

SSTSDS-Vol.4. N1. 097

Artificial intelligence and english language learning: enhancing education in the digital age

Inteligencia artificial y aprendizaje del idioma inglés: fortalecimiento de la educación en la era digital

Autores:

Leidy Silvana Curay Moreta
Unidad Educativa 12 de Febrero
Orellana – Ecuador
eidysilcuray024@gmail.com
<https://orcid.org/0009-0004-5919-8266>

Maritza Alexandra Jara Silva
Investigadora independiente
Chimborazo – Ecuador
mary.12eden@gmail.com
<https://orcid.org/0009-0001-5380-8324>

Jeison Paul Velásquez Fala
Universidad de las Fuerzas Armadas ESPE
Riobamba – Ecuador
jeison98velasquez@gmail.com
<https://orcid.org/0009-0007-4864-973X>

Gilma Alexandra Gordillo Obregón
Investigadora independiente
Chimborazo – Ecuador
gilmagordillo@gmail.com
<https://orcid.org/0009-0009-1682-4761>

Karina Susana Pineda Chávez
Investigadora independiente
Quito – Ecuador
pkary15@gmail.com
<https://orcid.org/0009-0002-4425-9869>

Autor de correspondencia: Leidy Silvana Curay Moreta, eidysilcuray024@gmail.com

Enviado: 09-05-2026
Aceptado: 26-07-2026

Revisado: 23-05-2026
Publicado: 29-07-2026



Cómo citar este artículo:

Curay Moreta, L. S., Jara Silva, M. A., Velásquez Fala, J. P., Gordillo Obregón, G. A., & Pineda Chávez, K. S. (2026). Inteligencia artificial y aprendizaje del idioma inglés: fortalecimiento de la educación en la era digital. *Sage Sphere of Technology, Sciences, Discoveries And Society*, 4(1), 1-28. <https://doi.org/10.63688/e8hj7g36>

© 2026; Los autores. Este es un artículo en acceso abierto, distribuido bajo los términos de una licencia Creative Commons (<https://creativecommons.org/licenses/by/4.0>) que permite el uso, distribución y reproducción en cualquier medio, siempre que la obra original sea correctamente citada.



ABSTRACT

Artificial intelligence has become a valuable resource for English language teaching and learning by providing interactive practice, immediate feedback, and personalized activities. This study analyzed the contribution of artificial intelligence-based applications to English language learning in the digital age. An integrative literature review was conducted following the identification, screening, eligibility, and inclusion stages of PRISMA 2020. Sixteen studies published between 2020 and 2026 were included: seven primary studies and nine secondary or conceptual sources. The findings addressed chatbots, conversational artificial intelligence, automatic speech recognition, pronunciation, speaking, listening, writing, vocabulary, grammar, affective factors, personalized learning, and ethical implications. Results indicated that artificial intelligence tools can improve pronunciation, fluency, vocabulary, grammatical accuracy, writing, motivation, autonomy, and willingness to communicate. However, their effectiveness depends on the type of tool, learners' proficiency, intervention duration, and teacher guidance. Limitations included inaccurate responses, privacy concerns, algorithmic bias, unequal access, and excessive dependence on technology. Artificial intelligence can enhance English language learning when ethically and systematically integrated as a complement to teacher instruction and human interaction.

Keywords: artificial intelligence; English language learning; digital education; language learning.

RESUMEN

La inteligencia artificial se ha convertido en un recurso valioso para la enseñanza y el aprendizaje del inglés al proporcionar práctica interactiva, retroalimentación inmediata y actividades personalizadas. El objetivo de este estudio fue analizar la contribución de las aplicaciones basadas en inteligencia artificial al aprendizaje del inglés en la era digital. Se realizó una revisión integrativa de la literatura siguiendo las etapas de identificación, cribado, elegibilidad e inclusión de PRISMA 2020. Se incluyeron dieciséis estudios publicados entre 2020 y 2026: siete primarios y nueve secundarios o conceptuales. Los hallazgos abordaron chatbots, inteligencia artificial conversacional, reconocimiento automático del habla, pronunciación, expresión oral, comprensión auditiva, escritura, vocabulario, gramática, factores afectivos, aprendizaje personalizado e implicaciones éticas. Los resultados indicaron que estas herramientas pueden mejorar la pronunciación, fluidez, vocabulario, precisión gramatical, escritura, motivación, autonomía y disposición para comunicarse. Sin embargo, su efectividad depende del tipo de herramienta, nivel lingüístico, duración de la intervención y orientación docente. También se identificaron respuestas inexactas, problemas de privacidad, sesgos algorítmicos, acceso desigual y dependencia tecnológica. Se concluye que la inteligencia artificial puede fortalecer el aprendizaje del inglés cuando se integra ética y sistemáticamente como complemento de la enseñanza docente y la interacción humana.

Palabras clave: inteligencia artificial; aprendizaje del inglés; educación digital; aprendizaje de lenguas.



1. INTRODUCTION

English language learning has become increasingly important in contemporary education because English enables participation in academic, professional, technological, social, and intercultural environments. However, learning a foreign language remains a complex process that requires the coordinated development of listening, speaking, reading, writing, vocabulary, grammar, pronunciation, and communicative competence. Traditional classroom instruction may not always provide sufficient time for individualized assistance, immediate feedback, repeated practice, or authentic interaction. In this context, artificial intelligence has emerged as a resource capable of extending learning beyond the classroom and offering more flexible and personalized experiences. AI applications have evolved from rule-based chatbots to systems supported by natural language processing, machine learning, speech recognition, and generative artificial intelligence (Fryer et al., 2020; Liang et al., 2023).

