ULUSLARARASI EĞİTİM BİLİMLERİ DERGİSİ
ISSN: 2149-0848
THE JOURNAL OF INTERNATIONAL EDUCATION SCIENCE

AI-Supported Flipped Learning in ELT in the Post-ChatGPT Period: A Systematic Review

ChatGPT Sonrasındaki İngilizce Dil Öğretiminde Yapay Zekâ Destekli Ters Yüz Öğrenme: Sistematik Bir Derleme

Mustafa DOLMACI*
Esra KÜÇÜKGÖK**

Geliş/Submitted: 16.06.2026
Kabul/Accepted: 28.09.2026

10.29228/INESJOURNAL.94960

Makale Türü: **Araştırma Makalesi** Article Information: **Research Article**

Citation / Atıf

Dolmacı, M. ve Küçükgök, E. (2026). AI-supported flipped learning in elt in the post-chatgpt period: A systematic review. *Uluslararası Eğitim Bilimleri Dergisi*, 13 (48), 167-192.

***Dr. Öğr. Üyesi, Selçuk Üniversitesi, dolmaci@selcuk.edu.tr**
****** Selçuk Üniversitesi, kucukgokesra@gmail.com**

ULUSLARARASI EĞİTİM BİLİMLERİ DERGİSİ
ISSN: 2149-0848
THE JOURNAL OF INTERNATIONAL EDUCATION SCIENCE

Öz: 2022 yılının sonlarında üretici yapay zekanın hızla yaygınlaşması, dil eğitimi alanında köklü pedagojik yaklaşımların yeniden ele alınmasını zorunlu kılmıştır. Bu süreçte sorgulamaya açılan modellerden biri, öğrencilerin içerikle ilk karşılaşmalarını sınıf dışında gerçekleştirdiği ters yüz öğrenme yaklaşımıdır. Bu çalışmada, ChatGPT sonrası dönemde İngilizce dil öğretiminde yapay zeka destekli ters yüz öğrenme uygulamalarını konu alan araştırmaları sistematik biçimde derleyerek incelenmesi amaçlanmaktadır. PRISMA 2020 ilkeleri çerçevesinde Web of Science ve Scopus veri tabanlarında 2023 yılı ve sonrasına ait açık erişimli yayınlar taranmış; belirlenen dahil edilme ölçütlerini karşılayan sekiz çalışma tematik analiz yöntemiyle çözümlenmiştir. Derleme kapsamında araştırma bağlamları ve katılımcı özellikleri, ters yüz öğrenme tasarımına yapay zekâ entegrasyonu, hedef alınan dil becerileri, elde edilen yararlar ve karşılaşılan uygulama güçlükleri ele alınmıştır. Alanyazında bildirilen yararlar; dil becerilerinde gelişim, öz güven, derse katılım, üstbilişsel farkındalık ve öğrenci özerkliği gibi boyutlarda yoğunlaşmaktadır. Öte yandan yapay zekâ çıktılarındaki hatalar, araçlara aşırı bağımlılık, akademik dürüstlük kaygıları, kişisel verilerin gizliliğine ilişkin sorunlar, teknolojiye erişimdeki eşitsizlikler ve öğretmenlerin artan iş yükü başlıca güçlükler olarak öne çıkmaktadır. Elde edilen bulgular, üretici yapay zekanın ters yüz dil öğretim uygulamalarına en somut katkıyı sınıf öncesi aşamada sunabileceğine işaret etmektedir; ancak bu potansiyelin hayata geçirilebilmesi için titizlikle kurgulanmış pedagojik tasarım, yapay zeka okuryazarlığının geliştirilmesi ve bilinçsiz kullanımı engelleyecek güvencelerin oluşturulması gerekmektedir.
Anahtar Kelimeler: ters yüz öğrenme; üretici yapay zeka; İngilizce dil öğretimi; sistematik derleme
Abstract: Since late 2022, generative artificial intelligence (AI) has prompted language educators to reconsider established pedagogical models, particularly flipped learning, where learners complete meaningful preparation before class. This systematic review examines empirical research on AI-supported flipped learning in ELT in the post-ChatGPT period. Following PRISMA 2020 guidance, open-access studies indexed in Web of Science and Scopus from 2023 onward were screened; eight studies met the inclusion criteria and were analyzed thematically. The review addressed research contexts and participants, AI integration in flipped design, targeted language skills, reported benefits, and implementation challenges. The evidence remains limited: all studies were conducted in higher education

EFL settings, mostly in Asia and the Middle East, with one from Ecuador. Generative AI, especially ChatGPT, was used mainly to support pre-class input, learner practice and feedback, and teacher planning. Reported benefits included gains in language learning, confidence, engagement, metacognition, and autonomy. However, the studies also identified inaccurate AI output, over-reliance, academic integrity concerns, privacy issues, unequal access, and increased teacher workload. The review suggests that generative AI may strengthen flipped ELT most clearly in the pre-class phase, but only with careful pedagogical design, AI literacy, and safeguards against uncritical use.
Keywords: Visual impairment, Mathematics education, Visually impaired students, Algebraic reasoning.

Introduction

The public release of ChatGPT in late 2022 changed the practical conditions of language education very quickly. In a short period, generative artificial intelligence (AI) moved from a specialist topic to something many learners and teachers meet in daily academic work, whether for drafting, feedback, conversation practice, or material preparation (Meniado, 2023; Moorhouse & Wong, 2025). Bibliometric work reflects this acceleration: research on AI in English language education had already been increasing since 2013, but it rose more sharply after large language models became widely available (Zhang & Umeanowai, 2025; Zhou et al., 2026). This has made older pedagogical questions newly unsettled. Flipped learning is especially affected, because the model asks learners to begin the work before class, and that is where AI can now intervene.

Flipped learning reverses the usual order of instruction. Learners first meet new content outside class, often through videos or digital materials, and class time is then used for guided practice, discussion, application, or production. Definitions vary: some stress technology, some the social nature of in-class work, and others the underlying pedagogy (Låg & Sæle, 2019). What they share is a shift toward active learning and learner responsibility. In ELT, this shift has usually been justified by the shortage of classroom time for meaningful language use. Moving input outside the lesson is expected to reserve class time for interaction and output (Kernagaran & Abdullah, 2022). Reviews in language education connect flipped learning with improvements in proficiency, speaking, writing, and reading (Jantakoon et al., 2024; Ling & Jan, 2024). Yet the evidence is not uniformly strong. A second-order meta-analysis reported only a modest advantage over conventional instruction (Hew et al., 2021), another meta-analysis warned that effects were small and open to publication bias (Låg & Sæle, 2019), and a recent controlled study found no advantage for flipping by itself (Wiesner et al., 2026). Reviews also note that flipped-learning research often rests on weak theoretical

grounding (Li et al., 2023). The model is therefore useful, but not by magic. Its value depends heavily on the quality of pre-class work and the support given to learners while they do it.

