Received: 30 September 2023 | Accepted: 15 March 2024

DOI: 10.1111/bjet.13460

ORIGINAL ARTICLE

British Journal of Educational Technology BERA

AI and English language teaching: Affordances and challenges

Helen Crompton[1,2] | Adam Edmett[3] | Neenaz Ichaporia[3] | Diane Burke[2]

[1]STEM-Professional Studies, Old Dominion University, Norfolk, Virginia, USA

[2]RIDIL ODUGlobal, Old Dominion University, Norfolk, Virginia, USA

[3]British Council, Manchester, UK

Correspondence
Helen Crompton, STEM-Professional Studies, Old Dominion University, Norfolk, VA, USA.
Email: crompton@odu.edu

Funding information
British Council

Abstract

English is one of the most used languages for jobs, markets, tourism, discourse and international connectivity. However, English learners face many challenges in gaining English language skills. Extant studies show that AI has affordances to support in English language teaching and learning ELT/L. This study answers the call to examine specific challenges and affordances for using AI in ELT/L. A systematic review method was used with PRISMA principles to identify 42 studies. Findings reveal the geographical locations of studies, learner ages and years of study. Grounded coding was then used to identify affordances of the use of AI in ELT/L in the areas of speaking, writing, reading, pedagogy and self-regulation. AI in ELT/L challenges uncovered were technology breakdowns, limited capabilities, fear and standardising language. Policymakers, funders, practitioners and educational leaders can use the information provided in this study to gain a holistic understanding of the current trend in the use of AI in ELT/L, and practical implications are provided to guide future use of AI.

KEYWORDS

AI, AIEd, artificial intelligence, education, English foreign language, English language learning, English language teaching, GenAI, generative AI

Practitioner notes

What is already known about this topic

- English is one of the most used languages for jobs, markets, tourism, discourse and international connectivity.
- Empirical evidence shows that pupils can often face difficulties when learning English, with challenges such as irregularity in English spelling.
- AI has supported language teaching and learning with studies showing that AI can support language-specific skills.

What this paper adds

- Provides the scholarly community with a unique systematic review in the use of AI in ELT/L across learner levels.
- Identifies affordances of AI in ELT/L in speaking, writing, reading, pedagogy and self-regulation.
- Identifies challenges of AI in ELT/L in technology breakdowns, limited capabilities, fear and standardising language.
- Provides researchers with a review of the field with identification of gaps and future research opportunities.

Implications for practice and/or policy

- Provides practical implications from the findings for educators, policy makers and program designers.
- Highlights the gaps in academic knowledge as a lack in the use of AI for assessment in ELT/L.

INTRODUCTION

English is one of the most used languages for jobs, markets, tourism, discourse and international connectivity (Lan et al., 2020). To achieve mastery of the English language, a pupil acquires proficiency in speaking, listening, reading and writing (Grabe & Stoller, 2002). However, empirical evidence shows that pupils can often face difficulties when learning English, with challenges such as limited connections to English outside the language courses, insufficient content knowledge (Tokoz-Goktepe, 2014) and irregularity in English spelling (Ulicheva et al., 2018). Pupils can also fear making embarrassing mistakes in front of peers (Cumming et al., 2018). AI has been used to remove or ameliorate some of those challenges (Baranwal, 2022; Kannan & Munday, 2018). AI are computer systems that simulate human intelligence (Sindermann et al., 2021) and can learn, understand and remember human language (Xiaohong & Yanzheng, 2021). These programs can help provide a safe environment for learning English.

AI has supported language teaching and learning with studies showing that AI can support language-specific skills, such as reading comprehension (Xu et al., 2019), practicing repetitive language skills on a machine (Kim, 2019) and correcting English pronunciation (Noviyanti, 2020). It can also help with broader teaching and learning activities, including automatic marking, providing feedback, adaptive educational experiences, intelligent tutoring and prediction systems (Pokrivčáková, 2019).

With the rise of AI use in education, it is important for educators to understand how AI can be used to support the teaching and learning of English to those whose first language is not English. Therefore, a review of extant studies is needed to support the academic and practitioner community in understanding how AI is supporting English language educators and pupils. This is echoed by scholars (eg, Sharadgah & Sa'di, 2022; Yang & Kyun, 2022) who called for an up-to-date systematic review of AI in English language education. Recent systematic reviews have examined AI in education, but not focused on English language education (viz., Crompton et al., 2022; Zawacki-Richter et al., 2019). Those targeting English language education have often focused on a single type of AI (eg, Baranwal, 2022), or examined studies learning language in general without a focus on English (eg, Yang & Kyun, 2022). Those systematic reviews that were focused on AI and English have not examined all learners including K-12, higher education and adult learners. Most importantly, the examination of many of the studies involved examining the data against a predetermined framework (eg, Sharadgah & Sa'di, 2022), which can lead to confirmation bias.

Therefore, the purpose of this study is to answer the call for an up-to-date systematic review on English language teaching and learning, while also addressing the gaps identified in the literature. Namely, that the review (1) includes learners of all levels, (2) includes all types of AI, (3) targets the learning of English and (4) uses an inductive method to examine the trends (rather than predetermined criteria). The findings of this study will provide an overview of the current state of AI's utilisation in English language teaching and learning. This information aims to provide valuable insights for both practitioners and academic audiences. Educators will gain a reference point for integrating AI into teaching English, encompassing pedagogical approaches and connections to language skills. The study will also uncover challenges, enabling educators to address issues, navigate misuses and understand limitations. Educational leaders, funders and policy makers can leverage this research for evidence-based policy and practice direction. Researchers will benefit from this overview of the state of AI in English language teaching and learning while also learning of the gaps in research to be addressed.

LITERATURE REVIEW

English language teaching and learning

As academics and practitioners focus efforts on the design and act of teaching and learning English language skills, many acronyms are often attached to various nuances and temporal trends. Computer-assisted language learning (CALL) and mobile-assisted language learning (MALL) focus on the technology used in language learning. English as a second language (ESL), English for speakers of other languages (ESOL) and English as a foreign language (EFL) focus on the learning of English as a language. Other terms such as TESL, TESOL and TEFL are focused on the teaching of English. For this study, the focus is on English as the target language for those whose first language is not English and investigating aspects of both teaching and learning. As English language teaching and learning (ELT/L) is the term used in other scholarly work (eg, Margana, 2016), it has been selected for use in this study.

Recent developments have seen a rise in AI, with empirical findings (eg, Crompton et al., 2022) revealing a trend towards using AI in language learning and writing skills more frequently than in other disciplines. Technology has been an integral support in ELT/L (Rivera Barreto, 2018). Technology provides teaching and learning resources, motivates pupils (Larsen-Freeman & Anderson, 2011), facilitates learning (Ahmadi, 2018) and provides new methods for learning (Gilakjani, 2017).

AI and ELT/L

AI is a multifaceted technology with three user categories: (1) learner facing, used by pupils to learn; (2) teacher facing, used by teachers to help in teaching activities, for example, grading; and (3) system facing, which is used by administrative staff to manage and examine pupil data (Pokrivčáková, 2019). Various AI technologies and systems provide affordances in ELT/L when targeting a specific user and objective. AI can mine large amounts of data, operate using natural language (speech, listening and writing) and follow rules and patterns of language.

Mobile devices have long been heralded as the tool for ELT/L with portability, universality, sharing and individuality (Ma, 2021). AI applications now available on those devices amplify ELT/L capabilities with speech synthesis, big data and intelligent systems (Luo & Cheng, 2020). While these AI technologies are far from human, they mimic humanness. Voice assistants offer an untiring chat partner with a human-like voice that can provide a variety of English language accents to choose from (Dizon & Tang, 2020). Amazon Alexa, Apple Siri, Samsung Bixby, Microsoft Cortina and Google Home Assistant are common voice assistants. In addition, intelligent tutors appear across a plethora of programs and can appear to the user as voice over assistance, text prompts and even take on the visual appearance in addition to voice (Vuong et al., 2023).

There are robots created to provide a physical embodiment of an AI. Erica and Sophie are humanoid robots capable of engaging in conversations with human-like expressions and body language (Sindermann et al., 2021). These qualities provide a more lifelike experience for English learners. In recent years, chatbots were a basic form of AI, as they were limited to answering simple questions through a limited text format. At the end of 2022, chatbot capabilities greatly expanded with AI transformers and large language models offering Microsoft's ChatGPT, Google's Bard, Anthropic's Claude and Apple's AppleGPT. These chatbots provide extensive possibilities for language education, including engaging learners in extensive intelligent conversations and providing writing models of various genera (Bozkurt et al., 2023). While these AI tools show great capabilities for ELT/L, these affordances can only be gained when understanding what tools are available and specifics on how pedagogies/andragogies match those tools to advance ELT/L.

It is also important for all educational stakeholders to recognise the limitations and misuses of AI in ELT/L. This can allow educators to plan accordingly to avoid or ameliorate those issues. Scholars using AI with ELT/L have noted that AI can lack emotion (Annamalai et al., 2023). While AI can appear to be showing emotions, such as text describing feelings and avatars showing facial expressions, AI does not have the capacity to feel emotions and can only mimic text and expressions. This may not always be correct from the data it has. It could also be that the AI program is not advanced enough to mimic accurately. Similar AI challenges to Annamalai were reported by other scholars (eg, Wang et al., 2023) that found that AI was behaving passively and mechanically. While AI has many advanced capabilities, problems with technical issues, such as programs breaking or not functioning correctly, have been reported (eg, Ericsson et al., 2023).

Extant systematic reviews

Scholars have started to build a base of collective insights into AI in education and ELT/L. Scholars conducted systematic reviews to examine AI across all educational disciplines. Chen et al. focused on the annual trends, leading journals, institutions, countries/regions, theories and technologies adopted in AI and education. Systematic reviews were conducted

on K-12 (Crompton et al., 2022) and higher education (Crompton & Burke, 2023; Zawacki-Richter et al., 2019). Both the K-12 and higher education systematic reviews examined topics, such as educational disciplines, levels of learning, research purposes, methodologies, annual trends, intended users, affordances and challenges. While these AI reviews across all subject disciplines, they do not target ELT/L. It is interesting to note that the findings of the two higher education studies both found language learning one of the most common disciplines for AI use and the K-12 study found writing and language learning the only disciplines showing a growing trend in the use of AI.

Using a narrower lens to focus on ELT/L, scholars pinpointed systematic reviews on pedagogical approaches, language skills and AI tools. The reviews of Chen (2016) and Loncar et al. (2023) studied how technology was used in ELT/L, which included AI, among other technologies and programs. Zhang and Zou (2023) focused their study on the development of writing skills using AI in ELT/L. Using activity theory as a framework, papers were examined on the various factors involved in using AI in ELT/L that influence the efficacy of the activity. Switching to a focus on AI tools, Baranwal (2022) examined the use of AI teachable agents in English learning and Crompton and Burke (2024) on ChatGPT.

Broadening the scope, Yang and Kyun (2022) conducted a systematic review of the use of AI in language learning. This study also included pupils whose first language was English and who were learning language skills. This could be problematic in generalising trends across both those who speak English as a first language and those who do not, as gaining a new language can involve different teaching and learning challenges. This review of formal education (K-12 and higher education) across 2007–21 also involved pupils learning different languages beyond English. Furthermore, Yang and Kyun (2022) also used activity theory as a predetermined framework to examine the studies. This approach leaves fewer opportunities to see what emerges from the literature.

