

<https://doi.org/10.17323/jle.2026.29839>

# A Bibliometric Analysis of Artificial Intelligence in EFL Education: Trends, Collaborations, Thematic Evolution, and Pedagogical Implications

Zhong Linling 

Zhejiang Shuren University, Hangzhou, China

## ABSTRACT

**Background.** Given the rapid advances in technology, artificial intelligence (AI) offers significant potential in educational contexts by enabling personalized learning and enhancing learner engagement. However, the development of AI-related scholarship in English as a Foreign Language (EFL) education, including its publication trajectory, collaborative networks, citation patterns, and thematic distribution, remains insufficiently explored.

**Purpose.** This study examines the extant research on AI in EFL education, spanning from the early developments in 1996 to mid-2025, with the aim of identifying structural patterns, thematic concentrations, and under-examined areas to inform future research priorities and pedagogical directions.

**Method:** By analyzing 829 publications from the Web of Science (WoS) and Scopus databases, screened by a single researcher, and utilizing the VOSviewer software and Microsoft Excel, the study identifies and interprets significant trends, including publication trends, international collaborative networks, citation impact and thematic developments across the literature. Bibliometric indicators, country collaboration networks, journal citation patterns, and keyword co-occurrence structures were analysed; results should be interpreted in light of English-language database coverage, single-researcher screening, and threshold-dependent keyword mapping.

**Results.** The findings reveal an exponential rise in research output since 2022, with China leading in publication volume, citation impact, and collaborative centrality. Highly impactful journals demonstrate substantial engagement across disciplinary areas, with Computer Assisted Language Learning emerging as the most influential source. The co-occurrence analysis of Keywords highlights a concentration on generative AI tools, particularly ChatGPT, and AI-enhanced writing, while areas such as speaking, listening, reading, and self-regulated learning appeared underrepresented in the keyword structure of the corpus, suggesting potential gaps in the thematic scope of current AI-related EFL research.

**Conclusion.** This study provides the first bibliometric account of AI in EFL education to encompass the generative AI period, simultaneously mapping collaboration networks, citation trajectories, and keyword structures. The findings offer evidence-based insights into research trends and thematic imbalances, with implications for future inquiry and pedagogical innovation in AI-enhanced EFL education.

## KEYWORDS

artificial intelligence (AI); bibliometric analysis; English as a Foreign Language (EFL); keyword co-occurrence; publication trends

## INTRODUCTION

In contemporary digitized society, the pursuit of optimizing language education has prompted educators, researchers, and policymakers to explore innova-

tive approaches that foster productive and effective learning environments (MacIntyre et al., 2019; Ryan & Liu, 2022). Alongside this pedagogical shift, rapid advancements in digital technologies, particularly the advent of artificial intel-

**Citation:** Zhong, L. (2026). A Bibliometric analysis of artificial intelligence in EFL education: Trends, Collaborations, thematic evolution, and pedagogical implications. *Journal of Language and Education*, 12(1), 175-194. <https://doi.org/10.17323/jle.2026.29839>

**Correspondence:**  
Zhong Linling  
[zhonglinling9@gmail.com](mailto:zhonglinling9@gmail.com)

**Received:** November 26, 2025

**Accepted:** March 16, 2026

**Published:** March 31, 2026



ligence (AI), have brought transformative changes across academia generally and language education specifically (Liu et al., 2025; Liu & Chang, 2024; Zawacki-Richter et al., 2019), thereby compelling a rethinking of established pedagogical paradigms (Al-kadi, 2025).

Since late 2022, the emergence of large-scale AI technologies has significantly influenced language education (Chang & Sun, 2024) and accelerated technology integration within English as a foreign (EFL) education. AI technologies address inherent limitations of conventional language instruction through tailored feedback, individualized pacing, and interactive content delivery (Akinsemolu & Onyeaka, 2025; Almelhes, 2023; Yesilyurt, 2023; Zou et al., 2023), demonstrating substantial potential for improving language acquisition and educational productivity (Liu & Chang, 2024). AI may further empower EFL learners to engage in informal, self-directed learning with greater creativity and flexibility (Almelhes, 2023; Liu et al., 2024a; Zou et al., 2024), and its integration has achieved sufficient depth to prompt a rethinking of established pedagogical approaches (Alshumaimeri & Alshememry, 2023). The present study focuses specifically on EFL rather than ESL or broader L2 contexts, a deliberate boundary reflecting meaningful differences in instructional goals, learner exposure, and the pedagogical role of AI across these settings; studies framed within general L2, or ESL frameworks were therefore excluded from the corpus.

Despite growing research interest, the field lacks a consolidated and mapped knowledge structure. Several recent reviews have examined AI-supported language education, but most remain broad in scope with limited domain-specific analysis of EFL learning. Kartal and Yeşilyurt (2024) conducted a bibliometric analysis of AI in L2 teaching and applied linguistics covering 1995–2022; however, their study terminates at the inflection point when generative AI began to influence EFL research, capturing neither the post-2022 publication surge nor the shift from rule-based systems toward generative, dialogic technologies. Lo et al. (2024) reviewed ChatGPT integration in ESL/EFL education, but their scope is bounded to a single tool and does not encompass the broader AI ecosystem, including intelligent tutoring systems, neural networks, and natural language processing applications. Moreover, their thematic synthesis differs fundamentally from the structural mapping that bibliometric analysis affords across collaboration networks, citation trajectories, and keyword co-occurrence clusters. Neither study simultaneously addresses the full temporal span of the field, the complete AI technological landscape, and the multi-dimensional bibliometric structure of the corpus.

Existing reviews do not simultaneously capture temporal growth, international collaboration, source-level influence, and thematic concentration within the post-generative AI period. The present study therefore maps these four dimensions: publication trends, country collaboration networks,

journal citation patterns, and keyword co-occurrence structures, to provide a structurally comprehensive and temporally current bibliometric account of AI in EFL education.

To address this gap, the present study conducts a bibliometric analysis to examine research on AI in EFL education. The study aims to provide both educators and related professionals with critical insights into significant trends in this research area, thereby supporting evidence-based decision making in research and pedagogical practice. In line with the purpose of the study, the following research questions are proposed:

- RQ1:** What are the publication trends in research on AI in EFL education?
- RQ2:** Which countries demonstrate the most extensive international collaborations and highest citation impact in research on AI in EFL education?
- RQ3:** Which journals have made the most impactful contributions based on citation metrics in research on AI in EFL education?
- RQ4:** What are the prevailing themes and under-researched areas in research on AI in EFL education?

## LITERATURE REVIEW

### AI-Supported EFL Learning and Pedagogical Applications

The integration of AI into educational contexts has profoundly impacted language education, indicating significant changes in pedagogical approaches and learning experiences (Huang et al., 2023; Sharadgah & Sa'di, 2022; Smutny & Schreiberova, 2020). The increasing accessibility of AI tools has encouraged more students to use these technologies for language learning purposes (Alharbi, 2023), contributing to positive learning experiences, especially in EFL classrooms (Yang et al., 2022). In EFL teaching and learning specifically, AI has shown substantial potential across several dimensions: it facilitates the refinement of instructional methodologies, enables the provision of individualized learning materials and immediate feedback, and demonstrates measurable impacts on learner performance outcomes (Jia et al., 2022; Mohamed, 2023; Qasem et al., 2023; Sun et al., 2021; Xia et al., 2022; Zhang & Zou, 2020). AI-powered applications have been applied to enhance language competencies, including writing, vocabulary acquisition, and grammar instruction (Liu et al., 2024b), while the ability of AI to process extensive datasets and operate through natural language interfaces contributes to supportive, low-anxiety learning environments that promote iterative skill development (Crompton et al., 2024). Taken together, these studies demonstrate that AI-supported EFL research has expanded considerably in pedagogical scope and application; howev-

er, they do not reveal how the field is structurally organized across time, countries, journals, and themes.

## From Earlier AI Tools to Generative AI

The development of AI in EFL education has not followed a linear trajectory. Earlier AI applications in language learning were predominantly rule-based, encompassing tools such as intelligent tutoring systems, automatic speech recognition, and grammar checkers, which operated within constrained parameters and offered limited adaptability to individual learner needs (Luckin et al., 2016; Zawacki-Richter et al., 2019). The gradual advancement of machine learning and natural language processing from the late 2010s onward enabled more sophisticated forms of AI-mediated interaction, including adaptive feedback systems and data-driven language assessment (Hwang et al., 2020). Since late 2022, the public release of large language models, most notably ChatGPT, has marked a qualitative shift in the field, substantially expanding the possibilities for personalized, dialogic, and scalable language learning (Dwivedi et al., 2021; Kasneci et al., 2023). This transition from static, tool-based interventions toward generative, interactional AI systems has altered both research agendas and pedagogical possibilities in EFL education (Chang & Sun, 2024). Taken together, these developments suggest that the field has undergone a paradigmatic reorientation rather than gradual evolution; yet the structural consequences of this shift, in terms of publication growth, collaborative patterns, and thematic concentration, have not been systematically mapped.