Generative AI enters the model at this vulnerable point. It can produce or adapt pre-class input, provide quick individual feedback, and act as a conversation partner when teachers cannot offer such practice to every learner at scale (Moorhouse & Wong, 2025). Recent reviews of generative AI in language education, however, show an uneven picture. The research is concentrated in higher education and writing, and it often relies on learners' perceptions rather than on measured learning outcomes (Jeon, 2025; Teng, 2024). The same literature also returns to several concerns: unreliable output, over-reliance, loss of authorial voice, academic integrity, and teachers' own anxiety about adoption (Al Hosni, 2025; Neff et al., 2024; Parviz & Arthur, 2025; Toncelli & Kostka, 2024). For that reason, the question is not simply whether AI can support a flipped classroom. The harder question is what happens when it is built into the design.

Some empirical work has already begun to place generative AI inside structured, student-centered designs that resemble the logic of flipped learning. In blended, project-based, and distributed-learning environments, AI has been used to scaffold independent work and to release class time for more demanding tasks (Hsu, 2025; Lai, 2026; Sağ & Kip-Kayabaş, 2026). Flipped learning is a particularly clear case because its two phases give AI obvious entry points. It may help prepare or personalise the pre-class input. It may also support interaction, practice, and feedback during or after class. What has been missing is a synthesis of studies where AI and the flipped model are examined together in ELT.

Two review traditions have therefore developed side by side. Reviews of flipped learning in ELT describe the model's record, but they largely predate the everyday availability of generative AI (Ling & Jan, 2024). Reviews of generative AI in language education describe rapid uptake and recurring risks, but they usually treat AI as a broad technology rather than as part of a specific instructional design (Jeon, 2025). To our knowledge, no systematic review has yet brought together empirical work on generative or broader AI-supported flipped learning in ELT during the post-ChatGPT period. Given the speed of adoption, such a review is needed before classroom practice outruns the evidence.

This study addresses that gap. It reviews empirical research on AI-supported flipped learning in ELT published from 2023 onward, after the public release of ChatGPT, and retrieved from Web of Science and Scopus under the filters described below. Because the eligible studies often combine generative AI with other AI-driven tools, the review uses the broader term AI-supported flipped learning rather than limiting the corpus to generative AI only. Through qualitative thematic synthesis, it asks:

What does research published since the emergence of generative AI reveal about AI-supported flipped learning in ELT contexts? The question is examined through four dimensions: (a) research contexts and participant profiles; (b) the ways AI is integrated into the flipped design and the language skills targeted; (c) reported benefits; and (d) challenges and limitations of implementation.

Literature Review

Flipped learning: foundations and evidence in ELT

Flipped learning moves direct instruction into individual pre-class study so that class time can be used for application, collaboration, and teacher-guided practice. It is usually linked to constructivist, active-learning, and student-centered traditions, and it also assumes a degree of self-regulation because learners must take responsibility for first contact with new material (Kozikoğlu, 2019). That theoretical grounding, however, is not always made explicit. Li et al. (2023) found in their scoping review that about two-thirds of flipped-learning studies did not clearly connect their work to a theory or conceptual framework. For the present review, active learning and self-regulated learning therefore serve as an interpretive lens, while the analysis also notes whether the reviewed studies state their own theoretical position.

The empirical record for flipped learning is wide but uneven. Across disciplines, reviews and meta-analyses report generally positive but moderate effects on achievement, together with effects on self-efficacy, motivation, autonomy, and other affective or self-regulatory outcomes (Algarni, 2025; Hanif et al., 2025; Hew et al., 2021). Within ELT, syntheses link the model with gains in proficiency and in specific skills, especially speaking and writing (Jantakoon et al., 2024; Ling & Jan, 2024), and with improved learner autonomy at university level (Adrianzén-Olano, 2026). Still, the benefits are conditional. Learners' acceptance and participation depend strongly on the quality of pre-class tasks (Lau et al., 2025). Younger learners need more scaffolding and often parental support (Ağırman & Ercoşkun, 2025). Controlled work also suggests that flipping alone has limited value unless learners are supported in regulating their independent study (Wiesner et al., 2026). In short, the flip itself is not the intervention. The decisive work lies in the materials, the task design, and the support around out-of-class learning.

Generative AI in English language teaching: functions and evidence

The release of ChatGPT in late 2022 prompted a rapid reorientation of language-education research and practice (Meniado, 2023; Moorhouse & Wong, 2025). Bibliometric studies show the scale of that shift. Research on AI in EFL had been accumulating since 2013, but output rose steeply after large language models became widely available, with China, the United States, and the United Kingdom among the

leading contributors (Zhang & Umeanowai, 2025; Zhou et al., 2026). Moorhouse and Wong (2025) describe generative AI as a tool that can support the whole instructional cycle: materials, modelling, feedback, and conversational practice. They also stress that its pedagogical value depends on how teachers frame, constrain, and monitor its use.

Synthesized evidence on generative AI in ELT points to a clear but narrow center. Reviews repeatedly identify writing as the dominant skill area, with ChatGPT used mainly for feedback, drafting, and revision (Jeon, 2025; Teng, 2024). Other work links AI use with autonomy, self-directed learning, and higher-order skills such as critical thinking (Slamet & Basthomi, 2025; Zakaria et al., 2025). At the same time, the evidence base remains thin in important places. Recent reviews report a concentration in higher education, a heavy dependence on self-reported perceptions, and limited evidence for skills other than writing (Jeon, 2025). The field is growing quickly, but its strongest claims still come from a narrow set of contexts and outcomes.

Concerns, ethics, and the teacher's role

Enthusiasm for generative AI in ELT is usually accompanied by caution. The most common concerns are the unreliability of AI output and the possibility that fluent, ready-made text will encourage over-reliance rather than independent effort (Slamet & Basthomi, 2025; Teng, 2024). Evidence from outside language education points in a similar direction: among knowledge workers, confidence in generative AI was associated with less critical thinking and reduced cognitive effort (Lee et al., 2025). Integrity concerns sit close to this. Heavy AI use may erode students' authorial voice (Al Hosni, 2025), while AI-detection tools remain unreliable and can misclassify both human and machine writing (Hadra et al., 2026). Ethical use of AI has therefore become a practical question for both students and teachers (Neff et al., 2024).

Teachers sit at the center of this problem, whether or not the technology seems learner-facing. Faculty attitudes toward AI are often mixed, combining interest in its usefulness with concern about its effects on learning (Toncelli & Kostka, 2024). Some teachers report active anxiety about integrating these tools (Parviz & Arthur, 2025). Work in language teacher education argues that meaningful integration requires more than technical familiarity; teachers need critical competence in judging, constraining, and contextualizing AI-generated text (Buendgens-Kosten, 2024; Karakaya et al., 2025). This is harder in under-resourced settings, where instructors may adopt AI for writing instruction despite uneven access and limited training (Ansaree & Uddin, 2026). For flipped designs, these concerns are sharper because much of the learning is deliberately moved into an unsupervised pre-class phase.

