call reflects a broader consensus in the field that AI is most effectively deployed not as a replacement for human judgment but as a collaborator within pedagogically designed environments.

2.2 Human-AI Interaction in Educational Contexts

The concept of human-AI interaction (HAI) has attracted growing theoretical attention in educational technology. Foundational work in this area draws on sociocultural learning theory, activity theory, and distributed cognition to argue that learning is always mediated — by tools, signs, social relationships, and institutional structures (Wang et al., 2021). AI systems enter this ecology as a distinctive class of mediating agent: one capable of responsive, adaptive, and generative behaviour that goes beyond static tool use.

Wang et al. (2023) offer an empirically grounded account of what matters in AI-supported learning, using cluster and epistemic network analysis to reveal that the quality of human-AI interaction is shaped not by AI capability alone but by the broader ecology of classroom practice, teacher design decisions, and learner orientations toward the technology. This finding has significant theoretical implications: it suggests that HAI cannot be understood as a dyadic relationship between learner and machine but must be situated within a triadic structure that includes the teacher as a purposeful mediating agent.

Molenaar (2022) theorises what she calls 'hybrid human-AI learning technologies', arguing that effective educational AI must be designed to support rather than supplant human agency. This hybridisation entails a careful orchestration of tasks between human and artificial intelligences, with the teacher responsible for the overall pedagogical architecture within which AI functions. Ou, Stohr, and Malmstrom (2024) extend this argument from a post-humanist perspective, proposing that AI-assisted academic communication involves a fundamental reconceptualisation of authorship, agency, and identity in educational settings.

2.3 AI in Language Teaching and Teacher Mediation

The teacher's role in AI-mediated language education has been theorised in terms that range from the highly optimistic — AI as a liberator of teacher time and creativity — to the deeply cautious — AI as a threat to professional expertise and relational pedagogy. Ji, Han, and Ko (2023), in a systematic review of conversational AI in language education, conclude that effective integration consistently depends on the collaboration between AI and human teachers, with teacher mediation serving as the critical variable that determines whether AI interaction produces educationally meaningful outcomes.

Chocarro, Cortinas, and Marcos-Matas (2023) examine teacher attitudes toward educational chatbots through a technology acceptance model lens, finding that social language use, proactiveness of the AI agent, and individual teacher characteristics all modulate the extent to which teachers embrace or resist AI integration. Critically, their findings suggest that teacher attitudes are not static but are shaped by experience, professional development, and the quality

of institutional support. This has direct implications for teacher education and professional learning.

Hwang and Chang (2023) review the opportunities and challenges presented by educational chatbots, concluding that while AI affords significant opportunities for differentiated instruction, immediate feedback, and extended practice opportunities, these affordances are only realised when teachers design learning environments with explicit pedagogical intentionality. The AI does not teach; rather, it is the teacher's design that teaches, with AI as one among several mediating resources.

2.4 Learner Autonomy, Personalisation, and Motivation

Learner autonomy — the capacity and disposition to take charge of one's own learning — has long been a central construct in language education theory. AI offers new possibilities for supporting autonomy through personalisation: the capacity to adapt instructional content, feedback, pacing, and task selection to the individual learner's needs, preferences, and developmental trajectory. Wei (2023) demonstrates that AI-assisted language instruction can positively impact English learning achievement, L2 motivation, and self-regulated learning, though the mechanisms of these effects are complex and context-dependent.

An et al. (2025) model students' perceptions of AI-assisted language learning, finding that perceived usefulness, ease of use, and social influence are significant predictors of behavioural intention to use AI tools. Crucially, their findings suggest that learner agency is not simply an outcome of AI personalisation but a prerequisite: learners who approach AI tools with greater metacognitive awareness and strategic intent derive greater learning benefit. This creates a productive paradox — AI tools designed to support learner autonomy may be most effective for learners who are already relatively autonomous.

Feng (2025) investigates the effects of AI-assisted learning strategies on cognitive load and learning outcomes, finding that well-designed AI scaffolding can reduce extraneous cognitive load and direct learner attention to target language features. However, this effect depends critically on alignment between AI design and learner characteristics — a misaligned AI scaffold can increase cognitive load and impede learning. These findings underline the importance of pedagogical design in mediating the relationship between AI affordances and learner outcomes.

2.5 AI-Assisted Feedback and Assessment

Feedback is widely recognised as one of the most powerful influences on learning. AI systems offer the possibility of immediate, detailed, and differentiated feedback at a scale that no human teacher can match. In language education, AI-generated feedback spans a range of forms: automated writing evaluation systems that identify syntactic and lexical errors; conversational AI that provides interactional feedback in speaking and writing tasks; and adaptive systems that track learner progress over time and adjust feedback accordingly.