In English language education, artificial intelligence encompasses conversational systems, speech-recognition applications, automated writing and feedback tools, adaptive platforms, intelligent personal assistants, and generative applications. Rather than treating each tool as an isolated innovation, these technologies can be understood as a group of digital resources that analyze learner responses, generate explanations and examples, recommend activities, identify linguistic errors, and adapt content to individual performance. They may also allow learners to request clarification, repeat tasks, review feedback, and progress at their own pace, while teachers retain responsibility for designing activities and interpreting the information produced by the systems (Kohnke et al., 2023; Lin & Chang, 2020).

Recent empirical and review-based research suggests that structured AI-mediated instruction may support several dimensions of English learning. Reported outcomes include improvements in oral performance, pronunciation, vocabulary, grammar, writing, language achievement, motivation, autonomy, and self-regulated learning. Nevertheless, these effects are not uniform and depend on factors such as the type of application, learners' proficiency, the duration of the intervention, the quality of feedback, and the pedagogical guidance provided by teachers (Du & Daniel, 2024; El Shazly, 2021; Wei, 2023). AI may also influence learners' affective experiences by providing a comparatively private environment in which they can practise, make mistakes, and repeat responses without the immediate social



pressure of classroom participation. At the same time, unclear feedback, unfamiliar interfaces, or recognition failures may produce frustration or maintain anxiety, showing that motivational and emotional outcomes vary according to learner characteristics, technological design, and implementation conditions (AlTwijri & Alghizzi, 2024; Lai & Lee, 2024).

Despite its educational potential, artificial intelligence also presents ethical, pedagogical, and technological limitations. AI-generated information may be inaccurate, biased, decontextualized, culturally inappropriate, or unsuitable for a learner's proficiency level. Automated systems may not adequately interpret humor, emotions, implied meaning, social relationships, or the cultural dimensions of authentic communication. Other concerns include personal-data protection, unequal access to devices and connectivity, academic dishonesty, excessive reliance on automated assistance, and the possible reduction of independent thinking. Consequently, AI should function as a complementary resource rather than as a replacement for teachers or human interaction. Effective integration requires clear objectives, carefully designed tasks, digital literacy, ethical guidelines, teacher supervision, and continued opportunities for communication with real people (Ji et al., 2023; Tang & Zhang, 2026).

Although the literature generally reports favorable outcomes, the available evidence remains heterogeneous. Results vary according to the technology employed, learners' age and proficiency, the language skill addressed, the educational context, the intervention period, and the assessment procedures. A recent meta-analysis identified a positive overall effect of artificial intelligence on English learning achievement, while also showing that educational level, sample characteristics, and implementation conditions may influence effectiveness (Xu & Wang, 2024). In addition, previous studies often focus on specific tools or isolated outcomes, which makes it difficult to obtain an integrated understanding of how artificial intelligence contributes simultaneously to linguistic development, affective factors, personalized instruction, autonomy, engagement, and access to additional practice.

Therefore, a broader synthesis is needed to examine the contribution of artificial intelligence to English language learning across different technologies, skills, outcomes, and educational contexts. This study aims to analyze the contribution of artificial intelligence-based applications to English language learning in the digital age. Accordingly, it addresses the following research question: How does the integration of artificial intelligence



influence the English language learning process and the educational experiences of EFL learners in the digital age?

2. THEORETICAL FRAMEWORK

Artificial intelligence in English language learning should not be understood only as the automation of educational activities. It represents a learning environment in which technology can analyze performance, provide support, and adapt instruction to learners' needs. Ouyang and Jiao (2021) describe three educational AI paradigms: AI-directed learning, in which technology controls instruction; AI-supported learning, in which learners collaborate with the system; and AI-empowered learning, in which learners use AI to make decisions and manage their progress. For English education, the last two paradigms are particularly relevant because they maintain learner participation, reflection, and autonomy instead of treating students as passive recipients of information.

Conversational AI and language interaction

Chatbots provide language practice through written or spoken interaction. They can present questions, simulate situations, recommend content, and respond immediately to learners. Nevertheless, not every chatbot possesses advanced artificial intelligence. Many systems still operate through predetermined responses and decision trees. Therefore, their educational value depends more on the quality of their activities, feedback, accessibility, and conversational design than on the simple use of the term *AI* (Smutny & Schreiberova, 2020). Effective chatbots should encourage learners to produce meaningful language rather than merely select fixed answers.

Interaction with chatbots can also be explained through the concept of negotiation of meaning. This process occurs when learners identify a communication problem, request clarification, reformulate an expression, or modify their original response to make it understandable. Yin and Satar (2020) found that chatbot interaction could generate opportunities for clarification and modified language production. However, the results depended on proficiency level: lower-level learners benefited from structured pedagogical chatbots, whereas more advanced learners sometimes considered their responses limited or repetitive. Consequently, AI activities must correspond to learners' linguistic level and communicative needs.



Speech recognition and oral-language development

Intelligent personal assistants such as Alexa, Siri, and Google Assistant use automatic speech recognition to process spoken language. In English learning, these systems can extend speaking and listening practice beyond the classroom. They may also offer indirect pronunciation feedback: when the application recognizes a learner's speech correctly, this can indicate that the message is sufficiently intelligible. Dizon (2020) explains that such assistants can support listening comprehension and spoken production by exposing learners to oral input and requiring them to formulate understandable responses.

Nevertheless, speech recognition does not always evaluate pronunciation accurately. Moussalli and Cardoso (2020) observed that intelligent assistants may experience difficulty processing non-native accents, especially when accentedness is strong. Thus, an unsuccessful interaction does not necessarily mean that the learner's pronunciation is incorrect; it may also indicate a technological limitation. For this reason, automatic speech recognition should complement, rather than replace, feedback from teachers and real interlocutors.