Sharadgah and Sa'di (2022) focused their systematic review on K-12 and higher education, but they focused their review specifically on the learning of English. Again, this review had predefined codes that the researchers connected the data to. To better identify the affordances and challenges of AI in ELT/L, a more inductive, grounded approach is needed to see what emerges from the research findings. Although the study by Sharadgah and Sa'di incorporates data up until 2021, it is important to note the rapid advancements in AI over the past year. These developments have led to the integration of large language models, such as ChatGPT, within generative AI tools, and have also resulted in enhanced backend systems. Consequently, these technological strides offer new potential avenues for the advancement of ELT/L.

To summarise the gaps in the systematic reviews, scholars (eg, Crompton et al., 2022; Zawacki-Richter et al., 2019) conducted reviews across learner levels, such as higher education, or K-12. They did not focus specifically on ELT/L. Others were too broad across all technologies in ELT/L (eg, Chen, 2016; Loncar et al., 2023). There are extant reviews that focus on AI and ELT/L only looked at one type of AI, such as Baranwal (2022), who examined teachable agents. This missed a review on all types of AI being used. There were studies examining AI used in teaching all languages (eg, Yang & Kyun, 2022) missing a focus on ELT/L. Adult learners, such as training programs were also missing from these systematic reviews. Finally, the systematic reviews used frameworks to examine the trends. Therefore, the researchers were looking for specific things, such as ELT/L skills (eg, reading, writing or types of pedagogies). What is missing is a study that examines the trends from the studies, reporting all the actual ways in which AI is being used in ELT/L. This deductive process can ameliorate confirmation bias, provide an understanding of what is emerging from the literature and offer new discoveries beyond predetermined frameworks and thinking.

Purpose

Sharadgah and Sa'di (2022) and Yang and Kyun (2022) called for an up-to-date systematic review of AI in English language education to understand how AI can be used in ELT/L. The extant systematic reviews advance what scholars know about AI in ELT/L. Nonetheless, the gaps in the literature show that an up-to-date study is needed to focus on ELT/L, across learners of all ages, and examine the literature from an inductive perspective to reveal what the data show. Therefore, the overarching question for this study is—How is AI being used for ELT/L? Four sub-questions provide more specifics with Questions 1 and 2 providing context for the studies and 3 and 4 going in-depth to examine the affordances and challenges.

1. Where are the AI studies taking place geographically and has there been a change in the trend in the number of studies across the years?
2. What educational level are the pupils in the studies?
3. How is AI being used for educational affordances in ELT/L?
4. What are the challenges of using AI in ELT/L?

METHOD

A systematic review methodology was conducted to answer the four questions guiding this study. The systematic review process involves two main parts. The first part is the search, identification and selection of the studies to be included in the study. The preferred reporting items for systematic reviews and meta-analysis (PRISMA principles; Page et al., 2021) were used for this. The second part of the systematic review is the analysis of those studies. This involved a mixed-method approach with quantitative methods to provide aggregate data with basic numbers, then qualitative inductive and deductive coding techniques using a priori and grounded coding (Strauss & Corbin, 1995) to provide a summary of how AI is being used in ELT/L and generate new theories from that collective review of the studies (Gough et al., 2017).

Search strategy

Only primary research was selected for this systematic review. Primary studies are studies in which the researcher has directly gathered data from participants and do not include systematic reviews or theoretical papers. In addition, to ensure a level of confidence in the quality of the research, only research published in a peer-reviewed journal was selected. Scholars (viz, Crompton & Burke, 2022, 2023) report an increase in the use of AI in education in the past decade. Therefore, the search parameters were set to include papers from 2014 to 2023 to keep current and include the latest additions to AI technologies.

The data retrieval protocol involved an electronic search and a hand search. For the electronic search, a Boolean string was used in the EBSCOhost platform of the educational databases and JSTOR, Science Direct and Web of Science. The Boolean search included ELT/L, AI and teaching/learning terms. The Boolean search is listed in Table 1. Searches were conducted with the various acronyms used in ELT/L, such as ESL, CLT and EFL, but the addition of the acronyms did not change the number of papers retrieved. On further examination, those acronyms (eg, ESL, CLT and EFL) were found in papers that also contained the terms in Part 3 of the search—'language learning', 'language teaching' and/or 'language acquisition'. Therefore, the acronyms did not further support finding relevant papers and were removed from the Boolean search.

TABLE 1 Boolean search terms.

Search section	Search terms
Part 1	'Artificial intelligence' OR 'AI'
Part 2	'high school' OR 'middle school' OR 'elementary school' OR 'secondary school' OR 'k-12' OR 'higher education' OR 'tertiary education' OR 'adult learner' OR 'professional development' OR 'continuing education' OR training
Part 3	'language learning' OR 'language teaching' OR 'language acquisition'

TABLE 2 Inclusion and exclusion criteria.

Inclusion	Exclusion
• Journal papers published between 2014 and 2023 • Peer-reviewed journal papers • Primary research • Involves teaching and learning English as another language • Uses AI in ELT/L • Journal papers written in English	• Conference proceedings • Editorials • Research that includes pupils learning about AI, such as a computer science class and not using AI for learning • Teaching and learning English to those who have English as their first language

To ensure the relevant research was located on ELT/L, a hand search was also conducted of ELT/L-specific journals that may not appear in databases due to their size or the age of the journal. The hand search included: TESL-EJ, Computer Assisted Language Learning, International Journal of Bilingual Education, and Bilingualism, Language Teaching Research, TESOL Quarterly and Studies in Second Language Acquisition.

Inclusion and exclusion criteria

The electronic and hand search resulted in the identification of 369 papers for possible inclusion in the systematic review. Of the 369, a total of 3 duplicates were removed, reducing the number of studies to 366. The 366 were then examined for a match against the inclusion/exclusion criteria listed in Table 2.

Two researchers independently reviewed each paper against the criteria, and the interrater reliability was calculated using percentage agreement (Belur et al., 2018). The two researchers gained an interrater agreement of 97%. However, a 100% agreement was achieved after discussing the misaligned papers. A total of 326 papers removed after applying the inclusion and exclusion criteria. At the end of the search selection process, 43 papers met the inclusion criteria for this study. Figure 1 presents the number of papers removed corresponding to the inclusion/exclusion criteria and the final number of studies for this systematic review.

Data analysis

Once the papers were identified, data were then extracted from the papers to match the four research questions guiding this study. A combination of 'a priori' and 'grounded' coding was used to examine that data. A priori coding is coding that is predetermined before examining the data. Geographical locations, years and education levels were all a priori coding as they used predefined categories. Geographical locations were coded

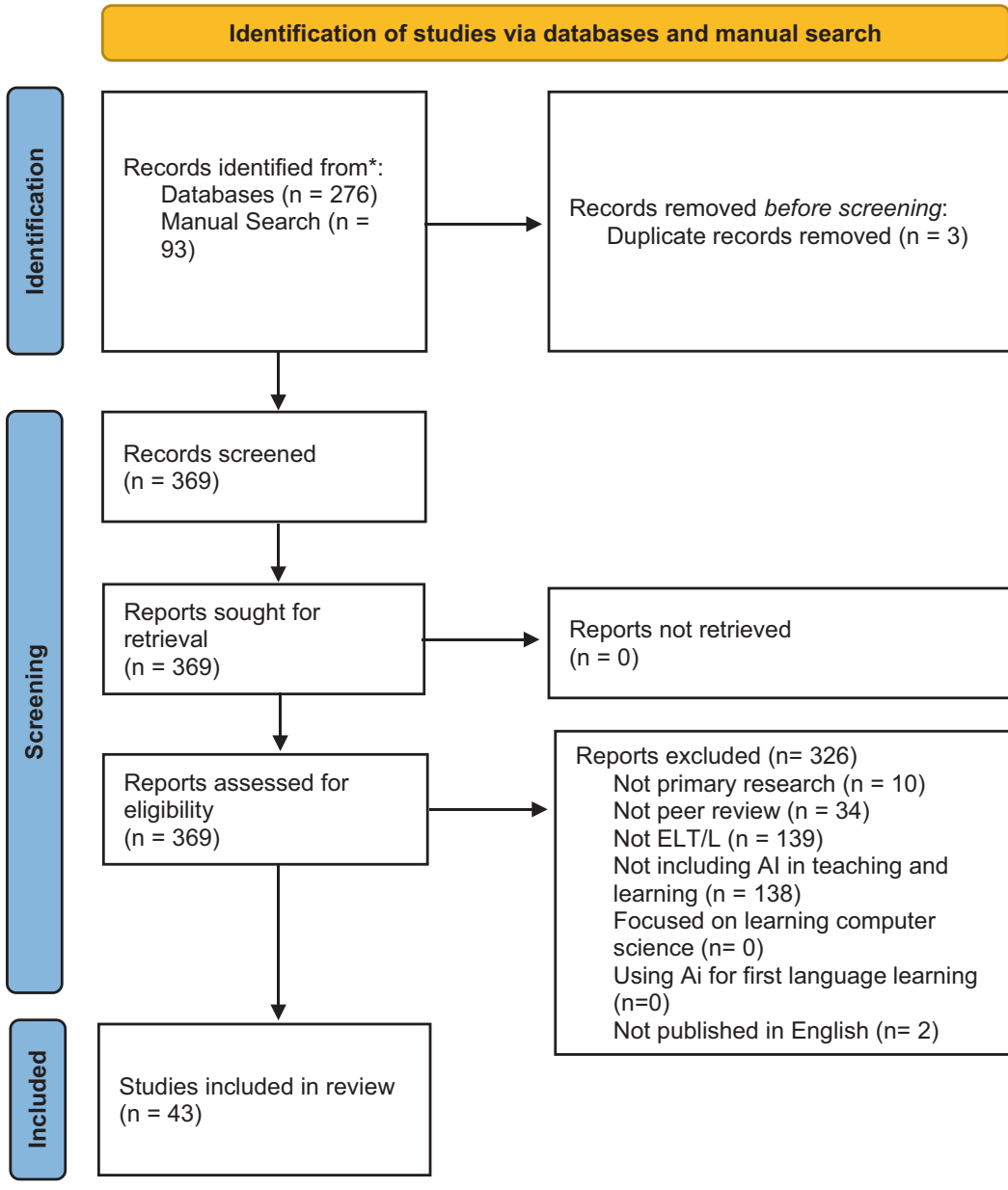

FIGURE 1 Adapted from Page et al. (2021).

by continent and country; years by the chronological year of publication; and educational levels as K-12 (learners 5 to 18 years of age), higher education (learners in formal university/tertiary settings) and adult (learners outside of a higher educational institution, eg, training).

Grounded coding (Strauss & Corbin, 1995) was then utilised to identify the affordances and challenges of AI in ELT/L. Grounded coding develops through an inductive method, not from predetermined categories or theories. This involved the researchers not specifically looking for a set of affordances or challenges from what prior researchers have found, but instead looking for any or all instances that are then highlighted and examined. Therefore,

you will notice that the literature review reported the state of the field but did not include a specific focus on any frameworks, models or theories. This is an important aspect of the grounded coding methodology to ameliorate bias.