Song and Song (2023) assess the efficacy of ChatGPT in AI-assisted language learning for EFL students, finding positive effects on academic writing skills and motivation when AI feedback is appropriately scaffolded by teacher guidance. However, they also identify risks: learners who rely exclusively on AI feedback may develop a narrowed understanding of quality that reflects the AI's evaluative criteria rather than the complex, contextual judgments that characterise expert human assessment.

Kuhail et al. (2023) conduct a systematic review of interactions with educational chatbots, finding that the most effective chatbot feedback is characterised by specificity,

timeliness, and dialogic responsiveness — qualities that require not only capable AI but thoughtful pedagogical design that structures the interaction in meaningful ways. This finding reinforces the centrality of teacher design in mediating the quality of AI-generated feedback.

2.6 AI in Academic Writing and Translation

Academic writing and translation represent two domains in which AI has produced both significant opportunity and significant anxiety. In academic writing, large language models can generate grammatically correct, rhetorically organised prose across a wide range of genres and registers. For second language writers, this capability offers substantial support — but also substantial temptation. Chen and Gong (2025) examine the role of AI-assisted learning in the academic writing of Chinese as a second language students, finding that AI tools can support writing development when used strategically but may impede it when substituted for the cognitive struggle involved in drafting and revising.

Niño (2009) identified the ambivalence that characterises pedagogical responses to machine translation in the first decade of the century, with both language learners and tutors recognising MT's practical utility while expressing concern about its effects on target language development. This ambivalence has only deepened as NMT systems have become vastly more capable. Lee (2020) examines the impact of MT use on EFL students' writing, finding that while MT can increase fluency and lexical sophistication, it may simultaneously reduce the learner's engagement with the generative processes through which language competence develops.

The question of what constitutes authentic authorship in AI-mediated writing contexts is emerging as one of the most pressing theoretical questions in applied linguistics and language education. Ou, Stohr, and Malmstrom (2024) propose a post-humanist framework that challenges the humanist assumption of a unitary, sovereign author, arguing instead for a distributed understanding of academic authorship as a relational and technological achievement. This theoretical move has significant implications for how language educators conceptualise, teach, and assess academic writing in AI-rich environments.

3. A Triadic Human-AI Interaction Framework

3.1 Framework Overview

Drawing on the theoretical traditions reviewed above, this article proposes a Triadic Human-AI Interaction (THAI) Framework for understanding the dynamics of AI integration in language education. The framework comprises three primary nodes — AI Systems, Teachers, and Learners — each constituted by distinctive capacities, orientations, and responsibilities, and connected by three relational dyads: AI-Teacher interaction, AI-Learner interaction, and Teacher-Learner interaction. These dyads intersect at a central nexus designated Language Development, which represents the educational purpose that gives the entire framework its direction and coherence.

The THAI Framework is triadic rather than dyadic because the dominant tendency in educational AI discourse has been to theorise human-AI interaction as a two-party affair between learner and machine, with the teacher relegated to a peripheral or purely facilitative role. This article argues that such a conception is theoretically inadequate and pedagogically dangerous. The teacher is not merely an optional intermediary between learner and AI but the architect of the learning environment within which both AI and learner operate. Without sustained attention to teacher mediation, AI-learner interaction risks becoming directionless, evaluatively impoverished, and ethically unmoored.

3.2 The AI Systems Node

Within the THAI Framework, AI Systems are understood not as neutral tools but as value-laden, designed artefacts that embody particular assumptions about language, learning, and communication. Current AI systems deployed in language education include: large language model-based writing assistants and tutors; neural machine translation systems; automated speech recognition and pronunciation feedback tools; conversational chatbots for speaking and writing practice; adaptive learning platforms that personalise content and pacing; and automated writing evaluation systems.

Each of these systems brings specific capabilities and specific limitations. LLMs, for example, produce fluent, contextually sensitive language but may generate factually incorrect content, reflect training data biases, and lack sensitivity to the pragmatic and cultural dimensions of language use. NMT systems achieve high translation quality at the sentence level but may struggle with discourse-level coherence and culturally specific content. These capabilities and limitations are not merely technical matters but pedagogically significant facts that teachers must understand in order to design learning environments in which AI functions appropriately.

Within the THAI Framework, the AI Systems node is characterised by four functional capacities: generation (the capacity to produce language), evaluation (the capacity to assess language against learned criteria), adaptation (the capacity to adjust output in response to learner input), and mediation (the capacity to intervene between learner and target language content in ways that shape the learner's linguistic experience). Each of these capacities carries both opportunity and risk, and each requires specific pedagogical design responses.

3.3 The Teacher Node

In the THAI Framework, the Teacher occupies a central and irreducible position. Three roles define teacher engagement with AI in language education. The first is the role of Designer: the teacher constructs the pedagogical environment within which AI operates, determining what tasks AI will support, how AI output will be used, what scaffolding will frame AI interaction, and how AI-generated content will be evaluated and elaborated. Effective pedagogical design requires both technological knowledge — an understanding of what AI systems can and cannot do — and pedagogical content knowledge — an understanding of how language develops and what kinds of instruction best support that development.

