

Artificial Intelligence in Language Learning: A Systematic Literature Review of Pedagogical Integration, User Acceptance, Assessment Challenges

Rina Azlin Abdul Rahman^{1*}, Norhayati Idris², Zaiton Md Isa³, Nik Nur Nadhirah Zakaria⁴, Amir Lukman Abd Rahman⁵, Chan Jie Yan⁶, Neo Yee Feng⁷, Venny Karolina⁸

^{1,2,3,4,5,6,7}Akademi Pengajian Bahasa, Universiti Teknologi MARA, Shah Alam, Malaysia

⁸Fakultas Keguruan dan Ilmu Pendidikan (FKIP), Universitas Tanjungpura, Kota Pontianak, Kalimantan Barat, Indonesia

*Corresponding Author

DOI: <https://doi.org/10.47772/IJRISS.2026.102100125>

Received: 02 September 2026; Accepted: 07 September 2026; Published: 01 October 2026

ABSTRACT

Artificial intelligence (AI) has rapidly transformed language learning by introducing innovative tools and applications that support personalized learning, communication, assessment, and academic writing. Despite the growing body of research on AI-enhanced language education, studies remain fragmented across pedagogical, psychological, and assessment-related perspectives. Therefore, this study presents a systematic literature review (SLR) examining the current state of AI in language learning, focusing on pedagogical integration, user acceptance, and assessment challenges in higher education. Guided by the PRISMA framework, relevant studies published between 2021 and 2026 were identified from the Scopus and Web of Science databases. A total of 24 eligible journal articles were selected and analyzed. The findings revealed three major themes. First, AI pedagogical integration demonstrated significant potential to enhance language acquisition, writing development, speaking proficiency, learner engagement, and personalized learning through tools such as ChatGPT, conversational AI, chatbots, and speech recognition systems. Second, studies on AI acceptance highlighted the importance of perceived usefulness, perceived ease of use, attitudes, and motivation in influencing learners' and educators' willingness to adopt AI technologies. Third, AI assessment challenges underscored concerns related to academic integrity, assessment validity, ethical use, overreliance on AI-generated content, and the evolving role of educators in AI-mediated learning environments. The review concludes that while AI offers substantial opportunities for improving language learning and academic writing, its sustainable implementation requires careful pedagogical design, ethical governance, and assessment reform. The findings provide valuable implications for educators, researchers, policymakers, and higher education institutions seeking to integrate AI effectively into language education.

Keywords: Artificial Intelligence, Language Learning, Academic Writing, Technology Acceptance, Assessment Challenges.

INTRODUCTION

Artificial intelligence (AI) has emerged as one of the most influential technological developments in contemporary education, transforming the ways in which teaching, learning, and assessment are conducted across disciplines. Within language learning, AI-powered tools such as ChatGPT, Copilot, Gemini, conversational chatbots, and speech-recognition systems are increasingly being adopted to support language acquisition, writing development, communication skills, and personalized learning experiences. Recent studies have shown that AI technologies can enhance learners' motivation, engagement, language proficiency, and self-regulated learning by providing immediate feedback, adaptive learning opportunities, and interactive learning environments (Chen, 2022; Karataş et al., 2024). Furthermore, the integration of AI

into language education has expanded beyond classroom instruction to include academic writing support, automated feedback, assessment assistance, and collaborative learning environments. These developments suggest that AI has the potential to reshape language education practices in higher education.

Despite these promising developments, the rapid adoption of AI in language learning has generated several concerns and challenges. Although many studies report positive outcomes from AI integration, issues related to technology acceptance, overreliance on AI-generated content, academic integrity, ethical use, assessment validity, and the changing role of educators remain unresolved. Studies by Muthmainnah et al. (2026), Yusri et al. (2026), and Al-Obaydi and Pikhart (2026) highlight concerns regarding assessment authenticity, learners' dependence on AI tools, and the need for responsible implementation strategies. Additionally, learners and educators may differ in their perceptions, attitudes, and willingness to adopt AI technologies, influencing the effectiveness of AI-enhanced learning environments. These challenges indicate that the successful implementation of AI requires not only technological advancement but also pedagogical, motivational, and assessment-related considerations.

Although research on AI in language learning has grown substantially, existing studies tend to focus on isolated aspects such as pedagogical applications, user acceptance, or technological effectiveness. A comprehensive synthesis that simultaneously examines AI pedagogical integration, AI acceptance, attitudes, perceptions and motivation, and AI assessment challenges remains limited. Consequently, there is a need for a systematic literature review that consolidates current evidence, identifies emerging trends, and highlights research gaps across these interconnected dimensions. By reviewing and synthesizing recent studies, this paper aims to provide a comprehensive understanding of the current state of AI in language learning and to offer directions for future research, policy development, and educational practice in higher education.

Research questions are important to a systematic literature review (SLR). The questions provide the foundation and direction for the entire review process. A well-defined research question ensures that the literature search is exhaustive and systematic, covering all relevant studies that address key aspects of the topic (Kitchenham, 2007). The research questions in this SLR are formulated based on the PICO framework. This framework is a mnemonic style used to formulate research questions, particularly in qualitative research, proposed by Lockwood et al. (2015). PICO stands for Population, Interest, and Context. This SLR is done to answer the four research questions below;

- RQ1: How is artificial intelligence integrated into language learning and academic writing practices in Higher Education?
- RQ2: What are the attitudes, perceptions, acceptance factors, and motivational influences associated with the use of artificial intelligence among learners and educators in higher education language learning environments?
- RQ3: What assessment challenges have been reported in the implementation of artificial intelligence for language learning and academic writing in higher education?