Automated analysis and English writing

AI can also support writing through computational tools that examine vocabulary, cohesion, sentence structure, readability, and discourse organization. Petchprasert (2021) demonstrated that automated textual analysis can provide information about linguistic and discourse characteristics in EFL students' compositions. Such information can help learners identify patterns in their writing, compare drafts, and monitor their development. However, numerical indicators should not be interpreted as complete evaluations of writing quality because creativity, purpose, argumentation, cultural meaning, and audience awareness still require human judgment.

Overall, this theoretical framework views AI as a pedagogical scaffold rather than an independent teacher. Its contribution to English learning depends on five central conditions: learner agency, meaningful interaction, level-appropriate activities, understandable feedback, and teacher mediation. AI becomes educationally valuable when learners analyze its responses, correct their production, make decisions, and apply their knowledge in authentic communication. Therefore, research on AI and English learning should examine not only improvements in test scores but also how learners interact with feedback, modify



language, develop autonomy, and transfer digitally practiced skills to real communicative contexts.

3. METHODOLOGY

Research Design

This study was conducted as an integrative literature review with a systematic search process to analyze the contribution of artificial intelligence-based applications to English language learning in the digital age. An integrative review allows empirical, mixed-methods, case-based, conceptual, and review-based evidence to be compared and synthesized within a common analytical framework (Lubbe et al., 2020).

A qualitative descriptive-analytical approach was used to organize and interpret the selected publications. The analysis considered the influence of artificial intelligence on speaking, listening, writing, pronunciation, vocabulary, grammar, language achievement, motivation, confidence, anxiety, learner autonomy, engagement, and self-regulated learning.

The corpus included primary evidence, such as empirical, case-based, mixed-methods, design-based, and applied studies, as well as secondary and conceptual evidence, including systematic reviews, a meta-analysis, bibliographic analyses, literature reviews, and conceptual or applied articles. Primary and secondary sources were examined separately to reduce the possibility of counting the same empirical findings more than once.

Information Sources and Search Strategy

The bibliographic search was conducted in Scopus, Web of Science, ScienceDirect, Taylor & Francis Online, SpringerLink, SAGE Journals, ERIC, SciELO, Redalyc, and Google Scholar. The searches were conducted and finalized on 1 July 2026.

Searches were performed in title, abstract, and keyword fields whenever these options were available. In platforms with limited advanced-search functions, the general search field was used. The strategy combined terms related to artificial intelligence, English language education, and EFL or ESL learners through the Boolean operators AND and OR.

The general search equation was:

(“artificial intelligence” OR “AI-based applications” OR “AI chatbots” OR “conversational artificial intelligence” OR “generative AI” OR “automatic speech recognition” OR



“automated feedback”) AND (“English language learning” OR “English language teaching” OR “EFL learning” OR “ESL learning”)

The equation was adapted to the fields, operators, filters, and interface characteristics of each database. The complete platform-specific equations, search fields, filters, execution date, modifications, and numbers of records identified are presented in Appendix A.

Google Scholar was used as a complementary academic source. Six searches were conducted, the results were ordered by relevance, and the first 100 results from each equation were reviewed. Because the number generated by each individual Google Scholar equation was not separately recorded during the original search, only the final number of 10 unique records retained after screening and overlap removal is reported. This incomplete equation-level documentation is acknowledged as a limitation of exact search reproducibility.

All references were organized in a Microsoft Excel spreadsheet. Duplicate records were identified by comparing authorship, title, publication year, journal, and DOI. Records containing incomplete or inconsistent information were manually verified before removal.

Eligibility Criteria

Publications were included when they met the following criteria: publication between 2020 and 2026; peer-reviewed empirical, experimental, mixed-methods, conceptual, or review design; direct relationship with artificial intelligence and English language learning or teaching; focus on EFL or ESL educational contexts; examination of linguistic, academic, affective, motivational, or pedagogical outcomes; publication in English or Spanish; accessible full text; DOI or verifiable bibliographic information; and sufficiently reported objectives, methods, evidence, or conclusions.

Publications were excluded when they were duplicates, editorials, opinion articles, informal essays, or nonacademic documents; were unrelated to English language education; addressed educational technology without a clearly identifiable artificial intelligence component; were published outside the established period; lacked accessible full text or reliable bibliographic information; or did not clearly report their methodology, participants, evidence, or results.

Study Selection Process

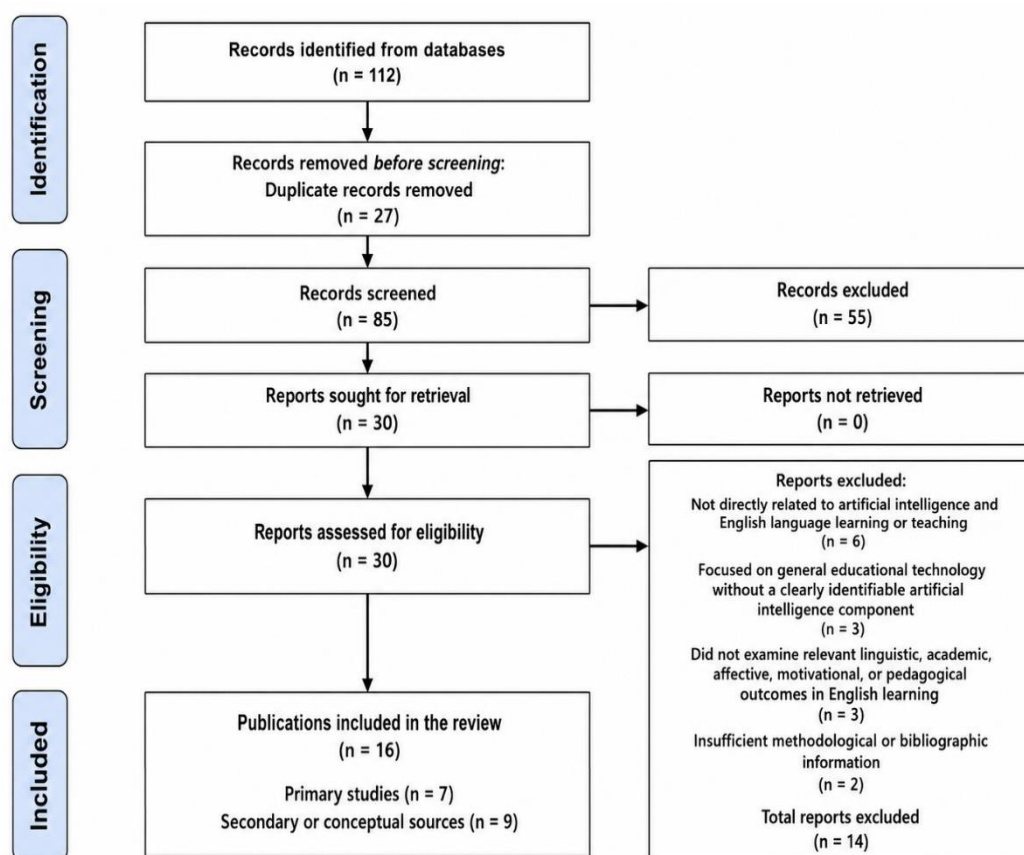
The selection process followed the identification, screening, eligibility, and inclusion stages established by the PRISMA 2020 statement (Page et al., 2021).