The second teacher role is that of Mediator: the teacher positions herself between the learner and the AI, interpreting AI output, directing learner attention to significant features of that output, and scaffolding the cognitive and linguistic processes through which learners make use of AI support. This mediating role is particularly critical in feedback contexts: AI-generated feedback is rarely self-explanatory, and learners typically require teacher guidance to interpret, evaluate, and act upon it in ways that support language development.

The third teacher role is that of Evaluator: the teacher monitors the quality and appropriateness of AI integration at both the macro level (Is this AI tool well-suited to my pedagogical goals?) and the micro level (Is this individual learner using AI in ways that support or impede her development?). This evaluative role requires ongoing attention and professional judgment, and it depends upon a level of AI literacy that is not yet systematically developed in most language teacher education programmes.

3.4 The Learner Node

Within the THAI Framework, the Learner is understood as an active, purposeful agent whose engagement with AI tools is shaped by metacognitive awareness, motivational orientation, prior linguistic and technological experience, and the pedagogical design of the learning environment. Learner engagement with AI is not a simple input-output process but a complex, recursive activity in which the learner continuously positions herself in relation to the AI's offerings — accepting, rejecting, modifying, and extending them in the service of her linguistic and communicative goals.

The Framework identifies three learner stances toward AI that carry distinct implications for language development. The strategic stance involves the learner using AI tools deliberately and reflectively, with clear awareness of her learning goals and the AI's capabilities and limitations. The dependent stance involves the learner delegating cognitive and communicative work to the AI, reducing the generative struggle through which language competence develops. The resistant stance involves the learner declining to engage with AI tools, either from principled pedagogical commitment or from technological anxiety.

The THAI Framework suggests that the goal of pedagogical design is to cultivate the strategic stance — not by prescribing how learners use AI, but by designing learning environments that make strategic use the path of most pedagogical reward. This requires explicit instruction in metacognitive strategies for AI use, transparent discussion of AI capabilities and limitations, and assessment designs that value the processes of language use rather than merely their products.

3.5 Relational Dynamics and the Language Development Nexus

The three dyadic relationships in the THAI Framework — AI-Teacher, AI-Learner, and Teacher-Learner — are not independent but mutually constitutive. The quality of AI-Learner interaction is shaped by the Teacher's design decisions; the Teacher's design decisions are informed by her knowledge of AI capabilities and learner characteristics; and the Learner's strategic orientation toward AI is cultivated through the Teacher-Learner relationship. These interdependencies mean that AI integration cannot be analysed in isolation from the broader ecology of classroom practice.

The Language Development nexus at the centre of the framework serves as both the integrating purpose and the evaluative criterion for all three nodes and their relational dyads. In language education, the ultimate measure of any pedagogical choice — including the choice to integrate AI — is its contribution to language development in the broad sense: not merely the acquisition of grammatical forms or vocabulary items, but the development of communicative competence, discursive identity, and critical linguistic awareness. The THAI Framework insists on this broader conception of language development as the touchstone against which AI integration must be evaluated.

4. Thematic Analysis: AI Across Language Education Domains

4.1 AI-Mediated Instructional Design and Classroom Practice

The integration of AI into instructional design represents a significant expansion of the pedagogical toolkit available to language teachers. AI-powered adaptive learning platforms can dynamically sequence content according to learner performance, providing more challenging material when learners demonstrate mastery and additional practice when gaps are identified. Intelligent tutoring systems can scaffold complex tasks — reading comprehension, grammar induction, writing — with graduated support that responds to individual learner needs. These

capabilities, if effectively designed and pedagogically situated, can substantially enhance the quality and efficiency of language instruction.

Son, Ruzic, and Philpott (2025) survey the range of AI technologies and applications in language teaching, noting that effective deployment requires not only access to capable AI tools but sustained professional development that equips teachers to integrate those tools with pedagogical intentionality. This finding resonates with Hwang and Chang's (2023) argument that AI affordances in education are always mediated by teacher design: the same AI tool can produce markedly different learning outcomes depending on the pedagogical framework within which it is embedded.

A critical dimension of AI-mediated instructional design is the management of cognitive load. Feng (2025) demonstrates that AI scaffolding can reduce extraneous cognitive load — the processing burden associated with unclear or poorly designed tasks — while preserving germane cognitive load, the productive effort associated with genuine linguistic and conceptual challenge. This distinction has direct implications for task design: AI should be used to remove unnecessary processing barriers while preserving the core cognitive challenge of language use. Tasks in which AI simply does the linguistic work for the learner produce minimal cognitive engagement and minimal learning.