Generative AI in student-centred designs and the research gap

A small line of research now embeds generative AI within structured, student-centred designs that share the flipped model's logic: learners prepare outside class, and class time is used for more interactive or higher-order work. In blended EFL instruction, AI-assisted designs have improved speaking performance and learning resilience (Lai, 2026). In project-based language learning, certain combinations of AI support, learner autonomy, and collaboration have been associated with stronger language development (Hsu, 2025). In distributed learning, AI scaffolding has supported argumentative writing and critical thinking (Sağ & Kip-Kayabaş, 2026). A ChatGPT-integrated BOPPPS framework, organised around pre-, during-, and post-class phases, has also been used to develop oral proficiency (Lai, 2025). These studies show that AI is already being placed inside phased, autonomy-oriented designs, even when the label flipped learning is not always foregrounded.

What is still missing is a synthesis of the flipped case on its own terms. Reviews of flipped learning in ELT establish the model's value but were conducted before generative AI became widely available (Ling & Jan, 2024). Reviews of generative AI in language education treat the technology broadly and do not isolate its use within flipped design (Jeon, 2025). The gap is therefore specific rather than general: individual studies of AI-supported flipped ELT have appeared, but they have not yet been brought together to show where the research has been conducted, how AI is placed in the flipped cycle, what outcomes are reported, and what problems recur. The present review addresses that gap through the four dimensions set out above.

Method

Research design

This study uses a systematic review design to synthesise empirical research on AI-supported flipped learning in English language teaching (ELT) published after the public emergence of generative AI. A systematic review requires an explicit and reproducible procedure for identifying, screening, and synthesising studies in response to a defined question. The review was conducted and reported with reference to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines (Page et al., 2021). Because the eligible studies differed substantially in design, tools, outcomes, and measures, the synthesis was qualitative and thematic rather than statistical. It was organized around the review's four guiding dimensions: research contexts and participant profiles; modes of AI integration and targeted language skills; reported benefits; and implementation challenges and limitations.

Scope and operational definition

The scope was defined as AI-supported flipped learning in ELT after the emergence of generative AI. This wording is deliberate. The eligible studies did not use generative AI alone. They included ChatGPT and other generative tools, but also AI chatbots, AI-generated videos, Duolingo, Quizlet, mobile AI-supported tools, and AI functions embedded in learning-management systems. A narrower definition would have excluded studies in which AI support was pedagogically real but not purely generative. The broader scope therefore reflects the literature as it is currently taking shape, while the analysis still distinguishes generative AI from other AI-supported tools where the studies make that distinction possible.

Search strategy

A systematic search was conducted in Web of Science (WoS) and Scopus because both databases cover peer-reviewed research in education and applied linguistics. The search was run on 1 June 2026 and limited to studies from 2023 onward, which was treated as the post-ChatGPT period and as the practical starting point for the wide availability of generative AI.

The search combined three concept blocks (flipped learning, the ELT context, and artificial intelligence) using Boolean operators and truncation. The Web of Science topic-field query was:

TS=(("flipped learning" OR "flipped classroom*" OR "flipped instruction" OR "flipped teaching" OR "flipped pedagogy" OR "inverted classroom*") AND ("English language teach*" OR "English language learn*" OR ELT OR EFL OR ESL OR TESOL OR "English as a foreign language" OR "English as a second language") AND ("artificial intelligence" OR AI OR "generative AI" OR GenAI OR ChatGPT OR "large language model*" OR LLM* OR chatbot* OR "AI-assisted" OR "AI-supported"))

Equivalent concept blocks were applied in Scopus using the TITLE-ABS-KEY field. In both databases, results were restricted to English-language, open-access publications. In Scopus, the search was also limited to journal articles in the subject areas of social sciences, computer science, and arts and humanities.

Inclusion and exclusion criteria

Inclusion and exclusion criteria were set before screening. A study was included if it:

- was published in the post-ChatGPT period covered by the search;
- was retrieved from Web of Science and/or Scopus under the applied filters and was available as an open-access journal article in English;

• addressed ELT, EFL, TESOL, English language education, or an English-related higher-education context involving EFL learners or English teachers;

• treated flipped learning, the flipped classroom, or a flipped learning strategy as a central pedagogical design rather than a passing reference;

• integrated AI, generative AI, AI chatbots, AI-generated materials, AI-powered language-learning tools, or mobile AI-supported technologies into the flipped learning process; and

• provided analysable evidence relevant to at least one of the four review dimensions.

A study was excluded if it:

• fell outside the post-generative-AI scope of the review;

• addressed AI algorithms, corpus modelling, or educational data modelling without examining AI-supported flipped ELT practice;

• discussed ChatGPT or generative AI in ELT without an implemented flipped learning design;

• was conceptual, theoretical, or scenario-based and lacked empirical, reflective, or classroom-based data; or

• did not provide enough evidence to be coded under the review's four dimensions.

Study selection

The Web of Science search produced 19 records; after applying the English-language and open-access filters, 7 remained. The Scopus search produced 29 records; after limiting the results to journal articles, English, relevant subject areas, and open access, 9 remained. The two sets gave 16 records in total. Five duplicates were removed, leaving 11 records for screening. Full-text screening excluded 3 studies and left 8 studies in the final corpus.

The three exclusions were handled as follows. Chen's study of a data-driven, multidimensional corpus-based English teaching model was excluded because it dealt with corpus modelling and AI algorithms rather than AI-supported flipped learning practice in ELT. Ljubojević's study was excluded because, although it concerned ChatGPT and English argumentative writing, it evaluated ChatGPT-generated essays rather than an AI-supported flipped classroom. Todd's conceptual article on generative AI as a disrupter of language education was excluded because it discussed possible future scenarios for TESOL, including flipped classrooms, without reporting empirical

classroom data, participant profiles, or implementation outcomes that could be coded for this review. Figure 1 summarizes the selection process.

