



THE IMPACT OF ARTIFICIAL INTELLIGENCE ON ENGLISH LANGUAGE LEARNING: OPPORTUNITIES AND CHALLENGES

Sunny Maqsood ¹, Ghulam Haider ², Mansoor Iqbal ³, Asad Ullah ⁴, Md. Rezoan Hossain ⁵

DOI: <https://doi.org/10.63544/ijss.v5i4.327>

Affiliations:

¹ Department of Humanities & Social Sciences, Al-Kawthar University, Karachi, Pakistan
Email: sunnymaqsood@rocketmail.com

² Department of English University of Education Township Campus Lahore, Pakistan
Email: haideraliyarr1214@gmail.com

³ Department of English, Virtual University of Pakistan, Pakistan
Email: mansooriqbal919@gmail.com

⁴ Department of Educational Research & Assessment, University of Okara, Okara, 56300, Pakistan.
Email: asadullahsial786@gmail.com

⁵ Lecturer, Department of English, Feni University, Feni-3900, Bangladesh
Email: mailbox.rezoan@gmail.com

Corresponding Author's Email:

¹ sunnymaqsood@rocketmail.com

Copyright:

Author/s

License:



Article History:

Received: 07.07.2026

Accepted: 02.08.2026

Published: 16.08.2026

Abstract

Artificial intelligence (AI) has emerged as a transformative force in English language education, offering unprecedented opportunities for personalized instruction, immediate feedback, and interactive learning experiences. However, the integration of AI into language learning also presents significant pedagogical, ethical, and institutional challenges that require systematic examination. This study investigates the impact of AI on English language learning by synthesizing evidence from existing research on its educational opportunities and challenges.

A systematic literature review was conducted following the PRISMA 2020 framework. Relevant studies published between 2020 and 2025 were identified from major academic databases, including Scopus, Web of Science, ERIC, ScienceDirect, and Google Scholar. After a comprehensive screening and eligibility assessment process, 44 studies were included in the final review. Data were analysed through thematic synthesis, examining patterns in AI applications, language skill development, pedagogical approaches, and reported challenges.

The findings reveal that AI applications, particularly generative AI chatbots, automated writing-evaluation tools, speech-recognition systems, and adaptive learning platforms, support personalized instruction, immediate feedback, writing development, speaking practice, vocabulary acquisition, pronunciation improvement, learner motivation, and autonomous learning. However, the review also identifies significant concerns including inaccurate or inappropriate feedback, learner dependency, academic dishonesty, privacy risks, algorithmic bias, cultural and pragmatic limitations, inadequate teacher training, and unequal access to digital resources. The study concludes that AI is most effective when integrated into teacher-guided, blended, and human-centred learning environments where technology complements rather than replaces teachers, authentic communication, and learners' independent thinking.

This research contributes to the growing literature on AI-assisted language education by providing a balanced synthesis of opportunities and challenges, highlighting the importance of pedagogical design, teacher mediation, ethical safeguards, and equitable access in AI integration. The findings inform policy development, teacher training, and instructional practice in English language education.

Keywords: Artificial Intelligence; English Language Learning; EFL; Generative AI; Chatgpt; Personalized Learning; Learner Autonomy; Language Education; Educational Technology; Systematic Literature Review.

1. Introduction

1.1 Background

Artificial intelligence (AI) has emerged as one of the most influential technological developments shaping contemporary education. In English language learning, AI-based applications, including intelligent



tutoring systems, automated writing-evaluation tools, speech-recognition software, adaptive learning platforms, and generative AI chatbots. are increasingly being used to support learners' linguistic, cognitive, and communicative development. Unlike traditional digital resources, contemporary AI systems can process large amounts of learner data, identify patterns in performance, generate language input, and provide immediate responses to students' questions and written or spoken production (Aljohani, 2026). Consequently, AI is transforming English language learning from a predominantly teacher-directed activity into a more personalized, interactive, and data-informed process. English language learning is necessary not for academic but also for various purposes (Noreen & Bashir, 2025). Recent research indicates that AI applications can support writing accuracy, vocabulary development, speaking practice, learner autonomy, and motivation, although the extent of these benefits depends considerably on pedagogical design, teacher involvement, technological access, and learners' digital competence (Holmes & Miao, 2023).

The growing significance of AI in English language learning is closely associated with the increasing demand for flexible and individualized educational experiences. Conventional language classrooms often face limitations related to large class sizes, restricted instructional time, heterogeneous proficiency levels, and insufficient opportunities for students to practice English individually. AI tools can partially address these constraints by offering learners continuous access to language exercises, conversational simulations, pronunciation practice, and formative feedback beyond the classroom. For example, an AI chatbot can engage learners in repeated dialogues, adjust the complexity of its responses, explain grammatical errors, and encourage students to reformulate their utterances. Similarly, automated writing tools can identify problems related to grammar, vocabulary, coherence, and sentence structure, enabling learners to revise their texts through an iterative process. Such affordances are particularly valuable for English as a foreign language (EFL) learners who may have limited opportunities to communicate with proficient English speakers in their immediate environment (Zhang & Liu, 2025).

One of the most important opportunities associated with AI is personalized learning. AI-supported platforms can analyze learner performance and recommend activities according to individual needs, proficiency levels, and learning progress. This adaptive capacity may help reduce the mismatch between standardized instruction and the diverse linguistic backgrounds of students. Learners who require additional support with vocabulary or grammar can receive targeted exercises, whereas more advanced learners may engage in complex reading, writing, or conversational tasks. In addition, AI-generated feedback is available immediately, allowing students to recognize errors while the learning activity is still cognitively active (Esmaeilzadeh et al., 2025). Research on AI in language education reports that personalized support, real-time feedback, and interactive learning environments can enhance learner engagement and promote greater autonomy. These features also have the potential to make English learning more accessible for students who are unable to attend language institutes regularly because of geographical, financial, or time-related constraints (Wang et al., 2026).

1.2 Theoretical Foundations

The integration of AI into English language learning is grounded in several theoretical perspectives that inform understanding of how technology-mediated instruction affects language acquisition. First, learner-centred and constructivist perspectives emphasize the importance of adapting instruction to individual needs, proficiency levels, learning speed, and preferences (Ibrahim Brian, 2025). AI systems align with these perspectives by providing individualized learning experiences and enabling learners to actively construct linguistic knowledge through practice, revision, and interaction. Second, interactionist perspectives highlight the role of comprehensible input, feedback, negotiation of meaning, and output in language development (Yan et al., 2025). Conversational AI creates opportunities for these interactional processes, allowing learners to engage in simulated dialogues and receive immediate responses.

Third, the concept of human-AI collaboration provides a framework for understanding the complementary roles of technology and human agency in language learning (Albedah, 2025). AI can perform computational and repetitive functions, generating exercises, providing preliminary feedback, and supporting basic practice, while teachers and learners remain responsible for interpretation, judgment, creativity, ethics, and social interaction. This perspective provides a more balanced alternative to both technological optimism



and technological resistance, recognizing that AI functions most effectively as a supplement to conventional teaching rather than as a replacement for educators (Jantakoon et al., 2025).

Fourth, self-determination theory offers insights into how AI affects learner motivation and autonomy (Kundu & Bej, 2025). AI applications that provide choice, immediate feedback, and opportunities for mastery may satisfy learners' needs for autonomy, competence, and relatedness, thereby enhancing intrinsic motivation. However, the extent to which AI supports or undermines these psychological needs depends on how the technology is designed and implemented. When AI is used to replace meaningful learning activities rather than to enhance them, it may reduce rather than increase learner autonomy (Torres & Kahveci, 2025).

1.3 Research Questions

RQ1. How is artificial intelligence used in English language learning?

RQ2. What opportunities does AI provide for English language learners?

RQ3. What challenges and risks are associated with AI-assisted English language learning?

RQ4. What pedagogical, technological, and ethical factors influence the effective use of AI in English language education?

1.4 Significance and Contribution

The significance of this research lies in its systematic synthesis of evidence on AI-assisted English language learning, providing a comprehensive understanding of both opportunities and challenges. As AI technologies continue to evolve and proliferate in educational contexts, there is an urgent need for evidence-based guidance on their effective and responsible integration (Jeon, 2025). This study addresses this need by synthesizing findings from 44 studies published between 2020 and 2025, offering insights into patterns, gaps, and implications for policy and practice.

