

Mapping the Terrain: A Scoping Review of Artificial Intelligence Tool Classification Frameworks in English Language Teaching

N. Polraj alias Goutham¹, K. Gurusamy², A. Pradheepa³

¹Research Scholar, Department of English, School of Liberal Arts and Special Education, Kalasalingam Academy of Research and Education, Krishnankovil, 626126, Virudhunagar, Tamil Nadu, India

²Associate Professor, Department of English, School of Liberal Arts and Special Education, Kalasalingam Academy of Research and Education, Krishnankovil, 626126, Virudhunagar, Tamil Nadu, India

³Assistant Professor, Department of English, AAA College of Engineering and Technology, Amathur, 626005, Virudhunagar, Tamil Nadu, India

ABSTRACT: While the number of artificial intelligence (AI) tools has been growing at an extraordinary pace in the field of English Language Teaching (ELT), it has been difficult to remain organised and clear in this field, where the tools come in all shapes and sizes. The need to capture a shared perspective on what a given AI tool is, how it is described, and how it is used has become clearer for researchers and educators alike. This is a scoping review based on the framework recommended by Arksey and O'Malley and guided by the PRISMA-ScR guidelines, covering AI tools in ELT and related language-education fields between 2016 and 2026. After systematically searching the open-access literature from Scopus, Wiley, Springer, and DOAJ, 22 studies were included. The review identifies five recurring classificatory logics across the studies: function-based - classifications based on the language skill supported; technique-based - classifications based on the primary AI technique, such as automatic speech recognition or natural language processing; educational-role-based - classifications based on the interaction between the learner and the AI, such as tutor or tool; ethical-spectrum-based - classifications based on the degree of learner agency preserved; and literacy-oriented - classifications based on the capabilities that learners and educators must build. The review shows that most frameworks are single-study proposals derived from a small number of tools, with limited testing across contexts. Few frameworks are ELT-specific, and there is little cross-testing between frameworks. It concludes with a synthesis that maps a shared spectrum of classificatory dimensions and outlines an agenda for a more empirically grounded ELT classification.

KEYWORDS: Artificial Intelligence, Classification Framework, Taxonomy, English Language Teaching, Scoping Review.

I. INTRODUCTION

Artificial intelligence (AI) has accompanied the field of English Language Teaching (ELT) for roughly a decade, and discussion of AI's role in ELT has continued across that period. Chatbots, automated writing assessment systems, and generative large language models (LLMs) such as ChatGPT are now expected to be used within classroom writing instruction. New systems and tools continue to be introduced - automated writing assessment, conversational chatbots, speech-recognition systems supporting pronunciation, and adaptive vocabulary systems - alongside ChatGPT itself. For teachers and researchers, the result is a perplexing array of tools that differ in use, function, and mechanism, yet are all marketed as "AI tools" (Godwin-Jones, 2024).

This is not merely an academic distinction. Tools that share surface themes - such as working with learner data or acting as a pronunciation trainer - often rest on very different designs. Alongside this diversity, several classification schemes have emerged: some grounded in technical features, some in pedagogical purpose, and some in the ethical implications raised by Shiri (2024) and Elmahdi et al. (2025). These schemes have largely developed within their own disciplinary communities - computer-assisted language learning, educational technology, information science, and composition studies - with limited interaction or integration across them. Three reasons justify approaching this review as a scoping review: the topic is still emerging, the relevant literature is heterogeneous in design (conceptual papers, systematic reviews, empirical validations), and few studies specifically address a general framework for classifying AI tools; instead, most reviews address the effects of AI on learning outcomes (AlTwijri & Alghizzi, 2024; Uygun & Demir, 2026) or contributions of AI to learning outcomes more broadly (Li et al., 2025).

Accordingly, this review is guided by three research questions: RQ1: What published classification schemes for AI tools exist in ELT and related language-learning literature between 2016 and 2026, and how might they be used or related across contexts? RQ2: What common classificatory dimensions recur across these schemes? RQ3: What are the relations and intersections among

Mapping the Terrain: A Scoping Review of Artificial Intelligence Tool Classification Frameworks in English Language Teaching

the five recurring classificatory logics identified (function, technique, educational role, ethics, and literacy)? The review does not evaluate the educational effectiveness of the tools themselves. Its conceptual value lies instead in clarifying the boundaries of "classification" as a foundation for future frameworks - a foundation that will itself need to be revisited as the field evolves.