LITERATURE REVIEW

THEORETICAL FRAMEWORK

Technological Pedagogical Content Knowledge (TPACK) Framework

The Technological Pedagogical Content Knowledge (TPACK) Framework developed by Mishra and Koehler (2006) provides a comprehensive theoretical foundation for understanding the effective integration of technology into teaching and learning. The framework proposes that successful technology-enhanced instruction occurs through the interaction of three core knowledge domains: Content Knowledge (CK), which refers to subject matter expertise; Pedagogical Knowledge (PK), which involves teaching methods and instructional strategies; and Technological Knowledge (TK), which concerns the effective use of digital

tools and technologies. According to TPACK, meaningful technology integration is achieved when educators are able to combine these domains to create learning experiences that support students' knowledge development and engagement. In the context of Artificial Intelligence (AI) pedagogical integration in language learning, the framework explains how educators can incorporate AI tools such as ChatGPT, Copilot, Gemini, conversational agents, chatbots, and speech-recognition systems into language instruction while maintaining pedagogical effectiveness and disciplinary relevance. AI technologies can facilitate personalized learning, instant feedback, language practice, writing support, and interactive communication activities. However, the effectiveness of these tools depends on teachers' ability to align technological affordances with pedagogical goals and language-learning content. The TPACK framework is therefore particularly relevant to studies examining AI-assisted language teaching because it highlights the critical role of educators in designing, implementing, and managing AI-enhanced learning environments. Consequently, TPACK provides a valuable lens for understanding how AI can be integrated into language learning practices to improve learner engagement, language proficiency, and overall educational outcomes.

Technology Acceptance Model (TAM)

The Technology Acceptance Model (TAM) proposed by Davis (1989) is one of the most influential frameworks for explaining users' acceptance and adoption of new technologies. TAM posits that an individual's intention to use a technology is primarily determined by two key factors: Perceived Usefulness (PU) and Perceived Ease of Use (PEOU). Perceived usefulness refers to the extent to which users believe that a technology can enhance their performance, productivity, or learning outcomes, while perceived ease of use refers to the degree to which users perceive the technology as easy to understand and operate. According to the model, these perceptions influence users' attitudes toward technology, which subsequently shape their behavioral intentions and actual usage. In the context of artificial intelligence (AI) in language learning and academic writing, TAM provides a valuable theoretical lens for understanding students' and educators' acceptance, attitudes, perceptions, and motivation toward AI tools such as ChatGPT, Copilot, Gemini, chatbots, and automated writing assistants. Users are more likely to adopt AI technologies when they perceive them as beneficial for improving writing quality, language proficiency, productivity, and learning efficiency. Similarly, ease of interaction with AI systems can enhance positive attitudes and increase motivation to engage with these tools. Numerous studies have shown that positive perceptions of AI usefulness and usability contribute significantly to learners' willingness to integrate AI into their academic activities. Therefore, TAM effectively explains why some users readily embrace AI technologies while others remain cautious, making it an appropriate framework for examining AI acceptance, attitudes, perceptions, and motivational factors in higher education language learning and academic writing contexts.

Assessment for Learning (Afl)

The Assessment for Learning (Afl) theory developed by Black and Wiliam (1998) emphasizes the role of assessment as a process that supports and enhances learning rather than merely measuring achievement. According to Afl, effective assessment practices should provide continuous and meaningful feedback that enables learners to identify their strengths and weaknesses, monitor their progress, and take appropriate actions to improve their performance. The theory highlights key elements such as formative feedback, self-assessment, reflection, and learner engagement in the assessment process. In the context of AI-assisted academic writing, Afl provides a valuable framework for understanding both the opportunities and challenges associated with AI-driven assessment tools. Technologies such as ChatGPT, Copilot, Gemini, Grammarly, and automated writing evaluation systems can provide immediate feedback on grammar, vocabulary, organization, coherence, and writing style, thereby supporting students' revision and learning processes. However, the increasing use of AI in assessment also raises significant challenges related to academic integrity, authenticity of student work, assessment validity, and overreliance on automated feedback. These concerns highlight the need for educators to balance technological efficiency with pedagogical objectives. From an Afl perspective, future directions should focus on developing AI-supported assessment practices that promote critical thinking, learner autonomy, ethical use of technology, and meaningful feedback while maintaining the integrity and quality of academic writing assessment in higher education.

CONCEPTUAL FRAMEWORK

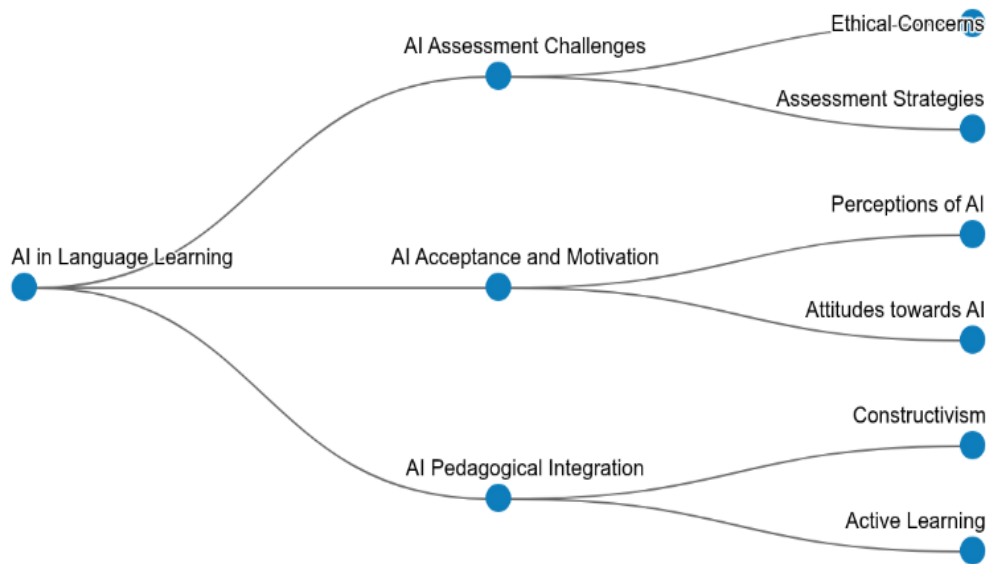


FIGURE 1. Conceptual Framework

The conceptual framework of this study is grounded in three interrelated dimensions of artificial intelligence (AI) in language learning and academic writing, namely AI pedagogical integration, AI acceptance and motivation, and AI assessment challenges and future directions, as illustrated in the proposed model. The framework positions AI in language learning as the central concept that influences these three dimensions. The first dimension, AI pedagogical integration, is anchored in the Technological Pedagogical Content Knowledge (TPACK) Framework developed by Mishra and Koehler (2006). TPACK posits that effective technology integration occurs when educators successfully combine technological knowledge, pedagogical knowledge, and content knowledge. In language learning contexts, AI tools such as ChatGPT, Copilot, Gemini, chatbots, and speech-recognition systems can support active learning, personalized instruction, language practice, and knowledge construction. Consequently, the framework suggests that successful AI implementation depends on educators' ability to align AI technologies with pedagogical strategies and language-learning objectives.