The specific grounded coding design used was a constant comparative method during which researchers examine and code text segments from the papers. This involved two researchers individually reading the papers. When text relating to an affordance or challenge was noticed in the text, that text would be highlighted, and a descriptive code was listed next to the text. For example, text that said, 'pupils received immediate feedback' may be listed as 'Affordance' and also 'Rapid Feedback'. Within the grounded coding process, 'In vivo' (Saldana, 2015) coding was also used, which involves using language from the original papers as codes. Therefore, one of the codes for 'pupils received immediate feedback' may be coded as 'Immediate Feedback' rather than rapid feedback. Using the author's language supports consistency with the author's original meaning of the text.

As the coding process evolves, trends are found in similar codes across papers. Codes are often broken into further axial codes as needed. For example, 'Affordance' may be the initial code and immediate feedback of an axial (sub) code. Once the two researchers coded half the papers, they met to discuss the codes. The codes were shared to look for agreement and develop an agreed set of codes. This process then continued to code the rest of the papers. Codes were deemed theoretically saturated when all the data on affordances/challenges fit with a code. The researchers met again and reached an interrater percentage agreement of 95%, reaching 100% after discussing the final misaligned papers. The final codes are reported in the findings.

FINDINGS AND DISCUSSION

The findings section of this study is organised around the four questions that guide this study. The first two questions provide the reader with an understanding of the context in which the studies took place. This covers the geographic location of the studies, publication year and educational level. Questions 3 and 4 then go into detail on the affordances and challenges of using AI in ELT/L.

1. Where are the AI studies taking place geographically, and has there been a change in the trend in the number of studies across the years?

Geographical locations

From the examination of the geographic locations in which the studies took place, there was a large trend in studies conducted in Asia with 31 of the 43 studies conducted in this continent, see Figure 2. In an examination of the locations with the largest number of publications, the three top locations were in Asia, with China (8), Taiwan (7) and Japan (4), see Figure 2. Extant systematic reviews on AI across all educational disciplines that took place prior to 2021 show a clear trend of at least 50% taking place in the United States with Asia following second (eg, Chen et al., 2020; Crompton et al., 2022). Reviews since 2021 show Asia leading the number of publications with China at the top of those numbers (eg, Crompton & Burke, 2023). Scholars report the rapid trend in AI publications and patents from China as more funding and incentives are provided (Li et al., 2021).

This study shows this trend increasing dramatically with 72% of the papers from Asia, with 19% from China. However, this study is focused on ELT/L, therefore, it is interesting to note that these studies are from countries that do not use English as the primary language.

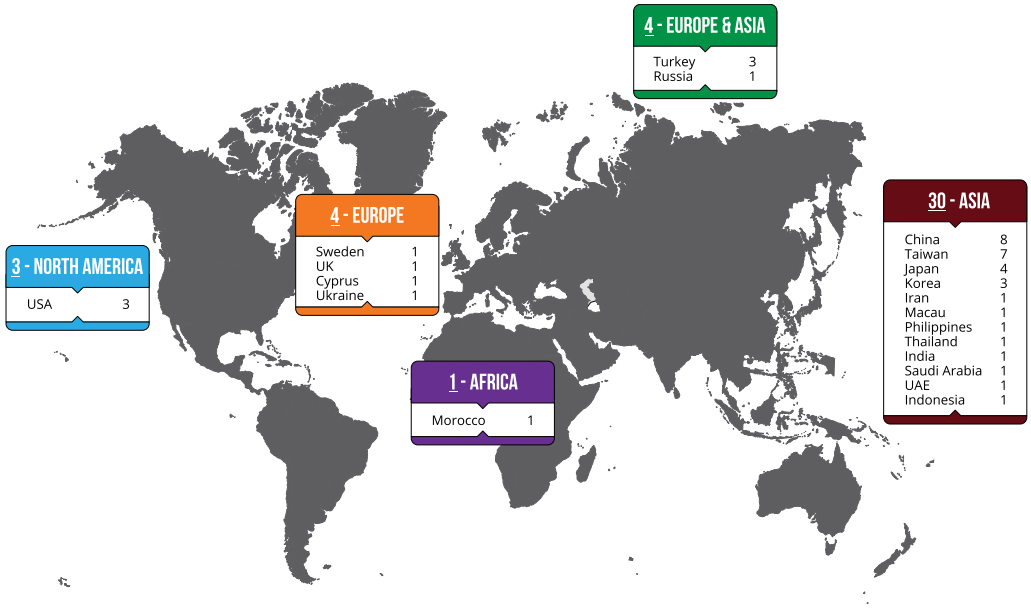

FIGURE 2 Geographical locations.

There are often migrants and displaced populations that are in countries, such as the UK, however, the majority of studies are not conducted in English-speaking countries. This may connect with Lan et al.'s (2020) posit that English is one of the most used languages for jobs, markets, tourism, discourse and international connectivity. From the data, in those regions where English is not the primary language, Asia appears to be leading the way in AI research. It would be interesting for future researchers to examine if Asia is producing more AI studies across disciplines or just targeting investigations to ELT/L.

Publication years

The findings show that the number of studies involving AI in ELT/L has greatly increased across the 10 years, see Figure 3. The first 5 years (2014–18) have a total of four publications with 2019 to July 2023 documenting a dramatic escalation to 39 studies. The rise in publications is a clear upward trend with the rest of 2023 expected to have the same trend. This increasing use of AI in ELT/L is not a surprise due to the rise in AI products, tools and educational offerings available and the media coverage of AI fostering a heightened awareness of the use of technology in education. Furthermore, the rise in 2022 may also be connected to the increase in AI large language models.

In January 2018, the International Society for Technology in Education (ISTE) started offering one of the first courses in the world for teaching K-12 educators how to integrate AI into the classroom (International Society for Technology in Education, (2018). This course offering reveals a growing trend in educator interest in using AI in the classroom. AI technologies have expanded since 2018 with an increase in Google AI experiments which are simple AI tools that require no advanced skills, to the expansion of generative AI tools including ChatGPT. While there were only a few studies published in 2014–18, these early studies may have sparked an interest in the use of AI in ELT/L. Sharadgah and Sa'di (2022) examined journal papers and conference papers from 2015 to 2021 and also noticed a similar

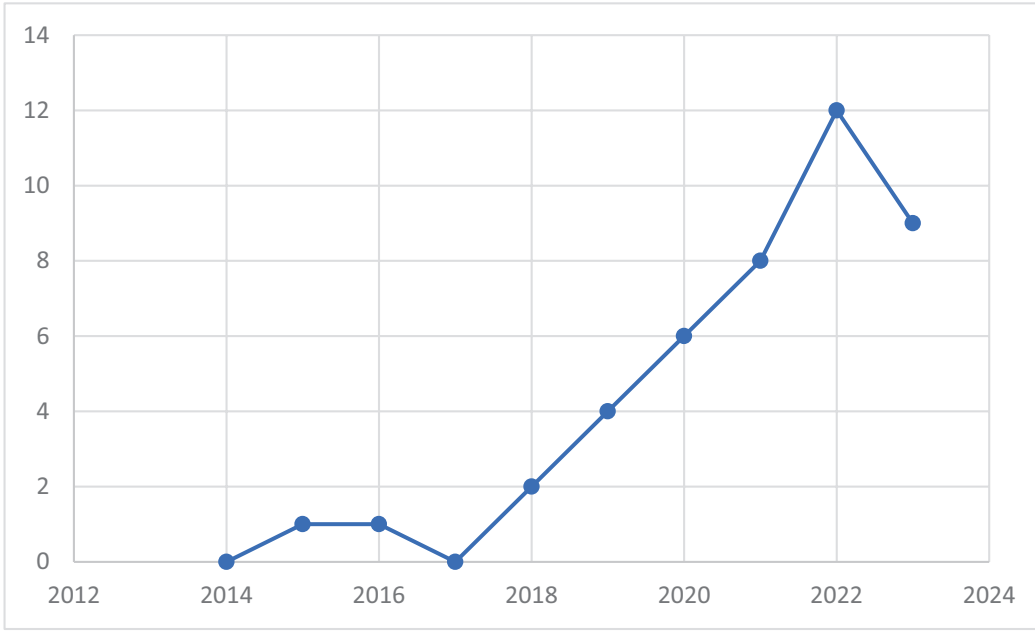

FIGURE 3 Publication years. 2023 represents half of the year.

rise in the number of papers with a large peak from 11 papers in 2020 to 30 papers in 2021. As interest grew in AI, the 2021 peak may have been due to an influx of conference papers for a targeted AI conference. This study also shows a jump in 2022, which may have been conference papers moving into journals.

2. What educational level are the pupils in the studies?

Levels of education

The data reveal that the majority of studies took place in higher education, see Figure 4. Yang and Kyun (2022) and Sharadghah and Sa'id (2022) examined K-12 and higher education and had similar findings that of the two, higher education had substantially more publications. K-12 may have used AI less due to AI tools' age restrictions. For example, OpenAI restricted the use of their products to 18-year-olds, only changing this to 13 years with parental guidance when they released ChatGPT at the end of 2022. This age restriction may account for lower numbers in K-12 in prior years. However, the age trend does not flow through into adult ELT/L. This systematic review is one of the first studies to examine AI and ELT/L across all three learner levels: K-12, higher education and adults. The numbers show a dramatic gap in peer-reviewed research in adult ELT/L.

In any studies involving pupils, the high numbers in higher education may also be due to accessibility of study participants. Faculty in higher education who publish their studies could have easier access to participants from their own institution, which also may require less stringent ethic approvals than working with young learners. However, it is important for researchers to have an understanding of how AI is being used in ELT/L across all learners. More research is needed to fill this gap.

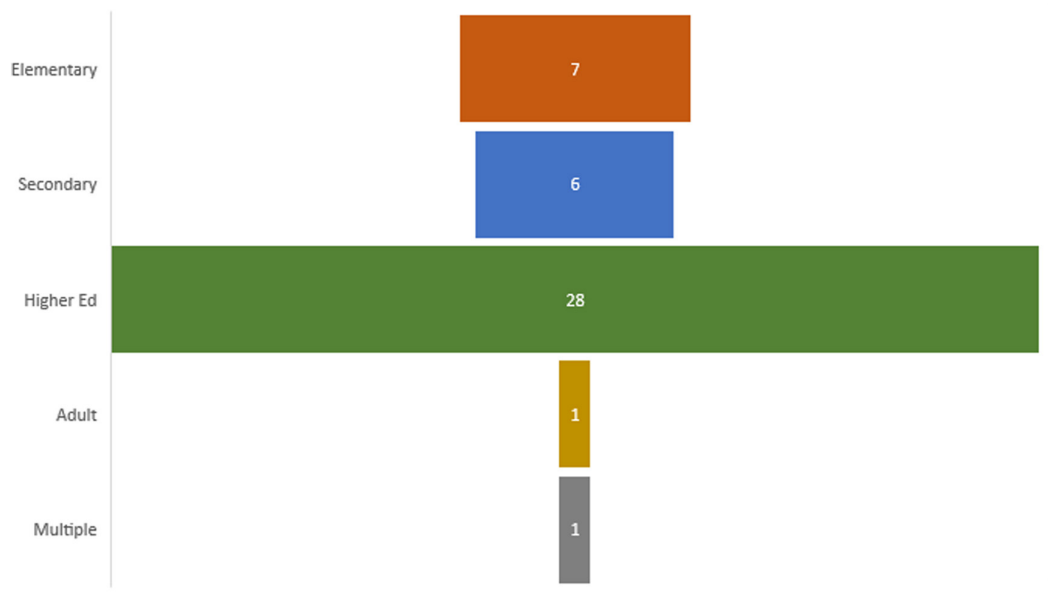

FIGURE 4 Learner levels.