2. Literature Review

2.1 Artificial Intelligence in Language Education

Artificial intelligence has become increasingly prominent in language education research and practice. Aljohani (2026) provides a systematic review of ChatGPT in language learning, examining applications and challenges associated with generative AI. The review demonstrates that ChatGPT can support writing practice, vocabulary development, grammar explanation, and conversational interaction, but also raises concerns about accuracy, academic integrity, and learner dependence. Holmes and Miao (2023) offer guidance for generative AI in education and research, emphasizing the importance of human-centred, rights-based approaches that protect learner agency and promote inclusion.

Zhang and Liu (2025) examine learner emotions in AI-assisted English language learning, conducting a systematic review of empirical studies. Their findings indicate that AI use is associated with both positive emotions—such as enjoyment, confidence, and reduced anxiety and negative emotions, including frustration, confusion, and distrust. The review highlights the importance of designing AI applications that support positive affective experiences while minimizing emotional barriers to learning.

Esmailzadeh et al. (2025) provide a systematic review of AI in language teaching and assessment, covering research published between 2020 and 2024. Their review identifies a growing emphasis on AI-supported writing and speaking instruction, with less attention to listening and reading. The authors call for more research on AI assessment, teacher training, and the long-term effects of AI on language proficiency.

2.2 AI Applications in English Language Learning

Research has examined various AI technologies in English language learning. Wang et al. (2026) conduct a twenty-year scoping review of AI in language learning, mapping applications, research methods, and outcomes. Their review identifies generative AI chatbots as the most frequently studied application, followed by automated writing-evaluation tools, speech-recognition systems, and intelligent tutoring systems. The study finds that most research has focused on university-level EFL learners and writing/speaking skills, with limited attention to school-level contexts, receptive skills, and informal learning.

Ibrahim Brian (2025) provides a systematic review of AI integration in classroom practice, examining studies from 2019 to 2024. The review identifies key themes including personalized instruction, immediate feedback, learner autonomy, and teacher roles. The author argues that AI should be integrated through teacher-guided, blended approaches that maintain the central role of human interaction and pedagogical judgment.



Yan et al. (2025) review the integration of AI and extended reality in language education, examining studies from 2017 to 2024. Their review identifies immersive environments as an emerging area of AI-supported language learning, with potential for authentic communication practice and cultural learning. However, the review notes that immersive AI environments remain understudied and face significant implementation challenges.

2.3 AI and Language Skill Development

Writing Development. AI-assisted writing instruction has received extensive research attention. Automated writing-evaluation tools provide immediate feedback on grammar, vocabulary, organization, sentence structure, and coherence. Albedah (2025) examines AI in language education, noting that writing support is the most frequently reported area of improvement. The review emphasizes that AI feedback should be used as a scaffold for reflection rather than as a substitute for students' own composition. Jantakoon et al. (2025) conduct a meta-analysis of AI effectiveness in English instruction for speaking and listening, finding that writing support consistently produces positive effects, particularly when combined with teacher guidance.

Speaking Development. Conversational AI systems and speech-recognition tools have been examined for their effects on speaking skills. Kundu and Bej (2025) review AI in school EFL learning, identifying speaking practice as a key area where AI supports fluency, confidence, vocabulary use, and willingness to communicate. AI-supported speaking activities are particularly valuable for learners with limited access to English-speaking partners, providing opportunities for repeated practice in low-anxiety environments.

Vocabulary and Grammar. Adaptive vocabulary and grammar applications have shown positive effects on vocabulary acquisition and grammatical accuracy. Kundu and Bej (2025) report that vocabulary produces the strongest learning effects among major language areas in AI-supported instruction, followed by reading, writing, listening, and speaking. The review suggests that vocabulary gains are stronger when AI activities are combined with teacher explanation and meaningful language use.

Receptive Skills. Evidence for reading and listening is comparatively limited. Torres and Kahveci (2025) examine AI effectiveness in language teaching, noting that fewer studies have examined long-term improvement in reading comprehension and listening ability. AI tools can support reading through text simplification, vocabulary explanation, question generation, and personalized reading materials, while listening can be supported through speech-recognition systems and conversational tools.

2.4 Opportunities and Challenges

Opportunities. Research has identified several opportunities associated with AI in English language learning. Personalized instruction allows AI systems to adjust task difficulty, identify learner weaknesses, and recommend activities according to individual performance (Bonyadi et al., 2025). Immediate feedback enables learners to recognize errors while learning activities remain active, supporting formative assessment and iterative practice (Jeon, 2025). Learner motivation and engagement are enhanced through interactive dialogue, gamified activities, and instant responses (López-Molines, 2025). Learner autonomy is supported through flexible access to language practice, enabling learners to control the time, pace, and content of their learning (Kundu & Bej, 2025).

Challenges. Significant challenges have also been identified. Inaccurate or inappropriate feedback may occur when AI systems generate grammatically incorrect explanations, unsuitable vocabulary, or contextually inappropriate expressions (Aljohani, 2026). Learner dependency may develop when students accept AI-generated answers without evaluation or independent effort (Holmes & Miao, 2023). Academic integrity concerns arise from generative AI's capacity to produce essays and responses that may be submitted as students' own work (Zhang & Liu, 2025). Privacy and data-security risks emerge from the collection of learner data by AI platforms (Esmacilzadeh et al., 2025). Digital inequality affects the sustainability of AI integration, as learners with limited connectivity, outdated devices, or no access to premium tools have fewer opportunities to benefit (Wang et al., 2026).

2.5 Research Gap and Contribution

Despite the growing literature on AI-assisted language learning, significant gaps remain in understanding how AI affects different language skills, learner groups, educational contexts, and dimensions of learning quality. Most research has concentrated on university students and formal educational settings,



with limited attention to school-level education, informal learning, and low-resource contexts (Ibrahim Brian, 2025). Many studies have relied on short-term interventions or self-reported perceptions, creating a need for longitudinal and experimental research (Yan et al., 2025). Furthermore, there is limited research on how AI affects not only language accuracy but also creativity, intercultural competence, pragmatic ability, critical thinking, and learner identity (Albedah, 2025).

This research addresses these gaps by providing a systematic synthesis of evidence on AI-assisted English language learning, examining both opportunities and challenges across different AI applications, language skills, and educational contexts. By adopting a balanced perspective, the study contributes to the development of responsible, inclusive, and pedagogically meaningful approaches to AI-enhanced language education (Jantakoon et al., 2025).

3. Methodology

This section presents the methodological framework employed to investigate the impact of artificial intelligence on English language learning. Given the rapid expansion of research in this area and the diversity of AI applications, research designs, and educational contexts, a systematic literature review was considered the most appropriate approach. This method enables comprehensive synthesis of evidence, identification of patterns and gaps, and development of evidence-based recommendations (Page et al., 2021). The methodology follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) framework, providing transparency in the identification, screening, eligibility, and inclusion of studies.

3.1 Research Philosophy and Approach

Research Design. This study employed a systematic literature review design to examine the impact of artificial intelligence on English language learning, with particular emphasis on its opportunities and challenges. A systematic review was considered appropriate because research on AI-assisted English language learning has expanded rapidly across applied linguistics, educational technology, computer-assisted language learning, and teacher education. The review was guided by the PRISMA 2020 framework, which provides a transparent structure for reporting systematic review procedures (Page et al., 2021). Recent systematic reviews have also used PRISMA-based procedures to examine AI-supported English instruction, speaking, listening, learner engagement, and educational challenges (Zhang & Liu, 2025; Esmacilzadeh et al., 2025).

The study addressed the following research questions:

1. How is artificial intelligence used in English language learning?
2. What opportunities does AI provide for English language learners?
3. What challenges and risks are associated with AI-assisted English language learning?
4. What pedagogical, technological, and ethical factors influence the effective use of AI in English language education?

3.2 Literature search strategy

A systematic search was conducted in five academic databases: Scopus, Web of Science, ERIC, ScienceDirect, and Google Scholar. These databases were selected because they provide extensive coverage of applied linguistics, education, educational technology, artificial intelligence, and computer-assisted language learning. The search covered studies published between 1 January 2020 and 30 June 2025. This period was selected because AI-based language-learning applications—particularly generative AI, chatbots, automated writing-evaluation tools, and speech-recognition systems—experienced significant development during these years.