The second dimension, AI acceptance and motivation, is supported by the Technology Acceptance Model (TAM) proposed by Davis (1989). TAM explains that users' perceptions of usefulness and ease of use influence their attitudes, intentions, and motivation to adopt technology. In AI-enhanced language learning, positive perceptions of AI tools can foster greater acceptance, favorable attitudes, and increased motivation among students and educators. These factors subsequently influence the extent to which AI technologies are integrated into learning environments. The framework also acknowledges that perceptions and attitudes toward AI are critical determinants of sustained engagement with AI-assisted learning tools.

The third dimension, AI assessment challenges and future directions, is anchored in the Assessment for Learning (AfL) theory developed by Black and Wiliam (1998). AfL emphasizes the importance of formative assessment and continuous feedback in supporting learning improvement. While AI-assisted assessment offers opportunities for immediate feedback, writing support, and self-regulated learning, it also raises concerns related to academic integrity, assessment validity, ethical use, and overreliance on technology. Therefore, the framework proposes that future developments in AI-assisted academic writing should focus on balancing technological innovation with sound assessment practices, ethical considerations, and learner development. Collectively, these three theories provide a comprehensive lens for understanding how AI is integrated into language learning, how users respond to AI adoption, and how emerging assessment challenges can shape future educational practices.

METHODOLOGY

Material and Methods

For conducting systematic literature reviews, the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach is a widely accepted standard that guarantees transparency, completeness, and consistency throughout the procedure by Page et al. (2021) guide how to systematically identify, screen, and include studies in their review. In the context of this SLR, the writers involved in this article act as researchers and work together to analyse the articles deemed relevant for the title. Two important databases, Web of Science and Scopus, were used in this analysis because of their wide coverage and robustness.

The PRISMA approach is organized into four key stages: (i) identification, (ii) screening, (iii) eligibility, and (iv) data abstraction. In the identification phase, databases are searched to locate all relevant studies. The screening phase then involves evaluating these studies against predefined criteria to eliminate irrelevant or low-quality research. During the eligibility phase, the remaining studies are thoroughly assessed to confirm they meet the inclusion criteria. Finally, data abstraction focuses on extracting and synthesizing data from the included studies, which is essential for deriving meaningful and reliable conclusions. This structured method ensures that the systematic review is conducted with rigor, leading to trustworthy results that can guide future research and practice. The PRISMA framework lists three main stages for analysis: (i) identification, (ii) screening, (iii) eligibility, (iv) data abstraction and analysis, as well as (v) quality of appraisal.

Identification

In accordance with the PRISMA framework, the identification phase represents a critical starting point in the systematic literature review process, as it is designed to ensure exhaustive coverage of relevant studies while reducing the risk of selection bias. In this study, a structured and comprehensive search was conducted across two leading bibliographic databases, Scopus and Web of Science (WoS), using the keywords “AI,” “foreign language learning,” and “university students.” Table 1 below shows the search string.

These databases were deliberately selected because of their extensive multidisciplinary coverage, stringent indexing criteria, and strong representation of high-quality, peer-reviewed journals. The search process resulted in the retrieval of 592 records from Scopus and 100 records from Web of Science, yielding a combined total of 692 records before the removal of duplicates.

The higher number of results obtained from Scopus can be attributed to its broader inclusion of journals and stronger coverage in the social sciences, whereas WoS contributed a more targeted collection of highly cited and methodologically rigorous studies. As shown in Table 1, the identification of 692 records reflects the growing academic interest in writing within higher education and highlights the conceptual diversity of research spanning linguistic, pedagogical, and psychological perspectives. Employing multiple databases enhances the methodological robustness of the review by minimizing database-specific bias and increasing the likelihood of capturing diverse research traditions and geographical contexts.

Additionally, the use of carefully designed search terms ensured sensitivity to terminological variation while maintaining alignment with the review's objectives. Overall, this transparent and systematic identification process establishes a strong foundation for subsequent screening and eligibility stages, consistent with PRISMA standards and expectations for rigorous, replicable research.

Table 1 : The search string.

Scopus	TITLE-ABS-KEY (("learning foreign languages" OR "foreign language learning" OR "language acquisition" OR "second language acquisition" OR "multilingual education") AND (students OR "university students" OR "college students" OR "higher education") AND (AI OR "artificial intelligence" OR "machine learning" OR "deep learning"))
Web of Science	TITLE-ABS-KEY (("learning foreign languages" OR "foreign language learning" OR "language acquisition" OR "second language acquisition" OR "multilingual education") AND (students OR "university students" OR "college students" OR "higher education") AND (AI OR "artificial intelligence" OR "machine learning" OR "deep learning")) Extracted July 2026

Screening

During the screening stage of the systematic literature review, potentially relevant studies were evaluated to ensure their alignment with the four established research questions. This phase focused on refining the initial pool of records by removing duplicates and applying predefined inclusion and exclusion criteria. Following the initial review, 597 publications were excluded for failing to meet the scope and relevance requirements of the study, leaving 95 (Scopus=78; Wos= 17) articles for further consideration, as outlined in Figure 2. The primary inclusion criterion was the type of literature, with priority given to scholarly sources that offer substantive and empirical insights. As a result, materials such as book reviews, book series, meta-syntheses, meta-analyses, conference proceedings, and book chapters not directly related to the core focus of the review were excluded. To ensure consistency and relevance to contemporary academic discourse, only peer-reviewed publications written in English and published between 2021 and 2026 were retained. Additionally, a duplication check led to the removal of 2 repeated records across the databases. This systematic and criteria-driven screening process enhanced the focus and quality of the evidence base, ensuring that only the most relevant and methodologically sound studies progressed to the subsequent stages of analysis.