A total of 112 records were identified through the selected databases and academic platforms. After removing 27 duplicate records, 85 publications remained for title and abstract screening. Fifty-five records were excluded because they did not correspond directly to the research objectives or eligibility criteria. Thirty full-text publications were retrieved and assessed for eligibility, of which 14 were excluded. The final corpus consisted of 16 publications: seven primary studies and nine secondary or conceptual sources.

Figure 1

PRISMA 2020 Flow Diagram of the Study-Selection Process



Note. The study-selection process followed the identification, screening, eligibility, and inclusion stages established by the PRISMA 2020 statement. A total of 112 records were identified, 27 duplicate records were removed, and 85 records were screened. After excluding 55 records during title and abstract screening, 30 full-text reports were sought and successfully retrieved. These 30 reports were assessed for eligibility, and 14 were excluded



for not meeting the established criteria. The final corpus consisted of 16 publications: seven primary studies and nine secondary or conceptual sources.

Data Extraction and Analysis

A standardized matrix was used to extract authorship, publication year, educational context, research design, participant or corpus characteristics, educational level, type of artificial intelligence application, English language skill or educational variable analyzed, main findings, reported limitations, bibliographic identifier, and overall contribution to English language education.

Each publication was compared with its original full text. When incomplete or inconsistent information was identified, the source was reviewed again before the data were incorporated into the matrix.

The final corpus was divided into two evidence groups. Primary evidence included empirical, case-based, design-based, mixed-methods, and applied studies. Secondary or conceptual evidence included systematic reviews, literature reviews, meta-analyses, bibliographic analyses, and conceptual articles. Both groups were analyzed separately to avoid treating findings summarized in secondary sources as additional independent empirical effects.

The evidence was analyzed through thematic categorization following Braun and Clarke (2021). The findings were organized into eight categories:

- Chatbots and conversational artificial intelligence.
- Automatic speech recognition and pronunciation.
- AI-supported speaking and listening development.
- Automated writing evaluation and linguistic feedback.
- Vocabulary, grammar, and general language achievement.
- Motivation, anxiety, confidence, autonomy, and engagement.
- Personalized and self-regulated English learning.
- Pedagogical opportunities, ethical concerns, and technological limitations.

The categories were not mutually exclusive because some publications examined more than one technological application, language skill, or educational outcome. A statistical meta-analysis was not conducted because the selected publications differed in their research designs, participants, educational levels, artificial intelligence tools, variables, interventions,



and assessment procedures. Therefore, the evidence was synthesized descriptively and thematically.

Methodological Quality Assessment

Due to the methodological diversity of the selected publications, a common eight-criterion framework was used to evaluate elements applicable across the different research designs.

The criteria were:

C1: Clarity of objectives.

C2: Adequacy of the theoretical foundation.

C3: Methodological transparency.

C4: Appropriateness of the research design.

C5: Description of participants, corpus, or evidence.

C6: Clarity of the analysis.

C7: Correspondence between findings and conclusions.

C8: Acknowledgment of limitations.

Each criterion was scored as follows:

0: The criterion was not met.

1: The criterion was partially met.

2: The criterion was fully met.

The maximum possible score was 16 points. The percentage was calculated by dividing the obtained score by 16 and multiplying the result by 100. Publications obtaining 80% or more were classified as high quality, those obtaining between 60% and 79% as moderate quality, and those below 60% as low quality.

The scores were used to guide the interpretation and weighting of the evidence rather than to establish direct methodological equivalence among publications with different research designs. Although the same eight criteria were applied to all included publications, the resulting scores should not be interpreted as demonstrating equivalence among empirical studies, mixed-methods research, systematic reviews, meta-analyses, bibliographic analyses, and conceptual articles. The framework was used only to assess shared reporting and methodological elements.