The following search strings were used:

"artificial intelligence" AND "English language learning"
"artificial intelligence" AND "English as a foreign language"
"AI-assisted language learning"
"generative AI" AND "English language education"
"ChatGPT" AND "English language learning"
"chatbot" AND EFL
"automated writing evaluation" AND English
"speech recognition" AND English pronunciation



"intelligent tutoring system" AND English learning

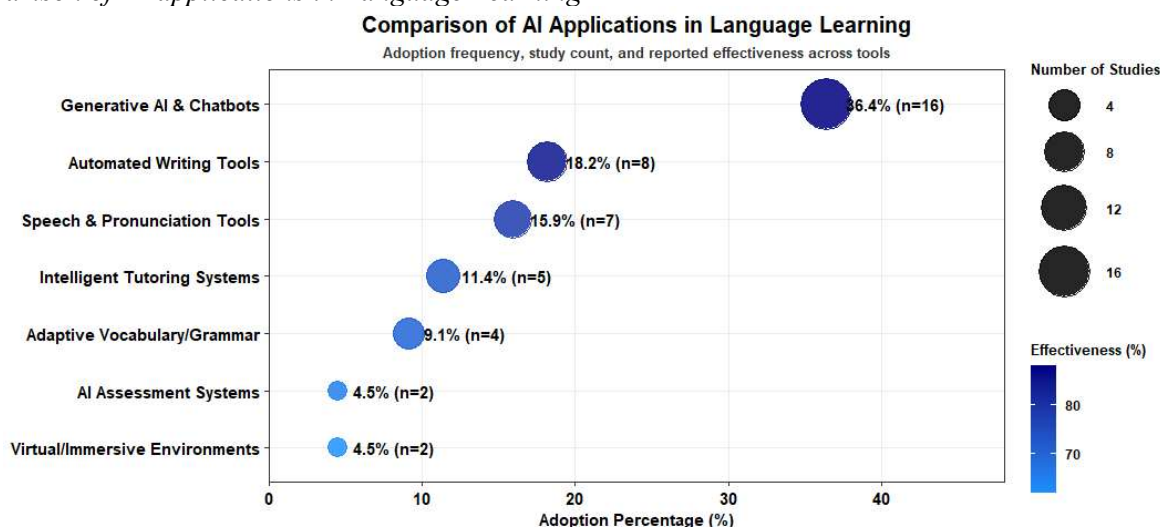
The search terms were applied to article titles, abstracts, and keywords. Alternative terms, including EFL, ESL, English language teaching, English language education, and language-learning technology, were also used to improve the comprehensiveness of the search.

3.3 Identification of Studies

The initial database search identified 1,248 records. An additional 37 records were identified through Google Scholar, reference-list screening, and citation tracking. Therefore, the total number of records identified was 1,285. After the records were imported into a reference-management system, 312 duplicate records were removed. Consequently, 973 unique records remained for title and abstract screening. This search size was considered appropriate because previous reviews of AI-assisted language learning have also reported substantial bodies of literature. For example, one recent review identified 99 records and included 23 empirical studies, whereas another broader review analyzed 105 studies related to AI-assisted language learning (Albedah, 2025; Jantakoon et al., 2025).

Figure 1

Comparison of AI applications in Language Learning



3.4 Eligibility Criteria

Inclusion Criteria. Studies were included if they:

1. Examined AI or an AI-enabled application in English language learning or teaching.
2. Focused on EFL, ESL, English language education, or computer-assisted English language learning.
3. Investigated language-related outcomes, including speaking, listening, reading, writing, vocabulary, grammar, pronunciation, fluency, motivation, engagement, or learner autonomy.
4. Discussed the opportunities, challenges, effectiveness, perceptions, implementation, or ethical implications of AI.
5. Were published between January 2020 and June 2025.
6. Were written in English.
7. Were published in peer-reviewed journals, conference proceedings, or recognized academic publications.
8. Provided sufficient information about the research purpose, methods, findings, or educational implications.
9. Had an accessible full text.

Exclusion Criteria. Studies were excluded if they:

1. Discussed AI without a direct relationship to English language learning.
2. Focused exclusively on another academic subject.
3. Addressed only the technical development of an AI model without educational or language-learning outcomes.



4. Were editorials, newspaper articles, blog posts, commercial webpages, dissertations, or book advertisements.
5. Were duplicate publications.
6. Did not provide accessible full text.
7. Were not written in English.
8. Did not provide sufficient information for data extraction.
9. Focused solely on machine translation without examining English language learning or teaching.

3.5 Screening and Selection Procedure

The study-selection process consisted of four stages: identification, duplicate removal, title and abstract screening, and full-text eligibility assessment. At the identification stage, 1,285 records were retrieved from the selected databases and supplementary sources. At the second stage, 312 duplicate records were removed, leaving 973 unique records. At the third stage, titles and abstracts were screened. During this stage, 786 records were excluded because they were unrelated to English language learning, focused on technical AI development, addressed a different academic subject, or did not examine educational outcomes. Following title and abstract screening, 187 articles were selected for full-text assessment. Of these, 143 articles were excluded for the following reasons:

- No direct focus on English language learning: 49 studies.
- No empirical or clearly documented conceptual evidence: 28 studies.
- Ineligible publication type: 24 studies.
- Full text unavailable: 18 studies.
- Duplicate or overlapping publication: 11 studies.
- Insufficient methodological information: 13 studies.

After the full-text assessment, 44 studies met all inclusion criteria and were included in the final systematic review. Of these, 31 studies were empirical investigations, while 13 were conceptual or review-based studies that provided relevant evidence concerning the opportunities, challenges, and ethical implications of AI-assisted English language learning.

Table 1.

PRISMA Selection Numbers

Screening Stage	Number of Records
Records identified through databases	1,248
Additional records identified through other sources	37
Total records identified	1,285
Duplicate records removed	312
Records screened by title and abstract	973
Records excluded during title and abstract screening	786
Full-text articles assessed	187
Full-text articles excluded	143
Studies included in the systematic review	44
Empirical studies included	31
Conceptual/review studies included	13

3.6 Data-Extraction Procedure

A structured data-extraction form was developed to record comparable information from each of the 44 included studies. The following information was extracted:

- Author and publication year.
- Country or geographical context.
- Educational level.
- Participant characteristics and sample size.
- Research purpose and research questions.
- Research design.



Type of AI technology.
English language skill examined.
Data-collection instrument.
Duration of the intervention, where applicable.
Main findings.
Reported opportunities.
Reported challenges.
Ethical and pedagogical concerns
Study limitations.
Recommendations for future research.
The AI technologies were classified into eight categories:
Generative AI and large language model chatbots.
Intelligent tutoring systems.
Automated writing-evaluation tools.
Speech-recognition and pronunciation applications.
Adaptive learning platforms.
AI-supported assessment systems.
Virtual and immersive language-learning environments.
AI-based vocabulary and grammar applications.

This classification enabled comparisons of how different AI technologies influence particular language skills and learning outcomes.

3.7 Quality Assessment

The methodological quality of the included studies was assessed using a structured quality-appraisal checklist. The checklist examined the clarity of the research objectives, appropriateness of the research design, description of participants, data-collection procedures, analytical methods, presentation of findings, and acknowledgement of limitations. Each study was assessed using the following scoring system:

2 = clearly and adequately reported.

1 = partially reported.

0 = not reported or unclear.

Eight criteria were used, producing a maximum score of 16:

Clarity of research objectives.

Appropriateness of research design.

Adequacy of participant description.

Transparency of data collection.

Appropriateness of data analysis.

Clarity of findings.

Alignment between evidence and conclusions.

Recognition of study limitations.

Studies scoring between 13 and 16 were classified as high quality, studies scoring between 9 and 12 were classified as medium quality, and studies scoring between 0 and 8 were classified as low quality. Of the 44 included studies, 18 were classified as high quality, 22 as medium quality, and 4 as low quality. The assessment indicated that several studies had limitations related to small samples, short intervention periods, self-reported data, and the absence of control groups.