Eligibility

For the third step, known as eligibility, a total of 27 articles have been prepared. All articles' titles and key content were thoroughly reviewed at this stage to ensure that the inclusion requirements were fulfilled and fit into the present study with the current research aims. Therefore, three reports were omitted because they were not pure science articles based on empirical evidence. Finally, 24 articles are available for review (see Table 3).

Table 2 :The selection criterion is searching

Criterion	Inclusion	Exclusion
Language	English	Non-English
Time line	2021 – 2026	< 2021
Literature type	Journal (Article)	Conference, Book, Review
Publication Stage	Final	In Press
Subject Area	AI	Besides AI

Data Abstraction and Analysis

An integrative analysis was used as one of the assessment strategies in this study to examine and synthesize a variety of research designs (quantitative, qualitative, and mixed methods). To develop themes, the authors carefully analyzed 24 publications for relevant material related to the study's topics. They examined the methodology and research results of significant studies on artificial intelligence in language learning. The authors collaborated to develop themes based on the evidence presented throughout the study. They kept a log of their data analysis process to document relevant perspectives, analyses, or ideas. The authors ensured consistency in theme design by comparing the results and discussing disagreements between themselves. This expert review phase established the validity of each subtheme by ensuring clarity, importance, and suitability (Refer to Appendix 1).

Quality of Appraisal

According to Kitchenham and Charters (Kitchenham, 2007), the researchers need to assess the quality of the research chosen. The next stage is applying the quality assessment suggested by Abouzahra et al., (2020), which consists of six questions complying to a scoring procedure for evaluating each criterion. The possible ratings are "Yes" (Y) with a score of 1 if the criterion is fully met, "Partly" (P) with a score of 0.5 if the criterion is somewhat met but contains some gaps or shortcomings, and "No" (N) with a score of 0 if the criterion is not met at all.

Table 3- Possible ratings for 6 questions by Abouzahra et al. (2020)

No	Question
QA1	Is the purpose of the study clearly stated?
QA2	Is the interest and the usefulness of the work clearly presented?
QA3	Is the study methodology clearly established?
QA4	Are the concepts of the approach clearly defined?
QA5	Is the work compared and measured with other similar work?

Each writer independently assesses the study according to these criteria, and the scores are then totalled across all experts to determine the overall mark. For a study to be accepted for the next process, the total mark, derived from summing the scores from all three experts, must exceed 3.0. This threshold ensures that only studies meeting a certain quality standard proceed further. The results are shown in Table 4.

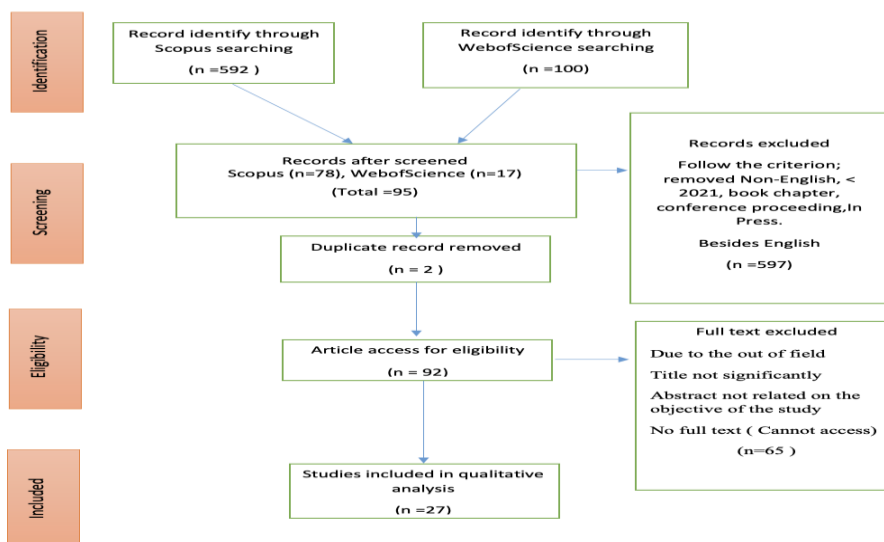


FIGURE 2. Flow diagram of the proposed searching study (Moher D, Liberati A, Tetzlaff J, 2009)

FINDINGS AND DISCUSSION

Quality of Assessment

Table 4 shows the results of the assessment performance for selected primary studies. The overall quality of the articles is as follows;

High quality ($\geq 80\%$): Majority of studies

Moderate quality (60–79%): Conceptual or descriptive studies

Lower quality ($<60\%$): Primarily survey-only or exploratory abstracts

Table 4- Quality Assessment for the Selected 24 articles

Study	Author(s) & Year	Article Title	QA1	QA2	QA3	QA4	QA5	QA6	Total Mark	%
PS1	Muthmainnah et al. (2026)	Human-robot interaction using ChatGPT for technology-driven language learning: Contextual insights from the higher education institution of Indonesia	Y	Y	Y	Y	P	Y	5.5	91.7
PS2	Ji et al. (2024)	Teaching Foreign Language with Conversational AI: Teacher-Student-AI Interaction	Y	Y	Y	Y	P	P	5.0	83.3
PS3	Nikonova et al. (2023)	Using Artificial Intelligence Tools in Teaching a Foreign Language in Higher Technical Institutions	Y	Y	Y	Y	P	P	5.0	83.3
PS4	An et al. (2025)	Modeling Students' Perceptions of Artificial Intelligence Assisted Language Learning	Y	Y	Y	Y	Y	P	5.5	91.7
PS5	Kiński et al. (2025)	Big Five Personality Traits, Attitudes Towards Artificial Intelligence and the Use of AI Solutions in Foreign Language Learners	Y	Y	Y	Y	P	P	5.0	83.3
PS6	Polakova & Klimova (2024)	Implementation of AI-Driven Technology into Education: A Pilot Study on the Use of Chatbots in Foreign Language Learning	Y	Y	Y	Y	P	P	5.0	83.3
PS7	Alsaedi (2025)	Exploring ChatGPT's Role in EFL Learning Through the Technology Acceptance Model:	Y	Y	Y	Y	Y	Y	6.0	100.0