Classroom practice involving AI also raises significant questions about the social and affective dimensions of language learning. Language is inherently intersubjective — it develops in and through social interaction — and AI interactions, however sophisticated, are qualitatively different from human-human communication. Ji, Han, and Ko (2023) argue that conversational AI is most pedagogically effective when it supplements rather than replaces human interaction, providing additional practice opportunities and immediate feedback that free up human interaction time for the more complex, socially rich communicative activities that AI cannot replicate.

4.2 Learner Autonomy and Personalisation

The relationship between AI personalisation and learner autonomy is theoretically complex. Personalisation — the adaptation of instructional content, pace, and feedback to individual learner characteristics — is frequently positioned in the literature as a mechanism for supporting autonomy, on the grounds that it removes the artificial homogenisation of traditional whole-class instruction and allows learners to engage with content appropriate to their level and interests. This argument has intuitive appeal but requires critical examination.

True learner autonomy involves not merely the capacity to choose from a range of AI-curated options but the metacognitive disposition to direct one's own learning, to evaluate one's progress critically, and to make informed decisions about the use of learning resources. An et al. (2025) demonstrate that perceived usefulness and ease of use significantly influence learners' engagement with AI tools, but they also find that learners with higher levels of self-regulated learning tend to engage more strategically and derive greater benefit from AI affordances. This suggests that AI personalisation, to be genuinely autonomy-supporting, must be accompanied by explicit instruction in self-regulated learning strategies.

Zou et al. (2023) examine AI-assisted speaking practice through social network-based interaction, finding that the combination of AI interlocutor and peer social network provides richer opportunities for autonomous practice than either alone. This finding suggests a productive design principle: AI is most effectively deployed not as a standalone personalisation engine but as a component of a broader learning ecology that includes human social interaction.

The AI provides the scalability and responsiveness that human teachers and peers cannot; human interaction provides the authenticity and relational richness that AI cannot.

4.3 AI-Assisted Feedback and Assessment

Feedback is the domain in which AI's educational potential is perhaps most immediately apparent. The chronic shortage of immediate, individualised feedback in language education — a consequence of teacher workload and class size constraints — has long been identified as a significant impediment to learning. AI systems offer, in principle, unlimited availability of immediate, detailed feedback. The question is whether AI-generated feedback supports learning in the way that effective human feedback does.

The answer, as the literature consistently suggests, is: it depends. Labadze, Grigolia, and Machaidze (2023) review the role of AI chatbots in education and conclude that chatbot feedback is most effective when it is specific, actionable, and dialogic — when it not only identifies an error or gap but explains the reasoning behind the correction and invites the learner to respond. This dialogic quality is difficult to achieve in automated feedback systems and requires careful design and, in many cases, teacher mediation to be fully realised.

Assessment represents a related but distinct challenge. AI-generated feedback and AI-based assessment are not equivalent. Assessment involves not only the identification of errors but the holistic evaluation of communicative performance in relation to contextually appropriate criteria. Current AI systems are considerably more capable in the former than the latter. Hockly (2023) identifies this as one of the central challenges of AI in English language teaching: AI can support formative feedback processes but is not yet well-suited to replace the contextual, relational, and holistic judgments involved in summative assessment of language proficiency.

Wang et al. (2023) find, through cluster and epistemic network analysis, that the patterns of human-AI interaction in language learning that are most strongly associated with positive outcomes are characterised by metacognitive engagement — learners who approach AI feedback reflectively, questioning, evaluating, and selectively applying it. This finding has important implications for pedagogy: AI feedback is not a passive input to be absorbed but an active invitation to metacognitive engagement that must be deliberately cultivated through instructional design.

4.4 Academic Writing and AI-Mediated Translation

Academic writing and translation-mediated tasks represent two domains in which AI has generated both substantial enthusiasm and substantial anxiety among language educators. Both domains involve the production of written language in high-stakes contexts; both have been transformed by AI tools of rapidly increasing capability; and both raise fundamental questions about authorship, authenticity, and the relationship between tool use and learning.

In academic writing, the pedagogical challenge is to leverage AI's capacity to support the development of writing skills — through suggestion, scaffolding, and feedback — without supplanting the cognitive and linguistic processes through which academic writing competence develops. Chen and Gong (2025) argue that AI-assisted writing can benefit Chinese L2 writers when tools are used to extend and refine the writer's own drafts rather than to generate text de novo. The pedagogical distinction between AI as writing scaffold and AI as writing substitute is crucial, and it is a distinction that must be made explicit in writing instruction.

Song and Song (2023) find that ChatGPT can enhance academic writing skills and motivation in EFL students when its use is structured by clear pedagogical goals and teacher

guidance. Their findings suggest that AI writing support is most beneficial when learners understand what the AI is doing and why — when they can engage critically with AI suggestions rather than simply accepting them. This requires not only technological transparency but pedagogical scaffolding that develops learners' capacity for critical engagement with AI output.