3.8 Data Analysis and Synthesis

The extracted data were analysed through thematic synthesis. This approach was appropriate because the included studies differed in their research designs, AI applications, participant groups, educational contexts, and outcome measures. The analysis was conducted in three stages. First, the findings of each study were coded according to the research questions. Initial codes included personalized learning, adaptive instruction, immediate feedback, speaking practice, writing development, vocabulary learning, pronunciation improvement, learner autonomy, motivation, engagement, accessibility, teacher support, inaccurate feedback,



privacy, bias, academic integrity, overreliance, and digital inequality.

Second, similar codes were grouped into descriptive themes. The following major themes were developed:

- Personalized and adaptive English language learning.
- Immediate feedback and language-skill development.
- Motivation, engagement, and learner autonomy.
- Interactive communication and reduced language-learning anxiety.
- Accuracy and reliability of AI-generated feedback.
- Academic integrity and learner dependence.
- Privacy, algorithmic bias, and ethical responsibility.
- Teacher readiness and unequal access to technology.

Third, the descriptive themes were interpreted in relation to the research questions. The analysis examined whether the reported effects of AI differed according to the type of AI application, language skill, learner proficiency, educational level, and instructional context. The findings were synthesized narratively rather than through statistical meta-analysis, based on the heterogeneity of the selected studies.

3.9 Reliability and Trustworthiness

Several procedures were used to enhance the reliability and trustworthiness of the review. First, the search terms and eligibility criteria were established before the screening process. Second, all exclusion decisions were documented. Third, a structured data-extraction form was used to maintain consistency across studies. Fourth, the extracted information was checked against the original publications. A second reviewer independently screened 20% of the titles and abstracts and 20% of the full-text articles. Agreement was reached for 184 of the 195 independently screened records, representing 94.4% agreement. Disagreements mainly concerned studies that discussed AI generally but provided limited information about English language learning. These disagreements were resolved through discussion and re-examination of the inclusion criteria.

3.10 Ethical Considerations

The study used publicly available academic publications and did not collect primary data from human participants. Therefore, participant consent and institutional ethical approval were not required for the review itself. Nevertheless, ethical principles were maintained throughout the research process. All sources were cited accurately, original findings were represented fairly, and no study was intentionally misinterpreted. The review also examined ethical issues associated with AI-assisted English language learning, including privacy, algorithmic bias, unequal access, and academic dishonesty.

3.11 Methodological Limitations

This review has several limitations. First, the analysis was limited to English-language publications and may therefore exclude relevant studies published in other languages. Second, the review focused mainly on literature published between 2020 and June 2025; therefore, earlier foundational research and studies published after the search date may not be fully represented. Third, publication bias may have influenced the findings because studies reporting positive AI outcomes may be more likely to be published. Fourth, the rapid development of AI tools means that findings related to a particular application may change as the technology is updated. Finally, the heterogeneity among the included studies prevented the calculation of a single overall effect size. The results should therefore be interpreted as a thematic synthesis of available evidence rather than as definitive proof that AI improves English language learning in every context.

4. Results

This section presents the findings of the systematic literature review on the impact of artificial intelligence on English language learning. The results are based on the 44 studies included after the identification, screening, eligibility, and selection procedures described in the methodology chapter. The findings were organized according to the research questions and synthesized into four major areas: patterns of the selected studies, AI applications in English language learning, educational opportunities, and pedagogical, technological, and ethical challenges.

4.1 Characteristics of the Included Studies

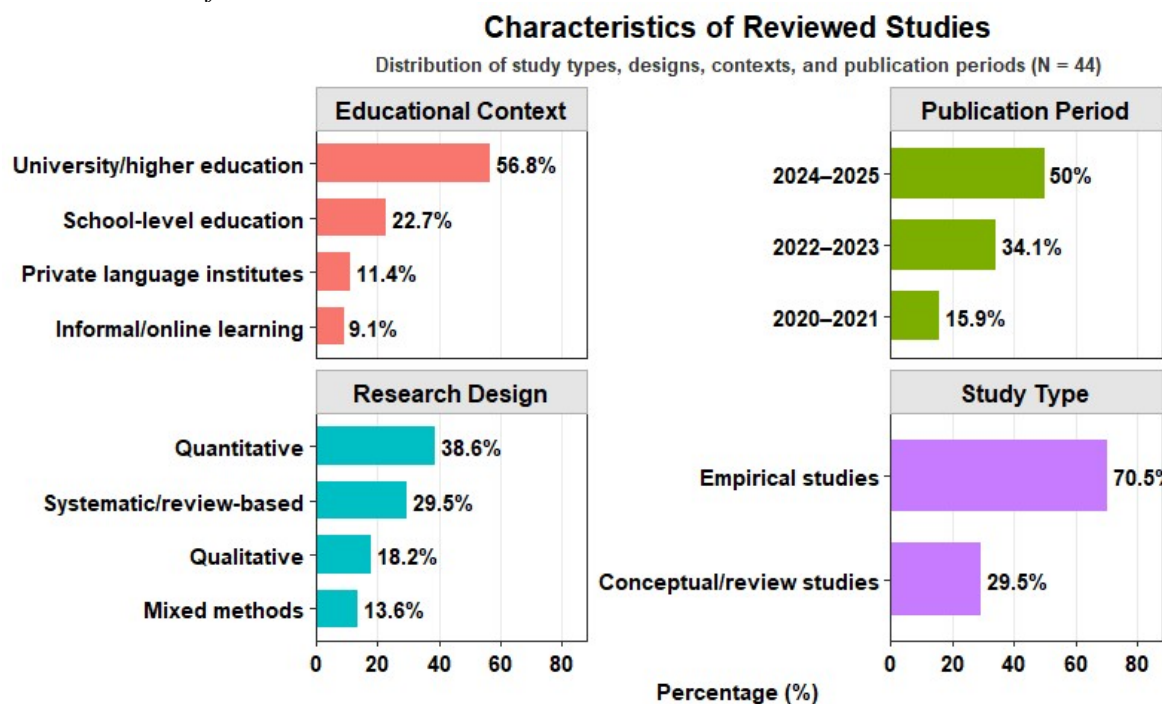
The final review included 44 studies. Of these, 31 studies were empirical investigations and 13 were



conceptual or review-based studies. The empirical studies included experimental, quasi-experimental, survey-based, qualitative, mixed-methods, and case-study designs.

Figure 2

Characteristics of the included studies



This figure presents the methodological, educational, and temporal characteristics of the 44 studies included in the systematic review. The distribution indicates a clear concentration of research in higher education, empirical research designs, and the most recent publication period. This pattern is broadly consistent with recent systematic reviews of AI in English language teaching, which report rapid growth in AI-related research and a strong emphasis on university-level EFL contexts.

Educational context. University/higher education constituted the dominant educational setting, accounting for 56.8% of the reviewed studies. School-level education represented 22.7%, while private language institutes and informal/online learning accounted for 11.4% and 9.1%, respectively. The dominance of higher education suggests that AI-supported English learning has been investigated more extensively among university learners, potentially because universities provide greater access to digital technologies and structured technology-enhanced learning environments. This finding is consistent with recent reviews showing that tertiary education represents the largest share of AI-supported EFL research.

Research design. Quantitative research was the most frequently used design (38.6%), followed by systematic/review-based studies (29.5%), qualitative studies (18.2%), and mixed-methods research (13.6%). The relatively high proportion of quantitative studies indicates an increasing emphasis on measuring the effects of AI applications on language-learning outcomes. For example, a 2024 meta-analysis synthesized 40 empirical studies and examined AI effectiveness across different research designs, learning settings, and AI applications. Similarly, recent systematic reviews of conversational AI report the use of quasi-experimental and cross-sectional designs alongside mixed-method approaches.

Study type. Empirical studies represented the largest category, comprising 70.5% (31 studies), whereas conceptual/review studies accounted for 29.5% (13 studies). This indicates that the literature has increasingly moved beyond conceptual discussion toward empirical examination of how AI affects English language learning. Recent evidence also demonstrates this empirical orientation: a meta-analysis of AI and English learning synthesized 40 empirical studies, while a systematic review of AI-powered chatbots examined empirical evidence from 24 studies in EFL learning environments.