		Perspectives from Saudi Students								
PS8	Neshkovska et al. (2025)	ChatGPT in Foreign Language Acquisition: A Close Ally or a Distant Concept for Macedonian University Students?	Y	Y	Y	Y	P	P	5.0	83.3
PS9	Bekdaş (2025)	The Pros and Cons of ChatGPT in Foreign Language Teaching and Its Impact on Student Motivation	Y	Y	Y	Y	Y	Y	6.0	100.0
PS10	Qin et al. (2026)	The Mediating Role of Motivation Between Generative Artificial Intelligence and Foreign Language Proficiency: A Technology Acceptance Model Perspective	Y	Y	Y	Y	Y	P	5.5	91.7
PS11	Zhang et al. (2024)	The Effect of Student Acceptance on Learning Outcomes: AI-Generated Short Videos Versus Paper Materials	Y	Y	Y	Y	P	P	5.0	83.3
PS12	Wang & Wang (2025)	College Students' Behavioural Intentions of AI-Assisted Language Learning: Based on the Technology Acceptance Model	Y	Y	Y	Y	P	Y	5.5	91.7
PS13	Dou & Sun (2025)	Artificial Intelligence Tools: Improvement of Motivation, Psychological Well-Being, and Psychological Capital of EFL Learners	Y	Y	Y	Y	P	P	5.0	83.3
PS14	Yusri et al. (2026)	Knowledge Resilience in AI-Mediated Foreign Language Learning: Challenges and Responses of Indonesian Lecturers	Y	Y	Y	Y	P	Y	5.5	91.7
PS15	Namaziandost (2025)	Integrating Flipped Learning in AI-Enhanced Language Learning: Mapping the Effects on Metacognitive Awareness, Writing Development, and Foreign Language Learning Boredom	Y	Y	Y	Y	P	P	5.0	83.3
PS16	Dolenc & Brumen	Exploring Social and Computer Science	Y	Y	Y	Y	P	Y	5.5	91.7

	(2024)	Students' Perceptions of AI Integration in (Foreign) Language Instruction								
PS17	Hasan & Mehnaz (2026)	Artificial Intelligence in Personalized Language Learning: Benefits, Attitudes, and Motivation of Bangladeshi Higher Education EFL Students	Y	Y	Y	Y	P	P	5.0	83.3
PS18	John (2025)	Exploring the Impact of Artificial Intelligence on Language Acquisition, Linguistic Development, and Language Use: A Case Study from India	Y	Y	Y	Y	P	P	5.0	83.3
PS19	Al-Obaydi & Pikhart (2026)	The Use of Copilot, Gemini and ChatGPT in the Context of Foreign Language Learning and Teaching: An Academic Technology Review	Y	Y	Y	Y	P	P	5.0	83.3
PS20	Zhou (2026)	Investigating L2 Learners' Task Motivation in Mandarin Chinese Essay Writing: The Role of ChatGPT in AI-Mediated Learning	Y	Y	Y	Y	P	P	5.0	83.3
PS21	Chen (2022)	Speech-to-Text Recognition in University English as a Foreign Language Learning	Y	Y	Y	Y	P	P	5.0	83.3
PS22	Karataş et al. (2024)	Incorporating AI in Foreign Language Education: An Investigation into ChatGPT's Effect on Foreign Language Learners	Y	Y	Y	Y	P	P	5.0	83.3
PS23	Çobanogullari & Özbek (2025)	AI-Powered Language Learning: Developing the ChatGPT Usage Scale for Foreign Language Learners	Y	Y	P	Y	P	P	4.5	75.0
PS24	Huang et al. (2023)	Trends, Research Issues and Applications of Artificial Intelligence in Language Education	Y	Y	P	Y	P	P	4.5	75.0

Table 4 presents the quality assessment results of the 24 studies selected for the systematic literature review on artificial intelligence in language learning and higher education. Overall, the findings indicate that the reviewed studies demonstrate good to excellent quality, with an average score of 5.17 out of 6.00 (86.1%). Eight studies achieved an excellent rating of 90% and above, including Alsaedi (2025), Bekdaş (2025), and Muthmainnah et al. (2026), reflecting strong methodological rigor, clearly stated purposes, well-defined concepts, and explicit discussions of limitations. Fourteen studies were rated as good, with scores ranging from 80% to 89%, indicating that most criteria were adequately addressed, although aspects such as comparison with previous studies (QA5) and discussion of limitations (QA6) were sometimes only partially reported. Only two studies, namely Çobanogullari and Özbek (2025) and Huang et al. (2023), received an acceptable rating of 75%, mainly due to limited methodological details and partial reporting of comparative and limitation-related aspects in their abstracts. Across the dataset, most studies clearly articulated their research objectives, significance, methodology, and conceptual frameworks, fulfilling QA1 to QA4. However, QA5 and QA6 received lower scores more frequently, suggesting that comparisons with related studies and explicit acknowledgment of limitations remain areas requiring further attention. Collectively, the assessment confirms that the selected studies provide a robust and credible evidence base for examining AI integration, assessment practices, and stakeholder perspectives in language education and academic writing.

Themes Table 5: Analysis for Themes

Theme	Number of Studies
Theme 1: AI Pedagogical Integration in Language Learning	9
Theme 2: AI Acceptance, Attitudes, Perceptions and Motivation	13
Theme 3: AI Assessment Challenges	7

Table 5 above presents the distribution of the reviewed studies across three major themes identified in the systematic literature review. The largest category is Theme 2: AI Acceptance, Attitudes, Perceptions and Motivation, which includes 13 studies, indicating strong scholarly interest in understanding learners' and educators' perceptions, technology acceptance, behavioral intentions, and motivational factors related to AI adoption. This is followed by Theme 1: AI Pedagogical Integration in Language Learning with 9 studies, highlighting the growing emphasis on integrating AI tools such as ChatGPT, chatbots, and other intelligent technologies into language teaching and learning practices. Finally, Theme 3: AI Assessment, Writing Development, Ethical Issues and Future Directions comprises 7 studies, reflecting increasing attention to AI-assisted assessment, writing enhancement, academic integrity, ethical concerns, and emerging challenges in educational contexts. Overall, the distribution suggests that current research primarily focuses on users' acceptance and motivation, while pedagogical implementation and assessment-related issues remain important and evolving areas of inquiry.