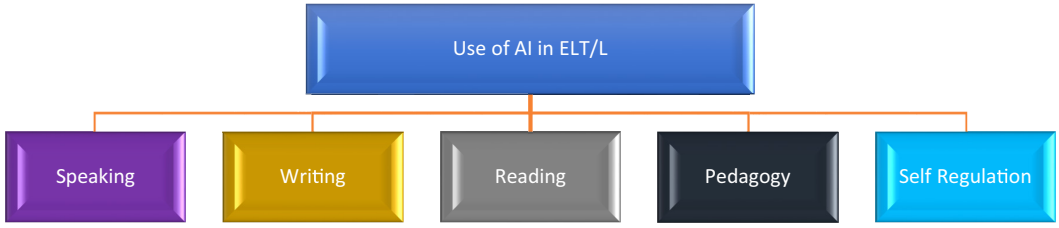

FIGURE 5 Use of AI in ELT/L.

3. How is AI being used for educational affordances in ELT/L?

From the grounded coding, six codes emerged from the way AI was being used, see Figure 5. Extant systematic reviews focused on a priori categories that had the researchers only looking for key aspects, such as language skills of speaking, writing, listening and reading (Sharadgah & Sa'di, 2022; Yang & Kyun, 2022). The grounded coding approach used in this study was used to reveal the trends that emerged from the literature on what was actually being studied. The grounded coding did reveal skills related to speaking, writing and reading, while the aspect of listening did not emerge from the data. With that grounded approach, discernible patterns also surfaced regarding the utilisation of AI in ETL/L with studies to enrich and broaden pedagogical practices, bolster pupil self-regulation and explore connections with affective objectives.

It is important to note that there was overlap in these codes, but the main use of AI in each study was what was identified by the code. To break down the codes further, axial codes were created to delineate further those uses involved in each main code. For each of the overarching codes, both the uses of AI in ELT/L and the affordances are presented together in the following sections.

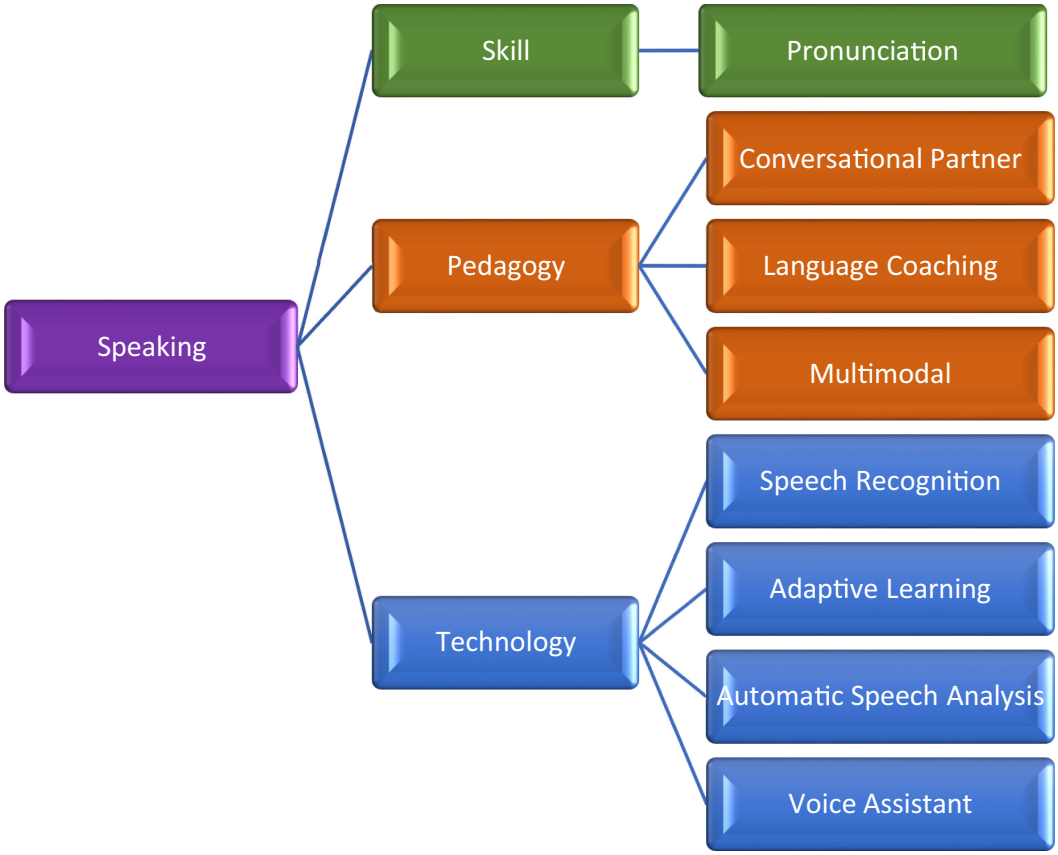

FIGURE 6 Axial codes for speaking.

Speaking

Further investigation of the studies which had pupils practising speaking skills as the main use of AI revealed three axial codes: subskill, pedagogy and technology, see Figure 6. The only speaking subskill revealed in the studies was a focus on pronunciation. Nonetheless, pronunciation was a common focus in these studies. There were a variety of AI systems and programs that helped pupils in this area. As Tokoz-Goktepe (2014) posited, a challenge for ELT/L is often the limited connections to English outside the language courses. AI can help to provide those connections. Liu and Hung (2016) used a system for teaching pronunciation to pupils in Taiwan and found that the AI significantly improved pupils' pronunciation by reducing the flatness of pitch and intonation patterns. The researchers found that the visual representation of the pitch as a spectrogram provided by the AI was helpful in supporting pronunciation.

Three codes emerged in the pedagogical AI methods for teaching English speaking. AI was used as a conversational partner and a language coach. Dizon and Tang (2020) had pupils hold conversations using Alexa, a personal voice assistant. The findings of this study show a plethora of affordances. Pupils found Alexa easy to use, and it promoted meaningful interactions, supported vocabulary acquisition, improved language skills and provided interesting, enjoyable learning. It is interesting that while conversations should typically involve speaking and listening skills, the listening aspect was not an area of focus that emerged from this systematic review. This finding is intriguing given that typical conversation should ideally engage both speaking and listening skills, yet the focus on listening skills did not emerge in

this systematic review. This insight is crucial for practitioners aiming to employ AI tools in an evidenced-based use of AI tools in supporting language skills development in ELT/L.

In the pedagogies code, studies highlighted the use of coaching and multimodal systems which employ multiple ways to present information, such as text, images, audio and video. Shivakumar et al. (2019) focused on both language coaching and a multimodal approach in a higher education setting. Pupils studying pronunciation were provided with a coach that tailored instruction to each pupil's learning patterns and needs. This coach presented information using text, images and audio. The findings show that the multimodal coaching approach was highly beneficial for young ELLs. Lin and Mubarok (2021) also used a multimodal system approach with videos, pictures, memes and songs to help with conversation. The findings show that this variety allowed higher education pupils to speak more fluently using consistently accurate language structures.

AI technologies were examined by scholars in their affordances for learning about spoken language. AI has a direct connection between speaking a language and natural language processing in that AI can recognise and produce speech. The axial codes reveal technology regarding speech recognition and automatic speech analysis as well as AI systems of adaptive learning and voice assistance. With high school pupils in Turkey, Kazu and Kuvvetli (2023) developed an AI-supported pronunciation model. This system was designed to help pupils practice, record and react to pupils pronouncing words. The findings show that the words remained in memory longer and it was most beneficial for learning consonant and vowel sounds.

Generative AI technologies, such as ChatGPT, now offer speech options where pupils can verbally converse with AI. More future research may see these programs used for developing speaking skills. These programs offer pupils an opportunity to practice their spoken English with an AI with an advanced vocabulary that can further advance the pupils' spoken skills with unlimited practice time. This may also avoid educators needing to purchase expensive programs. Nonetheless, generative AI may not offer the very specific targeted feedback support that these other programs, such as the one used by Kazu and Kuvvetli (2023), can provide.

Tai et al. (2022) used virtual reality with AI speech recognition with middle to high school pupils as they learned vocabulary. The study found that the linguistic cues provided by the agents in the VR making comments such as 'pardon me' and the kinesic signals (eg, gesturing and staring) were highly beneficial to the pupils' speech abilities. It appears from the studies that the AI providing context surrounding the vocabulary during interactions is beneficial in learning. This supports prior research findings supporting the importance of contextualising vocabulary learning (Godwin-Jones, 2018). Currently, these kinesic signals are not available to pupils using generative AI which highlights the benefit of programs that target language learning.

Writing

Similar to the speaking code, the studies revealed the axial codes of subskills, pedagogies and technologies. Two subskills of writing that emerged from the papers were vocabulary learning and grammar. Lo (2023) examined pupils' vocabulary improvement and retention in writing when pupils had access to neural machine translation programs. The researchers reported that this AI system helped to improve or expand vocabulary, especially when specialised or unambiguous expressions are involved. One of the common uses of AI in writing revealed by this study is the use of AI grammar checkers. This finding follows a trend across disciplines of the use of grammar feedback tools, for example, Grammarly (Koltovskaia, 2020). Dizon and Gayed (2021) specifically examined the impact of Grammarly when used in higher education ELT/L. He found that using this AI tool, pupils had fewer grammatical errors and wrote with more lexical variation than pupils without the AI. The predictive

text and real-time corrective feedback supported pupils studying English, especially novice writers. Educators have expressed apprehensions about pupils potentially becoming 'complacent' and utilising AI tools (such as grammar checkers or automatic translation tools) as a shortcut to learning (Bozkurt et al., 2023). However, the findings of this study suggest that AI-powered tools can be used to aid pupils in enhancing their writing skills. For practitioners looking to help pupils develop the vital skill of writing, it appears that there is value in real-time corrective feedback and the use of AI to extend vocabulary (see Figure 7).

Only one pedagogical focus emerged in the axial codes for writing. This was often connected back to AI tools providing feedback via spelling and grammar checkers, similar to that in Dizon and Gayed's (2021) study with Grammarly. Nazari et al. (2021) also studied the use of Grammarly as a feedback tool for ELLs and the findings again reported positive outcomes of statistically significant improvement in behavioural engagement, emotional engagement, cognitive engagement and self-efficacy in writing. With AI systems working well with procedural knowledge and systems within written text and languages (Crompton et al., 2022), it is discouraging to see feedback as the only writing pedagogy revealed in the grounded coding. Future researchers could further examine the pedagogical opportunities afforded by AI in teaching writing in English.