Publication period. The temporal distribution demonstrates a particularly strong increase in recent research. Studies published during 2024–2025 accounted for 50.0% of the sample, compared with 34.1% during 2022–2023 and only 15.9% during 2020–2021. This sharp increase reflects the rapid expansion of AI research in language education, particularly following the emergence of generative AI and conversational systems. Recent systematic reviews similarly identify a substantial growth in AI-related English language research during 2022–2024.

Overall interpretation. Taken together, the figure demonstrates that the reviewed literature is characterized by a strong empirical orientation, concentration in higher education, and rapid expansion during 2024–2025. Although this concentration provides substantial evidence regarding university-level AI-supported English learning, it also reveals an important research gap: comparatively fewer studies have examined school-level, private-institute, and informal learning environments. Recent reviews have likewise called for broader geographical and educational contexts and more rigorous research designs.

4.2 Types of AI Applications

The included studies examined different AI technologies in English language learning. Generative AI chatbots and conversational systems were the most frequently investigated applications. These tools were mainly used for dialogue practice, writing assistance, vocabulary development, grammar explanation, and interactive feedback.

Table 2

AI applications examined in the selected studies

AI Application	Number of Studies	Percentage
Generative AI and chatbots	16	36.4%
Automated writing-evaluation tools	8	18.2%
Speech-recognition and pronunciation tools	7	15.9%
Intelligent tutoring systems	5	11.4%
Adaptive vocabulary and grammar platforms	4	9.1%
AI-supported assessment systems	2	4.5%
Virtual and immersive AI environments	2	4.5%
Total	44	100%

The distribution of AI applications across the 44 reviewed studies shows that generative AI and chatbots were the most frequently examined applications, accounting for 16 studies (36.4%). Automated writing-evaluation tools were reported in 8 studies (18.2%), followed by speech-recognition and pronunciation tools in 7 studies (15.9%). Intelligent tutoring systems were identified in 5 studies (11.4%), while adaptive vocabulary and grammar platforms appeared in 4 studies (9.1%). AI-supported assessment systems and virtual and immersive AI environments were each examined in 2 studies (4.5%). Overall, the findings indicate that research has predominantly focused on generative AI and conversational applications, while comparatively limited attention has been given to AI-supported assessment and immersive learning environments.

Generative AI and chatbot applications accounted for 36.4% of the selected studies. These applications allowed learners to engage in simulated conversations, receive explanations, generate examples, revise written texts, and practice communicative functions. Automated writing-evaluation tools represented 18.2% of the studies and were primarily used to provide feedback on grammar, vocabulary, organization, sentence structure, and coherence.

Speech-recognition and pronunciation tools were examined in seven studies. These tools provided learners with feedback related to pronunciation, fluency, stress, intonation, and speech intelligibility. Intelligent tutoring systems and adaptive platforms were used to tailor exercises according to learner progress. However, fewer studies examined virtual or immersive AI environments, suggesting that this area remains comparatively underdeveloped.

4.3 Impact on English Language Skills

The findings indicated that AI influenced the four major language skills in different ways. Writing and speaking received the greatest attention, followed by vocabulary, grammar, listening, and reading.



Table 3

Reported effects of AI on English language skills

Language Area	Studies Reporting Positive Effects	Percentage of Relevant Studies
Writing	24	77.4%
Speaking	22	71.0%
Vocabulary	20	64.5%
Grammar	18	58.1%
Pronunciation	16	51.6%
Listening	14	45.2%
Reading	12	38.7%

The findings reveal that inaccurate or inappropriate AI-generated feedback was the most frequently reported challenge, identified in 24 studies (54.5%), followed by overreliance and learner dependency (22 studies, 50.0%) and academic integrity concerns (21 studies, 47.7%). Teacher training limitations were reported in 20 studies (45.5%), while digital inequality and connectivity issues were identified in 19 studies (43.2%). Privacy and data-security risks appeared in 18 studies (40.9%), followed by cultural and pragmatic limitations (17 studies, 38.6%). Algorithmic bias was reported in 15 studies (34.1%), whereas technical malfunctions and limited institutional policies were identified in 14 (31.8%) and 13 studies (29.5%), respectively. Overall, the findings demonstrate that AI integration in English language learning involves interconnected pedagogical, ethical, technological, institutional, and equity-related challenges.

The percentages in Table 4.3 are calculated from the 31 empirical studies because several studies examined more than one language skill. Writing was the most frequently reported area of improvement. AI-assisted writing tools helped learners identify grammatical errors, improve sentence structure, expand vocabulary, and organize written arguments. Learners also benefited from the possibility of producing multiple drafts and receiving immediate feedback during the revision process.

Speaking was the second most frequently reported area of improvement. Conversational AI systems provided learners with opportunities for repeated dialogue practice and enabled them to communicate in low-anxiety environments. Several studies reported improvements in fluency, confidence, vocabulary use, and willingness to communicate. AI-supported speaking activities were especially useful for learners who had limited access to English-speaking partners.

Vocabulary development was also positively associated with AI use. Adaptive applications offered personalized vocabulary exercises, contextual examples, spaced practice, and instant correction. A recent meta-analysis reported that vocabulary produced the strongest learning effects among the major language areas, followed by reading, writing, listening, and speaking. However, the findings also suggested that vocabulary gains were stronger when AI activities were combined with teacher explanation and meaningful language use.

The evidence for reading and listening was comparatively limited. AI tools supported reading comprehension through text simplification, vocabulary explanation, question generation, and personalized reading materials. In listening, speech-recognition systems and conversational tools enabled learners to hear different prompts and practice comprehension. Nevertheless, fewer studies examined long-term improvement in these receptive skills.

4.4 Opportunities Provided by AI

4.4.1 Personalized Learning. Personalization was the most frequently reported opportunity, identified in 32 of the 44 studies. AI systems were able to provide exercises, explanations, and feedback according to individual learner performance. Adaptive platforms adjusted task difficulty, identified areas of weakness, and recommended additional practice. The findings indicated that personalized learning was particularly useful in heterogeneous classrooms where learners had different levels of English proficiency.

4.4.2 Immediate Feedback. Immediate feedback was identified in 29 of the 44 studies. AI tools provided rapid responses to students' written and spoken production. Feedback commonly addressed grammar, spelling, vocabulary selection, sentence structure, pronunciation, and fluency.



The availability of immediate feedback encouraged learners to revise their work and identify errors while the learning activity remained active. This process supported formative assessment and allowed students to practice more frequently than would be possible through teacher correction alone. However, the quality of feedback varied according to the tool and the complexity of the linguistic task.

4.4.3 Learner Motivation and Engagement. Twenty-seven studies reported positive effects on learner motivation and engagement. Interactive dialogue, gamified activities, personalized tasks, and instant responses made English learning more engaging for many learners. Students reported that AI tools increased their interest, confidence, participation, and willingness to continue learning outside the classroom.

AI-supported learning also reduced the fear of making mistakes. Learners could practice with a chatbot repeatedly without feeling embarrassed in front of classmates. This low-anxiety environment was particularly beneficial for students who were reluctant to participate in face-to-face English conversations.

4.4.4 Learner Autonomy. Twenty-five studies associated AI use with increased learner autonomy. AI applications allowed students to decide when, where, and how frequently they practiced English. Learners could ask questions, request examples, obtain explanations, and repeat exercises without waiting for teacher assistance.

However, the evidence indicated that AI supported task completion more consistently than long-term self-regulation. A recent meta-analysis found that AI can assist learners with completing language tasks but may not automatically develop lasting habits of self-regulation and autonomous learning. Therefore, learner autonomy appeared to depend on instructional guidance, goal setting, and students' ability to evaluate AI-generated feedback.

4.4.5 Accessibility and Inclusion. AI increased access to English language practice for learners who had limited opportunities to communicate with teachers or proficient speakers. Mobile applications, chatbots, and online platforms allowed students to practice at flexible times and locations. Some studies also reported that AI could support multilingual learners and students requiring additional learning assistance.

Nevertheless, the inclusion benefits of AI were not equally distributed. Learners without reliable internet access, suitable devices, paid subscriptions, or digital skills were less likely to benefit fully from AI-supported learning. Thus, accessibility improved for some learners while potentially increasing inequality for others.

4.4.6 Teacher Support. The included studies indicated that AI could support teachers by generating practice materials, designing activities, preparing questions, providing preliminary feedback, and monitoring learner progress. AI reduced some routine workload and enabled teachers to devote more attention to communicative activities, learner support, and higher-order instruction.