## Existing Reviews and Their Limitations

Several prior reviews have examined AI applications in language education or EFL contexts, offering valuable but partial accounts of the field. Kartal and Yeşilyurt (2024) conducted a bibliometric analysis of AI in L2 teaching and applied linguistics covering 1995–2022; however, their study terminates precisely at the inflection point when generative AI began to influence EFL research, capturing neither the post-2022 publication surge nor the paradigmatic shift toward generative, dialogic technologies. Lo et al. (2024) reviewed ChatGPT integration in ESL/EFL education, contributing important evidence on the pedagogical applications of a single prominent tool, but their scope is bounded to ChatGPT and does not encompass the broader AI ecosystem, including intelligent tutoring systems, neural networks, and natural language processing applications. Moreover, as a thematic synthesis, their study differs fundamentally from the structural mapping that bibliometric analysis affords. Taken together, neither study simultaneously addresses the full temporal span of the field, the complete AI technological landscape in EFL education, and the multi-dimensional bibliometric structure of the corpus, leaving a significant gap in the structural and longitudinal understanding of how this research field has developed.

## Rationale for Bibliometric Analysis

Previous studies tend to focus on specific AI applications or pedagogical outcomes rather than mapping the intellectual landscape of the field as a whole. Addressing this gap requires a method capable of simultaneously tracing publication growth trajectories, quantifying international collaboration structures, identifying citation-based influence patterns, and detecting thematic clusters across a large and heterogeneous corpus, analytical dimensions that content-analytical approaches such as narrative syntheses or conventional reviews are not designed to capture at scale. Bibliometric analysis is uniquely suited to these research questions because it treats the corpus itself as the unit of analysis, revealing macro-level structural patterns that remain invisible when studies are examined individually or synthesized thematically. A comprehensive bibliometric analysis is therefore both methodologically appropriate and necessary to provide a structural and longitudinal account of how AI in EFL education has developed as a research field.

The present study addresses the identified gaps by mapping four complementary bibliometric dimensions. Publication trends address the temporal development of the field, revealing how research output has evolved and where discontinuities in orientation have occurred. Country collaboration networks address the geographical structure of knowledge production, identifying centers of scholarly influence and patterns of international engagement. Journal citation analysis addresses source-level influence, illuminating which publication venues shape the intellectual agenda of the field. Keyword co-occurrence analysis addresses thematic concentration and gaps, revealing both the prevailing research foci and the comparatively neglected areas within the corpus. Together, these four dimensions provide a structurally comprehensive account of AI in EFL education that existing reviews have not achieved.

## METHOD

### *Research Design*

This study employed a bibliometric analytical approach, utilizing quantitative methods to analyze an extensive corpus of scholarly literature in the field of artificial intelligence (AI) applied to English as a Foreign Language (EFL) education. Bibliometric analysis applies statistical and mathematical techniques to examine academic publications, providing a comprehensive overview and macro-level structural insights of the research field (Agarwal et al., 2016; Hallinger & Kovačević, 2019). This approach is particularly adept at synthesizing and critically evaluating accumulated research materials within specific thematic areas, enabling scholars to examine the intellectual framework of research topics, developmental trends, and interrelationships across extensive academic literature (Ellegaard & Wallin, 2015; Fellnhöfer, 2019; Hallinger & Kovačević, 2019).

The methodological procedure followed in this study comprised eight sequential stages: (1) database search using Boolean strings in WoS and Scopus; (2) title and abstract screening against inclusion and exclusion criteria; (3) deduplication across databases; (4) metadata cleaning including standardization of source titles, author names, country affiliations, and keywords; (5) bibliometric analysis using VOSviewer and Excel; (6) data synthesis aligning each research question with corresponding metrics and outputs; (7) interpretation of bibliometric patterns; and (8) acknowledgment of procedural limitations. Each stage is described in the subsections that follow.

## Data Sources and Search Strategy

The Web of Science (WoS) and Scopus databases were chosen for their extensive coverage of high-impact journals and recognized authority in academic research (Ahmi et al., 2019; Zhong, 2025b), as well as their widely cited collections of scholarly content (Sweileh et al., 2017). To examine research on AI in EFL education, a comprehensive search strategy was developed, incorporating two core concepts: 'artificial intelligence (AI)' and 'English as a Foreign Language (EFL),' supplemented by relevant synonyms. Boolean operators 'OR' and 'AND' were employed to optimize retrieval precision and coverage.

The search string was refined through a deliberate methodological decision to exclude "logistic regression" as a search term. Although logistic regression is a statistical technique occasionally referenced in AI-related studies, its inclusion in the search string would have introduced substantial conceptual and disciplinary noise. As detailed in Table 1, publications were retrieved using two Boolean search strings across two academic databases. To ensure full reproducibility, the complete search strings, including truncations, Boolean operators, filters, and database-specific settings, are provided in Appendix.

It should be acknowledged that the search string does not include several emerging generative AI terms, such as "large language model," "GPT," or "foundation model", that have gained prominence since 2022. This omission represents a limitation that may result in the under coverage of some recent generative AI literature not indexed under the terms included in the current string.

## Operational Definitions and Eligibility Criteria

"AI" was operationalized to include technologies characterized by machine-based learning, adaptive processing, and natural language capabilities, encompassing machine learning, deep learning, neural networks, natural language processing, intelligent tutoring systems, and generative AI tools such as chatbots. Generic statistical methods without an AI framing, such as logistic regression used purely as an analytical tool, were excluded, as their inclusion would have

introduced conceptual noise unrelated to the AI paradigm under investigation.

"EFL education" was defined as instructional contexts in which English is taught as a foreign language in non-Anglophone settings. Studies framed exclusively within ESL or broader L2 contexts without explicit EFL specification were excluded. Borderline cases were resolved as follows: machine learning applied to EFL writing assessment was included; generic language education studies without EFL specification were excluded; statistical prediction models without AI framing were excluded; ESL-only studies were excluded where the context fell outside the EFL scope. As presented in Table 2, the eligibility criteria were in line with the set of inclusion and exclusion criteria for the current review.

## Screening and Deduplication Procedure

As with other rigorous forms of evidence synthesis (Sutton et al., 2019), this review followed the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) 2020 guidelines as outlined in Figure 1, selecting all relevant, well-designed studies to ensure methodological rigor and transparency (Zawacki-Richter et al., 2020). PRISMA 2020 was used to report the identification, screening, and inclusion process; it was not used to frame the study as a systematic review of intervention effects.

The initial search retrieved 1,674 articles (WOS, n=555; Scopus, n=1119) (1 July 2025). To ensure the relevance and specificity of the data, several filters were applied. The emphasis on English language articles narrowed the selection to 1,666 publications (WOS, n=551; Scopus, n=1111). Consistent with established review protocols, non-peer-reviewed materials such as conference papers, editorials, opinion pieces, books and dissertations were excluded to ensure the reliability of the findings (Zhai & Wibowo, 2023; Zhong, 2025a). This decision was made to ensure the reliability and comparability of the bibliometric indicators analyzed; however, it should be noted that excluding conference proceedings may underrepresent early-stage AI and computational work, where findings are frequently disseminated before journal publication. The focus was then narrowed to peer-reviewed articles and article reviews, which reduced the total to 1,364.

Duplicate publications were identified using Microsoft Excel by comparing titles, authors, and publication years (Scopus and Web of Science). Checks were applied for title-level similarity, author-name format variation, and online-first versus print publication year discrepancies. This procedure allowed for detection of overlapping records, resulting in the removal of 471 duplicates. While the comparison of titles, authors, and publication years reduces the likelihood of substantial error, manual Excel-based deduplication cannot fully account for minor title variants, inconsistent author name formats, or online-first versus print publication year discrepancies; a small residual margin of misclassification therefore

**Table 1***The Boolean Expression*

Database	Search String (Boolean Expression)
Web of Science (WOS)	(TS=((«artificial intelligence» OR «AI» OR «machine intelligence» OR «machine learning» OR «deep learning» OR «deep network» OR «Intelligent agent*» OR «intelligent tutor*» OR «intelligent tutoring system*» OR «hybrid intelligent system*» OR «neural network» OR «natural language processing» OR «Chatbot*» OR «thinking computer system*»)) AND TS=(«EFL» OR «English as a foreign language»))
SCOPUS	TITLE-ABS-KEY((«artificial intelligence» OR «AI» OR «machine intelligence» OR «machine learning» OR «deep learning» OR «deep network» OR «Intelligent agent*» OR «intelligent tutor*» OR «intelligent tutoring system*» OR «hybrid intelligent system*» OR «neural network» OR «natural language processing» OR «Chatbot*» OR «thinking computer system*») AND («EFL» OR «English as a foreign language»))

**Table 2***Inclusion and Exclusion Criteria*

Inclusion criteria	Exclusion criteria
<ol style="list-style-type: none"> <li>Articles published in English (up to 1 July 2025).</li> <li>Peer-reviewed articles and article reviews.</li> <li>Articles focusing on AI in the context of EFL education.</li> </ol>	<ol style="list-style-type: none"> <li>Articles not published in English.</li> <li>Conference papers, conference reviews, book chapters, books, short surveys, notes, and editorials.</li> <li>Duplicate records across databases.</li> <li>Articles not focused on AI in EFL education.</li> </ol>

cannot be excluded. Future studies are encouraged to employ dedicated tools such as ASySD or Rayyan to enhance procedural precision.