Table 1

Results of the Methodological Quality Assessment



No	Author and year	Type of document	Total										Percentage	Quality level
			C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	1 score		
1	AlTwijri and Alghizzi (2024)	Systematic review	2	2	2	2	2	2	2	1		15/16	93.75%	High
2	Dizon (2020)	Empirical study	2	2	2	2	2	2	2	1		15/16	93.75%	High
3	Du and Daniel (2024)	Systematic literature review	2	2	2	2	2	2	2	1		15/16	93.75%	High
4	El Shazly (2021)	Case study	2	2	1	2	2	1	2	1		13/16	81.25%	High
5	Fryer et al. (2020)	Review/conceptua l article	2	2	1	1	1	1	2	1		11/16	68.75%	Moderat e
6	Ji et al. (2023)	Systematic review	2	2	2	2	2	2	2	2		16/16	100.00%	High
7	Kohnke et al. (2023)	Conceptual/applie d article	2	2	1	1	1	2	2	1		12/16	75.00%	Moderat e
8	Lai and Lee (2024)	Systematic review	2	2	2	2	2	2	2	2		16/16	100.00%	High
9	Liang et al. (2023)	Bibliographic analysis and systematic review	2	2	2	2	2	2	2	1		15/16	93.75%	High



No	Author and year	Type of document	C								Total score	Percentage	Quality level
			1	2	3	4	5	6	7	8			
10	Lin and Chang (2020)	Mixed-methods study	2	2	2	2	2	2	2	1	15/16	93.75%	High
11	Moussalli and Cardoso (2020)	Empirical study	2	2	2	2	2	2	2	1	15/16	93.75%	High
12	Petchprasert (2021)	Empirical study	2	2	1	2	2	1	2	1	13/16	81.25%	High
13	Tang and Zhang (2026)	Systematic review	2	2	2	2	2	2	2	2	16/16	100.00%	High
14	Wei (2023)	Empirical study	2	2	2	2	2	2	2	1	15/16	93.75%	High
15	Xu and Wang (2024)	Meta-analysis	2	2	2	2	2	2	2	2	16/16	100.00%	High
16	Yin and Satar (2020)	Empirical study	2	2	2	2	2	1	2	1	14/16	87.50%	High

Note. C1 = clarity of objectives; C2 = adequacy of the theoretical foundation; C3 = methodological transparency; C4 = appropriateness of the research design; C5 = description of participants, corpus, or evidence; C6 = clarity of the analysis; C7 = correspondence between findings and conclusions; and C8 = acknowledgment of limitations. Each criterion



was scored as 0 when it was not met, 1 when it was partially met, and 2 when it was fully met. The maximum score was 16 points. Publications obtaining 80% or more were classified as high quality, those obtaining between 60% and 79% as moderate quality, and those below 60% as low quality.

The methodological quality assessment showed that 14 of the 16 included publications were classified as high quality, while two publications were classified as moderate quality. No publication was classified as low quality. The highest scores were obtained by Ji et al. (2023), Lai and Lee (2024), Tang and Zhang (2026), and Xu and Wang (2024), each reaching 16 out of 16 points. Fryer et al. (2020) and Kohnke et al. (2023) were classified as moderate quality, mainly because their conceptual or applied designs provided less methodological and participant-related information than empirical studies and systematic reviews.

The high-quality publications generally presented clearly defined objectives, appropriate theoretical foundations, transparent methodological procedures, sufficiently described evidence, and consistency between their findings and conclusions. These publications provided comparatively stronger support for the educational use of chatbots, speech-recognition systems, generative artificial intelligence, automated writing tools, and personalized feedback.

The two moderate-quality publications were retained because they contributed relevant conceptual and pedagogical perspectives. Their lower scores were mainly associated with limited methodological detail and the absence of participant-related information, which is expected in conceptual or applied articles.

The quality assessment was not used as an automatic exclusion criterion. Instead, it supported the interpretation and weighting of the evidence while recognizing the methodological differences among empirical studies, systematic reviews, meta-analyses, mixed-methods studies, bibliographic analyses, and conceptual publications.

Review Process and Protocol

Study selection, full-text assessment, data extraction, and methodological quality evaluation were conducted by one reviewer. The extracted information was verified through repeated comparison with the original publications. Because no second independent reviewer participated, no inter-reviewer disagreement-resolution procedure or inter-rater agreement calculation was performed.



The review was not prospectively registered, and no public protocol was published before the search. Nevertheless, the databases, execution date, general search equation, eligibility criteria, selection process, data-extraction variables, analytical categories, quality-assessment criteria, and platform-specific strategies were documented to improve transparency and reproducibility. The complete search strategies are provided in Appendix A.

4. RESULTS

General Characteristics of the Included Publications

The final corpus consisted of 16 publications published between 2020 and 2026. Seven were classified as primary studies and nine as secondary or conceptual sources. The publications included empirical studies, mixed-methods research, systematic reviews, a meta-analysis, bibliographic analyses, and conceptual or applied articles.

The included studies examined conversational artificial intelligence, automatic speech recognition, intelligent personal assistants, generative AI, automated writing tools, adaptive platforms, and digital feedback systems. The main outcomes were related to oral communication, pronunciation, listening, writing, vocabulary, grammar, language achievement, motivation, anxiety, autonomy, engagement, personalization, and self-regulated learning.

Because the studies differed in design, participants, educational level, type of AI application, intervention duration, and assessment procedures, the results were synthesized descriptively. The findings were organized into four broader thematic blocks.

Table 2

General Matrix of the Sources Included in the Integrative Literature Review

No.	Study	Type of source	Main focus	Thematic category or categories
1	AlTwijri and Alghizzi (2024)	Systematic review	Integration of AI in EFL classes and its influence on learners' affective factors	Motivation, anxiety, confidence, and engagement
2	Dizon (2020)	Empirical study	Use of intelligent personal assistants for developing L2	Automatic speech recognition and pronunciation; AI-