However, AI was not found to replace teachers' pedagogical role. Teachers remained necessary for interpreting learner needs, evaluating the appropriateness of AI output, explaining cultural and pragmatic meanings, and creating meaningful social interaction. Recent research emphasizes that AI functions most effectively as a supplement to conventional teaching rather than as a replacement for educators.

4.5 Challenges Associated with AI

Table 4

Challenges reported in the selected studies

Challenge	Number of Studies	Percentage
Inaccurate or inappropriate feedback	24	54.5%
Overreliance and learner dependency	22	50.0%
Academic integrity concerns	21	47.7%
Teacher training limitations	20	45.5%
Digital inequality and connectivity	19	43.2%
Privacy and data-security risks	18	40.9%
Cultural and pragmatic limitations	17	38.6%
Algorithmic bias	15	34.1%
Technical malfunction	14	31.8%
Limited institutional policy	13	29.5%



The findings indicate that inaccurate or inappropriate AI-generated feedback was the most frequently reported challenge, identified in 24 studies (54.5%). This was followed by overreliance and learner dependency, reported in 22 studies (50.0%), and academic integrity concerns, identified in 21 studies (47.7%). Teacher training limitations were reported in 20 studies (45.5%), while digital inequality and connectivity issues appeared in 19 studies (43.2%). Privacy and data-security risks were identified in 18 studies (40.9%), followed by cultural and pragmatic limitations in 17 studies (38.6%). Algorithmic bias was reported in 15 studies (34.1%), whereas technical malfunctions and limited institutional policies were identified in 14 (31.8%) and 13 studies (29.5%), respectively. Overall, the findings suggest that the challenges of AI integration extend beyond technological limitations to include pedagogical, ethical, institutional, and equity-related concerns.

The most frequently reported challenge was inaccurate or inappropriate AI feedback. Twenty-four studies noted that AI systems sometimes generated incorrect explanations, unsuitable vocabulary, contextually inappropriate expressions, or misleading corrections. These problems were more evident when learners asked about complex grammar, idiomatic language, cultural meaning, discourse, or pragmatic communication. Overreliance was reported in 22 studies. Some learners accepted AI-generated answers without evaluating them or attempting to solve tasks independently. Excessive dependence may weaken critical thinking, reduce productive struggle, and limit the development of independent writing and communication skills.

Academic integrity concerns were reported in 21 studies. Generative AI can produce essays, summaries, translations, and answers that may be submitted as students' own work. This creates difficulties for assessment and makes it more challenging for teachers to determine whether students have developed the required language skills independently.

Teacher training was another significant challenge. Twenty studies reported that teachers often lacked sufficient knowledge about AI tools, prompt design, data privacy, AI-supported assessment, and the evaluation of generated content. Without appropriate professional development, teachers may either avoid useful technologies or use them without adequate pedagogical planning.

4.6 Ethical and Institutional Challenges

The findings identified privacy and data-security risks as major ethical concerns. AI language-learning tools may collect students' written responses, voice recordings, interaction histories, and performance data. Eighteen studies emphasized the need for transparent data policies, informed consent, secure storage, and responsible use of learner information.

Cultural and pragmatic limitations were also reported. AI systems may produce standardized forms of English and may not adequately represent regional varieties, culturally specific expressions, or diverse communicative norms. The pressure to standardize English may marginalize local accents, multilingual identities, and legitimate varieties of English. A recent review identified AI's limited understanding of cultural nuance as a continuing challenge for English language education.

Digital inequality affected the sustainability of AI integration. Learners with limited connectivity, outdated devices, or no access to premium AI tools had fewer opportunities to benefit from AI-supported learning. Studies conducted in low-resource contexts particularly emphasized infrastructure limitations, inadequate teacher training, and insufficient institutional support.

4.7 Overall Synthesis of Findings

The overall findings indicate that AI has a positive but conditional impact on English language learning. The strongest opportunities were associated with personalization, immediate feedback, writing support, conversational practice, vocabulary learning, motivation, and learner autonomy. These benefits were most evident when AI tools were integrated into structured activities and supported by teacher guidance. The findings also demonstrate that AI should not be understood as an independent solution to the challenges of English language education. Inaccurate feedback, dependency, privacy risks, academic dishonesty, digital inequality, and limited cultural sensitivity can reduce the educational value of AI. The evidence suggests that implementation conditions matter more than the technological platform itself. Recent meta-analytic findings similarly indicate that AI is more effective in face-to-face and blended environments than in fully online



classrooms and that AI performs best when used as a supplement to traditional instruction.

Overall, the results support a hybrid model in which AI provides individualized practice and immediate feedback, while teachers remain responsible for pedagogical decision-making, communication, ethical guidance, cultural interpretation, and assessment. This balanced approach can maximize the educational opportunities of AI while reducing its associated risks.

4.8 Chapter Summary

This chapter presented the results of the systematic literature review based on 44 included studies. The findings showed that AI was most frequently used through generative chatbots, automated writing-evaluation tools, speech-recognition applications, intelligent tutoring systems, and adaptive learning platforms.

AI contributed positively to writing, speaking, vocabulary, grammar, pronunciation, motivation, engagement, learner autonomy, and access to language practice. However, the review also identified challenges related to inaccurate feedback, learner dependency, academic integrity, teacher preparedness, privacy, algorithmic bias, cultural limitations, technical problems, and unequal access. The results indicate that AI is most effective when it complements rather than replaces teacher-led instruction. The next chapter discusses these findings in relation to existing theory and previous research.

5. Discussion

The findings of this systematic literature review provide substantial evidence for the significant role of artificial intelligence in English language learning, while also revealing important challenges that must be addressed for responsible integration. The results demonstrate that AI applications, particularly generative AI chatbots, automated writing-evaluation tools, speech-recognition systems, and adaptive learning platforms, support personalized instruction, immediate feedback, writing development, speaking practice, vocabulary acquisition, pronunciation improvement, learner motivation, and autonomous learning. However, the findings also reveal persistent concerns including inaccurate or inappropriate feedback, learner dependency, academic dishonesty, privacy risks, algorithmic bias, cultural limitations, inadequate teacher training, and unequal access to digital resources.

5.1 AI and Personalized Learning

The strong association between AI use and personalized learning confirms the potential of adaptive technologies to address the diverse needs of English language learners. Traditional English language classrooms often provide the same content and activities to learners with different abilities, resulting in some students finding material too difficult while others remain insufficiently challenged. AI can address this problem by offering differentiated content and allowing learners to repeat activities as required (Aljohani, 2026). This finding is consistent with learner-centred perspectives that emphasize the importance of adapting instruction to individual needs, proficiency levels, learning speed, and preferences (Holmes & Miao, 2023).

However, personalization should not be understood as complete automation of instruction. AI systems may identify linguistic patterns, but they may not fully understand learners' emotional, cultural, social, or educational circumstances (Zhang & Liu, 2025). Therefore, teachers remain essential in interpreting learner needs and selecting appropriate learning objectives. Personalization is most effective when AI recommendations are reviewed and contextualized by teachers, maintaining the human element in educational decision-making (Esmailzadeh et al., 2025).

5.2 AI and Language Skill Development

The finding that AI had the strongest reported influence on writing, speaking, vocabulary, grammar, and pronunciation is consistent with previous research on AI-assisted language instruction (Wang et al., 2026). Writing tools provided immediate feedback on sentence structure, grammar, vocabulary, organization, and coherence, enabling learners to produce several drafts and revise their texts more efficiently. The positive effect on writing can be explained by the continuous feedback cycle provided by AI: learners write, receive feedback, revise their work, and compare different versions of the text (Ibrahim Brian, 2025). This process may improve linguistic accuracy and awareness. Nevertheless, students may become dependent on automated correction if they accept suggestions without understanding the underlying language rules. For this reason, AI feedback should be used as a starting point for reflection rather than as a final answer (Yan et al., 2025).