Writing was an area using a variety of AI-targeted ELT/L technology supports and tools. An interesting find is the use of translation tools. It could be argued that AI translation tools could be used to bypass language learning with the translation readily available to the pupil. However, Chon et al. (2021) explored pupil use of an AI translation tool—Google Translate. Working with pupils studying English writing in a South Korean college, the study's findings revealed that using Google Translate enabled the pupils to produce the greatest number of complex words of more than three syllables. It also allowed the learners to retrieve a larger

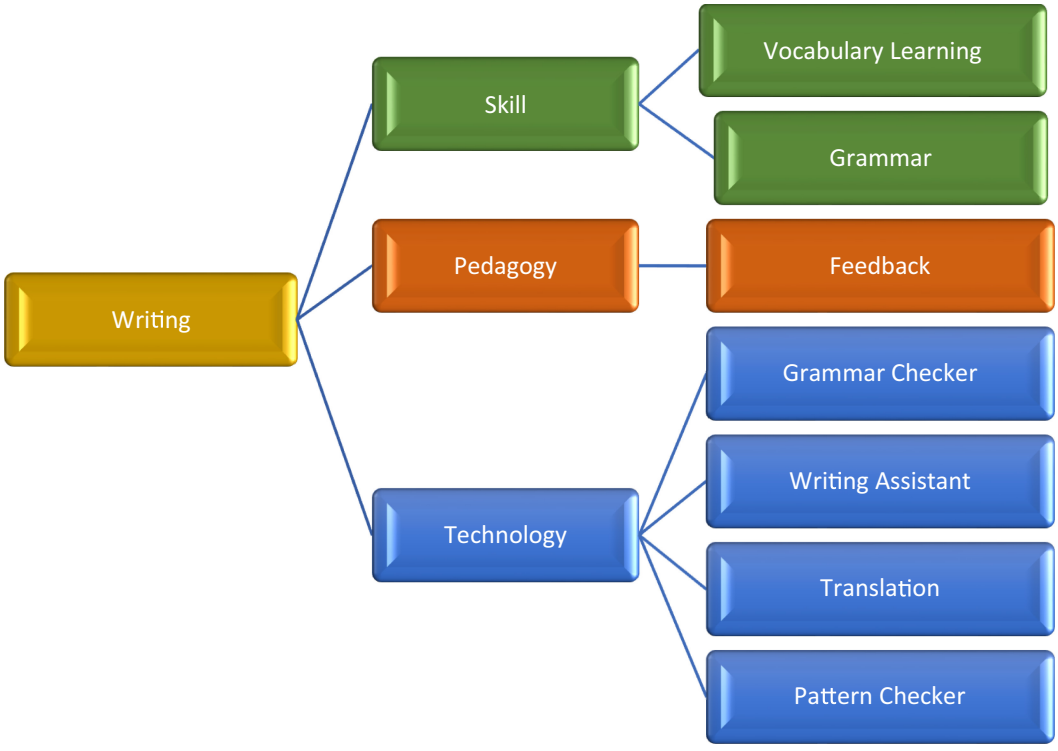

FIGURE 7 Axial codes for writing.

range of lexical items as the translation process directed the learners to reflect on the different lexical choices available to them to effectively express their intended message.

Reading

Reading did appear from the grounded coding. However, it was not as common in the data as speaking and writing. This reduced focus on reading is similar to the findings of Sharadgah and Sa'di (2022). This may be due to the extensive affordances of AI in natural language processing aligned to speaking and writing that researchers want to use, with reading less aligned to those AI affordances. Nonetheless, there were AI ELT/L studies focused on reading subskills, pedagogies and technology resulting in five axial codes, see Figure 8.

Vocabulary was the only subskill focus of studies in reading. Within pedagogy, gaming is the only code. It could be argued that gaming could just fit with the technology code. However, some gaming studies concentrated on the method of learning and others on developing AI gaming systems. Zheng et al. (2015) explored how vocabulary learning in reading occurs during gaming quest–play mediated in English with Japanese pupils playing with an English player. The pupils, embodied as avatars, used semiotic resources imbued in the game called World of Warcraft (WoW). For example, during the game, players can kill the enemy and then take items from the person killed with the word 'looting' appearing. The pupils then learn that word and read the explanation of the term provided by the other player via the chat, who then situates the word in the wider context. From the findings, Zheng and colleagues posit that pupils have opportunities to learn vocabulary and understand meaning via games beyond what a textbook or classroom can provide.

It appears that contextualising often decontextualised vocabulary is useful in learning and is offered by gaming that involves AI. WoW uses AI in the program to provide that context, through the inclusion of AI characters (that are not operated by humans), path-finding navigation algorithms and within the system to make the environment dynamic and engaging. For educators who have pupils who appear unengaged or unmotivated to learn, gaming could be an approach that supports those pupils in reaching their learning goals.

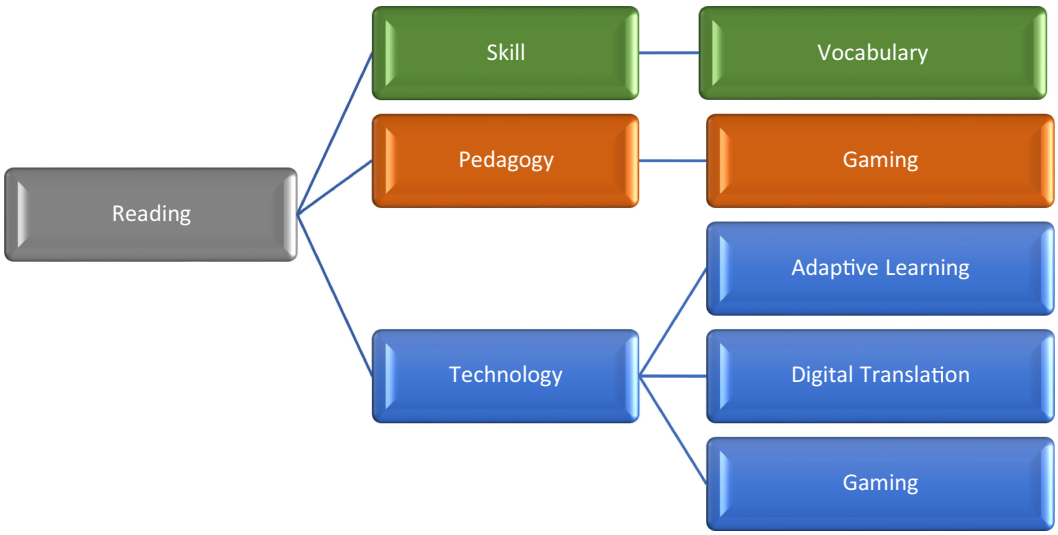

FIGURE 8 Axial codes for reading.

Pedagogy

Pedagogy was the code used to identify the methods, strategies and techniques used to facilitate ELT/L. Pedagogies were often connected to language skills and also appeared earlier in the axial codes for those skills. As a reminder to the reader, Table 3 provides a summary of those pedagogies that appeared in the earlier axial codes for language skills.

However, pedagogy here is a main code and not an axial (sub) code to denote that the researchers focused primarily on the pedagogical methods, strategies and techniques of AI in ELT/L. The pedagogical studies in this systematic review were in and across formal, informal and non-traditional learning settings. Figure 9 provides an overview of those targeted skills and pedagogies.

Coding from the pedagogy code revealed six axial codes, see Figure 9. Some studies examined multiple pedagogies which appear to provide more personalised learning approach. For example, Kim (2022) explored the effects of the pedagogical approaches of score predictions, lectures, explanations and practice tests on Korean pupils studying for their Test of English for International Communication (TOEIC). An AI system, *Soljam*, was used to provide these pedagogical approaches to learning. Pupils began with a diagnostic assessment in which the AI then used the data to provide lectures, explanations and practice tests at the level required by the pupil. In Kim's examination, there were two groups using AI: those using AI via a computer and those using a mobile device. The findings show that there was value in these pedagogies when AI was used either on a computer or a mobile device. However, they reported that some aspects of learning may work best on a particular device, such as a larger computer screen for more detailed work.

Lee et al. (2023) took a different look into pedagogy by exploring a context-based approach; specifically, learner-generated context (LGC) based. Lee et al. described LGC as the design and application of digital technology in a manner that assists learners in forming a 'learner-generated context' and learning within this context. This context is constructed from data that are gathered as the pupils perform actions and make choices. The system then learns about the pupils to provide them with more content to match their preferences. The researchers found that the LGC AI pedagogical approach catalysed learners' autonomous learning experiences.

As AI is rapidly extending its capabilities, it is positive to see that researchers are also focused on the pedagogical aspects of how effective practices connect with AI. Some of the pedagogies, such as context-based learning, use relatively recent ways of teaching, especially in targeting pupils' personal learning needs. However, it is interesting to note that there are a lot of traditional forms of pedagogies, including lectures and explanations being used. Scholars, such as Crompton and Burke (2020), lament 'using 21st Century technologies, with 20th Century teaching' (p. 2.). The new affordances of AI can provide new avenues of learning and it would be pertinent for researchers to explore and grow these areas. One area of AI value is the ability to look across large datasets and assess pupils' work at speed. Assessments were seen in the coding but were not explored in great detail in the extant literature. Again, further examination of AI assessment in ELT/L would be helpful to the academic community.

TABLE 3 Pedagogies for targeted language skill learning.

Speaking	Writing	Reading
Conversational partner	Feedback	Gaming
Language coaching		
Multimodal approach		

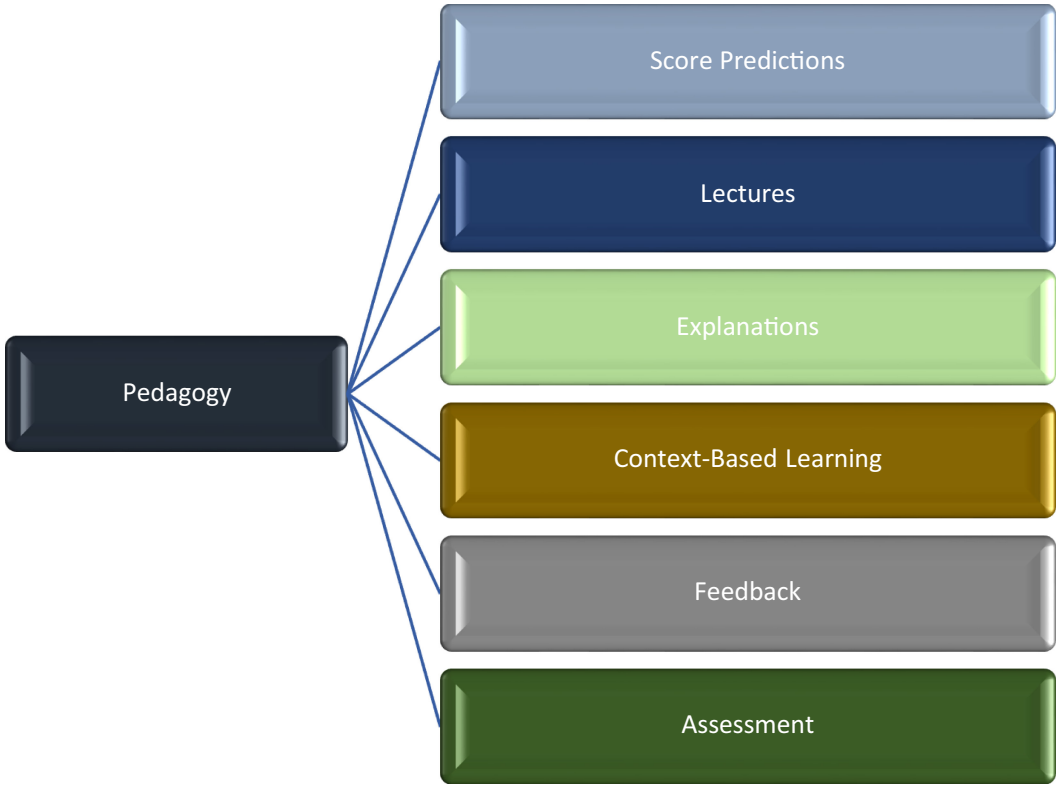

FIGURE 9 Axial codes for pedagogy.