No.	Study	Type of source	Main focus	Thematic category or categories
			listening and speaking	supported speaking and listening development
3	Du and Daniel (2024)	Systematic review	AI-powered chatbots for EFL speaking practice	Chatbots and conversational artificial intelligence; AI-supported speaking and listening development
4	El Shazly (2021)	Case study	Effects of AI on English-speaking anxiety and speaking performance	AI-supported speaking and listening development; motivation, anxiety, confidence, autonomy, and engagement
5	Fryer et al. (2020)	Review/conceptual article	Current use and future directions of bots for language learning	Chatbots and conversational artificial intelligence; pedagogical opportunities, ethical concerns, and technological limitations
6	Ji et al. (2023)	Systematic review	Conversational AI in language education and collaboration with teachers	Chatbots and conversational artificial intelligence; pedagogical opportunities, ethical concerns, and technological limitations
7	Kohnke et al. (2023)	Conceptual/applied article	Use of ChatGPT in language teaching and learning	Chatbots and conversational artificial intelligence; personalized and self-regulated English learning; pedagogical opportunities, ethical concerns, and technological limitations



No.	Study	Type of source	Main focus	Thematic category or categories
8	Lai and Lee (2024)	Systematic review	Trends, tools, methods, outcomes, and antecedents of conversational AI in ELT	Chatbots and conversational artificial intelligence; motivation, anxiety, confidence, autonomy, and engagement
9	Liang et al. (2023)	Bibliographic analysis and systematic review	Roles and research trends of AI in language education	Pedagogical opportunities, ethical concerns, and technological limitations
10	Lin and Chang (2020)	Mixed-methods study	Use of a chatbot to improve post-secondary students' writing skills	Chatbots and conversational artificial intelligence; automated writing evaluation and linguistic feedback
11	Moussalli and Cardoso (2020)	Empirical study	Ability of intelligent personal assistants to understand accented L2 speech	Automatic speech recognition and pronunciation
12	Petchprasert (2021)	Empirical study	Use of an automated analysis tool to evaluate EFL students' writing	Automated writing evaluation and linguistic feedback
13	Tang and Zhang (2026)	Systematic review	AI technologies, tools, and their effects on EFL learning	Vocabulary, grammar, and general language achievement; personalized and self-regulated English learning; pedagogical opportunities, ethical concerns, and technological limitations
14	Wei (2023)	Empirical study	Influence of AI-based instruction on English achievement,	Vocabulary, grammar, and general language achievement; motivation, anxiety, confidence,



No.	Study	Type of source	Main focus	Thematic category or categories
			motivation, and self-regulated learning	autonomy, and engagement; personalized and self-regulated English learning
15	Xu and Wang (2024)	Meta-analysis	Effectiveness of AI in improving English-language learning achievement	Vocabulary, grammar, and general language achievement
16	Yin and Satar (2020)	Empirical study	EFL learner interaction with chatbots and negotiation of meaning	Chatbots and conversational artificial intelligence; AI-supported speaking and listening development

Note. The final corpus consisted of 16 publications directly related to artificial intelligence and English language learning. Seven were classified as primary studies and nine as secondary or conceptual sources.

Oral Communication, Pronunciation, and Listening Comprehension

The reviewed evidence showed that artificial intelligence increased opportunities for oral practice, repetition, and interaction beyond regular classroom time. Conversational systems allowed learners to respond to prompts, reformulate messages, and request clarification, which supported negotiation of meaning and communicative practice. However, the quality of these interactions depended on learner proficiency and system flexibility, since more advanced learners sometimes considered chatbot responses repetitive or limited (Lin & Chang, 2020; Yin & Satar, 2020).

Automatic speech-recognition systems and intelligent personal assistants were associated with pronunciation, listening, and spoken production. Learners often modified their speech when the system did not recognize an utterance, using the response as indirect feedback. Nevertheless, recognition failures were not always caused by incorrect pronunciation, because some systems had difficulty processing non-native accents (Dizon, 2020; Moussalli & Cardoso, 2020).

Writing, Vocabulary, Grammar, and Language Achievement



The selected publications indicated that AI-supported tools could assist learners in revising written production, identifying linguistic patterns, and practising vocabulary and grammar. Automated writing analysis provided information about vocabulary use, sentence structure, cohesion, readability, and discourse organization, helping learners compare drafts and identify areas for revision (Petchprasert, 2021).

However, automated indicators did not represent a complete assessment of writing quality. Creativity, argumentation, audience awareness, communicative purpose, and cultural meaning still required human judgment. The usefulness of automated feedback therefore depended on whether learners understood the information and used it critically.

Vocabulary and grammar were generally integrated into conversational, generative, and adaptive activities. These applications increased access to examples, exercises, explanations, and corrections, although results varied according to the accuracy of the generated content and the quality of the task.

General English achievement showed a favorable but heterogeneous tendency. Improvements depended on educational level, learner characteristics, intervention duration, application type, and assessment procedures. The evidence therefore supported a conditional rather than universal positive effect of AI on language learning (Wei, 2023; Xu & Wang, 2024).

Motivation, Anxiety, Autonomy, and Self-Regulation

Artificial intelligence was associated with increased motivation, engagement, confidence, and willingness to participate in some learning contexts. Learners could practise privately, repeat tasks, and make mistakes without the immediate pressure of classroom evaluation. These conditions appeared to reduce fear of negative judgment for some students.

Nevertheless, the effects on anxiety were mixed. Unfamiliar interfaces, recognition failures, technical difficulties, and confusing corrections could maintain or increase anxiety. Therefore, affective outcomes varied according to learner characteristics, technological design, and pedagogical implementation (AlTwijri & Alghizzi, 2024; Lai & Lee, 2024).

Personalization, Teacher Mediation, and Technological Limitations

Personalization was one of the most frequently reported contributions of artificial intelligence. AI applications could adapt activities, recommend practice, generate



explanations, and provide feedback according to learner performance. This allowed students to work at different speeds and focus on specific needs.

However, personalization was not always accurate. Generated content could be unsuitable for the learner's level, contain errors, or lack sufficient context. The quality of personalized support therefore depended on technological reliability and appropriate task design.