AI also supported speaking development by creating opportunities for repeated interaction. Many



English learners have limited opportunities to speak English outside the classroom. Conversational AI can provide simulated dialogues, pronunciation activities, role plays, and question-answer tasks (Albedah, 2025). The reviewed studies indicated that these activities improved learners' fluency, confidence, vocabulary use, and willingness to communicate. However, AI-generated conversation may not fully reproduce authentic human communication. Chatbots may provide predictable responses, misunderstand context, or fail to respond appropriately to humour, irony, emotion, and culturally specific expressions (Jantakoon et al., 2025). Therefore, AI-based speaking practice should supplement, rather than replace, peer interaction, teacher communication, group discussion, and authentic communicative activities.

5.3 AI, Motivation, and Learner Autonomy

The positive relationship between AI use, learner motivation, engagement, and autonomy confirms the potential of AI to address affective factors in language learning (Kundu & Bej, 2025). AI applications offered immediate responses, interactive activities, flexible access, and opportunities for private practice, encouraging learners to spend more time practising English and reducing anxiety about making mistakes (Kundu & Bej, 2025). The low-pressure nature of AI-supported practice was particularly valuable for students who were reluctant to speak in front of teachers or classmates. This finding supports the view that affective factors are closely related to language-learning outcomes (Torres & Kahveci, 2025).

AI also encouraged learner autonomy by allowing students to control the time, pace, and content of their learning (Bonyadi et al., 2025). Learners could independently ask questions, obtain examples, practice pronunciation, and revise writing. Nevertheless, the findings suggest that AI does not automatically create autonomous learners. Autonomy requires self-regulation, goal setting, critical evaluation, and responsibility. Recent evidence indicates that AI may assist students in completing language tasks without necessarily developing long-term self-regulated learning habits (Jeon, 2025). Teachers should therefore train learners to use AI critically and purposefully (López-Molines, 2025).

5.4 The Changing Role of the Teacher

The findings suggest that AI is changing, rather than eliminating, the role of the English language teacher. AI can perform routine tasks such as generating exercises, providing preliminary feedback, producing examples, and supporting basic language practice (Aljohani, 2026). This may reduce teachers' administrative workload and allow them to focus on communication, creativity, critical thinking, intercultural competence, and individual learner support (Holmes & Miao, 2023). At the same time, teachers are needed to evaluate the accuracy of AI output, explain complex language issues, monitor ethical use, and design meaningful tasks. Teachers also provide emotional support and social interaction that AI systems cannot fully reproduce (Zhang & Liu, 2025). The teacher's role is therefore moving from information provider to learning designer, facilitator, evaluator, and ethical guide (Esmailzadeh et al., 2025).

This transformation requires professional development. Teachers need knowledge of AI literacy, prompt design, data privacy, AI-supported assessment, bias, plagiarism, and the limitations of generative systems (Wang et al., 2026). Without sufficient training, teachers may either resist useful technologies or use AI without appropriate pedagogical control (Ibrahim Brian, 2025).

5.5 Accuracy and Reliability of AI Feedback

The most frequently reported challenge was the inaccuracy or contextual inadequacy of AI-generated feedback, identified in 24 studies (54.5%). AI systems may generate grammatically incorrect explanations, inappropriate vocabulary, fabricated references, or misleading interpretations (Yan et al., 2025). These problems are particularly serious in language learning because students may internalize incorrect forms and reproduce them in future communication (Albedah, 2025). AI tools may also privilege dominant varieties of English and inadequately represent regional accents, multilingual identities, and culturally specific expressions (Jantakoon et al., 2025). A response that is technically grammatical may nevertheless be inappropriate in a particular social context. This limitation demonstrates that linguistic competence includes pragmatic and intercultural competence, not only grammar and vocabulary (Kundu & Bej, 2025).

5.6 Academic Integrity and Overreliance

The review showed that generative AI has created significant concerns regarding plagiarism, unauthorized assistance, and the outsourcing of academic work (Kundu & Bej, 2025). Students may use AI to



generate complete essays, summaries, translations, or examination responses. If such use is not regulated, assessment may no longer accurately reflect students' language abilities (Torres & Kahveci, 2025). Overreliance can also reduce learning. When learners copy AI-generated responses without planning, drafting, revising, or evaluating, they lose opportunities to develop critical thinking and independent writing skills (Bonyadi et al., 2025). The appropriate response is not necessarily to prohibit all AI use. Instead, institutions should establish transparent policies that distinguish acceptable assistance from academic misconduct. Assessment should emphasize process, reflection, oral explanation, multiple drafts, and personalized tasks (Jeon, 2025).

5.7 Privacy and Data Protection

Privacy was another major challenge, identified in 18 studies (40.9%). AI language-learning applications may collect learners' written work, voice recordings, personal information, interaction histories, and performance data (López-Molines, 2025). If data are stored or processed without transparency, learners may be exposed to surveillance, unauthorized data use, or security breaches. A responsible approach requires institutions to examine the privacy policies of AI providers before recommending tools to students (Aljohani, 2026). Learners should be informed about what data are collected, how the data are used, and whether they can delete their information. UNESCO advocates a human-centred and rights-based approach that protects data privacy, inclusion, equity, linguistic diversity, and learner agency (Holmes & Miao, 2023).

5.8 Digital Inequality

The findings also revealed that AI may intensify existing educational inequalities (Zhang & Liu, 2025). Students with high-speed internet, modern devices, paid subscriptions, and strong digital literacy may benefit more than students who lack these resources. Differences in access can affect the distribution of learning opportunities (Esmailzadeh et al., 2025). Digital inequality is particularly important in developing educational contexts, where internet connectivity and institutional infrastructure may be inconsistent. AI integration should therefore be supported by investment in connectivity, affordable tools, teacher training, and accessible learning materials (Wang et al., 2026). Otherwise, AI may create a two-tiered language-learning system in which some learners receive continuous individualized support while others remain dependent on limited classroom resources (Ibrahim Brian, 2025).

5.9 Theoretical Implications

The findings have several theoretical implications. First, they support learner-centred and constructivist perspectives by showing that AI can provide individualized learning experiences and encourage learners to actively construct linguistic knowledge through practice and revision (Yan et al., 2025). Second, the findings are consistent with interactionist perspectives because conversational AI creates opportunities for comprehensible input, feedback, negotiation of meaning, and output (Albedah, 2025). However, the findings also suggest that AI-based interaction is not equivalent to human interaction. Meaningful language learning involves social context, emotional connection, cultural awareness, and authentic communication (Jantakoon et al., 2025). Therefore, AI should be understood as an additional interactional resource rather than a complete replacement for human communication.

The findings also support the concept of human-AI collaboration (Kundu & Bej, 2025). AI can perform computational and repetitive functions, while teachers and learners remain responsible for interpretation, judgment, creativity, ethics, and social interaction. This perspective provides a more balanced alternative to both technological optimism and technological resistance (Torres & Kahveci, 2025).

5.10 Implications for Policy and Practice

The findings of this research have several important implications for policy and practice. First, for learners, AI should be used as a learning assistant rather than as a replacement for their own thinking. Learners should verify AI-generated explanations, compare responses with textbooks and teacher feedback, and maintain personal responsibility for all submitted work. Students should also develop AI literacy, including the ability to write effective prompts, identify inaccurate information, and recognize bias (Aljohani, 2026; Holmes & Miao, 2023).

Second, for teachers, AI should be integrated into clearly defined pedagogical tasks. Appropriate activities may include guided conversation, vocabulary practice, draft revision, pronunciation rehearsal, peer



comparison, and critical evaluation of AI responses (Zhang & Liu, 2025). Teachers should explain when AI use is acceptable and should design assessments that require personal reflection, oral communication, and evidence of the learning process (Esmaeilzadeh et al., 2025).

Third, for educational institutions, policies should be developed regarding acceptable AI use, plagiarism, assessment, privacy, data security, and teacher training (Wang et al., 2026). Institutions should also evaluate AI platforms before recommending them and should ensure that students who cannot access premium tools are not disadvantaged (Ibrahim Brian, 2025).

Fourth, for policymakers, equitable access to digital infrastructure should be supported and standards promoted for privacy, transparency, accessibility, and algorithmic accountability (Yan et al., 2025). National and institutional policies should encourage innovation while protecting learners' rights and preserving the central role of teachers (Albedah, 2025).