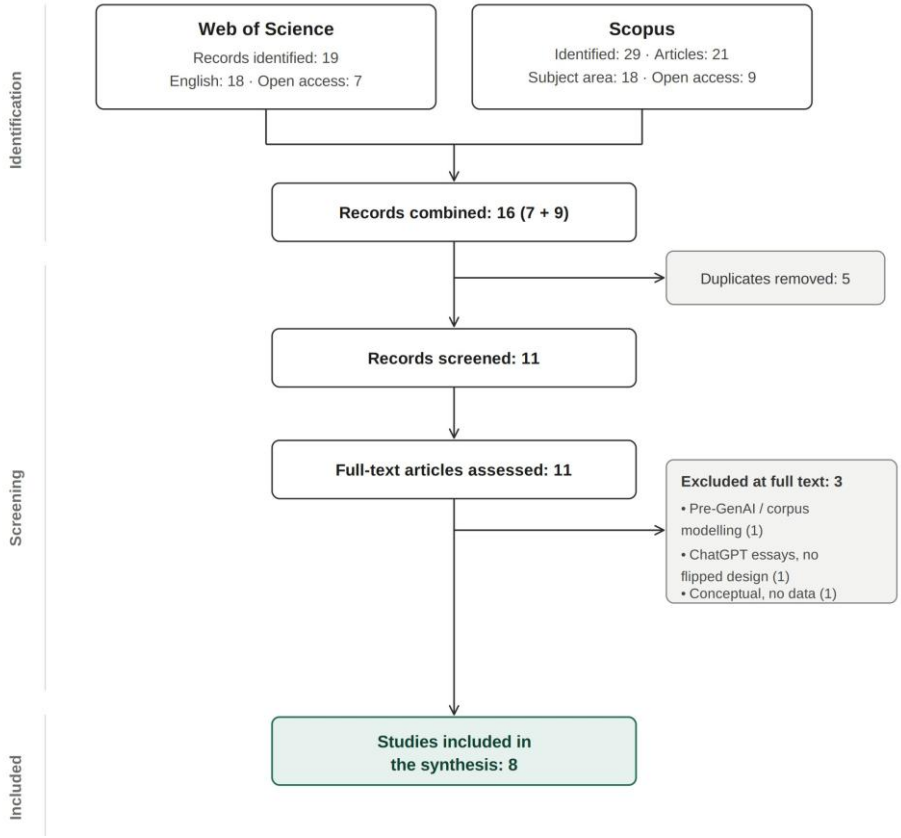

Figure 1. PRISMA 2020 flow diagram of the study selection process.

Data extraction and synthesis

Data were extracted from each included study along five dimensions: research context, participant characteristics, the AI tool used and its location or function within the flipped design, the language skill or content area targeted, and reported outcomes, including both benefits and challenges. The extracted data were analysed through qualitative thematic synthesis. Recurring patterns were grouped into themes and then mapped onto the four guiding dimensions of the review.

Findings

Overview of the included studies

The eight eligible studies were published in 2025 and 2026, which confirms that empirical work on AI-supported flipped learning in ELT is very recent. All eight were set in higher education and all concerned English as a foreign language (EFL). None examined primary or secondary education, and none was located in a second-language or native-English-speaking environment. Most studies were located in Asia and the

Middle East: two in Iran (Moulavinafchi et al., 2026; Namaziandost, 2025), two in Thailand (Anggoro, 2025; Phanwiriyarat et al., 2025), and one each in Iraq (Jadwe et al., 2025), Malaysia (Ling & Jan, 2025), and Indonesia (Mudra, 2025). The only case outside this cluster was conducted in Ecuador (Castillo-Cuesta et al., 2026).

The studies were methodologically mixed. Four used quasi-experimental designs that compared an AI-supported condition with a comparison group (Castillo-Cuesta et al., 2026; Jadwe et al., 2025; Moulavinafchi et al., 2026; Namaziandost, 2025), and one of these also collected survey and interview data (Castillo-Cuesta et al., 2026). Phanwiriyarat et al. (2025) used a mixed-methods design built around twelve weeks of classroom work. The remaining studies consisted of a survey of learner preferences (Mudra, 2025), a qualitative case study of teachers (Ling & Jan, 2025), and an autoethnography of one instructor's practice (Anggoro, 2025). Sample sizes ranged from one instructor to 279 learners. Because the designs and measures varied too widely for statistical synthesis, the findings were analysed thematically.

The AI tools also varied. ChatGPT appeared most often, either alone or together with other systems. Moulavinafchi et al. (2026) paired it with Quizlet; Jadwe et al. (2025) combined ChatGPT, Gemini, and DeepSeek; and Mudra (2025) reported ChatGPT among the mobile tools learners selected. Other studies used AI video-generation platforms (Castillo-Cuesta et al., 2026), AI chatbots (Ling & Jan, 2025), and the AI-powered application Duolingo (Phanwiriyarat et al., 2025). One study described its AI components by function rather than by platform, referring to an AI-powered learning-management system, adaptive pre-assessments, and an AI writing assistant, but not naming the specific tools (Namaziandost, 2025). The skill focus was strongest for productive skills and grammar: writing (Namaziandost, 2025), grammar (Moulavinafchi et al., 2026), speaking (Phanwiriyarat et al., 2025), and literary reading or interpretation (Jadwe et al., 2025). Receptive and integrated skills were much less visible. Table 1 summarises the studies.

Table 1. Characteristics of the Included Studies (N = 8)

Study	Country	Level / learners	Design	Generative AI tool(s)	Skill focus	Key reported outcome(s)
Namaziandost (2025)	Iran	University, intermediate EFL (N = 70)	Quasi-experimental (ANCOVA)	AI-powered LMS, adaptive pre-assessment, AI writing assistant (specific tools not documented)	Writing	Writing and metacognitive awareness ↑; boredom ↓

Moulavinafchi et al. (2026)	Iran	University, English Literature (N = 20)	Quasi-experimental, 16 wks	ChatGPT, Quizlet	Grammar	Grammatical competence ↑; anxiety ↓
Jadwe et al. (2025)	Iraq	University, 4th-year EFL (N = 27)	Quasi-experimental, 9 wks	ChatGPT, Gemini, DeepSeek	Literature/reading	Achievement ↑; broader content coverage
Ling & Jan (2025)	Malaysia	University English teachers (N = 6)	Qualitative case study (TPACK)	AI chatbots	Vocabulary, grammar, speaking	Autonomy, confidence, interaction; teacher agency
Phanwiriyarat et al. (2025)	Thailand	University, 1st-year (N = 48)	Mixed methods, 12 wks	Duolingo (AI-powered)	Speaking	Speaking and confidence ↑; anxiety ↓
Anggoro (2025)	Thailand	University instructor (N = 1)	Autoethnography	GenAI platforms	General EFL	Lesson planning, personalization, engagement
Castillo-Cuesta et al. (2026)	Ecuador	University, B2 (N = 156)	Quasi-experimentalmixed methods	AI video tools (Fliki, Steve AI, etc.)	Active learning/comprehension	Comprehension, retention, engagement ↑
Mudra (2025)	Indonesia	University EFL, 13 universities (N = 279)	Survey	ChatGPT (+ other mobile tools)	Mixed (oral, writing)	Flexibility, self-pacing, collaboration

Note. ↑ = reported increase; ↓ = reported decrease. Outcomes are reported as stated in each study and were not pooled statistically.

Research contexts and participant profiles

The small corpus is uniform in two basic ways. Every study was conducted in higher education, and every study targeted EFL. There is no evidence here from primary or secondary schools, and no study comes from a second-language or native-English-speaking setting. On the evidence available, AI-supported flipped learning in ELT is

therefore a university-level EFL phenomenon. That matters because adult university learners are more likely than younger learners to have the autonomy, self-regulation, and device access that both flipped and AI-supported designs assume. The transferability of these findings to school-age learners cannot be claimed from the present corpus.