Teacher mediation remained essential. AI applications were more effective when teachers selected suitable tools, established clear objectives, helped learners interpret feedback, and verified generated content. The reviewed evidence supported the use of AI as a complementary educational resource rather than as a substitute for teacher instruction (Ji et al., 2023; Tang & Zhang, 2026).

The main limitations included inaccurate responses, algorithmic bias, difficulty interpreting non-native accents, repetitive interaction, privacy concerns, unequal access to technology, academic dishonesty, and excessive dependence on automated systems.

Overall, the findings showed that AI can expand opportunities for language practice, feedback, motivation, autonomy, and personalized learning. However, its effects depend on learner characteristics, educational context, technological quality, intervention duration, and teacher guidance.

5. DISCUSSION

This integrative review examined the contribution of artificial intelligence to English language learning and the educational experiences of EFL learners. The findings indicate that AI can expand access to practice, feedback, revision, interaction, and independent learning. However, its educational value does not result from the technology alone, but from the way it is integrated into meaningful activities and used critically by learners and teachers.

Linguistic Performance

The reviewed evidence suggests that artificial intelligence can support oral communication, pronunciation, listening, writing, vocabulary, grammar, and general English achievement. Its main contribution lies in increasing opportunities to practise, repeat tasks, identify difficulties, and modify linguistic production.

In oral communication, conversational systems encouraged learners to respond, reformulate expressions, and request clarification. Yin and Satar (2020) found that chatbot interaction



could promote negotiation of meaning, while Lin and Chang (2020) associated these tools with additional opportunities for English practice. These findings indicate that conversational AI is most useful when it encourages meaningful production rather than repetitive or predetermined exchanges.

Pronunciation and listening results were favorable but should be interpreted cautiously. Dizon (2020) showed that intelligent personal assistants can extend oral and listening practice, whereas Moussalli and Cardoso (2020) found that automatic speech-recognition systems may have difficulty processing non-native accents. Therefore, recognition failure does not necessarily demonstrate incorrect pronunciation. Automated responses can function as indirect feedback, but broader assessment by teachers and human listeners remains necessary.

In writing, automated analysis helped identify vocabulary use, sentence structure, cohesion, readability, and discourse organization. Petchprasert (2021) demonstrated that these indicators can assist learners in comparing drafts and directing attention toward areas requiring revision. Nevertheless, automated systems cannot fully evaluate creativity, argumentation, audience awareness, cultural meaning, or communicative purpose. Their contribution is therefore stronger for targeted linguistic feedback than for complete judgments of writing quality.

General English achievement showed a positive but heterogeneous tendency. Wei (2023) linked AI-supported instruction with achievement, motivation, and self-regulation, while Xu and Wang (2024) identified a positive overall effect. However, the evidence also showed variation among learner groups and educational contexts. AI is therefore best understood as a pedagogical scaffold whose effectiveness depends on clear objectives, suitable tasks, reliable feedback, and active learner participation.

Affective Factors

Artificial intelligence was also associated with motivation, confidence, engagement, and willingness to communicate. Learners could practise privately, repeat responses, and make mistakes without the immediate pressure of classroom evaluation. This may be especially valuable for students who fear negative judgment or hesitate to participate publicly.

AlTwijri and Alghizzi (2024) identified affective benefits in AI-supported language learning. However, Lai and Lee (2024) showed that technological unfamiliarity, recognition



failures, confusing corrections, and doubts about system accuracy can produce anxiety or frustration. These findings are not necessarily contradictory. AI may strengthen confidence when learners understand the system and receive useful feedback, but it may increase anxiety when interactions are unclear or unreliable.

Motivation should also be distinguished from temporary interest. A new application may initially encourage participation, but repetitive conversations or predictable activities can reduce engagement over time. Consequently, short-term enthusiasm should not automatically be interpreted as sustained motivation.

Affective and linguistic outcomes may also influence each other. Increased confidence can encourage additional practice, while observable progress can strengthen motivation. Conversely, repeated technological failure may reduce participation and limit opportunities for language development.

Personalization and Learner Autonomy

Personalization was one of the principal contributions identified in the literature. AI-supported systems can recommend activities, generate explanations, adapt practice, and provide feedback according to learner responses. These possibilities allow students to work at different speeds and focus on particular difficulties.

This finding is consistent with the AI-supported and AI-empowered learning paradigms described by Ouyang and Jiao (2021), in which technology assists learners in making decisions and managing their progress. Wei (2023) also connected AI-supported learning with motivation and self-regulation.

Nevertheless, personalized content is not always accurate or appropriate. A system may produce an individualized response that is too simple, too difficult, incorrect, or insufficiently contextualized. Therefore, personalization requires critical evaluation rather than automatic acceptance.

Autonomy should not be understood simply as learning without a teacher. It involves planning, selecting strategies, monitoring performance, evaluating information, and assuming responsibility for decisions. AI can support these processes, but it may weaken autonomy when learners depend on complete generated answers or accept every recommendation without reflection.



This tension is particularly evident in generative AI. The same tool can be used to request explanations, revise original work, compare examples, or produce an entire assignment. Kohnke et al. (2023) emphasized its pedagogical possibilities, but these possibilities depend on digital literacy, linguistic competence, and clear academic expectations.

Pedagogical and Ethical Challenges

The effects of artificial intelligence were influenced by interconnected factors, including learner proficiency, task design, educational context, technological reliability, feedback quality, digital competence, intervention duration, and teacher guidance. These conditions help explain why similar tools may produce different outcomes across studies.

Teacher mediation remains essential. Ji et al. (2023) and Tang and Zhang (2026) highlighted both the opportunities and limitations of AI in education. Teachers are responsible for selecting appropriate technologies, defining objectives, designing activities, verifying generated content, and helping learners interpret feedback.