Theme 1 (T1): AI Pedagogical Integration in Language Learning

Table 6 : AI Pedagogical Integration in Language Learning

Study	Author(s) & Year	Article Title
PS1	Muthmainnah et al. (2026)	Human-robot interaction using ChatGPT for technology-driven language learning: Contextual insights from the higher education institution of Indonesia
PS2	Ji et al. (2024)	Teaching Foreign Language with Conversational AI: Teacher-Student-AI Interaction
PS3	Nikonova et al. (2023)	Using Artificial Intelligence Tools in Teaching a Foreign Language in Higher Technical Institutions
PS6	Polakova and	Implementation of AI-Driven Technology into Education: A Pilot

	Klimova (2024)	Study on the Use of Chatbots in Foreign Language Learning
PS15	Namaziandost (2025)	Integrating Flipped Learning in AI-Enhanced Language Learning: Mapping the Effects on Metacognitive Awareness, Writing Development, and Foreign Language Learning Boredom
PS18	John (2025)	Exploring the Impact of Artificial Intelligence on Language Acquisition, Linguistic Development, and Language Use: A Case Study from India
PS19	Al-Obaydi and Pikhart (2026)	The Use of Copilot, Gemini and ChatGPT in the Context of Foreign Language Learning and Teaching: An Academic Technology Review
PS21	Chen (2022)	Speech-to-Text Recognition in University English as a Foreign Language Learning
PS22	Karataş et al. (2024)	Incorporating AI in Foreign Language Education: An Investigation into ChatGPT's Effect on Foreign Language Learners

The studies presented in Table 6 above focus on the pedagogical integration of artificial intelligence (AI) in foreign language learning and teaching contexts. Collectively, these studies demonstrate the growing adoption of AI-powered technologies such as ChatGPT, conversational AI, chatbots, speech-to-text recognition, and AI-enhanced instructional models to support language acquisition, communication skills, and learner engagement. Several studies reported positive outcomes, including improvements in writing development, speaking proficiency, vocabulary acquisition, metacognitive awareness, motivation, and willingness to communicate (Chen, 2022; Namaziandost, 2025).

Additionally, researchers emphasized the evolving role of teachers in facilitating human-AI interactions and integrating AI tools into classroom practices (Ji et al., 2024; Muthmainnah et al., 2026). While the findings highlight the pedagogical potential of AI to enhance language learning experiences, concerns regarding overreliance on AI, ethical use, and the continued importance of human guidance were also noted (Al-Obaydi & Pikhart, 2026; Muthmainnah et al., 2026). Overall, the studies indicate that AI serves as a valuable educational tool that can complement, rather than replace, traditional language teaching practices in higher education.

Theme 2 (T2): AI Acceptance, Attitudes, Perceptions and Motivation

Table 7: AI Acceptance, Attitudes, Perceptions and Motivation

Study	Author(s) & Year	Article Title
PS4	An et al. (2025)	Modeling Students' Perceptions of Artificial Intelligence Assisted Language Learning
PS5	Kiński et al. (2025)	Big Five Personality Traits, Attitudes Towards Artificial Intelligence and the Use of AI Solutions in Foreign Language Learners
PS7	Alsaedi (2025)	Exploring ChatGPT's Role in EFL Learning Through the Technology Acceptance Model: Perspectives from Saudi Students
PS8	Neshkovska et al. (2025)	ChatGPT in Foreign Language Acquisition: A Close Ally or a Distant Concept for Macedonian University Students?
PS9	Bekdaş (2025)	The Pros and Cons of ChatGPT in Foreign Language Teaching and Its Impact on Student Motivation
PS10	Qin et al. (2026)	The Mediating Role of Motivation Between Generative Artificial Intelligence and Foreign Language Proficiency: A Technology

		Acceptance Model Perspective
PS11	Zhang et al. (2024)	The Effect of Student Acceptance on Learning Outcomes: AI-Generated Short Videos Versus Paper Materials
PS12	Wang and Wang (2025)	College Students' Behavioural Intentions of AI-Assisted Language Learning: Based on the Technology Acceptance Model
PS13	Dou and Sun (2025)	Artificial Intelligence Tools: Improvement of Motivation, Psychological Well-Being, and Psychological Capital of EFL Learners
PS16	Dolenc and Brumen (2024)	Exploring Social and Computer Science Students' Perceptions of AI Integration in (Foreign) Language Instruction
PS17	Hasan and Mehnaz (2026)	Artificial Intelligence in Personalized Language Learning: Benefits, Attitudes, and Motivation of Bangladeshi Higher Education EFL Students
PS20	Zhou (2026)	Investigating L2 Learners' Task Motivation in Mandarin Chinese Essay Writing: The Role of ChatGPT in AI-Mediated Learning
PS14	Yusri et al. (2026)	Knowledge Resilience in AI-Mediated Foreign Language Learning: Challenges and Responses of Indonesian Lecturers

The studies (Table 7) classified under Theme 2: AI Acceptance, Attitudes, Perceptions and Motivation primarily examine how learners and educators perceive, adopt, and respond to artificial intelligence in language learning contexts. Several studies employed theoretical frameworks such as the Technology Acceptance Model (TAM) and related motivational theories to investigate factors influencing AI adoption and behavioral intentions (Alsaedi, 2025; Qin et al., 2026; Wang & Wang, 2025). The findings consistently indicate that perceived usefulness, ease of use, and positive attitudes significantly enhance learners' willingness to engage with AI-supported language learning.

Other studies highlighted the role of personality traits, learner motivation, psychological well-being, and educational background in shaping AI acceptance and usage patterns (Kiński et al., 2025; Dou & Sun, 2025; Hasan & Mehnaz, 2026). Furthermore, students generally viewed AI tools such as ChatGPT positively for supporting language learning, although concerns regarding overreliance, academic integrity, and the replacement of human instruction were also reported (Bekdaş, 2025; Neshkovska et al., 2025). Collectively, these studies suggest that successful AI integration depends not only on technological capabilities but also on users' perceptions, motivation, confidence, and readiness to embrace AI-enhanced learning environments.