After deduplication, 893 unique publications remained for subsequent bibliometric analysis. The remaining 893 publications were then manually screened for eligibility. Because screening was conducted by one researcher, no inter-rater reliability coefficient could be calculated; therefore, eligibility decisions should be interpreted with caution. Each document's title and abstract were carefully assessed for relevance to the subject of AI in EFL education by a single researcher, and 64 publications that did not explicitly address AI in EFL education were subsequently excluded. This process yielded 829 publications for final analysis.

## Data Preparation

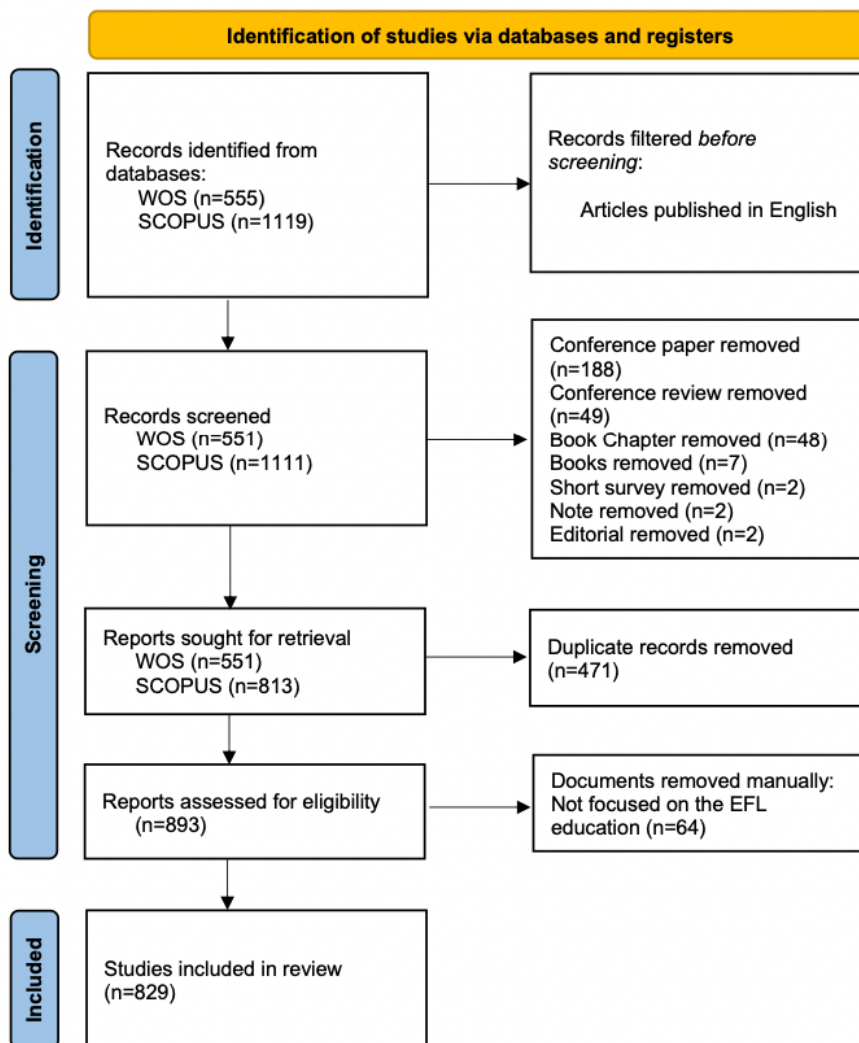
Data retrieved from WoS and Scopus, including author details, article and source titles, publication years, abstracts, and citation counts, were exported in Microsoft Excel and comma-separated values (.csv) formats, which were manually consolidated into a single editable .csv file to facilitate bibliometric analysis. After screening, the final dataset of 829 records was consolidated into a single CSV file. Metadata fields were checked for missing values, inconsistent source titles, author-name variants, country-affiliation variants, and keyword synonyms. Source title inconsistencies between database exports were standardized manually prior to analysis.

## Bibliometric Analysis and Visualization Parameters

VOSviewer was employed in this study for its robust text mining and bibliometric visualization capabilities, particularly its ability to generate co-word maps and network diagrams that reveal conceptual relationships within the literature (Kartal & Yeşilyurt, 2024; Moral-Muñoz et al., 2020; Van Eck & Waltman, 2020). The parameters applied across analyses are summarized in Table 3. For the keyword co-occurrence analysis, a minimum occurrence threshold of ten was applied, yielding 45 keywords for inclusion. Full counting was used for both keyword co-occurrence and country collaboration analyses. Country-level data were coded by institutional affiliation as listed in the metadata exported from WoS and Scopus, using full counting rather than fractional counting, meaning that each country listed in a multi-authored publication was counted once per publication. No VOSviewer thesaurus file was applied to merge variant keyword forms; as a result, terminologically fragmented clusters may appear in the co-occurrence network, and semantically equivalent terms indexed under different labels may be represented as separate nodes. This represents a limitation of the keyword analysis that readers should bear in mind.

For journal-level analysis, three distinct indicators were examined: most productive sources (ranked by total publication count), most cited sources (ranked by total citation count), and highest average citation sources (ranked by citations per article). Bibliographic coupling was additionally examined using VOSviewer to identify sources with shared reference profiles.

**Figure 1**  
The PRISMA Flow Diagram for Literature Selection



**Table 3**  
VOSviewer Parameters by Analysis

Analysis	Type	Unit	Counting Method	Normalization	Min. Threshold	Items Included
Keyword co-occurrence	Co-occurrence	Author keywords	Full counting	Association strength	10 occurrences	45 keywords
Country collaboration	Co-authorship	Countries	Full counting	Association strength	8 documents	28 countries

## Data Synthesis

The analysis process involved examining publication trends over time, to understand the development trajectory of AI in EFL education research, countries with the most extensive international collaborations and highest citation impact to reveal global research dynamics and influential contributions, and journals that have made the most impactful contributions based on citation metrics. Through keyword co-occurrence analysis, the study also identified prevailing research themes and under-researched areas, thereby offering a comprehensive overview of current trends and emerging directions in this research field. The analysis addressed four research questions through the following bibliometric dimensions and corresponding outputs as shown in Table 4.

Bibliometric indicators were then generated in relation to each research question. Thematic gaps were identified as keyword domains absent from or weakly represented in the co-occurrence network relative to their established importance in the SLA and EFL literature, rather than as definitive evidence of empirical neglect.

## RESULTS

### Publication Trends

The publication trends in research on AI applications in EFL education reveal two clearly differentiated developmental eras rather than a smooth or continuous evolution, namely a prolonged pre-generative AI period characterized by sporadic and exploratory studies, and a post-generative AI period marked by exponential growth and structural transformation in research output (as shown in Figure 2).

#### Pre-generative AI era (1996–2020)

Figure 2 suggests that initial explorations in this research area can be traced back to 1996 and 1997, when a small number of studies focused on intelligent tutoring systems, establishing an early conceptual foundation for AI-assisted

language learning. However, these early efforts were limited in scale, frequency, and technological sophistication, reflecting both the nascent state of AI technologies and their constrained applicability in language education at the time. Throughout the late 1990s and early 2000s, publication output remained sporadic. Isolated contributions appeared in 2001 on automatic assessment mechanisms, in 2005 on accent reduction software for EFL learners, and in 2006 on online chat tools. These studies, while occasionally employing the label of “AI,” were largely rule-based or technology-driven interventions rather than data-intensive or adaptive systems associated with contemporary AI paradigms.

From 2007 to 2017, the annual number of publications rarely exceeded one to three articles per year, indicating that AI in EFL education remained a marginal and underexplored research topic for more than two decades. Although a modest increase can be observed between 2018 and 2020, with annual publications rising gradually from four to eight, this growth reflects incremental expansion rather than a fundamental shift in research orientation. During this extended pre-generative AI era, research was predominantly exploratory, tool-oriented, and constrained by limited computational capacity and pedagogical integration.

#### Post-generative AI era (2021–mid-2025)

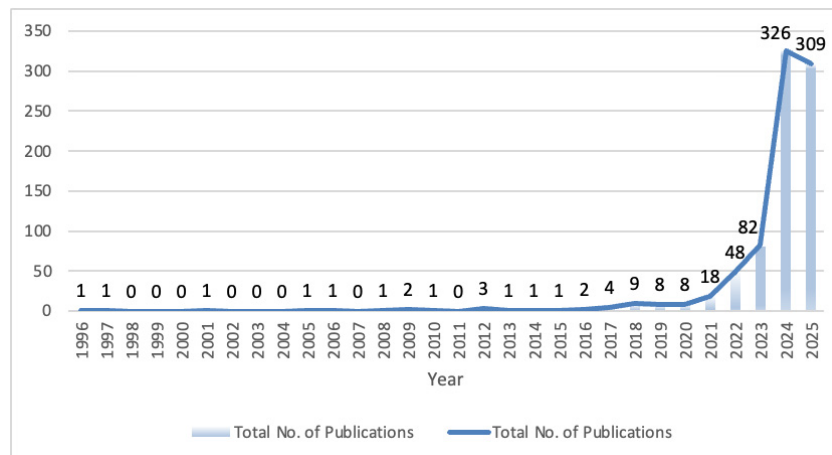
A marked acceleration occurs from 2021 onward, corresponding temporally with the public emergence of large language models and, most notably, the release of ChatGPT in late 2022. While this temporal correspondence is suggestive, the bibliometric data cannot confirm that ChatGPT alone accounts for the observed growth, as other structural factors, including increased institutional support and broader recognition of AI’s educational potential, may have contributed concurrently. The publication output increased sharply from 18 studies in 2021 to 48 in 2022, before more than tripling to 162 publications in 2023. This dramatic escalation suggests not a continuation of earlier trends, but the onset of a qualitatively different research phase that appears to be driven by generative AI, particularly ChatGPT and related tools.