Translation-mediated tasks represent a particularly rich site for examining the pedagogical implications of AI. The history of MT in language education — from the early scepticism documented by Niño (2009) and Clifford, Merschel, and Munne (2013) through the growing acceptance documented by Lee (2020) — reflects a progressive recognition that MT cannot be banished from language classrooms and that pedagogical responses must engage constructively with its presence. The emergence of NMT systems capable of high-quality translation across many language pairs has accelerated this reckoning.

Garcia and Pena (2011) demonstrated early on that MT-assisted writing could support language development for beginners when tasks were designed to engage learners with the MT output rather than simply reproduce it. This design principle — MT as input for active linguistic engagement rather than as a bypass for linguistic effort — remains foundational. He et al.'s (2016) dual learning framework for NMT, while technically oriented, carries a pedagogical resonance: learning through translation involves a bidirectional process of meaning negotiation that is impoverished when automated translation eliminates the need for the learner's own engagement in that process.

5. Ethical Challenges in AI-Mediated Language Education

5.1 Privacy and Data Governance

AI systems in education are data-hungry. Adaptive learning platforms, conversational tutors, and automated assessment systems all depend on the collection, storage, and processing of learner data — including linguistic performance data, interaction patterns, and in some cases biometric data such as voice recordings. The pedagogical affordances of these systems are inseparable from their data practices, and this creates significant privacy challenges that are not adequately addressed by either existing data protection legislation or current institutional practices in most educational contexts.

The language educational context adds a distinctive dimension to privacy concerns. Language learning involves the expression of personal meaning, cultural identity, and social positioning. Learner writing and speaking samples are not merely data points about linguistic performance but windows into the learner's intellectual and personal development. When these data are collected by commercial AI providers operating under opaque terms of service, the learner's privacy interests are at risk in ways that extend beyond the standard concerns of data protection law.

Ou, Stohr, and Malmstrom (2024) address the implications of post-humanist theories of authorship for data governance in AI-assisted academic writing, noting that when AI systems contribute to the production of learner texts, questions of intellectual property and data ownership become genuinely complex. Institutional policies on AI use in education typically focus on academic integrity but rarely address the prior question of who owns the data generated in AI-learner interactions — a gap that requires urgent attention from educators, institutions, and policymakers alike.

5.2 Algorithmic Bias and Equity

AI systems are not neutral. They are trained on data that reflects historical patterns of language use, and these patterns embed social inequalities, cultural biases, and epistemological assumptions that may disadvantage particular groups of learners. In language education, the most immediately significant form of algorithmic bias concerns the representation of language varieties and the implicit privileging of standard, prestigious varieties over regional, non-native, and marginalised forms.

Automated speech recognition systems, for example, have been shown to perform significantly less accurately for speakers of non-standard language varieties, speakers with accents, and speakers from under-represented demographic groups. This performance gap is not merely a technical inconvenience; it creates a systematically unequal learning environment in which some learners receive more accurate feedback than others on account of factors entirely unrelated to their linguistic competence. Hockly (2023) identifies equity as one of the central ethical concerns of AI in English language teaching, arguing that uncritical adoption of AI tools may reinforce existing inequalities in access and outcomes.

Labadze, Grigolia, and Machaidze (2023) note that the training data for most educational AI systems is dominated by text from English-dominant, Western, and formally educated sources — a pattern that produces systems implicitly calibrated to certain cultural and linguistic norms. For language learners from diverse cultural and linguistic backgrounds, this calibration may produce feedback that is technically correct but culturally alienating — feedback that positions the learner's own linguistic and cultural repertoire as deficiency rather than resource.

5.3 Transparency and Explainability

A recurring concern in the literature on educational AI is the opacity of AI systems — the difficulty of understanding how they arrive at their outputs, what assumptions underlie their evaluations, and what criteria govern their recommendations. This opacity is particularly problematic in educational contexts, where the rationale for feedback and assessment is itself an important object of learning. When a learner cannot understand why an AI system has flagged a particular construction as incorrect or suboptimal, the feedback cannot easily be integrated into the learner's developing metalinguistic understanding.

Transparency in AI educational systems operates at multiple levels. At the technical level, explainability refers to the capacity to provide human-understandable accounts of AI decision-making processes. At the pedagogical level, transparency requires that teachers understand the design assumptions and evaluative criteria of the AI tools they deploy. At the institutional level, transparency requires that learners, teachers, and administrators have access to information about how AI systems use their data and how those systems' outputs influence consequential decisions about learners.

Mageira et al. (2022) argue that educational AI chatbots for content and language integrated learning must be designed with pedagogical transparency as a core principle — that learners must be able to understand the basis for AI responses and evaluate them critically. This argument extends to all AI tools in language education: the pedagogical value of AI is enhanced, not diminished, when its workings are made visible to learners and teachers alike.