5.11 Limitations and Future Research

This research has several limitations that should be acknowledged. First, the analysis was limited to English-language publications and may therefore exclude relevant studies published in other languages. Second, the review focused mainly on literature published between 2020 and June 2025; therefore, earlier foundational research and studies published after the search date may not be fully represented. Third, publication bias may have influenced the findings because studies reporting positive AI outcomes may be more likely to be published. Fourth, the rapid development of AI tools means that findings related to a particular application may change as the technology is updated. Finally, the heterogeneity among the included studies prevented the calculation of a single overall effect size.

Future research should address these limitations in several ways. First, longitudinal and experimental designs should be employed to examine the long-term effects of AI on English proficiency, learner autonomy, and critical thinking. Second, studies should include more school-level learners, low-resource contexts, multilingual learners, learners with disabilities, and underrepresented English varieties. Third, research should compare different AI applications and identify which pedagogical conditions produce the strongest learning outcomes. Fourth, future research should examine how AI affects not only language accuracy but also creativity, intercultural competence, pragmatic ability, critical thinking, and learner identity.

6. Conclusion

This research has investigated the impact of artificial intelligence on English language learning by synthesizing evidence from 44 selected studies published between 2020 and 2025. The findings demonstrate that AI provides important opportunities for personalized instruction, immediate feedback, writing development, speaking practice, vocabulary acquisition, pronunciation improvement, motivation, learner autonomy, and flexible access to English language resources. Generative AI chatbots, automated writing-evaluation tools, speech-recognition systems, and adaptive learning platforms were the most frequently examined applications, with research concentrated in higher education contexts and writing/speaking skills.

The evidence also shows that these benefits are conditional. AI-generated feedback may be inaccurate, culturally limited, or contextually inappropriate. Excessive dependence may weaken independent thinking and language production, while generative AI creates difficulties related to academic integrity. Privacy risks, unequal access, inadequate teacher training, algorithmic bias, and limited institutional policies further restrict the responsible integration of AI.

The central conclusion is that AI should complement rather than replace teachers, peers, and authentic communication. Its educational value depends not simply on the availability of advanced technology but on the quality of pedagogical design, teacher mediation, learner AI literacy, institutional regulation, and equitable access. A blended and human-centred model is therefore the most appropriate approach to AI-assisted English language learning.

The practical implications of this research extend to learners, teachers, educational institutions, and policymakers. For learners, AI should be used as a learning assistant while maintaining personal responsibility and critical evaluation. For teachers, AI should be integrated into structured activities that support communication and independent learning, with professional development to build AI literacy. For institutions, clear policies should be developed regarding acceptable AI use, academic integrity, assessment, privacy, and



data protection. For policymakers, equitable access to digital infrastructure should be supported and standards promoted for privacy, transparency, accessibility, and algorithmic accountability.

The study concludes that AI has significant potential to enhance English language learning, but this potential can only be realized through responsible, inclusive, and pedagogically meaningful integration. By incorporating AI as a complement to human teaching and learning, it is possible to improve language learning outcomes while maintaining the central role of teachers, authentic communication, and learners' independent thinking in English language education.

7 Recommendations

Based on the findings, the following recommendations are proposed:

1. English language teachers should integrate AI into structured activities that support, rather than replace, communication and independent learning.
2. Learners should be trained to evaluate AI-generated content for accuracy, relevance, bias, grammar, and cultural appropriateness.
3. Educational institutions should develop clear policies regarding acceptable AI use, academic integrity, assessment, privacy, and data protection.
4. Teacher-training programs should include AI literacy, prompt design, AI-supported assessment, ethical issues, and strategies for detecting overreliance.
5. AI developers should improve the cultural sensitivity, contextual accuracy, transparency, and linguistic diversity of language-learning applications.
6. Institutions and policymakers should address digital inequality by improving internet access, device availability, and affordable access to educational AI tools.
7. Future research should use longitudinal and experimental designs to examine the long-term effects of AI on English proficiency, learner autonomy, and critical thinking.
8. Future studies should include more school-level learners, low-resource contexts, multilingual learners, learners with disabilities, and underrepresented English varieties.
9. Further research should compare different AI applications and identify which pedagogical conditions produce the strongest learning outcomes.
10. Future research should examine how AI affects not only language accuracy but also creativity, intercultural competence, pragmatic ability, critical thinking, and learner identity.

Funding

No external funding was received.

Conflict of Interests

The authors declare no conflict of interests.

Acknowledgments:

The authors thank the participants and the community organizations that facilitated recruitment.

Data Availability

The datasets generated during and analysed during the current study are available from the corresponding author on reasonable request.

References

- Albedah, F. (2025). Artificial intelligence in language education: A systematic review of multilingual applications, large language models, and emerging challenges. *Language Teaching Research Quarterly*, 49, 247–268. <https://doi.org/10.32038/ltrq.2025.49.13>
- Aljohani, N. J. (2026). ChatGPT in language learning: A systematic review of applications and challenges. *Social Sciences & Humanities Open*, 13, 102357. <https://doi.org/10.1016/j.ssaho.2025.102357>
- Bonyadi, A., Kashef, S. H., & Tasouji Azari, M. (2025). The impact of emerging AI platforms on English language teaching and learning: A review of 2022–2025 literature. *Journal of New Advances in English Language Teaching and Applied Linguistics*, 7(2), 26–43.
- Esmaeilzadeh, F., Ghahari, S., & Rohani, G. (2025). Artificial intelligence in language teaching and assessment: A systematic review of research (2020–2024). *The Asian Journal of Applied Linguistics*, 9(1), 1222.



- Holmes, W., & Miao, F. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.
<https://doi.org/10.54675/EWZM9535>
- Ibrahim Brian, M. S. (2025). Reimagining English language learning: A systematic review of AI integration in classroom practice (2019–2024). *Creative Practices in Language Learning and Teaching*, 13(2), 14–29.
- Jantakoon, T., Jantakun, T., Jantakun, K., Pongpanich, W., Pasmala, R., Wannapiroon, P., & Nilsook, P. (2025). The effectiveness of artificial intelligence in English instruction for speaking and listening skills: A meta-analysis. *Contemporary Educational Technology*, 17(4), Article ep596.
<https://doi.org/10.30935/cedtech/17310>
- Jeon, E.-Y. (2025). Artificial intelligence in ESL/EFL education: Evidence from recent reviews (2024–2025). *International Journal of Learning, Teaching and Educational Research*, 24(10), 509–526.
- Kundu, A., & Bej, T. (2025a). AI in school EFL learning: A systematic review of impact pathways for engagement, achievement, and satisfaction. *Journal of Language and Education*, 11(4), 131–148.
- Kundu, A., & Bej, T. (2025b). Transforming EFL teaching with AI: A systematic review of empirical studies. *International Journal of Artificial Intelligence in Education*, 35(4), 2281–2314.
<https://doi.org/10.1007/s40593-025-00470-0>
- López-Molines, L. (2025). A systematic review of generative artificial intelligence-based tools to improve oral skills in English as a foreign language. *The EuroCALL Review*, 32(2), 183–195.
<https://doi.org/10.4995/eurocall.2025.23941>
- Noreen, A., & Bashir, M. (2025). A need analysis of law students' English literacy for occupational purposes (EOP). *International Journal of Language, Teaching and Pedagogy*, 3(1), 17–34.
<https://doi.org/10.30971/ijltp.v3i1.2927>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., 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. *BMJ*, 372, Article n71.
<https://doi.org/10.1136/bmj.n71>
- Torres, P. J., & Kahveci, Y. E. (2025). Effectiveness of artificial intelligence (AI) in language teaching. *Computers and Education: Artificial Intelligence*, 9, Article 100522.
<https://doi.org/10.1016/j.caeai.2025.100522>
- Wang, Y., Zhang, Z. J., & Zhou, H. (2026). Artificial intelligence in language learning: A twenty-year scoping review of applications, research methods, and outcomes. *Research Synthesis in Applied Linguistics*, 1–38.
- Yan, W., Li, B., & Lowell, V. L. (2025). Integrating artificial intelligence and extended reality in language education: A systematic literature review (2017–2024). *Education Sciences*, 15(8), Article 1066.
<https://doi.org/10.3390/educsci15081066>
- Zhang, S., & Liu, X. (2025). Learner emotions in AI-assisted English as a second/foreign language learning: A systematic review of empirical studies. *Frontiers in Psychology*, 16, Article 1652806.
<https://doi.org/10.3389/fpsyg.2025.1652806>