**Table 4**

*Research Question, Metadata Fields, Metrics, Software, and Outputs*

RQ	Metadata Fields	Metric	Software	Output
<b>RQ1: Publication trends</b>	Publication year	Annual publication count	Excel	Figure 2
<b>RQ2: Country collaboration and citation impact</b>	Author affiliations, citation counts	Co-authorship links, total link strength, total citations, publications per country	VOSviewer, Excel	Figure 3, Figure 4
<b>RQ3: Journal citation impact</b>	Source title, citation counts	Total citations, total publications per source, average citations per article	VOSviewer, Excel	Table 5
<b>RQ4: Thematic concentration and gaps</b>	Author keywords	Co-occurrence frequency, cluster membership, network centrality	VOSviewer	Figure 6

**Figure 2**  
*Publication Trends*



The most pronounced growth is observed in 2024, with 326 publications, nearly a tenfold increase compared to 2021, which can be interpreted as indicative of the catalytic effect of generative AI on scholarly activity in EFL education. Although data for 2025 were collected only up to July 1st, the year had already yielded 309 publications, suggesting that research output in this domain continues to accelerate. These trends collectively indicate a sustained and accelerating growth in research output across the post-2021 period.

### International Collaborations and Country Citation Impact

The analysis of international collaboration and country citation impact reveals a highly concentrated and uneven distribution of research on AI in EFL education, characterized by significant disparities in research output and collaborative engagement across countries (as shown in Figure 3 & Figure 4). In VOSviewer, links indicate co-authorship connections between countries, while total link strength represents the cumulative strength of these connections. Figure 3 demonstrates the dominant position of China and its strong internal cluster (size/influence), alongside the key collaborative links it maintains with countries like South Korea, Japan, and the United States, indicating the global distribution and specific regional focus of research collaboration, while Figure 4 suggests a significant correlation between high document output and citation impact, with China dominating both metrics, suggesting a concentration of high-productivity and high citation visibility within specific regions.

The network visualization identified 28 countries demonstrating the most extensive international collaborations in this research area. As the central node in the network, China exhibits the highest level of connectivity and productivity, with 26 international links and a total link strength of 308, indicating broad and frequent research collaborative activity across countries, while maintaining the highest citation impact contributing 291 publications with 4,743 citations.

South Korea ranks second in collaborative network intensity, with 19 international links and a total link strength of 131, and third in citation impact with 46 publications and 786 citations. These links are particularly strong with East Asian and Anglophone countries, reflecting South Korea’s active participation in international research networks.

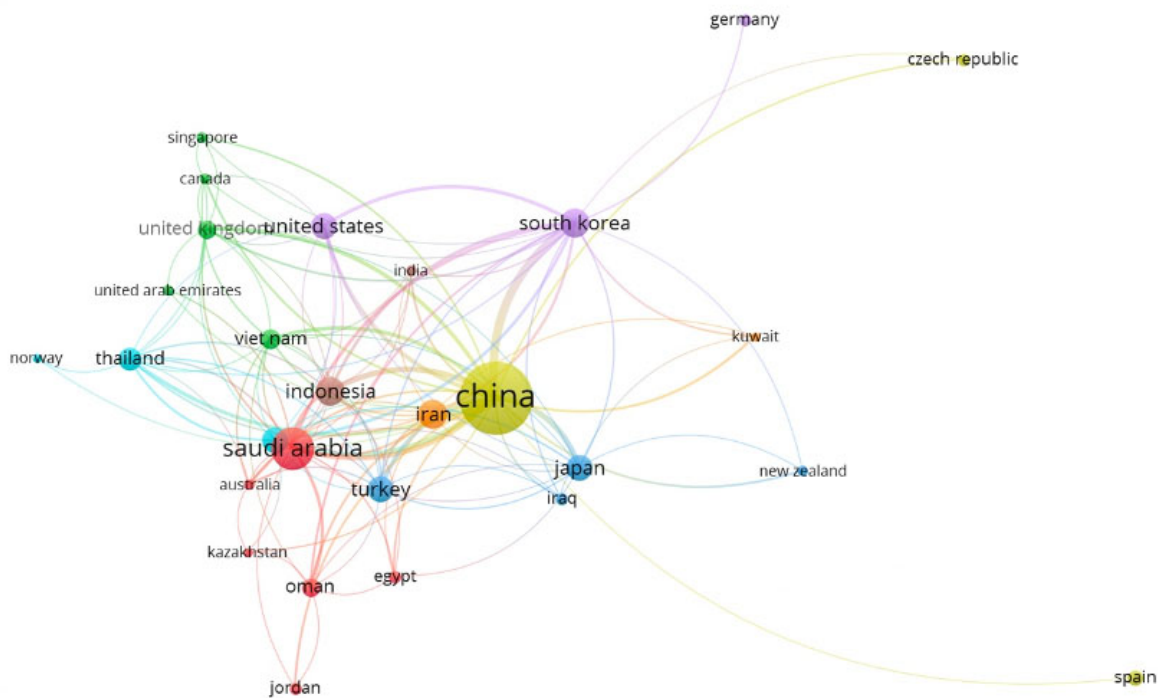
While South Korea produced more publications than the United States (38 publications), its total citation count was lower than the U.S. (874 citations). The United States exhibits the second highest citation impact, resulting in an average of 23 citations per article. However, this strong citation performance contrasts with the United States’ relatively low level of international collaboration with 10 international links and a total link strength of 34, suggesting high citation impact relative to its collaborative engagement. Similarly, the United Kingdom, with just 21 publications, achieves 350 citations, further indicating that its contributions carry considerable scholarly weight despite limited collaborative breadth.

Saudi Arabia demonstrates the third-highest level of connectivity with an extensive collaborative network featuring 18 international links and a total link strength of 73, alongside by a notable volume of 99 publications, but its total citation count is comparatively modest (581 citations). Countries including Iran, Indonesia, Japan, and Vietnam also display moderate levels of productivity and citation impact, though with comparatively limited international connectivity, as reflected in regionally concentrated research practices and potential additional constraints such as language barriers.

The network visualization also illustrates the formation of regional research clusters. East Asia, the Middle East, and parts of Southeast Asia form dense collaborative nodes. In contrast, several European countries, such as Germany, the Czech Republic and Spain, while included in the network, occupy relatively peripheral positions in terms of collaboration and citation impact.

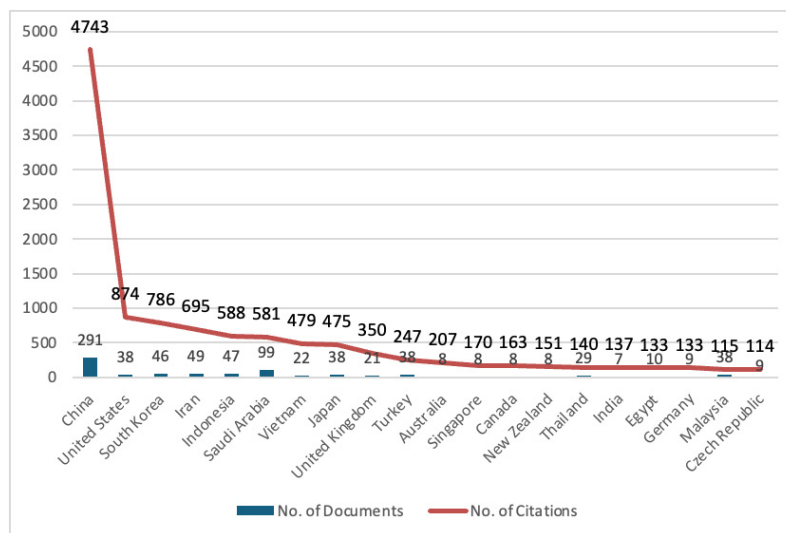
**Figure 3**

The Network Visualization of International Collaborations (minimum threshold set to 2)



**Figure 4**

Country Citation Impact



Note. Minimum number of documents set to 8.

**Journal Impact**

The distribution of research on AI in EFL education across academic journals indicates engagement across applied linguistics, educational technology, and computer science. While traditional CALL (Computer-Assisted Language Learning)-focused journals maintain centrality, interdisciplinary platforms are demonstrating increasing influence (as

shown in Figures 5). Figure 5 presents the journal co-citation network, which illustrate the intellectual structure and interdisciplinary foundations of AI research in EFL education.