These pedagogical findings have implications for educators as they navigate the intersection of effective pedagogical practices and AI. The research uncovered some innovative applications of AI for pedagogy. These pedagogies often highlighted ways to provide more of a student-centred focus with personalised learning and learner-generated context. However, practitioners should be aware that more traditional pedagogies, such as lectures and explanations, continue to prevail and it would be beneficial for educators to explore new possibilities for learning from those AI capabilities.

Self-regulation

Self-regulation emerged as a code from the grounded coding, with six axial codes, see Figure 10. Self-regulation refers to the ability to manage and control one's thoughts, emotions, behaviours and physiological responses to achieve personal goals and maintain well-being. Affect can influence the choices and actions the pupil takes. For example, someone in a negative mood might be more prone to impulsive decisions or avoidance behaviours.

It is interesting that self-regulation has appeared in this study. Scholars (viz, He, 2021) posit that there is value in AI in language learning, and they also warn of the dangers of pupils becoming over-reliant on AI to perform tasks and avoid mastering content themselves. It is positive to see that of the 43 studies in this systematic review, there was a trend towards self-regulation in relation to AI and ELT/L by having pupils actively thinking, especially in relation to developing goals and learning independence. The study's findings concerning the affordances of AI-powered tools in promoting student self-regulation also have substantial

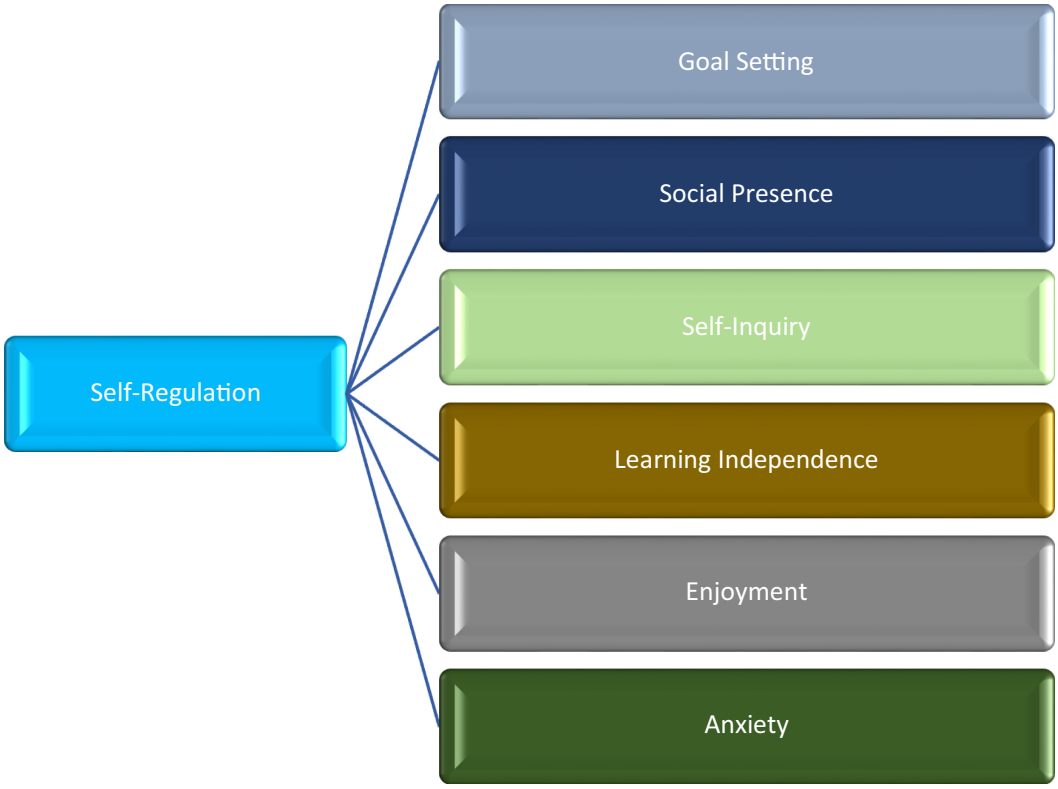

FIGURE 10 Axial codes for self-regulation.

implications for educators. The use of AI-powered tools such as chatbots to foster self-regulation could aid pupils in becoming more autonomous learners, setting more effective goals and alleviating anxiety about speaking in English.

For example, Hew et al. (2023) used chatbots in ELT/L to support pupil goal setting and social presence in fully online activities. Hew's study purports that pupils often lack self-regulation in online language courses with pupils feeling isolated and disengaged. However, using a grounded approach, the results of this study revealed three primary affordances of AI in goal setting as the chatbot helped pupils with (1) clarifying their learning goals, (2) developing techniques for setting goals and (3) raising awareness of learning strategies in goal setting. The findings also report that the pupils enjoyed using the chatbots and had a positive learning experience.

In another study, Chen, Hsu, et al. (2022), examined robot-assisted language learning. In Chen's study, AI and virtual reality were combined to create a system to use robots as a tool for training English language tour guides to develop a sense of autonomy. Pupils played the role of tour guide or traveller and interacted with various exercises including proper working, matching words, collocation and sentence structure. As the pupils worked with the AI system, the findings of the study identified affordances including developing autonomy, while also increasing motivation and engagement.

Anxiety was the final axial code from the systematic review. This a priori term was used in regard to pupils feeling anxious while learning English. This included anxiety about speaking in public, making mistakes with vocabulary and interacting with others. From the 43 studies, researchers were working on studies that used AI to reduce anxiety as pupils tried to grasp

another language and AI reducing anxiety was often referred to in the findings of the affordances of the AI. Chen, Koong, and Liao (2022) used an AI automatic speech recognition technique to support pupils' English-speaking skills and learning anxiety with fifth-grade pupils in Taiwan. Both Çakmak (2022) and Chen, Koong et al. found that AI raised the pupils' skills and lowered anxiety.

4. What are the challenges of using AI in ELT/L?

From across the studies, the challenges did not appear to be well reported as the researchers focused instead on the affordance of the AI systems in ELT/L. This can be a common problem with *positive publication bias* as journals often have more of an interest in publishing positive study findings than reporting challenges and problems (Mlinarić et al., 2017). Of the 43 studies in this systematic review, 27 (67%) did not report challenges. Those that did report challenges often were only brief statements following positive comments. Nonetheless, there are trends that emerged from the grounded coding to reveal four axial codes: technology breakdowns, limited capabilities, fear and standardising language, see Figure 11.

Technology breakdowns are a common problem across all technologies, such as connectivity issues, or program or computer functions that fail. One specific AI technology breakdown was that incorrect answers were given by the AI. Limited capabilities were an interesting type of challenge that was also revealed. This code referred to people wanting the AI systems to do more of what the system can do in a more advanced way. Thompson et al. (2018) reported that pupils wanted more improvements with the AI chatbots they were using. Some pupils found chatbot interactions unnatural. This led to those pupils becoming disinterested in using the chatbot. Ericsson et al. (2023) worked with pupils in Sweden and reported many affordances in using virtual humans as conversational agents. However, similar to Thompson's study, there were reports that sometimes pupils were not understood or heard as expected. Pupils in the study not only enjoyed using the AI for learning English but also called for more advancement in AI capabilities to avoid those mistakes or limitations.

With powerful technologies, fear is often connected for a variety of reasons. The axial codes found pupils experiencing three types of fear when using AI. The first fear was in providing personal information as it was not clear how that data were stored and who had access to it. Pupils also were unsure of how AI was operating, and this resulted in fear of the unknown. A fear also appeared in a perception study where pupils reported fear of losing a natural environment and real emotions when interacting with AI (eg, Viktorivna et al., 2022).

One of perhaps the most compelling challenges found in the studies in this systematic review is that AI may be contributing to standardising languages and ideologies. Rowe (2022) examined one second-grade (7- to 8-year-olds) classroom in the USA. Of the 24 pupils in the class, 17 did not speak English as a first language and Google Translate was used to support the emergent bilingual pupils' English development. The findings of the study revealed that Google Translate's programming appeared to carry messages about *appropriate* and *standard language* use. For example, one pupil in the class spoke Tagalog with her family. However, Tagalog was listed by Google Translate only as Filipino and when translating her own language to English, she had to follow the AI's chosen name choice for her family's language. While AI systems are becoming commonplace, it is essential that educators and educational leaders remain cognisant of these challenges to avoid or ameliorate them where possible. AI developers also need to consider the challenges identified in the literature and indeed adjust to ensure language represents all pupils' cultures and needs in ELT/L. Furthermore, it would be prudent to have pupils become digitally literate English learners and navigate AI tools with care and consideration.

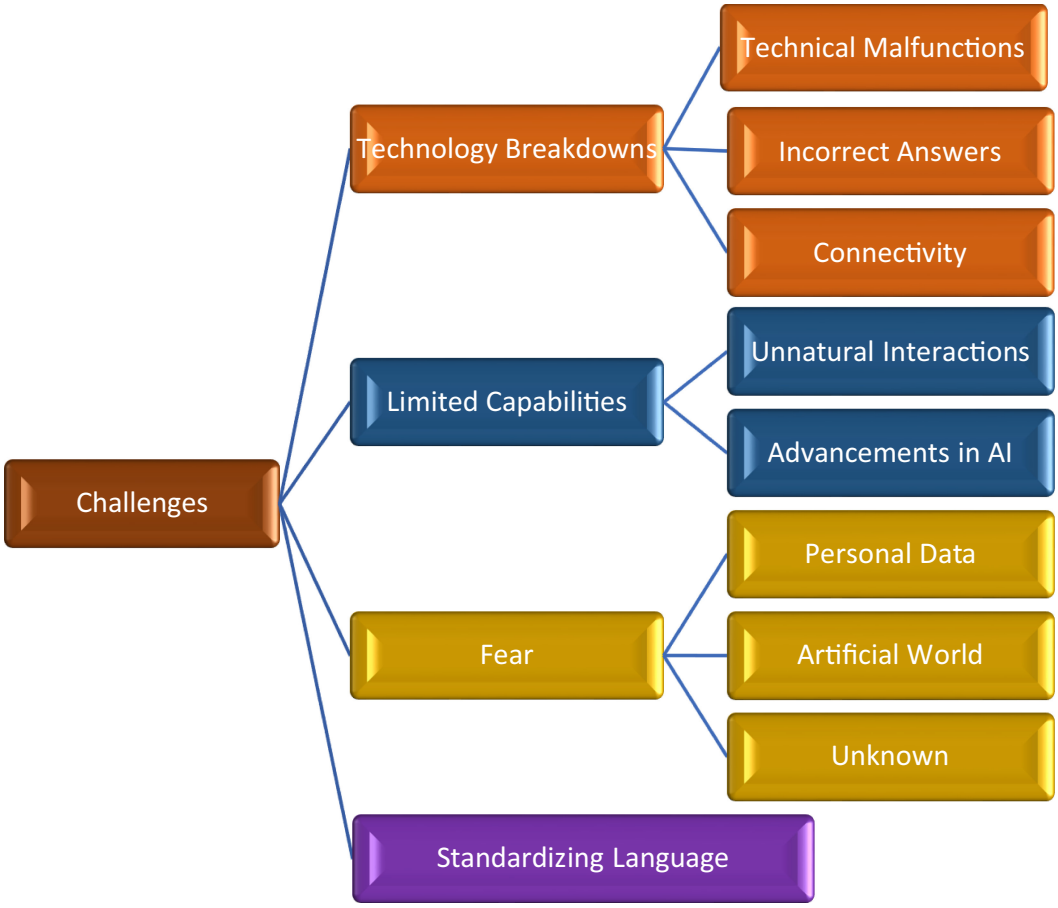

FIGURE 11 Challenges with the use of AI in ELT/L.