The evidence does not support replacing teachers or authentic interaction. AI provides availability, repetition, and rapid responses, but it cannot fully reproduce empathy, cultural understanding, contextual judgment, or spontaneous social communication. Human guidance remains especially important when systems provide inaccurate information, misunderstand accents, or fail to recognize communicative intention.

6. CONCLUSIONS

This integrative review analyzed the contribution of artificial intelligence-based applications to English language learning in the digital age. The evidence indicates that AI can support linguistic performance, affective engagement, personalized learning, and access to additional practice. Its most consistent contribution lies in expanding opportunities for learners to produce language, repeat tasks, receive feedback, revise their work, and manage learning beyond conventional classroom time. However, the evidence does not demonstrate a uniform effect across all learners, technologies, skills, and educational contexts.

The principal benefits included support for oral communication, pronunciation, listening, writing revision, vocabulary and grammar practice, motivation, confidence, autonomy, and self-regulation. These benefits were strongest when AI was incorporated into purposeful activities that required learners to interpret feedback, modify their production, and make



decisions. Appropriate implementation requires reliable tools, clear learning objectives, understandable feedback, critical digital literacy, teacher mediation, and opportunities to transfer digitally practised skills to authentic human communication.

The conclusions are limited by the methodological heterogeneity of the 16 publications, the combination of seven primary studies with nine secondary or conceptual sources, the participation of a single reviewer, the absence of prospective protocol registration, and incomplete equation-level documentation for Google Scholar. Future research should prioritize longer and more comparable interventions, diverse educational contexts, transparent reporting, authentic communication outcomes, and ethical issues related to privacy, bias, academic integrity, and equitable access. Artificial intelligence should therefore be used as a critically evaluated pedagogical support that strengthens, rather than replaces, learner responsibility, teacher guidance, and meaningful human interaction.

BIBLIOGRAPHIC REFERENCES

- AlTwijri, L., & Alghizzi, T. M. (2024). Investigating the integration of artificial intelligence in English as foreign language classes for enhancing learners' affective factors: A systematic review. *Heliyon*, 10(10), e31053. <https://doi.org/10.1016/j.heliyon.2024.e31053>
- Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE Publications.
- Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology*, 24(1), 16–26. <https://doi.org/10.64152/10125/44705>
- Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, 6, 100230. <https://doi.org/10.1016/j.caeai.2024.100230>
- El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3), e12667. <https://doi.org/10.1111/exsy.12667>



- Fryer, L. K., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). Bots for language learning now: Current and future directions. *Language Learning & Technology*, 24(2), 8–22. <http://hdl.handle.net/10125/44719>
- Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, 55(1), 48–63. <https://doi.org/10.1080/15391523.2022.2142873>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Lai, W. Y. W., & Lee, J. S. (2024). A systematic review of conversational AI tools in ELT: Publication trends, tools, research methods, learning outcomes, and antecedents. *Computers and Education: Artificial Intelligence*, 7, 100291. <https://doi.org/10.1016/j.caeai.2024.100291>
- Liang, J.-C., Hwang, G.-J., Chen, M.-R. A., & Darmawansah, D. (2023). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*, 31(7), 4270–4296. <https://doi.org/10.1080/10494820.2021.1958348>
- Lin, M. P.-C., & Chang, D. (2020). Enhancing post-secondary writers' writing skills with a chatbot: A mixed-method classroom study. *Educational Technology & Society*, 23(1), 78–92. [https://doi.org/10.30191/ETS.202001_23\(1\).0006](https://doi.org/10.30191/ETS.202001_23(1).0006)
- Lubbe, W., ten Ham-Baloyi, W., & Smit, K. (2020). The integrative literature review as a research method: A demonstration review of research on neurodevelopmental supportive care in preterm infants. *Journal of Neonatal Nursing*, 26(6), 308–315. <https://doi.org/10.1016/j.jnn.2020.04.006>
- Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning*, 33(8), 865–890. <https://doi.org/10.1080/09588221.2019.1595664>
- Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, 2, 100020. <https://doi.org/10.1016/j.caeai.2021.100020>



- 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>
- Petchprasert, A. (2021). Utilizing an automated tool analysis to evaluate EFL students' writing performances. *Asian-Pacific Journal of Second and Foreign Language Education*, 6, Article 1. <https://doi.org/10.1186/s40862-020-00107-w>
- Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for Facebook Messenger. *Computers & Education*, 151, 103862. <https://doi.org/10.1016/j.compedu.2020.103862>
- Tang, W. K.-W., & Zhang, X.-Y. (2026). Artificial intelligence in English as a foreign language learning: A systematic review of technologies, tools, and their impacts. *Computers and Education: Artificial Intelligence*, 10, 100595. <https://doi.org/10.1016/j.caeai.2026.100595>
- Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14, 1261955. <https://doi.org/10.3389/fpsyg.2023.1261955>
- Xu, T., & Wang, H. (2024). The effectiveness of artificial intelligence on English language learning achievement. *System*, 125, 103428. <https://doi.org/10.1016/j.system.2024.103428>
- Yin, Q., & Satar, M. (2020). English as a foreign language learner interactions with chatbots: Negotiation for meaning. *International Online Journal of Education and Teaching*, 7(2), 390–410.

CONFLICTS OF INTEREST

None.

FUNDING

Not applicable.



AUTHOR CONTRIBUTIONS (CRediT)

Conceptualization: Leidy Silvana Curay Moreta.

Methodology: Leidy Silvana Curay Moreta, Jeison Paul Velásquez Fala.

Investigation: Leidy Silvana Curay Moreta, Maritza Alexandra Jara Silva, Gilma Alexandra Gordillo Obregón.

Writing – Original Draft: Leidy Silvana Curay Moreta.

Writing – Review and Editing: Maritza Alexandra Jara Silva, Jeison Paul Velásquez Fala, Gilma Alexandra Gordillo Obregón, Karina Susana Pineda Chávez.