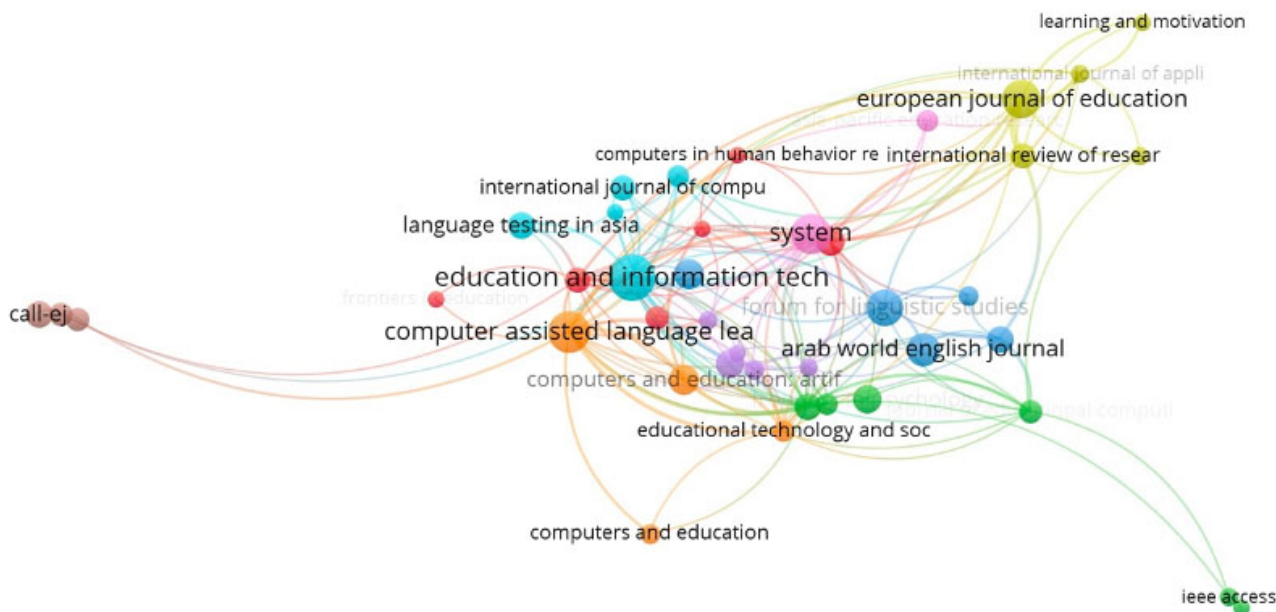
The bibliometric network identified 42 journals as impactful journals in this research field, with 41 of them displaying co-citation, indicating shared citation patterns and overlapping intellectual foundations. Although the journal Assess-

ing Writing is not connected to the network, it still exhibits notable scholarly impact with 5 publications and 206 citations (illustrated in Table 5).

Among the leading journals by citation metrics, Computer Assisted Language Learning emerges as the most influential source, with 27 articles and a total of 711 citations. This

journal’s longstanding focus on technology-mediated language learning contribute to its sustained scholarly visibility. Following closely is Education and Information Technologies and Computers and Education: Artificial Intelligence, suggesting strong crossover interest between AI-enhanced learning environments and second language acquisition research.

**Figure 5**  
Impactful Journals



Note. Minimum threshold set to five.

**Table 5**  
Impactful Journals

Rank	Journal	No. C	No. Pub
1	Computer Assisted Language Learning	711	27
2	Education and Information Technologies	688	34
3	Computers and Education: Artificial Intelligence	571	15
4	Interactive Learning Environments	527	12
5	Frontiers in Psychology	508	13
6	System	482	26
7	Computers and Education	390	7
8	European Journal of Education	309	23
9	Cogent Education	301	13
10	Educational Technology and Society	286	8
11	Innovation in Language Learning and Teaching	256	13
12	Assessing Writing	206	5
13	Computers in Human Behavior	203	8

Note. No. Pub=Number of Publications; No. C= Number of Citations. More than 200 citations.

Notably, *Interactive Learning Environments* demonstrates remarkable citation efficiency, achieving an average of 43.9 citations per article, which is the highest rate among the analyzed journals. The journal's focus on interactive educational technologies aligns closely with contemporary AI applications in language learning. Several journals with fewer publications nonetheless demonstrate high citation intensity. For instance, *Computers and Education* published only seven articles, but had 390 citations, reflecting strong scholarly impact for each contribution.

## Prevailing Themes and Under-Researched Areas

The keyword co-occurrence analysis offers insights into current research priorities and potential future directions for research on AI in EFL education. The network visualization illustrates complex interconnections among 45 frequently occurring keywords from a total of 2,938 indexed terms based on a minimum occurrence threshold of 10 (as shown in Figure 6). Figure 6 demonstrates the keyword co-occurrence network, which reveals thematic concentrations and research imbalances in AI-related EFL studies.

Within this network, "Artificial Intelligence" (261 occurrences) and "EFL" (162 occurrences) represent the two core search concepts in this study and are retained in the network as anchor terms to confirm thematic alignment of the corpus; their high frequency reflects search string construction rather than independent research emphasis and should be interpreted accordingly. Among AI technologies, "ChatGPT" (147 occurrences) emerges as the most frequently studied AI tool, occurring considerably more frequently than related terms such as "AI chatbot", "generative AI", and "machine learning".

The salience of the keyword "EFL learners" (123 occurrences) as a research focus indicates strong emphasis on learner-centered inquiry in current research, while the keyword "EFL teachers" (20 occurrences) received moderate attention, reflecting comparatively limited consideration of educator perspectives in the integration of AI into EFL education. Other prevailing themes, including "motivation", "engagement", and "technology acceptance model", underscore the significance of learner psychology and behavioral responses in research on AI in EFL education.

Among the specific language skills, the keyword "EFL writing" (53 occurrences) stands out as the most extensively explored research area. Closely associated topics such as "automated writing evaluation", "academic writing", and "writing skills" further highlight the concentrated focus on writing-related research on AI in EFL education.

The findings also reveal limited representation of other fundamental language skills in the research landscape. Specifically, keywords related to speaking, listening, and reading

appear underrepresented in the keyword co-occurrence structure of the corpus; confirmation of whether these skills are genuinely neglected in the literature would require abstract or full-text content analysis. Additionally, keywords associated with learner autonomy and communication willingness, such as "self-regulated learning" and "willingness to communicate", appear infrequently in the dataset, suggesting that affective and autonomous learning dimensions appear underrepresented in the keyword co-occurrence structure, though this pattern may reflect author labeling practices rather than the absence of such themes within the studies themselves.

## DISCUSSION

### Generative AI and the Acceleration of AI-EFL Research

The publication trends on AI in EFL education suggest a non-linear and discontinuous developmental trajectory, rather than a steady or cumulative evolution of artificial intelligence applications. This pattern is consistent with what Zawacki-Richter et al. (2019) describe as the early developmental phase of AI in education, where research was primarily constrained by technical limitations and elementary conceptual frameworks. During this period, AI-related interventions in EFL predominantly centered on basic tools, such as automatic assessment mechanisms and accent-reduction software, which lacked adaptive or context-aware intelligence. The gradual acceleration onwards may suggest that the researchers began recognizing the potential of machine learning and neural networks to address longstanding challenges in EFL education. However, researchers were still engaging in resolving theoretical and methodological uncertainties regarding AI's pedagogical efficacy (Holmes et al., 2019).

The sharp increase in publications from 2022 onward appears to represent a qualitative shift rather than a continuation of earlier trends, corresponding temporally with the public emergence of large language models and generative AI systems (Dwivedi et al., 2021; Kasneci et al., 2023). While generative AI tools, particularly ChatGPT, contributed to this surge, other structural factors also played a role, including increased computational resources, institutional support, and broader recognition of AI's pedagogical potential.

In this sense, the observed trends in the 2021–mid-2025 period do not reflect the gradual evolution of artificial intelligence as a whole, but rather the considerable influence of generative AI technologies that appears to have substantially altered, and pedagogically relevant AI-supported language learning. Without careful conceptual differentiation between generative AI and earlier AI paradigms, there is a danger of overgeneralizing findings derived from a narrow



countries occupy central positions while others remain underrepresented.

China leads in both collaborative intensity and research output, likely reflecting its substantial investment in educational technology and strategic emphasis on EFL innovation (Knox, 2020). South Korea ranks second in collaborative network intensity and third in citation impact, with strong ties to East Asian and Anglophone research communities, consistent with findings associating cross-national collaboration with increased research visibility (Gazni et al., 2012).

The United States and United Kingdom produce fewer publications but achieve higher citation impact per article, a pattern attributable to linguistic accessibility, journal prominence, and network positioning rather than inherently superior research quality (Holmes et al., 2019). By contrast, Saudi Arabia demonstrates extensive collaborative reach but comparatively modest citation impact, suggesting that collaboration volume does not automatically translate into citation visibility and that scholarly influence is shaped by multiple interacting factors including publication venue and research focus.

While international collaboration of research on AI in EFL education is expanding, it remains geographically concentrated in a limited number of countries. To address this imbalance, these patterns suggest that the current corpus may not fully represent the diversity of global EFL contexts.

### Interdisciplinary Journal Influence

The journal impact analysis may be associated with a growing interdisciplinary engagement across applied linguistics, educational technology, and computer science, reflecting broader shifts in academic dissemination and knowledge production. The dominance of Computer Assisted Language Learning as the most cited journal is consistent with previous bibliometric findings that Computer Assisted Language Learning journal is critical for integrating language learning and technological advancement (Linling, 2025). This finding also aligns with previous research that computer-assisted language learning is foundational in technology-mediated language learning research (Hubbard & Levy, 2016).

The substantial citation counts and publication volume of journals such as *Education and Information Technologies* and the AI-focused journal *Computers and Education* indicate that AI-enhanced language learning should be understood not merely as an application of computational tools, but as a complex socio-technical process requiring interdisciplinary collaboration (Dwivedi et al., 2021).