The geographic pattern is also narrow. The studies are concentrated in Iran, Thailand, Iraq, Malaysia, Indonesia, and Ecuador. No study in the corpus was based in Europe, North America, or Africa. This is consistent with broader reviews of generative AI in language education, which also report a strong higher-education and EFL orientation (Jeon, 2025). The implication is simple but important: current claims about AI-supported flipped learning rest on a regionally and institutionally specific body of evidence.

The studies also differ in whose experience they foreground. Six centered on learners (Castillo-Cuesta et al., 2026; Jadwe et al., 2025; Moulavinafchi et al., 2026; Mudra, 2025; Namaziandost, 2025; Phanwiriyarat et al., 2025). Two foregrounded teachers: six university instructors in a qualitative case study (Ling & Jan, 2025) and a single instructor in an autoethnography (Anggoro, 2025). This imbalance is worth noting. Flipped learning places heavy design and facilitation demands on teachers, but teacher-focused evidence in this corpus comes from only seven practitioners. Several learner samples were also specialized, including English-literature students (Jadwe et al., 2025; Moulavinafchi et al., 2026) and pre-service English-education majors (Mudra, 2025), rather than general EFL learners.

Sample sizes ranged from one instructor (Anggoro, 2025) to 279 learners across thirteen universities (Mudra, 2025). The quasi-experimental studies, which provide the corpus's strongest causal claims, were still modest in size: 20 participants (Moulavinafchi et al., 2026), 27 (Jadwe et al., 2025), 70 (Namaziandost, 2025), and 156 (Castillo-Cuesta et al., 2026). No single effect claim should therefore carry too much weight on its own. The safer reading is across studies, not from any one study.

Learner proficiency was reported unevenly. Where it was specified, levels ranged from A1 in Anggoro's (2025) classes, to intermediate learners in Namaziandost (2025), and B2 learners in Castillo-Cuesta et al. (2026). Other studies did not report proficiency in a comparable form, or reported it only loosely. This matters because proficiency is likely to affect how learners manage self-directed AI-supported pre-class tasks. The uneven reporting is therefore a limitation of the evidence base itself.

Integration of generative AI into the flipped design and the language skills targeted

Across the studies, AI entered the flipped model in three main ways: it prepared or delivered pre-class input, acted as a partner for practice and feedback, and supported the teacher's own planning or mediation. The roles overlap in a few studies, but they are useful because they show where the AI is doing pedagogical work. Reporting of the flipped phase was not always clear. Three studies stated explicitly when AI was used in the cycle (Castillo-Cuesta et al., 2026; Namaziandost, 2025; Phanwiriyarat et al., 2025), while the others described AI's function without always locating it precisely.

In the first role, AI produced or delivered the material learners used before class. The clearest example is Castillo-Cuesta et al. (2026), where AI video-generation tools were used to create instructional videos that students reviewed before synchronous sessions. Namaziandost (2025) followed a similar logic through an AI-powered learning-management system that supplied multimodal materials and adaptive pre-assessments. In Jadwe et al. (2025), learners used free tools such as ChatGPT, Gemini, and DeepSeek to prepare literary modules before class, helping the course cover more texts within a semester than conventional instruction allowed. These examples show AI being used to address a familiar bottleneck in flipped teaching: the time and effort needed to prepare good pre-class input.

The second role positioned AI as a practice and feedback partner. Phanwiriyarat et al. (2025) asked students to complete daily gamified speaking and vocabulary tasks on Duolingo as pre-class work, with the application generating progress and accuracy reports. Mudra's (2025) learners selected ChatGPT and other mobile tools for drafting, revising, and rehearsing oral output before class, suggesting that learners may bring AI into the flipped cycle on their own when the design allows it. Ling and Jan (2025) documented teachers using chatbots for vocabulary support, grammar explanation, and dialogue simulation across preparation and in-class collaboration. Namaziandost's (2025) study also belongs partly in this group because its AI writing assistant gave automated, criterion-referenced feedback on drafts.

The third role turned AI toward the teacher. Anggoro's (2025) autoethnography describes how generative AI changed his own practice by streamlining lesson planning and supporting more personalised and interactive activities across the flipped cycle. Ling and Jan (2025) also belong here because their teachers did not simply hand chatbots to students. They selected, limited, and explained chatbot use, taking on the role of facilitators of AI-mediated learning. This teacher-facing dimension is thinly represented, but it is important because flipped learning has always depended heavily on design and facilitation.

Moulavinafchi et al. (2026) sits at the edge of these categories and needs careful handling. In that study, the experimental treatment was an AI-powered classroom using ChatGPT for interactive and personalised learning, while the comparison condition was a traditional flipped classroom. AI and flipping were therefore treated partly as alternatives rather than as a fully integrated design. The study is useful for this review because it shows how close AI-supported learning and flipped learning can become, but it also marks a boundary. For the term AI-supported flipped learning to be precise, AI should operate inside the flipped structure, not simply stand in for it.

The targeted language areas clustered around production and grammar. Writing was the most direct focus (Namaziandost, 2025), followed by grammar (Moulavinafchi et al., 2026) and speaking (Phanwiriyarat et al., 2025). Jadwe et al. (2025) focused on the comprehension and interpretation of literary texts. Several studies worked across skills rather than isolating one: Ling and Jan's (2025) chatbots supported vocabulary, grammar, and spoken interaction; Mudra's (2025) learners pursued oral and written aims; and Castillo-Cuesta et al. (2026) framed the outcome as active learning and comprehension rather than as one discrete skill. Listening, reading outside the literary context, pronunciation, and integrated skills were largely absent. The evidence therefore speaks most clearly to productive and grammatical development, and much less clearly to other areas central to language education.

Reported benefits

The reported benefits can be grouped into three overlapping areas: language or content learning, affective outcomes such as confidence and reduced anxiety, and cognitive or self-regulatory outcomes such as metacognition, autonomy, and engagement. The strength of these claims differs by study design, so the findings below should be read with that variation in mind.

The most direct evidence concerns language and content learning. In the quasi-experimental studies, AI-supported conditions were associated with measurable gains over comparison groups. Namaziandost (2025) reported improvement in writing development, Castillo-Cuesta et al. (2026) reported better comprehension and retention of course content, and Jadwe et al. (2025) reported higher achievement across literary modules. Phanwiriyarat et al. (2025) found gains in speaking, especially in topic-based communication and discussion. Moulavinafchi et al. (2026) reported improved grammatical competence, although that finding should be read with the caution noted in Section 4.3 because AI was compared with a flipped condition rather than embedded in it. Overall, the studies point in a positive direction, but the samples, skills, and instruments differ too much for the gains to be treated as one pooled effect.

The affective findings are also positive. Moulavinafchi et al. (2026) reported lower foreign-language anxiety in the AI condition, Namaziandost (2025) reported a reduction in foreign-language learning boredom, and Phanwiriyarat et al. (2025) found that learners felt more confident in casual conversation and presentations. A plausible explanation is that AI gives learners a private, low-pressure space for rehearsal before they enter the more public classroom. Still, the evidence is not strong enough to make this a settled claim, since much of it depends on small samples and self-report.