IMPLICATIONS FOR PRACTICE

The study reveals several key findings that have implications for practitioners. Previous systematic reviews into use of AI (Crompton et al., 2022) indicate that English language learning AI use is more prevalent than in any other taught subject. English language teachers need to make principled decisions to enable the effective use of AI in teaching, but the rapid development of the technology will provide a constant challenge for up-to-date AI literacy. This indicates that English language teachers particularly need to be equipped to make principled and effective use of AI in teaching and raises the question of how initial teacher education and in-service teacher training can help prepare teachers for teaching with AI.

The findings of this study revealed that AI can carry messages about appropriate and standard language use. This has serious implications around a 'neutral' AI choice of language that promotes a particular ideology or privileges certain groups while excluding or deprioritising others. For instance, it is important for practitioners to carefully consider how pronunciation models are chosen and who decides which models are considered desirable or acceptable. This will be key in avoiding the kind of implicit bias that elevates one variety of English over world Englishes.

The results of this research support the idea that AI can reduce anxiety. AI can provide opportunities for language practice with a neutral AI partner outside class and help alleviate pupils' anxiety about speaking and making mistakes in English. The question remains whether

AI effectively prepares pupils for real-life conversations, or whether their anxiety would persist when faced with human interlocutors. Another important practical implication is the continuing need to improve teachers' and pupils' digital literacy. There is an urgent need to broaden the understanding of AI literacy, including issues of transparency, trust, surveillance and privacy. Practitioners will need help and guidance from policymakers, in 'building awareness about the limitations and risks of AI and debating the various aspects of AI ethics' (Ziesche & Kumar Bhagat, 2022). Ignoring this would run a high risk of widening the digital divide.

This study raises important questions about pupils' fear when using AI, for example, due to opaqueness in how data are being processed by AI systems could lead to a lack of trust in the system. The implication is that a lack of clear data privacy statements and lack of clarity around the ethical deployment of AI in ELT/L could have adverse effects on learning and teaching and therefore, there is a need for adequate regulatory and ethical frameworks to protect pupils and promote trust in AI-powered systems. A key priority for policy makers should therefore be to develop and commit to an accessible and unambiguous ethics statement for AI in ELT/L, which could promote user confidence in the adoption of AI systems. When doing so, there is also the issue of digital divides to be considered, that is, the implications if certain national education systems introduce regulation and ethical alignment, whereas others are not in a position to do so. As with most emerging educational technology, there is a gap between the expectations around AI and its current capabilities. Practitioners should be cautious of 'AI hype' and be realistic about its capabilities.

GAPS AND FUTURE RESEARCH

This research has identified various gaps in need of further investigation. Notably, the majority of studies published on AI in ELT/L were from Asia (72.09%). This should be of interest not only to researchers in other geographies but also to policy makers and commercial leaders who seek to excel in this field. The findings will be of interest to researchers in other geographies, particularly in the US and UK, which are prominent players in the global ELT industry. Another trend identified is most of the studies taking place in higher education contexts, even though ELT/L occurs in a range of different contexts. Taken together, these results suggest a need for future studies conducted in a wider range of geographies and contexts, making the results more widely generalisable, particularly the need to rectify the lack of studies in adult learning and K-12.

This study gained only limited information on the challenges of AI in ELT/L, which did not appear to be as well reported as its affordances. It is posited that this could be due to positive publication bias, which leads researchers to focus more on positive findings rather than issues and problems. Therefore, further efforts are needed to make explicit the challenges of AI in ELT/L. For the relatively newer types of AI, such as ChatGPT or similar large language models, more is needed on how these types of AI can support ELT/L. However, it is expected that an uptake in the trend towards research on these specific forms of AI will be seen in 2024 onwards.

The findings of this systematic review show that speaking and writing emerged as the main focus of existing research. The question raised by this study is whether AI will be better at developing certain skills in ELT/L, namely productive skills, or if it will be useful across all language skills. Moreover, there is a noticeable lack of focus in the available body of research on subskills typically required to develop these language skills. Therefore, future studies could explore the use of AI in the subskills involved in each language skill.

Another interesting finding is that in writing, 'feedback' was the only pedagogical focus that emerged. What remains to be explored is the whole area of procedural knowledge

in writing that AI could potentially assist with. This would be a fruitful area for further research work. Research is also needed to determine the place for AI-powered gaming in educational settings for ELT/L. More research could explore explicit design features of multi-player online games for language learning. Yet another interesting finding for practitioners is how translation tools could help pupils by giving them access to a larger range of lexical items. While such tools clearly aid output, it remains to be seen what happens when the tool is taken away. The question raised is whether this also results in an improvement in pupils' skills independent of these tools. If the debate is to be moved forward, a better understanding of the role of such machine translation tools in ELT/L needs to be developed.

CONCLUSION

This systematic review examined the use of AI in ELT/L to provide an overview of the context in which AI and ELT/L are being studied. Then, most importantly, how AI is being used and the affordances and challenges. This study is unique on three points, in providing (1) one of the first reviews across all learner levels: K-12, higher education and adult; (2) a grounded examination to examine AI in ELT/L; and (3) an up-to-date review to ensure inclusion of recent advances in AI.

From this systematic review, the contextual data show that studies from the continent of Asia were the most common with over half from this location. The anticipated trend showed an increase in AI publications. However, this study was able to reveal that the clear upward trajectory started in 2017. AI and ELT/L studies took place across all educational levels. Nonetheless, 65% of the studies took place in higher education. Interestingly, the second highest context was young children followed by secondary school.

These data revealed five areas in which AI is being used in ELT/L: speaking, writing, reading, pedagogy and self-regulation. Each of these provided further axial codes to delineate further. This highlighted gaps such as a paucity of research in receptive skills and subskills generally and on specific AI tools. The examination of the challenges of using AI in ELT/L was surprisingly sparse. Five challenges were identified: technology breakdowns, limited capabilities, fear and standardising language. Some of the challenges could be described as typical technology problems, such as technology breakdowns. However, new technology challenges are appearing with AI, such as concerns that AI is fitting languages into a standardised set and may be omitting lesser-known languages forcing users to conform to the choices and naming system provided by the AI systems.

This study highlights a need to prepare English language teachers to have an understanding of what AI is, how to exploit the many benefits of these tools with English learners and knowledge of what to avoid. It is important to note the limitations of this study that only studies published in English were included in this study. Examination of further databases may also provide further data. This study can be used as a springboard for future researchers to address the gaps highlighted to provide further depth into AI and ELT/L. Policy makers, funders, practitioners and educational leaders can use the information provided in this study to gain a holistic understanding of the current trend in the use of AI in ELT/L.

ACKNOWLEDGEMENTS

The authors would like to thank the British Council for their vision in highlighting the need for this systematic review at such a pertinent time in educational history and providing funding to conduct this work.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data are gathered from Open Access Databases that can be accessed through university subscriptions or from access through local libraries.

ETHICS STATEMENT

Institutional Review Board approval is not needed for this systematic review.

ORCID

Helen Crompton https://orcid.org/0000-0002-1775-8219
Adam Edmett https://orcid.org/0000-0002-4048-1043
Diane Burke https://orcid.org/0000-0002-8214-0386

REFERENCES

Ahmadi, M. R. (2018). The use of technology in English language learning: A literature review. *International Journal of Research in English Education*, *3*(2), 115–125. https://doi.org/10.29252/ijree.3.2.115

Annamalai, N., Eltahir, M. E., Zyoud, S. H., Soundrarajan, D., Zakarneh, B., & Al Salhi, N. R. (2023). Exploring English language learning via chabot: A case study from a self determination theory perspective. *Computers and Education: Artificial Intelligence*, *5*, 100148. https://doi.org/10.1016/j.caeai.2023.100148

Baranwal, D. (2022). A systematic review of exploring the potential of teachable agents in English learning. *Pedagogical Research*, *7*(1). https://doi.org/10.29333/pr/11553

Belur, J., Tompson, L., Thornton, A., & Simon, M. (2018). Interrater reliability in systematic review methodology: Exploring variation in coder decision-making. *Sociological Methods & Research*, *50*(2), 837–865. https://doi.org/10.1177/0049124118799372

Bozkurt, A., Xiao, J., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., Farrow, R., Bond, M., Nerantzi, C., Honeychurch, S., Bali, M., Dron, J., Mir, K., Stewart, B., Costello, E., Mason, J., Stracke, C. M., Romero-Hall, E., Koutropoulos, A., … Jandrić, P. (2023). Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education*, *18*(1), 53–130. https://doi.org/10.5281/zenodo.7636568

Çakmak, F. (2022). Chatbot-human interaction and its effects on EFL pupils' L2 speaking performance and anxiety. *Novitas-ROYAL (Research on Youth and Language)*, *16*(2), 113–131.

Chen, C.-H., Koong, C.-S., & Liao, C. (2022). Influences of integrating dynamic assessment into a speech recognition learning design to support pupils' English speaking skills, learning anxiety and cognitive load. *Educational Technology & Society*, *25*(1), 1–14.

Chen, T. (2016). Technology-supported peer feedback in ESL/EFL writing classes: A research synthesis. *Computer Assisted Language Learning*, *29*(2), 365–397. https://doi.org/10.1080/09588221.2014.960942

Chen, X., Xie, H., & Hwang, G.-J. (2020). A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. *Computers and Education: Artificial Intelligence*, *1*, 100005. https://doi.org/10.1016/j.caeai.2020.100005

Chen, Y.-L., Hsu, C.-C., Lin, C.-Y., & Hsu, H.-H. (2022). Robot-assisted language learning: Integrating artificial intelligence and virtual reality into English tour guide practice. *Education Sciences*, *12*, 437. https://doi.org/10.3390/educsci12070437

Chon, Y. V., Shin, D., & Kim, G. E. (2021). Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System*, *96*, 102408. https://doi.org/10.1016/j.system.2020.102408

Crompton, H., & Burke, D. (2020). Mobile learning and pedagogical opportunities: A configurative systematic review of PreK-12 research using the SAMR framework. *Computers & Education*, *156*, 103945. https://doi.org/10.1016/j.compedu.2020.103945

Crompton, H., & Burke, D. (2022). Artificial intelligence in K-12 education. *SN Social Sciences*, *2*(113). https://doi.org/10.1007/s43545-022-00425-5

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The State of the field. *International Journal of Educational Technology in Higher Education*, *20*, 22. https://doi.org/10.1186/s41239-023-00392-8

Crompton, H., & Burke, D. (2024). The educational affordances and challenges of ChatGPT: State of the field. *TechTrends*, *68*, 380–392. https://doi.org/10.1007/s11528-024-00939-0

Crompton, H., Jones, M. V., & Burke, D. (2022). Affordances and challenges of artificial intelligence in K-12 education: A systematic review. *Journal of Research on Technology in Education*. https://doi.org/10.1080/15391523.2022.2121344

Cumming, A., Yang, L., Qiu, C., Zhang, L., Ji, X., Wang, J., Wang, Y., Zhan, J., Zhang, F., Xu, C., Cao, R., Yu, L., Chu, M., Liu, M., Cao, M., & Lai, C. (2018). Pupils' practices and abilities for writing from sources in English at universities in China. *Journal of Second Language Writing*, *39*, 1–15. https://doi.org/10.1016/j.jslw.2017.11.001

Dizon, G., & Gayed, J. M. (2021). Examining the impact of Grammarly on the quality of mobile L2 writing. *JALT CALL Journal*, *17*(2), 74–92. https://doi.org/10.29140/jaltcall.v17n2.336

Dizon, G., & Tang, D. (2020). Intelligent personal assistants for autonomous second language learning: An investigation of Alexa. *JALT CALL Journal*, *16*(2), 107–120. https://doi.org/10.29140/jaltcall.v16n2.273

Ericsson, E., Hashemi, S. S., & Lundin, J. (2023). Fun and frustrating: Pupils' perspectives on practising speaking English with virtual humans. *Cogent Education*, *10*, 2170088. https://doi.org/10.1080/2331186X.2023.2170088

Gilakjani, A. P. (2017). A review of the literature on the integration of technology into the learning and teaching of English language skills. *International Journal of English Linguistics*, *7*(5), 95–106. https://doi.org/10.5539/ijel.v7n5p95

Godwin-Jones, R. (2018). Contextualized vocabulary learning. *Language Learning & Technology*, *22*(3), 1–19.