II. METHOD

The five-stage scoping review framework of Arksey and O'Malley (2005) - research question, identification of studies, study selection, charting the data, and reporting the findings - guided the review, which followed the PRISMA Extension for Scoping Reviews (PRISMA-ScR) checklist recommended by Tricco et al. (2018). A scoping, rather than systematic, design was selected because the aim was to map the current breadth and diversity of classification frameworks rather than to establish certainty about an intervention effect, consistent with the guidance of Munn et al. (2018). The electronic search covered four pools: Scopus-indexed journals, Wiley Online Library, Springer Open, and the Directory of Open Access Journals (DOAJ), restricted to gold and green open-access sources to ensure independent verifiability. Three concept blocks structured the search: (1) "artificial intelligence", "generative AI", "chatbot", or "large language model"; (2) "classification", "framework", or "typology"; and (3) "English language teaching", "EFL", "ESL", "language learning", or "language education".

Following the citation-tracking recommendations of Neca et al. (2022), the search was conducted in three phases. In Phase 1, the general concept blocks were applied across all four source pools, producing an initial candidate set. In Phase 2, studies whose titles did not explicitly mention classification but referenced a classification scheme were identified through comparison with the compound terms used in Phase 1 (e.g., "classification EFL", "language learning", "generative AI framework writing instruction"). In Phase 3, the reference lists and citing-article networks of studies remaining after Phase 2 were analysed to identify further relevant frameworks not surfaced by the keyword search. Inclusion criteria were: (1) published between 2016 and 2026; (2) freely available as a full-text article from an open-access source indexed in Scopus, Wiley, or a DOAJ-listed language-education or language-learning journal; (3) proposed a classification scheme or typology for AI-based tools, technologies, or applications in language education (ELT/EFL/ESL) or language learning generally; and (4) reported the classificatory dimensions in sufficient detail to be charted. Studies were excluded if they were paywalled, preprints, or patents for which no open-access version could be obtained.

Data were extracted into a standardised table recording authorship and year, publication venue, methodology, classificatory dimension(s), and up to six tool-category labels. No formal risk-of-bias assessment was conducted, consistent with the breadth-oriented aim of a scoping review; open-access status for each source was verified against the publisher's website before inclusion.

III. FINDINGS AND DISCUSSION

A. Overview of Included Studies

The literature search identified 22 studies meeting the inclusion criteria: ten published in gold open-access journals, eleven in green open-access journals indexed in Scopus and listed in DOAJ, and one further study identified through citation tracking. Table I presents a representative sample of the included studies and the classificatory focus each proposes.

TABLE I: REPRESENTATIVE OVERVIEW OF INCLUDED STUDIES

Author(s)/Year	Journal	Classificatory Focus	Key Categories Proposed
Elmahdi et al. (2025)	International Journal of Artificial Intelligence	Teaching role	AI as semiotic mediator; analysis tools vs. production assistants across a Listening-Composing Cycle
Uygun & Demir (2026)	Discover Education	Function	Personalization tools, adaptive-feedback tools, writing-support tools
Li et al. (2025)	Computers & Education: AI	Research-focus classification	Perceptions/attitudes, cognitive-psychological, teaching strategies, skills development, writing/feedback, implementation
Huang, Hew & Fryer (2022)	Journal of Computer Assisted Learning	Social presence	Chatbots as discussor, simulation, transmission channel, helpline, recommender
Chiu (2025)	Smart Learning Environments	Activity-theory role	Subject, tool, rules, community, division of labor mapped onto chatbot functions
Lyu et al. (2025)	International Journal of Applied Linguistics	Deployment origin	Publicly available chatbots, researcher-built custom chatbots, intelligent personal assistants

Mapping the Terrain: A Scoping Review of Artificial Intelligence Tool Classification Frameworks in English Language Teaching