Cognitive and self-regulatory benefits were also reported. Namaziandost (2025) found gains in metacognitive awareness and linked them to diagnostic feedback and self-assessment prompts in the pre-class phase. Autonomy appeared repeatedly: Castillo-Cuesta et al. (2026) found that AI videos supported independent learning, Ling and Jan (2025) observed that chatbot integration enhanced learner autonomy and confidence, and Mudra's (2025) learners valued the flexibility and self-pacing of mobile AI tools. Engagement was reported by several studies (e.g., Anggoro, 2025; Castillo-Cuesta et al., 2026; Phanwiriyarat et al., 2025). These outcomes fit the basic rationale of flipped learning, which is to shift some responsibility for learning onto the student. AI may make that responsibility easier to manage, but the studies also show that it must be scaffolded.

Two studies reported benefits for teaching practice rather than direct learning outcomes. Anggoro (2025) described generative AI as easing lesson planning and helping him design more personalised activities. Ling and Jan (2025) found that teachers used chatbots deliberately, adapting them to specific pedagogical purposes and repositioning themselves as facilitators. These findings are based on qualitative and self-reported evidence, so they should be treated cautiously. They are still important because teacher design work is one of the conditions on which flipped learning depends.

Challenges and limitations of implementation

The benefits were accompanied by a consistent set of problems. This is important because the same studies that reported gains often reported risks as well. The studies describe a promising design that needs careful handling, not a tool that simply works.

The most common concern was the reliability of AI output. Ling and Jan (2025) found that teachers were wary of vague or incorrect chatbot responses, and Anggoro (2025) described the need to verify AI-generated content before using it with students. In a flipped design, this problem has higher stakes. If a learner meets an inaccurate explanation before class, it may remain uncorrected until the lesson, or may already have shaped the learner's understanding by then.

A related concern was over-reliance. Ling and Jan (2025) reported teachers' fear that students might lean on chatbot-generated content instead of practising or thinking

independently. Anggoro (2025) raised the same issue for teachers, noting that the convenience of generative tools could weaken human judgement and connection. This cuts against one of the aims of flipped learning. A design meant to build autonomy can produce dependence if AI replaces effort rather than scaffolding it.

Ethics and integrity formed a third concern. Anggoro (2025) raised data privacy and academic integrity as practical issues, especially because generative tools can produce work that looks ready to submit. These problems are especially difficult in the pre-class phase, where learners work independently and where the teacher cannot easily see whether AI has supported the work or replaced it.

The studies also reported learner-side barriers that predate AI but are not solved by it. Mudra (2025) identified technical, pedagogical, financial, and psychological barriers to mobile technology use in pre-class flipped work, with access and cost among the most influential. These barriers matter for equity. A flipped design that assumes reliable devices, connectivity, and access to AI tools may widen gaps between learners. Self-regulation remained a problem as well: Anggoro (2025) reported that students initially struggled with the self-discipline required for self-directed pre-class learning, even though AI and gamification were introduced partly to help with that problem.

The fit between AI tools and proficiency levels was also uneven. Phanwiriyarat et al. (2025) found that the gamified AI application supported foundational speaking skills, but intermediate and advanced learners saw it as less useful. A single AI tool may therefore not serve a mixed-proficiency class equally well. This is a small point in the corpus, but a large one for classroom design.

Finally, the corpus itself has limits. As shown above, the studies are confined to higher education and EFL contexts, clustered in a small number of countries, and often based on small samples and short interventions. Several rely on self-reported perceptions. Proficiency levels were inconsistently reported, and at least one study described AI functions without naming the tools used (Namaziandost, 2025). The boundary case of Moulavinafchi et al. (2026) also shows that researchers are not yet using a stable operational definition of AI-supported flipped learning. These problems do not cancel the positive findings, but they do mean that the evidence is early and still somewhat loose at the edges.

Discussion

This review examined what post-ChatGPT research reveals about AI-supported flipped learning in ELT. The eight eligible studies show a field that is clearly emerging but still thin. Its present shape tells us as much about absent evidence as about existing findings. The discussion interprets the synthesis in three areas: how the corpus fits the

wider literature, what AI changes in the flipped model, and what the findings imply for practice and teacher education.

Patterns in the evidence and their fit with prior research

The clearest pattern is that the corpus reproduces several biases already visible in the broader generative-AI-in-ELT literature. Reviews of that wider field report a concentration in higher education, a dominance of writing and other productive skills, a heavy presence of ChatGPT, and reliance on self-reported perceptions rather than measured outcomes (Jeon, 2025; Teng, 2024). The present corpus mirrors these tendencies. Every study was set in higher education, productive skills and grammar predominated, ChatGPT was the most frequent named tool, and several studies foregrounded perceptions. Research on AI-supported flipped learning has therefore inherited the contours of its parent field rather than yet forming a distinct evidence base. The complete absence of school-age learners is the sharpest example, especially because flipped-learning research more broadly has examined primary settings and shown that they raise different developmental and support issues (Ağırman & Ercoşkun, 2025).

The reported benefits also fit both parent literatures. Gains in writing, grammar, speaking, and content achievement echo the positive but moderate effects reported for flipped learning in ELT (Jantakoon et al., 2024; Ling & Jan, 2024) and for generative AI in writing instruction (Teng, 2024). The self-regulatory gains in autonomy and metacognition align with flipped-learning research on self-efficacy and autonomy (Adrianzén-Olano, 2026; Algarni, 2025) and with AI-in-ELT work on self-directed learning and higher-order thinking (Slamet & Basthomi, 2025; Zakaria et al., 2025). The most repeated finding was affective: lower anxiety or boredom and higher confidence. One plausible reading is that AI lowers the affective cost of rehearsal by moving some language practice into a private and self-paced space before class. This interpretation should remain cautious, because the affective claims still rest on small samples and self-report, the same weakness noted in the wider field (Jeon, 2025).

How generative AI reshapes the flipped model

The central contribution of this review is to show where AI sits in the flipped design. AI is not simply added on top of the model. In the studies reviewed, AI prepared or delivered pre-class input, acted as a feedback and practice partner, or supported teacher planning and mediation. These roles map onto the structure of flipped learning itself. That matters because previous research has located the model's weakness in the quality of pre-class materials and in learners' ability to regulate independent study (Lau et al., 2025; Wiesner et al., 2026). Generative AI is being used to work on those pressure points. Its promise, then, is not novelty alone. It is the fit between what AI can do and where the flipped model often struggles.

The same fit also increases the risks. This is also where the risks become more visible. In flipped learning, students do much of the first work alone before class. If they use generative AI at that stage, problems such as inaccurate output, over-reliance, academic integrity, and loss of authorial voice may not be noticed until later, if they are noticed at all (Al Hosni, 2025; Hadra et al., 2026; Neff et al., 2024; Teng, 2024). A wrong explanation met before class may not be corrected quickly. A tool meant to support autonomy may instead become a shortcut around effort. Lee et al. (2025) raise a similar concern outside language education, linking confidence in generative AI with reduced critical thinking and cognitive effort. SAM (2025) goes further, warning that generative AI may be counter-productive for language acquisition if it removes the productive struggle through which learning happens. The flipped classroom is therefore a natural place for AI, but also an exposed one. The boundary case of Moulavinafchi et al. (2026), where AI and flipping were treated partly as alternatives, reinforces the need for a clearer definition: AI inside the flip is not the same as AI instead of the flip.