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews* (2nd ed.). Sage.

Grabe, W. P., & Stoller, F. L. (2002). *Teaching and researching reading*. Pearson Education. https://doi.org/10.4324/9781315833743

He, Y. (2021). Challenges and countermeasures of translation teaching in the era of artificial intelligence. *Journal of Physics: Conference Series 1881*, *2*, 022086. https://doi.org/10.1088/1742-6596/1881/2/022086

Hew, K. F., Huang, W., Du, J., & Jia, C. (2023). Using chatbots to support pupil goal setting and social presence in fully online activities: Learner engagement and perceptions. *Journal of Computing in Higher Education*, *5*, 40–68. https://doi.org/10.1007/s12528-022-09338-x

International Society for Technology in Education (ISTE). (2018). *Artificial intelligence explorations and their practical use in schools*. https://www.iste.org/professional-development/iste-u/artificial-intelligence

Kannan, J., & Munday, P. (2018). New trends in second language learning and teaching through the lens of ICT, networked learning, and artificial intelligence. *Circle of Linguistics Applied to Communication*, *76*, 13–30. https://doi.org/10.5209/CLAC.62495

Kazu, I. Y., & Kuvvetli, M. (2023). The influence of pronunciation education via artificial intelligence technology on vocabulary acquisition in learning English. *International Journal of Psychology and Educational Studies*, *10*(2), 480–493. https://doi.org/10.52380/ijpes.2023.10.2.1044

Kim, N.-Y. (2019). A study on the use of artificial intelligence chatbots for improving English grammar skills. *Journal of Digital Convergence*, *17*(8), 37–46. https://doi.org/10.14400/JDC.2019.17.8.037

Kim, N.-Y. (2022). AI-integrated mobile-assisted language learning: Is it an effective way of preparing for the TOEIC test in classroom environments? *English Teaching*, *77*(3), 79–102. https://doi.org/10.15858/engtea.77.3.202209.79

Koltovskaia, S. (2020). Pupil engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, *44*, 100450. https://doi.org/10.1016/j.asw.2020.100450

Lan, P.-S., Liu, M.-C., & Baranwal, D. (2020). Applying contracts and online communities to promote pupil self-regulation in English learning at the primary-school level. *Interactive Learning Environments*, *31*(1), 468–479. https://doi.org/10.1080/10494820.2020.1789674

Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and principles in language teaching* (3rd ed.). Oxford University Press.

Lee, D., Kim, H.-H., & Sung, S.-H. (2023). Development research on an AI English learning support system to facilitate learner-generated-context-based learning. *Educational Technology Research and Development*, *71*, 629–666. https://doi.org/10.1007/s11423-022-10172-2

Li, D., Tong, T. W., & Xiao, Y. (2021, February 18). Is China emerging as the global leader in AI? *Harvard Business Review*. https://hbr.org/2021/02/is-china-emerging-as-the-global-leader-in-ai

Lin, C.-J., & Mubarok, H. (2021). Learning analytics for investigating the min map-guided AI chatbot approach in an EFL flipped speaking classroom. *Educational Technology & Society*, *24*(4), 16–35.

Liu, S.-C., & Hung, P.-Y. (2016). Teaching pronunciation with computer assisted pronunciation instruction in a technological university. *Universal Journal of Educational Research*, *4*(9), 1939–1943. https://doi.org/10.13189/ujer.2016.040902

Lo, S. (2023). Neural machine translation in EFL classrooms: Learners' vocabulary improvement, immediate vocabulary retention and delayed vocabulary retention. *Computer Assisted Language Learning*, 1–20. https://doi.org/10.1080/09588221.2023.2207603

Loncar, M., Schams, W., & Liang, J.-S. (2023). Multiple technologies, multiple sources: Trends and analyses of the literature on technology-mediated feedback for L2 English writing published from 2015–2019. *Computer Assisted Language Learning*, *36*(4), 722–784. https://doi.org/10.1080/09588221.2021.1943452

Luo, M., & Cheng, L. (2020). Exploration of interactive foreign language teaching mode based on artificial intelligence. In *Proceedings of 2020 International Conference on Computer Vision, Image and Deep Learning* (pp. 285–290). IEEE. https://doi.org/10.1109/CVIDL51233.2020.00-84

Ma, G. (2021). The current situations of mobile assisted language learning. In J. MacIntyre, J. Zhao, & X. Ma (Eds.), *Proceedings of the International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy* (pp. 675–679). Springer. https://doi.org/10.1007/978-3-030-62746-1_99

Margana, M. (2016). Voices of English teachers and pupils on blended culture as a model of English language teaching and learning at vocational high schools in Yogyakarta. *Mediterranean Journal of Social Sciences*, *7*(3), 459–466. https://doi.org/10.5901/mjss.2016.v7n3p459

Mlinarić, A., Horvat, M., & Smolčić, V. Š. (2017). Dealing with the positive publication bias: Why you should really publish your negative results. *Biochemia Medica*, *27*(3), 447–452. https://doi.org/10.11613/BM.2017.030201

Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of artificial intelligence powered digital writing assistance in higher education: Randomized controlled trial. *Heliyon*, *7*(5), e07014. https://doi.org/10.1016/j.heliyon.2021.e07014

Noviyanti, S. D. (2020). Artificial intelligence (AI)-based pronunciation checker: An alternative for independent learning in pandemic situation. *The Journal of English Language Teaching in Foreign Language Context*, *5*(2), 162. https://www.syekhnurjati.ac.id/jurnal/index.php/eltecho/article/view/7246

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal*, *372*, n71. https://doi.org/10.1136/bmj.n71

Pokrivčáková, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, *7*(3), 135–153. https://doi.org/10.2478/jolace-2019-0025

Rivera Barreto, A. M. (2018). Motivating English language use by using the benefits of technology. *Gist Education and Learning Research Journal*, *16*, 117–140. https://doi.org/10.26817/16925777.428

Rowe, L. W. (2022). Google translate and biliterate composing: Second-graders' use of digital translation tools to support bilingual writing. *TESOL Quarterly*, *56*(3), 883–905. https://doi.org/10.1002/tesq.3143

Saldana, J. (2015). *The coding manual for qualitative researchers*. Sage.

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

Shivakumar, A., Shukla, S., Vasoya, M., Kasrani, I. M., & Pei, Y. (2019). AI-enabled language speaking coaching for dual language learners. *IADIS International Journal on WWW/Internet*, *17*(1), 66–78.

Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H. S., Li, M., Sariyska, R., Stavrou, M., Becker, B., & Montag, C. (2021). Assessing the attitude towards artificial intelligence: Introduction of a short measure in German, Chinese, and English language. *KI-Künstliche Intelligenz*, *35*(1), 109–118. https://doi.org/10.1007/s13218-020-00689-0

Strauss, A., & Corbin, J. (1995). Grounded theory methodology: An overview. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273–285). Sage.

Tai, T.-Y., Chen, H. H., & Todd, G. (2022). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Computer Assisted Language Learning*, *35*(4), 892–917. https://doi.org/10.1080/09588221.2020.1752735

Thompson, A., Gallacher, A., & Howarth, M. (2018). Stimulating task interest: Human partners or chatbots? In P. Taalas, J. Jalkanen, L. Bradley, & S. Thouësny (Eds.), *Future-proof CALL: Language learning as exploration and encounters—Short papers from EUROCALL 2018* (pp. 302–306). Research-publishing.net. https://doi.org/10.14705/rpnet.2018.26.854

Tokoz-Goktepe, F. (2014). Speaking problems of 9th grade high school Turkish learners of L2 English and possible reasons for those problems: Exploring the teachers and pupils' perspectives. *Procedia—Social and Behavioral Sciences*, *116*, 1875–1879. https://doi.org/10.1016/j.sbspro.2014.01.487

Ulicheva, A., Harvey, H., Aronoff, M., & Rastle, K. (2018). Skilled readers' sensitivity to meaningful regularities in English writing. *Cognition*, *195*, 103810. https://doi.org/10.1016/j.cognition.2018.09.013

Viktorivna, K. L., Oleksandrovych, V. A., Oleksandrivna, K. I., & Oleksandrivna, K. N. (2022). Artificial intelligence in language learning: What are we afraid of. *Arab World English Journal (AWEJ) Special Issue on CALL*, *8*, 262–273. https://doi.org/10.24093/awej/call8.18

Vuong, Q.-H., La, V.-P., Nguyen, M.-H., Jin, R., La, M.-K., & Le, T.-T. (2023). AI's humanoid appearance can affect human perceptions of its emotional capability: Evidence form self-reported data in the U.S. *International Journal of Human-Computer Interaction*, 1–12. https://doi.org/10.1080/10447318.2023.2227828

Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI0supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, *194*, 104703. https://doi.org/10.1016/j.compedu.2022.104703

Xiaohong, W., & Yanzheng, W. (2021, February). The application of artificial intelligence in modern foreign language learning. In *Proceedings of the 2021 4th International Conference on Big Data and Education* (pp. 34–37). ACM. https://doi.org/10.1145/3451400.3451406

Xu, Z., Wijekumar, K., Ramirez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K–12 pupils' reading comprehension: A meta-analysis. *British Journal of Educational Technology*, *50*(6), 3119–3137. https://doi.org/10.1111/bjet.12758

Yang, H., & Kyun, S. (2022). The current research trend of artificial intelligence in language learning: A systematic empirical literature review from an activity theory perspective. *Australasian Journal of Educational Technology*, *38*(5), 180–210. https://doi.org/10.14742/ajet.7492

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

Zhang, R., & Zou, D. (2023). A review of research on technology-enhanced peer feedback for second language writing based on the activity theory framework. *Education and Information Technologies*, *28*, 6727–6753. https://doi.org/10.1007/s10639-022-11469-8

Zheng, D., Bischoff, M., & Gilliland, B. (2015). Vocabulary learning in massively multiplayer online games: Context and action before words. *Educational Technology Research and Development*, *63*, 771–790. https://doi.org/10.1007/s11423-015-9387-4

Ziesche, S., & Kumar Bhagat, K. (2022). *Here, there, and everywhere—AI in education: State of the education report for India 2022*. UNESCO. https://en.unesco.org/news/state-education-report-india-2022-artificial-intelligence-education

How to cite this article: Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*, *00*, 1–27. https://doi.org/10.1111/bjet.13460