Author(s)/Year	Journal	Classificatory Focus	Key Categories Proposed
Gabay et al. (2026)	IJEMST	Application domain	Idea generation, drafting, revision, feedback, and citation-support applications in academic writing
Shiri (2024)	Digital Library Perspectives	Faceted literacy taxonomy	13 qualities across technical, ethical, and application-oriented AI literacy competencies
Godwin-Jones (2024)	Language Learning & Technology	Agency-distribution	Tools distinguished by the degree of learner vs. system control over language production
Godwin-Jones (2023)	Language Learning & Technology	Environment	Chatbots, ambient/IoT systems, and immersive metaverse tools classified by spatial embodiment
Pedersen (2023)	Frontiers in Artificial Intelligence	Cultural / ethical viewpoint	Writing tools positioned along a spectrum from facilitative aid to substitutive authorship
Chan & Tang (2025)	World Journal of English Language	Teacher-knowledge (TPACK)	AI tools mapped onto technological, pedagogical, and content-knowledge intersections
Peña-Acuña & Corga Fernandes Durão (2024)	Frontiers in Education	Skill-domain	Chatbots, pronunciation apps, adaptive-content platforms for prospective EFL teachers

Note: Table I lists a representative subset of the 22 included studies, illustrating the range of classificatory approaches identified.

B. Five Recurring Classificatory Judgments

1) Function-based classifications: AI tools are largely classified by the language skill or task they target - listening, speaking, reading, writing, vocabulary, or grammar (Elmahdi et al., 2025; Gabay et al., 2026; Peña-Acuña & Corga Fernandes Durão, 2024). This logic is familiar to educators because it preserves the four-skills approach long emphasised in ELT pedagogy, but it becomes harder to sustain once learners use general-purpose generative AI tools such as ChatGPT, Claude, or Grok across multiple skills and purposes simultaneously (Li et al., 2025).

2) Technique-based classifications: A second logic classifies tools by the underlying digital technique: natural language processing (NLP), automatic speech recognition (ASR), machine translation, and generative large language models (LLMs). This framing is common in computer-science-oriented reviews and describes what a system does - for example, parsing a string into a syntax tree - but offers little guidance on how the underlying capability might be taught, since a single technique (such as automated scoring) can serve very different pedagogical purposes.

3) Educational-role classifications: A third logic classifies tools by the relational position the AI occupies with respect to the learner. Huang et al. (2022) classify chatbot functions into five roles: interlocutor, simulation, transmission channel, helpline, and recommender. Chiu (2025), drawing on activity theory, distinguishes chatbots acting as semiotic mediators for receptive versus productive activities (in relation to rules, community, and division of labour). Elmahdi et al. (2025) similarly distinguish semiotic mediation for receptive versus productive activities. This logic is useful for comparing tools across skill areas and underlying techniques.

4) Ethical-spectrum classifications: A fourth logic, prominent in composition-studies and writing-studies literature, classifies tools by the degree of learner agency preserved, ranging from editorial assistance to near-complete delegation of authorship (Pedersen, 2023). This dimension is less concerned with academic-integrity policy than with the kind of learning activity a learner is engaged in when using a specific tool for a specific task.

5) Literacy-oriented classifications: A fifth logic concerns the capabilities teachers and learners need to engage critically with AI tools, exemplified by a thirteen-facet AI-literacy taxonomy (Shiri, 2024) and a teacher-readiness model (Chan & Tang, 2025). Tool-centred classifications are complemented by these frameworks, since questions of "what is the tool?" are not equivalent to questions of "what human capability does using the tool require?" - a distinction also emphasised by Stracke et al. (2023).

C. Junctions, Separations, and Shared Elements

Table II relates the five classificatory logics to their guiding questions and representative sources. No single reviewed study applies all five dimensions simultaneously; most studies draw on only one or two and leave the remainder unaddressed. This suggests that

Mapping the Terrain: A Scoping Review of Artificial Intelligence Tool Classification Frameworks in English Language Teaching

the field does not share one unified vocabulary for classification but rather several partially overlapping vocabularies, echoing the fragmentation seen in other areas of L2 writing research.