5.4 Academic Integrity

The challenge of academic integrity in AI-mediated language education is simultaneously practical and philosophical. At the practical level, AI writing tools make it easier than ever before for learners to produce text that does not reflect their own linguistic competence — to submit AI-generated prose as the product of their own writing effort. At the

philosophical level, the widespread adoption of AI writing tools raises fundamental questions about what it means to be the author of a text and what the relationship between authorship and learning should be in educational contexts.

Hockly (2023) identifies academic integrity as among the 'ugly' aspects of AI in language teaching — not because the challenge is intractable but because it requires educators to make difficult and contested judgments about the legitimate and illegitimate uses of AI tools in learning contexts. The appropriate response to this challenge is not the blanket prohibition of AI tools — which is both practically unenforceable and pedagogically counterproductive — but the thoughtful redesign of assessments and learning tasks in ways that are robust to AI assistance and that evaluate the processes of language use as well as their products.

This redesign requires a reconceptualisation of what academic writing tasks are for. If the purpose of a writing task is to produce a text that conforms to genre conventions and displays command of the target language, AI assistance can trivially discharge that purpose. If the purpose is to develop the learner's capacity for sustained, purposeful, and critically reflective engagement with ideas in the target language, AI assistance may support or impede that purpose depending on how it is designed and deployed. Language educators must be clear about which purposes their writing tasks serve — and must design tasks, assess accordingly, and teach learners to engage with AI in ways that serve those purposes.

5.5 Overreliance and the Question of Cognitive Offloading

One of the most significant and least theorised risks of AI in language education is overreliance — the tendency of learners to delegate to AI cognitive and linguistic work that, if performed by the learner herself, would contribute to language development. This risk is particularly acute in language education because language competence develops precisely through the effortful, generative engagement with the target language that AI tools are designed to make easier.

Huang, Hew, and Fryer (2022) critically examine the question of whether language learning chatbots are 'really useful', finding that the evidence for their effectiveness is mixed and that the benefits claimed for AI-assisted practice are frequently confounded by design variables related to the quality of pedagogical integration. Their analysis suggests that chatbot interaction, in the absence of explicit pedagogical scaffolding, may support surface-level practice without contributing to the deeper linguistic processing associated with durable acquisition.

The concept of cognitive offloading — the use of external resources to reduce the cognitive demands of a task — is theoretically relevant here. Cognitive offloading is not inherently problematic; it is a ubiquitous feature of human cognition and a foundation of tool use. In educational contexts, however, the question is whether the cognitive work being offloaded is the work that education is trying to develop. In language education, the work of generating, monitoring, revising, and reflecting upon language in use is precisely the work that develops linguistic competence. When AI assumes that work, the pedagogical value of the task is diminished or eliminated. Language educators must therefore distinguish carefully between AI use that supports and scaffolds linguistic cognitive work and AI use that simply replaces it.

5.6 Teacher AI Literacy

No discussion of ethical challenges in AI-mediated language education can be complete without attention to the question of teacher AI literacy — the knowledge, skills, and critical dispositions that language teachers need in order to engage productively and responsibly with AI tools. The literature consistently identifies teacher AI literacy as a significant gap: teachers

are increasingly expected to integrate AI into their instructional practice but are rarely provided with the professional development they need to do so with genuine pedagogical intentionality.

Chocarro, Cortinas, and Marcos-Matas (2023) identify teacher characteristics as a significant predictor of attitudes toward educational chatbots, finding that more experienced and technologically confident teachers are more likely to view AI tools favourably. However, favourable attitudes toward AI are not equivalent to AI literacy — teachers may embrace AI tools without critically evaluating their assumptions, limitations, or effects on learner development. Teacher education programmes must therefore develop not only technological confidence but critical AI literacy: the capacity to interrogate AI tools, to evaluate their pedagogical appropriateness, and to design learning environments in which AI is integrated with genuine pedagogical intention.

Kannan and Munday (2018) situate this challenge within the broader context of ICT integration in second language learning, arguing that effective technology integration requires a reconceptualisation of the teacher's professional role — not as a technologist or an AI operator but as a reflective practitioner who uses all available resources, including AI, in the service of well-theorised pedagogical goals. This reconceptualisation has implications for pre-service teacher education, in-service professional development, and the institutional cultures within which language teachers work.

6. Implications for Pedagogical Theory and Practice

6.1 Redesigning Language Learning Environments

The THAI Framework developed in this article has direct implications for the design of language learning environments. If effective AI integration requires triadic rather than dyadic human-AI interaction — with the teacher as designer, mediator, and evaluator — then learning environments must be structured to support all three teacher roles. This means investing in teacher professional development, in collaborative design processes that bring together teachers with AI design expertise, and in institutional cultures that value pedagogical innovation and reflective practice.

Kim, Cha, and Kim (2019) propose that the future of English learning lies in the integration of chatbots and AI within a broader reconceptualisation of language learning environments — one that leverages AI's capacity for personalised, always-available practice while preserving the social and relational dimensions of language learning that AI cannot replicate. This vision resonates with the THAI Framework's insistence on the complementarity of AI and human elements within a purposefully designed pedagogical ecology.