Implications for practice and teacher education

Two practical points follow from the synthesis. First, pre-class AI use should be structured, not simply allowed. The studies that presented the most convincing uses of AI were those in which teachers curated, limited, and explained the tools (Anggoro, 2025; Ling & Jan, 2025). Unguided access is more likely to produce the problems reported above. Second, equity has to be designed for. Mudra (2025) reported technical, financial, and access-related barriers to learners' pre-class use of AI and mobile tools, and Ansaree and Uddin (2026) report similar difficulties in under-resourced teaching contexts. A flipped design that assumes devices, connectivity, and tool access may widen the gap between learners. Proficiency also matters. Phanwiriyarat et al. (2025) showed that one AI tool did not serve a mixed-ability cohort equally well, so tool choice and task design need to match learner level.

These findings put the teacher, not the tool, at the centre of integration. Teacher attitudes toward generative AI are often ambivalent (Toncelli & Kostka, 2024), and some teachers experience anxiety about adopting it (Parviz & Arthur, 2025). Yet the evidence here suggests that meaningful AI-supported flipped learning depends on teacher judgement: selecting tools, setting limits, designing tasks, and helping learners decide what counts as responsible use. Language teacher education therefore needs to go beyond tool training. It should prepare teachers to design and mediate learning environments in which students use AI critically and productively (Buendgens-Kosten, 2024; Karakaya et al., 2025).

Limitations of this review

Several limitations qualify these conclusions. The corpus is small, with only eight studies, and it was drawn from two databases under an open-access restriction. The review protocol was not preregistered, and no independent risk-of-bias or quality appraisal was conducted. The open-access restriction also matters because publishing models differ across countries and journals; the geographic pattern reported here may partly reflect what was accessible rather than the whole field. The studies were heterogeneous in design and outcome measures, which prevented statistical pooling and required qualitative thematic synthesis. Outcomes were taken as reported by the original authors. Finally, because recent flipped designs combine generative AI with other AI-driven tools, the review used a broad definition of AI support. This allowed the corpus to reflect current practice, but it also means that studies where generative AI was central sit beside studies where AI was more peripheral. The findings should therefore be read as an early characterization of a forming field, not as a settled account of effectiveness.

Conclusion

This review asked what research published since the emergence of generative AI reveals about AI-supported flipped learning in ELT. The eight studies from 2025 and 2026 show a field that is real but still forming. The evidence is active enough to show patterns, but too narrow to support firm claims about effectiveness. The research has so far been conducted almost entirely in higher-education EFL settings, within a limited set of countries and with modest samples. No study examined school-age learners. AI was integrated into the flipped model in three main roles: preparing or delivering pre-class input, serving as a practice and feedback partner, and supporting teachers' planning and mediation. ChatGPT was the most frequently named tool, while other AI-supported tools such as Duolingo, AI videos, chatbots, and AI-enabled learning systems also appeared. The skill focus was strongest for writing, grammar, and speaking. Reported benefits included language gains, confidence, reduced anxiety or boredom, engagement, autonomy, and metacognitive awareness. Reported challenges centered on reliability, over-reliance, academic integrity, privacy, unequal access, proficiency fit, and teacher workload.

The main point is that AI fits the flipped model at the pre-class phase, the point where the model has always been most fragile. Flipped learning depends on good preparatory materials and on learners' capacity to manage independent study. The reviewed studies suggest that AI is being used to strengthen that phase by making input, feedback, and rehearsal more available. But the same fit intensifies the risks, because inaccurate output or over-reliance is hardest to notice when learners are working alone. AI-supported flipped learning should therefore not be treated as a straightforward

upgrade. It is a demanding combination whose value depends on deliberate teacher design, learner guidance, and realistic attention to access.

Future research needs to widen the field. Studies are needed beyond higher education and beyond the EFL contexts that currently dominate the evidence, including school-age learners, more countries, and more language settings. Receptive and integrated skills need more attention, as the present corpus is weighted toward writing, grammar, and speaking. Designs should measure learning outcomes more directly and rely less on perception data. The field also needs a clearer operational definition of AI-supported flipped learning, especially to distinguish AI used within the flipped structure from AI used instead of it. Finally, more attention should be paid to teachers: how they design AI-supported pre-class tasks, how they scaffold responsible use, and how they decide when AI helps learning and when it merely makes the work look finished.

References

Adrianzén-Olano, I. (2026). Influence of the flipped classroom on autonomous learning at the university level: A systematic review. *Journal of Technology and Science Education, 16*(1), 315–332. https://doi.org/10.3926/jotse.3799

Ağırman, N., & Ercoşkun, M. H. (2025). Mapping the first twenty years of the flipped classroom model in primary school education: A systematic review. *International Electronic Journal of Elementary Education, 17*(5). https://doi.org/10.26822/iejee.2025.410

Al Hosni, J. (2025). Preserving authorial voice in academic texts in the age of generative AI: A thematic literature review. *Arab World English Journal, 16*(3), 244–258. https://doi.org/10.24093/awej/vol16no3.14

Algarni, B. (2025). Impact of flipped learning on student self-efficacy across age groups and subject domains: A systematic review of 44 studies (2015–2024). *Journal of Information Technology Education: Research, 24*, Article 4. https://doi.org/10.28945/5442

Anggoro, K. J. (2025). Navigating the integration of generative AI in EFL teaching: An autoethnographic journey with the flipped classroom model. *Studies in Self-Access Learning Journal, 16*(1), 237–250. https://doi.org/10.37237/160112

Ansaree, F. N., & Uddin, M. M. (2026). University instructor agency in GenAI-supported ELT writing instruction: Practices and perceived challenges in under-resourced Asia–Pacific contexts. *Asian-Pacific Journal of Second and Foreign Language Education, 11*, Article 23. https://doi.org/10.1186/s40862-026-00399-4

Buendgens-Kosten, J. (2024). "Just a pocket knife, not a machete": Large language models in TEFL teacher education and digital text sovereignty. *Technology in Language Teaching & Learning, 6*(1), 1–16. https://doi.org/10.29140/tltl.v6n1.1192

Castillo-Cuesta, L., Cabrera-Solano, P., & Ochoa-Cueva, C. (2026). Integrating AI-generated videos and the flipped classroom approach to enhance active English language learning. *Studies in English Language and Education, 13*(1), 45–60. https://doi.org/10.24815/siele.v13i1.129

Hadra, M., Cambridge, K., & Mesbah, M. (2026). Evaluating the accuracy and reliability of AI content detectors in academic contexts. *International Journal for Educational Integrity, 22*, Article 4. https://doi.org/10.1007/s40979-026-00213-1