TABLE II: FIVE CLASSIFICATORY DIMENSIONS AND THEIR GUIDING QUESTIONS

Dimension	Guiding Question	Representative Sources
Linguistic function	Which skill or subskill does the tool support?	Elmahdi et al. (2025); Gabay et al. (2026)
Digital technique	What AI method underlies the tool (NLP, ASR, generative model)?	Li et al. (2025); Godwin-Jones (2024)
Educational role	What relational position does the AI occupy (tutor, device, partner)?	Huang et al. (2022); Chiu (2025)
Learner agency / ethics	How much control over output does the learner preserve?	Pedersen (2023)
Required literacy	What capabilities must the user hold to use the tool critically?	Shiri (2024); Chan & Tang (2025)

D. Implications for ELT Practice

This lack of a unified classification is not only a theoretical challenge; it also affects the selection, application, and evaluation of AI tools at the institutional and classroom level. When considering a generative writing assistant, for example, the decision is better guided by the five questions raised in Table II - linguistic function, underlying technique, pedagogical role, learner agency, and required literacy - than by a binary "AI or no AI" judgement. Chan and Tang (2025) and Elmahdi et al. (2025) make this point explicitly. Because teachers show uneven readiness across the dimensions of TPACK, a tool that is effective for one purpose may not be effective for another (Chan & Tang, 2025); tools that build listening skills, for instance, may not transfer to speaking skills and should be evaluated separately (Elmahdi et al., 2025). Teacher-education programmes may benefit from these classifications (Stracke et al., 2023; Shiri, 2024), which suggest that teacher preparation should emphasise transferable capabilities rather than tool-specific training, since AI tools change faster than curricula can be revised. This capability-first approach is consistent with recent research on generative AI in language learning, which likewise emphasises building durable evaluative skills over tool-specific competence (Li et al., 2025; Daniel et al., 2025).

IV. CONCLUSION

This scoping review drew on 22 open-access studies published between 2016 and 2026 that propose classification frameworks for AI tools in English Language Teaching and adjacent language-education contexts. Five recurring classificatory logics were identified - function-based, technique-based, educational-role-based, ethical-spectrum-based, and literacy-based - each capturing a distinct, non-overlapping facet of what it means to classify an AI tool for teaching purposes. No single reviewed framework integrates all five, and the literature remains distributed across disciplines - educational technology, composition studies, information science, and applied linguistics - that rarely cite one another.

This review carries the limitations inherent to scoping methodology. Because the objective was breadth rather than exhaustiveness, and because inclusion was restricted to open-access sources from Scopus-indexed, Wiley, Springer, and DOAJ outlets to ensure independent verifiability, some relevant paywalled frameworks may have been excluded, and no formal quality appraisal was performed. Future systematic or meta-synthesis work, informed by the map offered here, would be needed to weigh the relative empirical support behind each classificatory logic. Nonetheless, by consolidating a distributed literature into a single conceptual map, this review provides a foundation on which a validated, ELT-specific AI-tool classification can be built.

REFERENCES

- 1) 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>
- 2) Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19-32. <https://doi.org/10.1080/1364557032000119616>
- 3) Chan, K. K.-W., & Tang, W. K.-W. (2025). Evaluating English teachers' artificial intelligence readiness and training needs with a TPACK-based model. *World Journal of English Language*, 15(1), 129. <https://ideas.repec.org/a/jfr/wjel11/v15y2025i1p129.html>

Mapping the Terrain: A Scoping Review of Artificial Intelligence Tool Classification Frameworks in English Language Teaching