6.2 Towards Principled AI Integration

This article proposes four principles for the pedagogically principled integration of AI in language education, derived from the theoretical analysis above.

The principle of pedagogical intentionality holds that AI tools should be integrated in response to clearly articulated pedagogical goals, not simply because they are available. Every decision to introduce an AI tool into a language learning environment should be preceded by a clear account of the pedagogical purpose the tool is expected to serve and the evidence base for expecting it to serve that purpose effectively.

The principle of human primacy holds that teacher mediation and learner agency are not optional supplements to AI-assisted language learning but its essential conditions. AI tools

should be designed and deployed in ways that support and enhance human judgment, creativity, and relational engagement rather than replacing or marginalising them.

The principle of generative challenge holds that AI integration should be designed to scaffold rather than substitute for the cognitive and linguistic work through which language competence develops. This requires careful attention to the distinction between AI use that reduces unnecessary processing barriers and AI use that eliminates educationally necessary cognitive challenge.

The principle of critical transparency holds that learners, teachers, and institutions should have access to accurate, comprehensible information about how AI tools work, what assumptions they embody, and what data they collect and process. Critical transparency is not merely a technical requirement but a pedagogical one: learners who understand how AI tools work are better positioned to use them strategically, evaluate their outputs critically, and develop the metacognitive skills essential for lifelong language learning.

7. Future Research Directions

The theoretical analysis presented in this article points toward several priority areas for future empirical and theoretical research.

First, longitudinal studies of AI integration in language education are urgently needed. The existing literature is dominated by short-term experimental and quasi-experimental studies that measure immediate effects on performance or attitude. Understanding the effects of sustained AI integration on language development, learner autonomy, and the teacher-learner relationship requires research conducted over months and years rather than weeks.

Second, the THAI Framework requires empirical validation. The framework's core claims — about the centrality of teacher mediation, the conditions for strategic learner engagement, and the relationship between AI design and language development outcomes — are theoretically grounded but empirically underspecified. Future research should develop and test operationalisations of the framework's key constructs and investigate their predictive relationships.

Third, equity research in AI-mediated language education is a significant gap. Existing studies have focused predominantly on learner populations in well-resourced educational settings with access to high-quality AI tools and reliable internet connectivity. Research is needed that examines the differential effects of AI integration for learners from diverse linguistic, cultural, and socioeconomic backgrounds — and that takes seriously the risk that AI tools, uncritically adopted, may reinforce existing inequalities.

Fourth, the ethics of AI in language education requires sustained philosophical attention. The ethical challenges identified in this article — privacy, bias, transparency, academic integrity, overreliance, and teacher literacy — are interconnected, and addressing them requires not only empirical research but normative theoretical work. How should language educators conceptualise the boundaries of legitimate AI assistance? What does authentic authorship mean in AI-rich writing environments? What obligations do educational institutions have toward the learners whose data AI systems collect and process? These questions do not have straightforward empirical answers; they require the kind of careful conceptual analysis that philosophy of education and applied ethics can provide.

Fifth, there is a need for design-based research that develops, implements, and iteratively refines pedagogical models for AI integration in language education. The principles

proposed in Section 6 provide a starting point, but principled practice requires working through the particular design challenges of specific educational contexts, learner populations, and AI tools. Such research should involve close collaboration between language educators, AI designers, and learners — and should be conducted with the kind of attention to ethical dimensions that this article has argued is essential.

Finally, teacher education and professional development for AI-literate language teaching requires urgent research attention. How can pre-service and in-service teacher education develop not only technological proficiency but the critical AI literacy that effective and ethical AI integration demands? What institutional structures best support the development and ongoing renewal of teacher AI literacy? What does excellent AI-integrated language teaching look like in practice, and how can models of such teaching be developed, documented, and shared? These questions are at the frontier of the field, and their investigation will define the next generation of scholarship in computer-assisted and AI-mediated language education.

8. Conclusion

Artificial intelligence is reshaping language education in ways that are simultaneously profound and incompletely understood. The preceding analysis has sought to provide a theoretically grounded account of this reshaping — one that takes seriously both the pedagogical possibilities of AI and the ethical challenges it presents. The Triadic Human-AI Interaction Framework developed here offers a conceptual architecture for understanding the dynamics of AI integration: not as a bilateral encounter between learner and machine, but as a triadic relationship in which teacher, learner, and AI system are mutually constitutive agents in the service of language development.

The analysis has shown that AI's pedagogical potential in language education — whether in instructional design, personalisation, feedback, academic writing, or translation — is consistently mediated by the quality of teacher design and learner agency. AI that functions within a carefully designed, pedagogically intentional learning environment can extend the possibilities of language education in significant ways. AI that is adopted without pedagogical reflection risks creating new forms of dependence, inequity, and superficiality in learning.