Theme 3 (T3): AI Assessment Challenges

Table 8 :AI Assessment Challenges

Study	Author(s) & Year	Article Title
PS1	Muthmainnah et al. (2026)	Human-robot interaction using ChatGPT for technology driven language learning: Contextual insights from the higher education institution of Indonesia
PS14	Yusri et al. (2026)	Knowledge Resilience in AI-Mediated Foreign Language Learning: Challenges and Responses of Indonesian Lecturers
PS15	Namaziandost (2025)	Integrating Flipped Learning in AI-Enhanced Language Learning: Mapping the Effects on Metacognitive Awareness, Writing Development, and Foreign Language Learning Boredom

PS19	Al-Obaydi and Pikhart (2026)	The Use of Copilot, Gemini and ChatGPT in the Context of Foreign Language Learning and Teaching: An Academic Technology Review
PS23	Çobanogullari and Özbek (2025)	AI-Powered Language Learning: Developing the ChatGPT Usage Scale for Foreign Language Learners
PS24	Huang et al. (2023)	Trends, Research Issues and Applications of Artificial Intelligence in Language Education
PS16	Dolenc and Brumen (2024)	Exploring Social and Computer Science Students' Perceptions of AI Integration in (Foreign) Language Instruction

The studies grouped (Table 8) under Theme 3: AI Assessment, Challenges, examine the broader implications of artificial intelligence in language education beyond classroom implementation and learner acceptance. Several studies highlighted the use of AI tools such as ChatGPT, Copilot, and Gemini for supporting writing development, language assessment, and metacognitive learning (Muthmainnah et al., 2026; Namaziandost, 2025; Al-Obaydi & Pikhart, 2026). At the same time, concerns regarding assessment validity, academic integrity, ethical use, overreliance on AI, and the changing role of educators were frequently reported (Muthmainnah et al., 2026; Yusri et al., 2026).

Other studies focused on evaluating AI usage through measurement scales and reviewing emerging trends and applications of AI in language education (Çobanogullari & Özbek, 2025; Huang et al., 2023). Furthermore, research on students' perceptions of AI integration emphasized the need for responsible implementation, training, and policy development to maximize educational benefits while addressing potential risks (Dolenc & Brumen, 2024). Collectively, these studies suggest that while AI offers significant opportunities for enhancing assessment and writing development, careful consideration of ethical and pedagogical challenges is essential for its sustainable integration into higher education.

CONCLUSION

Summary of Findings

This systematic literature review examined the current state of artificial intelligence (AI) in language learning through three major themes: AI pedagogical integration in language learning, AI acceptance, attitudes, perceptions and motivation, and AI assessment challenges. The findings reveal that AI technologies such as ChatGPT, Copilot, Gemini, conversational AI, chatbots, and speech-recognition systems are increasingly integrated into language learning environments to support personalized learning, writing development, speaking proficiency, vocabulary acquisition, and learner engagement.

Studies by Chen (2022), Karataş et al. (2024), and Namaziandost (2025) consistently demonstrated the positive impact of AI on language acquisition and learning outcomes. Furthermore, AI was frequently positioned as a complementary learning tool that enhances, rather than replaces, the role of teachers in facilitating meaningful learning experiences.

The review also found that learners' and educators' acceptance of AI is strongly influenced by perceived usefulness, perceived ease of use, attitudes, and motivational factors. Studies grounded in the Technology Acceptance Model (e.g., Alsaedi, 2025; Qin et al., 2026; Wang & Wang, 2025) reported generally positive attitudes toward AI adoption. However, concerns regarding overreliance, academic dishonesty, assessment validity, and ethical implications remain significant challenges. Studies by Muthmainnah et al. (2026), Yusri et al. (2026), and Al-Obaydi and Pikhart (2026) highlighted the need for responsible AI implementation and the development of robust assessment practices that maintain academic integrity while leveraging the benefits of AI-assisted learning.

Implications and Suggestions for Future Research

The findings of this review have important theoretical, pedagogical, and practical implications. From a theoretical perspective, the integration of TPACK, TAM, and Assessment for Learning (AfL) provides a comprehensive framework for understanding AI implementation, acceptance, and assessment practices in language learning. Practically, educators and institutions should develop AI literacy programmes, ethical guidelines, and professional development initiatives to ensure effective and responsible use of AI technologies. Curriculum designers should also incorporate AI-supported learning activities that promote critical thinking, learner autonomy, and meaningful engagement rather than dependence on automated systems.

Future research should extend beyond learners' perceptions and technology adoption by investigating the long-term effects of AI on language proficiency, writing quality, critical thinking, and self-regulated learning. More studies involving educators, assessment specialists, and policymakers are needed to examine the evolving role of AI in academic writing and assessment. Additionally, future research should explore emerging AI technologies, cross-cultural implementation, and ethical frameworks that can support sustainable and equitable AI integration in higher education language learning environments. Such efforts will contribute to a deeper understanding of how AI can be effectively harnessed to enhance teaching, learning, and assessment in the rapidly evolving educational landscape.

ACKNOWLEDGEMENTS

The authors acknowledge the support from the Akademik Pengajian Bahasa, Universiti Teknologi MARA, Shah Alam, which made this study possible. Additionally, we would like to express my sincere gratitude to Prof. Dr. Noor Hanim Rahmat for her invaluable guidance throughout this study. Her insightful feedback and expertise were instrumental in shaping this research.