The high citation impact of *Interactive Learning Environments* highlights the scholarly value placed on research that integrates technical innovation with learner interaction and personalized intelligent learning systems. As noted by Kabu-

di et al. (2021), interactive learning frameworks enhanced by AI are increasingly recognized for providing adaptive, communication-rich learning environments, which contribute to pedagogical effectiveness.

The presence of journal *Assessing Writing* among the most cited journals, despite its lack of bibliographic coupling in the VOSviewer network, is also noteworthy. Despite contributing relatively few publications, it demonstrates the field's growing emphasis on AI-supported assessment and feedback mechanisms in language education, a trend consistent with established scholarship on automated writing evaluation and AI-driven scoring systems (Warschauer & Grimes, 2008).

Furthermore, the inclusion of journals such as *Frontiers in Psychology* and *Cogent Education* in the list of high-impact journals underscores the growing relevance of psychological and cognitive frameworks in understanding AI-enhanced language learning.

### Thematic Narrowing around ChatGPT and Writing

The keyword co-occurrence analysis reveals a noteworthy imbalance in research priorities within AI in EFL education, one possible explanation is that tool-focused research is more readily operationalizable, while broader pedagogical dimensions may be underrepresented in how authors label their work and warrant further scholarly investigation.

The keyword "ChatGPT", as the most frequently studied AI tool, underscores the transformative impact of LLMs on educational research priorities. This aligns with Baskara's (2024) emphasis on the pivotal role of generative AI, particularly LLMs, in reshaping pedagogical discourse. Lo et al. (2024) have also identified the rapid rise of ChatGPT-focused research in language education since its emergence in late 2022. However, this emphasis on a specific AI technology may also suggest a phenomenon where research attention becomes disproportionately concentrated on a single tool, potentially hindering the exploration of broader theoretical or pedagogical implications.

The prevalence of "EFL learners" as a key theme points to a continued emphasis on learner-centered research. This focus is consistent with the promising potential of generative AI to create adaptive learning environments in a modern education context (Chris & Sherifdeen, 2024). However, teacher-related perspectives appear less visible in the keyword structure of the corpus, suggesting a need for closer investigation of teacher agency, professional development, and classroom decision-making in AI-mediated EFL contexts. Tan et al. (2024) noted that teacher perspectives and professional development needs often receive insufficient attention despite their crucial role in determining the successful implementation of AI in teaching.

Another noteworthy thematic trend is the concentration on AI-enhanced writing research, which emerges as the most extensively studied skill domain in this corpus. This dominance is not merely a function of tool availability or methodological convenience but reflects deeper epistemological and disciplinary dynamics. Writing has long occupied a privileged position in language assessment traditions, where performance can be discretized, scored, and benchmarked against codified standards — properties that align closely with the logic of machine learning architectures, which depend on large volumes of structured, labelled textual data to generate meaningful outputs (Warschauer & Grimes, 2008). The compatibility between text-based language production and computational processing has effectively rendered writing the most datafied of the language skills, making it both technically tractable and institutionally legible within AI research frameworks. Furthermore, automated writing evaluation systems have accumulated decades of development, establishing an infrastructure of tools, datasets, and evaluation metrics that incentivizes continued research concentration in this domain (Hegelheimer et al., 2015). This pattern therefore reveals not simply a preference for readily implementable applications, but a structural alignment between writing's disciplinary centrality, its amenability to quantification, and the technical affordances of contemporary AI systems.

### Underrepresented Pedagogical and SLA Dimensions

The underrepresentation of speaking, listening, and reading in the keyword co-occurrence map reflects a set of epistemological challenges that extend beyond methodological inconvenience. Oral language skills are inherently multimodal, temporally unfolding, and deeply context-dependent, characteristics that resist the kind of discrete, scalable annotation that machine learning pipelines require (Chapelle & Voss, 2016). Unlike written text, spoken interaction involves prosody, gesture, turn-taking, and real-time pragmatic negotiation, dimensions that computational frameworks have historically struggled to capture within educationally meaningful contexts. This difficulty is compounded by longstanding disciplinary debates within applied linguistics about what constitutes valid and reliable assessment of oral competence, debates that have not yet been resolved in ways that translate straightforwardly into AI-compatible operationalizations. This pattern may partly reflect the epistemological challenges of capturing multimodal and temporally unfolding oral interaction within computational frameworks, though keyword-based evidence alone cannot confirm the full extent of this gap.

Similarly, because affective and learner-centered constructs such as 'self-regulated learning' and 'willingness to communicate' appear underrepresented in the keyword co-occurrence structure, the corpus may partly reflect a conceptualization of AI primarily as an instructional delivery mechanism

rather than as a scaffold for learner autonomy and affective engagement. Given that self-regulated learning has been identified as a robust predictor of language acquisition in digital environments (Yu, 2023), and that willingness to communicate represents a well-established affective construct in SLA (Kirkpatrick et al., 2024), this pattern suggests that future theoretical and empirical work may benefit from more systematic integration of these dimensions.

### Implications for SLA, CALL, and Future Research Methods

The bibliometric trends identified in this study offer several important theoretical implications for second language acquisition (SLA) and computer-assisted language learning (CALL). As writing-related keywords dominate the co-occurrence network, the keyword analysis points to a theoretical narrowing within current AI-in-EFL research. The strong emphasis on writing and automated feedback suggests that AI is most frequently examined in form-focused, surface-level domains, while core SLA constructs related to oral communication, listening comprehension, and affective engagement remain undertheorized. AI is still predominantly framed as an instructional delivery mechanism rather than as a scaffold for learner autonomy and metacognitive development. This bibliometric pattern points to a gap between current research labeling practices and established SLA models that foreground learner agency, strategic regulation, and motivation, suggesting that future theoretical work in this area remains underdeveloped relative to tool-focused inquiry.

The findings of this bibliometric analysis also have important pedagogical implications for EFL teachers and learners. Because writing-related keywords dominate the corpus and speaking, listening, and reading are comparatively absent from the co-occurrence network, future AI-EFL pedagogy should move beyond text-based applications toward a more balanced integration of AI across all core language skills. For learners, this trend offers opportunities to engage in iterative practice, receive timely feedback, and develop linguistic accuracy in a low-anxiety environment. However, the predominance of writing-focused applications also raises concerns about an imbalanced skill development trajectory if AI tools are adopted uncritically or exclusively. For EFL teachers, the comparatively low frequency of teacher-related keywords in the co-occurrence map, points to a structural asymmetry in how the field conceptualizes human agency in AI-mediated instruction. This disproportion suggests that teachers are more often positioned as passive implementers of AI tools than as active pedagogical decision-makers, a framing that may underestimate the extent to which effective AI integration depends on teachers' capacity to align algorithmic affordances with specific learner needs, curricular goals, and contextual constraints.

Moreover, because affective and learner-centered constructs appear infrequently in the keyword map, current AI

applications in EFL education may be insufficiently aligned with frameworks that foreground communicative competence and learner autonomy. This bibliometric pattern implies that current AI applications in EFL education may be insufficiently aligned with frameworks that foreground communicative competence, affective engagement, and learner autonomy, though it should be noted that keyword frequency reflects research labeling practices and cannot directly confirm the absence of such orientations within individual studies.

From a methodological perspective, the findings of this bibliometric analysis point to several directions for future research on AI in EFL education. Because application-focused keywords dominate the corpus while methodological and theoretical terms remain scarce in the co-occurrence map, future research should shift from immediate tool evaluation toward longitudinal and theory-driven inquiry. The uneven thematic distribution revealed by the keyword co-occurrence analysis further underscores the need for research designs that capture the complexity of language learning beyond easily measurable outcomes. Future studies should adopt mixed-methods and longitudinal approaches to examine how AI influences learner development, strategy use, and affective engagement across different proficiency levels and instructional contexts. Greater methodological attention to underexplored language skills, such as speaking and listening, would also help address current gaps in the literature.

Finally, because the collaboration network reveals that research output and citation impact are concentrated within a limited number of countries, comparative and cross-cultural research designs are needed to capture how AI-supported EFL learning is shaped by diverse local educational practices and policy environments.

## Limitations

Several limitations of this study should be acknowledged. First, data were drawn exclusively from WoS and Scopus; the exclusion of databases such as ERIC, Google Scholar, and the ACM Digital Library may omit relevant studies in educational technology and applied linguistics. The restriction to English-language publications further limits coverage of non-Anglophone scholarship. Future research should incorporate a broader range of databases and languages. Second, the search string does not include several prominent generative AI terms that have emerged since 2022, such as “large language model,” “GPT,” or “foundation model.” Re-running the search with these additional terms was not feasible within the revision timeline, and the current corpus may therefore underrepresent some recent generative AI literature.

Then, no validation check was conducted to confirm whether known relevant papers appeared in the retrieved corpus, which limits confidence in the completeness of the search

results, and the screening process was conducted by a single researcher, precluding the calculation of formal inter-rater reliability. Because no inter-rater reliability coefficient could be calculated, eligibility decisions should be interpreted with caution. Future bibliometric studies are encouraged to employ dual-coder verification and report reliability coefficients. Furthermore, the exclusion of conference proceedings may underrepresent early-stage AI and computational work frequently disseminated before journal publication.