- 4) Chiu, T. K. F. (2025). Design language learning with artificial intelligence (AI) chatbots based on activity theory from a systematic review. *Smart Learning Environments*, 12, Article 26. <https://doi.org/10.1186/s40561-025-00379-0>
- 5) Daniel, B. K., et al. (2025). Can generative AI revolutionise academic skills development in higher education? A systematic literature review. *European Journal of Education*. <https://doi.org/10.1111/ejed.70036>
- 6) Elmahdi, O. E. H., Balla, A. A. S., Abdelrady, A. H., Osman, E., & Ahmed, A. O. A. (2025). A comprehensive taxonomy and pedagogical framework for AI-driven tools in EFL listening and speaking instruction: Integrating multimodal composing and adaptive learning. *International Journal of Artificial Intelligence*. <https://www.academicpublishers.org/journals/index.php/ijai/article/view/7550>
- 7) Gabay, R. A. E., Funa, A. A., & Ricafort, J. D. (2026). Generative artificial intelligence (GenAI) for academic writing in higher education: A scoping review of applications, challenges, and implications. *International Journal of Education in Mathematics, Science and Technology*, 14(1), 200-232. <https://files.eric.ed.gov/fulltext/EJ1494380.pdf>
- 8) Godwin-Jones, R. (2010). Emerging technologies. *Language Learning & Technology*, 14(2), 4-11. <https://doi.org/10.125/44208>
- 9) Godwin-Jones, R. (2023). Emerging spaces for language learning: AI bots, ambient intelligence, and the metaverse. *Language Learning & Technology*, 27(2), 6-27. <https://scholarspace.manoa.hawaii.edu/items/0bcd730d-6027-420c-b57a-b8f52c92af6e>
- 10) Godwin-Jones, R. (2024). Distributed agency in language learning and teaching through generative AI. *Language Learning & Technology*, 28(2), 5-30. <https://doi.org/10.125/73570>
- 11) Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning-Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237-257. <https://doi.org/10.1111/jcal.12610>
- 12) Jeon, J. (2025). Artificial intelligence in ESL/EFL education: Evidence from recent reviews (2024-2025). *International Journal of Learning, Teaching and Educational Research*. <https://ijlter.org/index.php/ijlter/article/view/14625>
- 13) Li, B., Tan, Y. L., Wang, C., & Lowell, V. (2025). Two years of innovation: A systematic review of empirical generative AI research in language learning and teaching. *Computers and Education: Artificial Intelligence*, 9, 100445. <https://doi.org/10.1016/j.caeai.2025.100445>
- 14) Lyu, Q., et al. (2025). Effectiveness of chatbots in improving language learning: A meta-analysis of comparative studies. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.12668>
- 15) Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18, Article 143. <https://doi.org/10.1186/s12874-018-0611-x>
- 16) Neca, P., Borges, M. L., & Pinto, P. C. (2022). Teachers with disabilities: a literature review. *International Journal of Inclusive Education*. <https://doi.org/10.1080/13603116.2020.1776779>
- 17) Peña-Acuña, B., & Corga Fernandes Durão, R. (2024). Learning English as a second language with artificial intelligence for prospective teachers: A systematic review. *Frontiers in Education*, 9, Article 1490067. <https://doi.org/10.3389/feduc.2024.1490067>
- 18) Pedersen, I. (2023). The rise of generative AI and enculturating AI writing in postsecondary education. *Frontiers in Artificial Intelligence*, 6, Article 1259407. <https://doi.org/10.3389/frai.2023.1259407>
- 19) Shi, X. (2025). Insights into the effectiveness of artificial intelligence-integrated speaking instruction in enhancing speaking skills and social-emotional competence as well as reducing demotivation and shyness. *European Journal of Education*. <https://doi.org/10.1111/ejed.70174>
- 20) Shiri, A. (2024). Artificial intelligence literacy: A proposed faceted taxonomy. *Digital Library Perspectives*, 40(4), 681-699. <https://doi.org/10.1108/DLP-04-2024-0067>
- 21) Stracke, C. M., Chounta, I. A., Holmes, W., Tlili, A., & Bozkurt, A. (2023). A standardised PRISMA-based protocol for systematic reviews of the scientific literature on Artificial Intelligence and education (AI&ED). *Journal of Applied Learning and Teaching*, 6(2), 64-70. <https://doi.org/10.37074/jalt.2023.6.2.38>
- 22) Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467-473. <https://doi.org/10.7326/M18-0850>
- 23) Uygun, E., & Demir, B. (2026). A PRISMA-based systematic review of artificial intelligence in English as a foreign and second language education (2023-2025). *Discover Education*, 5. <https://doi.org/10.1007/s44217-026-01504-y>