REFERENCES

1. Bekdaş, M. (2025). The pros and cons of ChatGPT in foreign language teaching and its impact on student motivation. *Arab World English Journal, AI Special Issue*, 139–156. <https://doi.org/10.24093/awej/AI.8>
2. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
3. Buivydienė, V., Rutkienė, L., & Žemienė, A. (2026). The social inclusion of international students through language and their motivation to learn it: The case of VILNIUS TECH. *Sustainability*, 18(1), 1–17. <https://doi.org/10.3390/su18010434>
4. Chen, K. T. C. (2022). Speech-to-text recognition in university English as a foreign language learning. *Education and Information Technologies*, 27(7), 9857–9875. <https://doi.org/10.1007/s10639-022-11016-5>
5. Çobanogullari, F., & Özbek, Ö. (2025). AI-powered language learning: Developing the ChatGPT usage scale for foreign language learners. *Education and Information Technologies*, 30(9), 12517–12534. <https://doi.org/10.1007/s10639-025-13342-w>
6. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
7. Dolenc, K., & Brumen, M. (2024). Exploring social and computer science students' perceptions of AI integration in (foreign) language instruction. *Computers and Education: Artificial Intelligence*, 7, 1–13. <https://doi.org/10.1016/j.caeai.2024.100285>
8. Dou, W., Sun, X., Li, W., Li, H., & Li, J. (2025). Artificial intelligence tools: Improvement of motivation, psychological well-being, and psychological capital of EFL learners: A self-determination theory perspective. *Learning and Motivation*, 92. <https://doi.org/10.1016/j.lmot.2025.102169>

9. Hasan, M. M., & Mehnaz, M. (2026). Artificial intelligence in personalized language learning: Benefits, attitudes, and motivation of Bangladeshi higher education EFL students. *LLT Journal: Journal on Language and Language Teaching*, 29(1), 318–334. <https://doi.org/10.24071/llt.v29i1.373>
10. Huang, X. Y., Zou, D., Cheng, G., Chen, X. L., & Xie, H. R. (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](https://doi.org/10.30191/ETS.202301_26(1).0009)
11. Ji, H., Han, I., & Park, S. (2024). Teaching foreign language with conversational AI: Teacher-student-AI interaction. *Language Learning & Technology*, 28(2), 91–108. <https://doi.org/10.64152/10125/73573>
12. John, A. (2025). Exploring the impact of artificial intelligence on language acquisition, linguistic development, and language use: A case study from India. *Forum for Linguistic Studies*, 7(3), 1104–1117. <https://doi.org/10.30564/fls.v7i3.8671>
13. Karatas, F., Abedi, F. Y., Gunyel, F. O., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, 29(15), 19343–19366. <https://doi.org/10.1007/s10639-024-12574-6>
14. Kiński, C., Kiński, B., & Przyborowska, A. (2025). Big Five personality traits, attitudes towards artificial intelligence and the use of AI solutions in foreign language learners. *Neofilolog*, 65(1), 65–83. <https://doi.org/10.14746/n.2025.65.1.4>
15. Kitchenham, B. (2007). Guidelines for performing systematic literature reviews in software engineering (Version 2.3 EBSE Technical Report). EBSE.
16. Lockwood, C., Munn, Z., & Porritt, K. (2015). Qualitative research synthesis: Methodological guidance for systematic reviewers utilizing meta-aggregation. *International Journal of Evidence-Based Healthcare*, 13(3), 179–187. <https://doi.org/10.1097/XEB.0000000000000062>
17. Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
18. Muthmainnah, M., Asad, M. M., Yakin, A. A., & Almusharraf, N. M. (2026). Human-robot interaction using ChatGPT for technology-driven language learning: Contextual insights from the higher education institution of Indonesia. *Asian Education and Development Studies*, 15(3), 637–657. <https://doi.org/10.1108/AEDS-01-2025-0038>
19. Namaziandost, E. (2025). Integrating flipped learning in AI-enhanced language learning: Mapping the effects on metacognitive awareness, writing development, and foreign language learning boredom. *Computers and Education: Artificial Intelligence*, 9, 1–11. <https://doi.org/10.1016/j.caeai.2025.100446>
20. Neshkovska, S., Ivanovska, L., Trajkovska, V., & Kasaposka-Chadlovskaa, M. (2025). ChatGPT in foreign language acquisition: A close ally or a distant concept for Macedonian university students? *Respectus Philologicus*, 47(52), 50–65. <https://doi.org/10.15388/RESPECTUS.2025.47.4>
21. Nikonova, E., Yakhyayeva, K., Pivkina, N., & Schetinina, A. (2023). Using artificial intelligence tools in teaching a foreign language in higher technical institutions. *European Journal of Contemporary Education*, 12(2), 578–589. <https://doi.org/10.13187/ejced.2023.2.578>
22. Page, M. J., McKenzie, J. E., Bossuyt, P., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Medicina Fluminensis*, 57(4), 444–465. https://doi.org/10.21860/medflum2021_264903
23. Polakova, P., & Klimova, B. (2024). Implementation of AI-driven technology into education: A pilot study on the use of chatbots in foreign language learning. *Cogent Education*, 11(1), 1–15. <https://doi.org/10.1080/2331186X.2024.2355385>
24. Qin, F., Liu, B., Li, W., Li, H., & Li, J. (2026). The mediating role of motivation between generative artificial intelligence and foreign language proficiency: A technology acceptance model perspective. *SAGE Open*, 16(1). <https://doi.org/10.1177/21582440261430425>

25. Wang, W., & Wang, W. (2025). College students' behavioural intentions of AI-assisted language learning: Based on the technology acceptance model. *Journal of Computer Assisted Learning*, 41(4). <https://doi.org/10.1111/jcal.70075>
26. Yusri, Y., Haliq, A., Nurfadhilah, A. S., R. M., Nur, M. S., Umar, N. F., Syaputra, A. F., Angreany, F., Kudus, R., & Junaeny, A. (2026). Knowledge resilience in AI-mediated foreign language learning: Challenges and responses of Indonesian lecturers. *Journal of Applied Research in Higher Education*, 1–12. <https://doi.org/10.1108/JARHE-02-2026-0254>
27. Zhang, Y., Lucas, M., Bem-haja, P., & Pedro, L. (2024). The effect of student acceptance on learning outcomes: AI-generated short videos versus paper materials. *Computers and Education: Artificial Intelligence*, 7, 1-10. <https://doi.org/10.1016/j.caeai.2024.100286>
28. Zhou, X. (2026). Investigating L2 learners' task motivation in Mandarin Chinese essay writing: The role of ChatGPT in AI-mediated learning. *AERA Open*, 12. <https://doi.org/10.1177/2332858426143>