The co-word analysis is sensitive to the minimum occurrence threshold applied. The threshold of ten keywords adopted here follows established conventions, but alternative thresholds would yield networks of different density and composition; thematic patterns should therefore be interpreted as threshold-dependent rather than definitive. As no VOSviewer thesaurus file was applied, terminologically fragmented clusters may appear, and semantically equivalent terms may be represented as separate nodes. Additionally, author-assigned keywords are terminologically inconsistent across databases, meaning that a skill domain’s apparent absence from the co-occurrence map may reflect variation in indexing practices rather than a genuine gap in research attention.

Lastly, country-level collaboration data were coded using full counting by institutional affiliation, meaning that each country listed in a multi-authored publication was counted once regardless of the number of co-authors from that country. This may inflate the collaborative weight of countries with large research teams relative to those with smaller teams. Citation counts are influenced by factors independent of research quality, including journal age, open-access status, and indexing coverage. Cross-journal citation comparisons should therefore be interpreted with caution.

## CONCLUSION

This study presents a bibliometric analysis of 829 publications on AI in EFL education spanning 1996 to mid-2025, mapping structural patterns across publication trends, international collaboration networks, citation impact, and thematic developments. The central finding is that the field’s growth is not a gradual accumulation but a discontinuous transformation: research remained limited until 2022, when generative AI triggered an exponential publication surge and a qualitative shift in research orientation from rule-based tools toward dialogic, learner-facing technologies. This distinction matters theoretically, as findings derived from the post-2022 generative AI period should not be generalized to the broader historical arc of AI in EFL education.

The study makes three specific empirical contributions. First, it provides one of the first bibliometric accounts focused specifically on the post-2022 generative AI period in EFL education. Second, it simultaneously maps collaboration

networks, citation trajectories, and keyword co-occurrence structures within a single analytical framework, rather than examining these dimensions in isolation. Third, the keyword analysis reveals a structural imbalance: writing and ChatGPT dominate the research landscape while speaking, listening, reading, and learner-centered constructs such as self-regulated learning and willingness to communicate remain comparatively neglected, a misalignment between technological adoption and pedagogical breadth that the field has not yet systematically addressed.

For future research, three priorities emerge from these findings: longitudinal and mixed-methods design that move beyond short-term outcome measurement; greater empirical attention to oral skills and affective constructs grounded in established SLA frameworks; and more inclusive international collaboration that reflects the linguistic and cultural diversity of global EFL contexts.

## REFERENCES

- Abramo, G., Ciriaco Andrea D'Angelo, & Murgia, G. (2013). The collaboration behaviors of scientists in Italy: A field level analysis. *Journal of Informetrics*, 7(2), 442–454. <https://doi.org/10.1016/j.joi.2013.01.009>
- Agarwal, A., Durairajanayagam, D., Tatagari, S., Esteves, S., Harlev, A., Henkel, R., Roychoudhury, S., Homa, S., Puchalt, N., Ramasamy, R., Majzoub, A., Ly, K., Tvrda, E., Assidi, M., Kesari, K., Sharma, R., Banihani, S., Ko, E., Abu-Elmagd, M., & Gosalvez, J. (2016). Bibliometrics: Tracking research impact by selecting the appropriate metrics. *Asian Journal of Andrology*, 18(2), 296. <https://doi.org/10.4103/1008-682x.171582>.
- Ahmi, A., Elbardan, H., & Raja Mohd Ali, R. H. (2019). Bibliometric analysis of published literature on industry 4.0. *2019 International Conference on Electronics, Information, and Communication (ICEIC)* (pp. 1-6). IEEE. <https://doi.org/10.23919/elinfo-com.2019.8706445>
- Akinsemolu, A. A., & Onyeaka, H. (2025). The role of artificial intelligence in transforming language learning: Opportunities and ethical considerations. *Journal of Language and Education*, 11(1), 148–152. <https://doi.org/10.17323/jle.2025.22118>
- Alharbi, W. (2023). AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools. *Education Research International*, 2023(1), 4253331. <https://doi.org/10.1155/2023/4253331>
- Al-kadi, A. (2025). Teacher-student dynamics in AI-driven language education in the post-truth era. *Journal of Language and Education*, 11(3), 152–159. <https://doi.org/10.17323/jle.2025.24343>
- Almelhes, S. A. (2023). A review of artificial intelligence adoption in second-language learning. *Theory and Practice in Language Studies*, 13(5), 1259–1269. <https://doi.org/10.17507/tpls.1305.21>
- Alshumaimeri, Y. A., & Alshememry, A. K. (2023). The extent of AI Applications in EFL learning and teaching. *IEEE Transactions on Learning Technologies*, 17, 1–11. <https://doi.org/10.1109/tlt.2023.3322128>
- Baskara, F. R. (2024). ChatGPT and personalised learning: Reshaping pedagogical approaches in the VUCA age. *Jurnal Educativa: Journal of Educational Studies*, 8(2), 167–167. <https://doi.org/10.30983/educative.v8i2.7620>
- Chang, W.-L., & Sun, J. C.-Y. (2024). Evaluating AI's impact on self-regulated language learning: A systematic review. *System*, 126, 103484. <https://doi.org/10.1016/j.system.2024.103484>
- Chapelle, C. A., & Voss, E. (2016). 20 years of technology and language assessment in *Language Learning & Technology*. *Language Learning & Technology*, 20(2), 116–128. <https://doi.org/10.64152/10125/44464>
- Chris, E., & Sherifdeen, K. (2024). *Student-Centered Learning: Leveraging Generative AI for Adaptive Learning Environments*. International Journal of Science and Research Archive.
- Crompton, H., Edmett, A., Neenaz Ichaporia, & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology (Print)*, 55(6), 2503–2529. <https://doi.org/10.1111/bjet.13460>

## DECLARATION OF COMPETING INTEREST

None declared.

## ACKNOWLEDGMENTS

The author would like to thank the anonymous reviewers and the editor for their helpful comments and suggestions.

## DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES

During the preparation of this manuscript, the author used ChatGPT for grammar checking. The author reviewed and edited all outputs and takes full responsibility for the content of the published article

- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., & Medaglia, R. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57(101994), 101994. <https://doi.org/10.1016/j.ijinfo.2019.08.002>
- Ellegaard, O., & Wallin, J. A. (2015). The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics*, 105(3), 1809–1831. <https://doi.org/10.1007/s11192-015-1645-z>
- Fellnhöfer, K. (2019). Toward a taxonomy of entrepreneurship education research literature: A bibliometric mapping and visualization. *Educational Research Review*, 27, 28–55. <https://doi.org/10.1016/j.edurev.2018.10.002>
- Hallinger, P., & Kovačević, J. (2019). A bibliometric review of research on educational administration: Science mapping the literature, 1960 to 2018. *Review of Educational Research*, 89(3), 335–369. <https://doi.org/10.3102/0034654319830380>
- Hegelheimer, V., Dursun, A., & Li, Z. (2015). Automated writing evaluation in language teaching: Theory, development, and application. *CALICO Journal*, 33(1), i–v. <https://doi.org/10.1558/cj.v33i1.29251>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, 26(1), 112–131. [https://doi.org/10.30191/ETS.202301\\_26\(1\).0009](https://doi.org/10.30191/ETS.202301_26(1).0009)
- Hubbard, P., & Levy, M. (2016). Theory in computer-assisted language learning research and practice. In *The Routledge handbook of language learning and technology* (pp. 24–38). Routledge.
- Hwang, G.-J., Xie, H., Wah, B., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, 1(1), 100001. <https://doi.org/10.1016/j.caeai.2020.100001>
- Jia, F., Sun, D., Ma, Q., & Looi, C. (2022). Developing an AI-based learning system for L2 learners' authentic and ubiquitous learning in English language. *Sustainability*, 14, Article 15527. <https://doi.org/10.3390/su142315527>
- Kartal, G., & Yeşilyurt, Y. E. (2024). A bibliometric analysis of artificial intelligence in L2 teaching and applied linguistics between 1995 and 2022. *ReCALL*, 1–17. <https://doi.org/10.1017/S0958344024000077>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Stadler, M. (2023). ChatGPT for good? on opportunities and challenges of Large Language Models for education. *Learning and Individual Differences*, 103. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence*, 2(100017), 100017. <https://doi.org/10.1016/j.caeai.2021.100017>
- Kirkpatrick, R., Vafadar, H., & Mohebbi, H. (2024). A scoping review of willingness to communicate in language education: Premises to doubt, lessons to learn, and future research questions to examine. *Language Testing in Asia*, 14(1). <https://doi.org/10.1186/s40468-024-00284-2>
- Knox, J. (2020). Artificial intelligence and education in China. *Learning, Media and Technology*, 45(3), 1–14. <https://doi.org/10.1080/17439884.2020.1754236>
- Lantolf, J., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.
- Linling, Z. (2025). A Ten-year perspective of bibliometric analysis of flipped classroom in English as a Foreign Language (EFL) Context. *International Journal of Education and Information Technologies*, 19, 10–20. <https://doi.org/10.46300/9109.2025.19.2>
- Liu, G. L., Darvin, R., & Ma, C. (2024a). Exploring AI-mediated Informal Digital Learning of English (AI-IDLE): A mixed-method Investigation of Chinese EFL Learners' AI Adoption and Experiences. *Computer Assisted Language Learning*, 1–29. <https://doi.org/10.1080/09588221.2024.2310288>
- Liu, Y., & Chang, P. (2024). Exploring EFL teachers' emotional experiences and adaptive expertise in the context of AI advancements: A positive psychology perspective. *System*, 126, 103463. <https://doi.org/10.1016/j.system.2024.103463>
- Liu, Y., Zhang, H., Jiang, M., Chen, J., & Wang, M. (2024b). A systematic review of research on emotional artificial intelligence in English language education. *System*, 126, 103478–103478. <https://doi.org/10.1016/j.system.2024.103478>
- Liu, Z., Zhang, W., & Yang, P. (2025). Can AI chatbots effectively improve EFL learners' learning effects? A meta-analysis of empirical research from 2022–2024. *Computer Assisted Language Learning*, 1–27. <https://doi.org/10.1080/09588221.2025.2456512>
- Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, 11(1), 50. <https://doi.org/10.1186/s40561-024-00342-5>