The ethical challenges identified — privacy, bias, transparency, academic integrity, overreliance, and teacher literacy — are not peripheral to the pedagogical project but constitutive of it. A language education that deploys AI powerfully but unjustly, or efficiently but opaquely, is not a language education worth having. The integration of AI into language education must therefore be understood not merely as a technological or pedagogical challenge but as a moral one — a question of what kind of education, and what kind of society, we are committed to building.

This article has argued that answering this question requires theory adequate to the relational, ethical, and pedagogical complexity of AI-mediated language education. The THAI Framework is offered as a contribution to that theoretical project — as a starting point for the sustained scholarly conversation that this historically significant moment demands.

Declarations

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

Author Contributions: Conceptualization, investigation, writing – original draft, and writing – review & editing: Z.S.

How to cite:

Sadigzade, Z. (2026). Artificial intelligence in language education: Human-AI interaction, pedagogical design, and ethical challenges. *Global Spectrum of Research and Humanities, 3*(3), 5–22. https://doi.org/10.69760/gsrh.0260303001

References

An, X., Chai, C. S., Li, Y., Zhou, Y., & Yang, B. (2025). Modeling students' perceptions of artificial intelligence assisted language learning. *Computer Assisted Language Learning, 38*(5–6), 987–1008.

Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences, 12*(17), 8427.

Chen, C., & Gong, Y. (2025). The role of AI-assisted learning in academic writing: A mixed-methods study on Chinese as a second language students. *Education Sciences, 15*(2), 141.

Chocarro, R., Cortiñas, M., & Marcos-Matás, G. (2023). Teachers' attitudes towards chatbots in education: A technology acceptance model approach. *Educational Studies, 49*(2), 295–313.

Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the landscape: What is the role of machine translation in language learning? *@tic. Revista d'Innovació Educativa*, (10), 108–121.

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.

Feng, L. (2025). Investigating the effects of artificial intelligence-assisted language learning strategies on cognitive load and learning outcomes: A comparative study. *Journal of Educational Computing Research, 62*(8), 1741–1774.

García, I., & Peña, M. I. (2011). Machine translation-assisted language learning: Writing for beginners. *Computer Assisted Language Learning, 24*(5), 471–487.

Haristiani, N. (2019). Artificial intelligence (AI) chatbot as language learning medium: An inquiry. *Journal of Physics: Conference Series, 1387*(1), 012020.

He, D., Xia, Y., Qin, T., Wang, L., Yu, N., Liu, T. Y., & Ma, W. Y. (2016). Dual learning for machine translation. *Advances in Neural Information Processing Systems, 29*, 820–828.

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

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.

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.

Hwang, G. J., & Chang, C. Y. (2023). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments, 31*(7), 4099–4112.

Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education, 55*(1), 48–63.

Kannan, J., & Munday, P. (2018). New trends in second language learning and teaching through the lens of ICT, networked learning, and artificial intelligence. *Círculo de Lingüística Aplicada a la Comunicación, 76*, 13–30.

Kim, N. Y., Cha, Y., & Kim, H. S. (2019). Future English learning: Chatbots and artificial intelligence. *Multimedia-Assisted Language Learning, 22*(3), 32–53.

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies, 28*(1), 973–1018.

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education, 20*(1), 56.

Lee, S. M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning, 33*(3), 157–175.

Mageira, K., Pittou, D., Papasalouros, A., Kotis, K., Zangogianni, P., & Daradoumis, A. (2022). Educational AI chatbots for content and language integrated learning. *Applied Sciences, 12*(7), 3239.

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education, 57*(4), 632–645.

Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL, 21*(2), 241–258.

Ou, A. W., Stöhr, C., & Malmström, H. (2024). Academic communication with AI-powered language tools in higher education: From a post-humanist perspective. *System, 121*, 103225.

Schank, R. C., & Childers, P. (1983). *The cognitive computer: On language, learning, and artificial intelligence*. Addison-Wesley.

Schmidt, T., & Strasser, T. (2022). Artificial intelligence in foreign language learning and teaching: A CALL for intelligent practice. *Anglistik: International Journal of English Studies, 33*(1), 165–184.

Son, J. B., Ružić, N. K., & Philpott, A. (2025). Artificial intelligence technologies and applications for language learning and teaching. *Journal of China Computer-Assisted Language Learning, 5*(1), 94–112.

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*, 1260843.

Wang, Q., Saha, K., Gregori, E., Joyner, D., & Goel, A. (2021). Towards mutual theory of mind in human-AI interaction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). ACM.

Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education, 194*, 104703.

Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology, 14*, 1261955.

Zou, B., Guan, X., Shao, Y., & Chen, P. (2023). Supporting speaking practice by social network-based interaction in artificial intelligence (AI)-assisted language learning. *Sustainability, 15*(4), 2872.

Received: 9 June 2026
Accepted: 24 June 2026
Published: 02 July 2026