- Long, M. H. (1981). Input, interaction, and second-language acquisition. *Annals of the New York Academy of Sciences*, 379(1), 259–278. <https://doi.org/10.1111/j.1749-6632.1981.tb42014.x>
- Luckin, R., Holmes, W., Griffiths, M., & Pearson, L. (2016). *Intelligence unleashed an argument for AI in education*. Pearson.
- MacIntyre, P. D., Gregersen, T., & Mercer, S. (2019). Setting an agenda for positive psychology in SLA: Theory, practice, and research. *The Modern Language Journal*, 103(1), 262–274. <https://doi.org/10.1111/modl.12544>
- Mohamed, A. M. (2023). Exploring the potential of an AI-based chatbot (ChatGPT) in enhancing English as a foreign language (EFL) teaching: Perceptions of EFL faculty members. *Education and Information Technologies*, 29(3), 3195–3217. <https://doi.org/10.1007/s10639-023-11917-z>.
- Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., & Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. *El Profesional De La Información*, 29(1), 4. <https://doi.org/10.3145/epi.2020.ene.03>.
- Qasem, F., Ghaleb, M., Mahdi, H. S., Khateeb, A. A., & Fadda, A. H. (2023). Dialog chatbot as an interactive online tool in enhancing ESP vocabulary learning. *Saudi Journal of Language Studies*. <https://doi.org/10.1108/SJLS-10-2022-0072>
- Ryan, S., & Liu, H. (2022). Introducing Positive Psychology in Applied Linguistics: Asian Perspectives. *Chinese Journal of Applied Linguistics*, 45(1), 3–6. <https://doi.org/10.1515/cjal-2022-0101>
- Sharadgah, T. A., & Sa'di, R. A. (2022). A systematic review of research on the use of artificial intelligence in English language teaching and learning (2015-2021): What are the current effects? *Journal of Information Technology Education: Research*, 21, 337–377. <https://doi.org/10.28945/4999>
- Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook messenger. *Computers & Education*, 151, Article 103862. <https://doi.org/10.1016/j.compedu.2020.103862>
- Sun, Z., Anbarasan, M., & Praveen Kumar, D. J. C. I. (2021). Design of online intelligent English teaching platform based on artificial intelligence techniques. *Computational Intelligence*, 37(3), 1166–1180. <https://doi.org/10.1111/coin.12351>
- Sutton, A., Clowes, M., Preston, L., & Booth, A. (2019). Meeting the review family: Exploring review types and associated information retrieval requirements. *Health Information and Libraries Journal*, 36(3), 202–222. <https://doi.org/10.1111/hir.12276>
- Sweileh, W. M., Al-Jabi, S. W., AbuTaha, A. S., Zyoud, S. H., Anayah, F. M. A., & Sawalha, A. F. (2017). Bibliometric analysis of worldwide scientific literature in mobile - health: 2006–2016. *BMC Medical Informatics and Decision Making*, 17(1). <https://doi.org/10.1186/s12911-017-0476-7>
- Tan, X., Cheng, G., & Ling, M. H. (2024). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education Artificial Intelligence*, 8, 100355–100355. <https://doi.org/10.1016/j.caeai.2024.100355>
- Van Eck, N. J., & Waltman, L. (2020). Visualizing bibliometric networks. In *Measuring scholarly impact: Methods and practice* (pp. 285–320). Springer International Publishing.
- Wagner, C. S. (2018). *The collaborative era in science: governing the network*. Palgrave Macmillan.
- Warschauer, M., & Grimes, D. (2008). Automated Writing Assessment in the Classroom. *Pedagogies: An International Journal*, 3(1), 22–36. <https://doi.org/10.1080/15544800701771580>
- Xia, Q., Chiu, T. K., Lee, M., Sanusi, I. T., Dai, Y., & Chai, C. S. (2022). A self-determination theory (SDT) design approach for inclusive and diverse artificial intelligence (AI) education. *Computers & Education*, 189, 104582. <https://doi.org/10.1016/j.compedu.2022.104582>
- Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL*, 34(3), 327–343. <https://doi.org/10.1017/S0958344022000039>
- Yesilyurt, Y. E. (2023). AI-Enabled Assessment and Feedback Mechanisms for Language Learning. *Advances in Educational Technologies and Instructional Design Book Series* (pp. 25–43). IGI Global Scientific Publishing. <https://doi.org/10.4018/978-1-6684-9893-4.ch002>
- Yu, B. (2023). Self-regulated learning: A key factor in the effectiveness of online learning for second language learners. *Frontiers in Psychology*, 13(13). <https://doi.org/10.3389/fpsyg.2022.1051349>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1–27. <https://doi.org/10.1186/s41239-019-0171-0>
- Zawacki-Richter, O., Kerres, M., Bedenlier, S., Bond, M., & Buntins, K. (Eds.). (2020). *Systematic Reviews in Educational Research*. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-27602-7>
- Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, 4. <https://doi.org/10.1016/j.caeai.2023.100134>. Article 100134.

- 
- Zhang, R., & Zou, D. (2020). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning*, 33, 1–47. <https://doi.org/10.1080/09588221.2020.1744666>
- Zhong, L. L. (2025a). A ten-year perspective of bibliometric analysis of flipped classroom in English as a Foreign Language (EFL) context. *International Journal of Education and Information Technologies*, 19, 10–20. <https://doi.org/10.46300/9109.2025.19.2>
- Zhong, L. L. (2025b). An examination of the literature on self-regulated learning in EFL education: A bibliometric analysis. *Cogent Education*, 12(1). <https://doi.org/10.1080/2331186x.2025.2464363>
- Zou, B., Du, Y., Zhi Heng Wang, Chen, J., & Zhang, W. (2023). An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *SAGE Open*, 13(3). <https://doi.org/10.1177/21582440231193818>
- Zou, B., Liviero, S., Ma, Q., Zhang, W., Du, Y., & Xing, P. (2024). Exploring EFL learners' perceived promise and limitations of using an artificial intelligence speech evaluation system for speaking practice. *System*, 126, 103497. <https://doi.org/10.31219/osf.io/ec68h>

## APPENDIX

### Full Search Strategies and Reproducibility Details

#### **Database: Web of Science Core Collection**

---

Search fields	Topic (TS)
Timespan	1996–2025 (1, July)
Language	English
Document types	Article, Review
Subject scope	Studies addressing artificial intelligence applications in English as a Foreign Language (EFL) education (No additional subject-area filters were imposed to avoid excluding interdisciplinary research relevant to AI-enhanced EFL education).
Complete Boolean search string	(TS=((“artificial intelligence” OR “AI” OR “machine intelligence” OR “machine learning” OR “deep learning” OR “deep network” OR “Intelligent agent*” OR “intelligent tutor*” OR “intelligent tutoring system*” OR “hybrid intelligent system*” OR “neural network” OR “natural language processing” OR “Chatbot*” OR “thinking computer system*”)) AND TS=(“EFL” OR “English as a foreign language”))
Controlled vocabulary	No controlled vocabulary (e.g., thesaurus or subject heading terms) was applied, as the Web of Science database does not rely on a standardized controlled vocabulary system. Searches were conducted using author keywords, Keywords Plus, and indexed terms within the Topic field.

---

#### **Database: Scopus**

---

Search fields	TITLE-ABS-KEY (Title, Abstract, and Author Keywords)
Timespan	1996–2025 (1, July)
Language	English
Document types	Article, Review
Subject scope	Studies addressing artificial intelligence applications in English as a Foreign Language (EFL) education (No additional subject-area filters were imposed to avoid excluding interdisciplinary research relevant to AI-enhanced EFL education).
Complete Boolean search string	TITLE-ABS-KEY((“artificial intelligence” OR “AI” OR “machine intelligence” OR “machine learning” OR “deep learning” OR “deep network” OR “Intelligent agent*” OR “intelligent tutor*” OR “intelligent tutoring system*” OR “hybrid intelligent system*” OR “neural network” OR “natural language processing” OR “Chatbot*” OR “thinking computer system*”) AND (“EFL” OR “English as a foreign language”))
Controlled vocabulary	No controlled vocabulary was employed. The Scopus search relied on free-text terms within the TITLE-ABS-KEY fields, which index author-provided keywords and database-generated terms.

---

All search strategies were executed using the exact Boolean strings, truncations, parentheses, and filters described above. These details are provided to ensure full transparency and reproducibility, allowing other researchers to replicate or extend the bibliometric analysis under identical search conditions.