Hanif, M., Mustaji, M., & Arianto, F. (2025). A systematic review of the impact of flipped learning on the emotional attributes of college students. *Journal of Learning for Development, 12*(2), 386–402. https://doi.org/10.56059/jl4d.v12i2.1495

Hew, K. F., Bai, S., Huang, W., Dawson, P., Du, J., Huang, G., Jia, C., & Thankrit, K. (2021). On the use of flipped classroom across various disciplines: Insights from a second-order meta-analysis. *Australasian Journal of Educational Technology, 37*(2), 132–151. https://doi.org/10.14742/ajet.6475

Hsu, L. (2025). A configurational exploration on the promise and perils of generative AI in project-based language learning. *International Journal of Computer-Assisted Language Learning and Teaching, 15*(1), 1–20. https://doi.org/10.4018/IJCALLT.382182

Jadwe, N. I., Al Dilaimy, H., Sayil, A. A., Hakkoosh, M. H., & Jameel, A. S. (2025). The role of integrating AI tools with flipped learning strategy to improve students' English literature competence. *Arab World English Journal, 16*(4), 3–18. https://doi.org/10.24093/awej/vol16no4.1

Jantakoon, T., Jantakun, K., Jantakun, T., Noibuddee, A., Pasmala, R., Wannapiroon, P., & Nilsook, P. (2024). The effectiveness of flipped classroom in English language learning: A meta-analysis. *World Journal of English Language, 15*(3), 50. https://doi.org/10.5430/wjel.v15n3p50

Jeon, E.-Y. (2025). Artificial intelligence in ESL/EFL education: Evidence from recent reviews (2024–2025). *International Journal of Learning, Teaching and Educational Research, 24*(10), 509–526. https://doi.org/10.26803/ijlter.24.10.24

Karakaya, K., Alpat, M. F., Uçar, H., Karakaya, Ö., & Bozkurt, A. (2025). Preparing teachers for the algorithmic educational landscape: A critical mapping of generative AI integration in language teacher education. *Technology in Language Teaching & Learning, 7*(2), Article 102841. https://doi.org/10.29140/tltl.v7n2.102841

Kernagaran, V., & Abdullah, A. (2022). A systematic literature review of flipped learning in English as second language (ESL) context. *International Journal of Instruction, 15*(2), 793–814. https://doi.org/10.29333/iji.2022.15243a

Kozikoğlu, İ. (2019). Analysis of the studies concerning flipped learning model: A comparative meta-synthesis study. *International Journal of Instruction, 12*(1), 851–868. https://doi.org/10.29333/iji.2019.12155a

Låg, T., & Sæle, R. G. (2019). Does the flipped classroom improve student learning and satisfaction? A systematic review and meta-analysis. *AERA Open, 5*(3). https://doi.org/10.1177/2332858419870489

Lai, Z. C. (2025). Enhancing EFL oral proficiency through a ChatGPT-integrated BOPPPS learning framework. *International Journal of Online Pedagogy and Course Design, 15*(1), 1–21. https://doi.org/10.4018/IJOPCD.383301

Lai, Z. C.-C. (2026). AI in EFL blended learning: Impact on speaking performance and learning resilience. *Asian-Pacific Journal of Second and Foreign Language Education, 11*, Article 11. https://doi.org/10.1186/s40862-025-00367-4

Lau, K.-L., Qian, Q., & Chiu, T. K. F. (2025). Effects of instructional factors on students' acceptance of and participation in flipped language learning. *Interactive Learning Environments, 33*(2), 1755–1769. https://doi.org/10.1080/10494820.2024.2375642

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery. https://doi.org/10.1145/3706598.3713778

Li, R., Lund, A., & Nordsteien, A. (2023). The link between flipped and active learning: A scoping review. *Teaching in Higher Education, 28*(8), 1993–2027. https://doi.org/10.1080/13562517.2021.1943655

Ling, Y., & Jan, J. M. (2024). Flipped classroom in college English language teaching in China: A systematic review. *AsTEN Journal of Teacher Education*. https://doi.org/10.56278/asten.vi.2748

Ling, Y., & Jan, J. M. (2025). Voices from the flip: Teacher perspectives on integrating AI chatbots in flipped English classrooms. *Education Sciences, 15*(9), Article 1219. https://doi.org/10.3390/educsci15091219

Meniado, J. C. (2023). The impact of ChatGPT on English language teaching, learning, and assessment: A rapid review of literature. *Arab World English Journal, 14*(4), 3–18. https://doi.org/10.24093/awej/vol14no4.1

Moorhouse, B. L., & Wong, K. M. (2025). *Generative artificial intelligence and language teaching* (Elements in Generative AI in Education). Cambridge University Press. https://doi.org/10.1017/9781009618823

Moulavinafchi, A., Alhaideri, S. A. A., & Alfawwaz, S. N. M. (2026). The impact of an AI-powered flipped classroom on grammar competence and language anxiety in EFL learners. *International Journal of Instruction, 19*(1), 105–130. https://doi.org/10.29333/iji.2026.1916a

Mudra, H. (2025). EFL learners' preferences, activities, rationales, and barriers in utilising mobile learning in the context of flipped learning. *Journal of Learning for Development, 12*(3), 560–572. https://doi.org/10.56059/jl4d.v12i3.1611

Namaziandost, E. (2025). Integrating flipped learning in AI-enhanced language learning: Mapping the effects on metacognitive awareness, writing development, and foreign language learning boredom. *Computers and Education: Artificial Intelligence, 9*, Article 100446. https://doi.org/10.1016/j.caeai.2025.100446

Neff, J., Arciaga, K., & Burri, M. (2024). EFL students' and teachers' perceptions of the ethical uses of artificial intelligence tools. *Technology in Language Teaching & Learning, 6*(3), 1–19. https://doi.org/10.29140/tltl.v6n3.1714

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

Parviz, M., & Arthur, F. (2025). AI anxiety in English language education: A study of Iranian EFL teachers' perceptions and demographic influences. *International Journal of Computer-Assisted Language Learning and Teaching, 15*(1), 1–21. https://doi.org/10.4018/IJCALLT.386135

Phanwiriyarat, K., Anggoro, K. J., & Chaowanakritsanakul, T. (2025). Exploring AI-powered gamified flipped classroom in an English-speaking course: A case of Duolingo. *Cogent Education, 12*(1), Article 2488545. https://doi.org/10.1080/2331186X.2025.2488545

Sağ, İ., & Kip-Kayabaş, B. (2026). AI as a pedagogical scaffold: Enhancing English as a foreign language argumentative writing and critical thinking in a distributed

learning environment. *International Review of Research in Open and Distributed Learning, 27*(2), 83–108. https://doi.org/10.19173/irrodl.v27i2.9499

SAM, A. (2025). The end of English language learning: GenAI as a tool for counter-productivity. *Environment and Social Psychology, 10*(2), Article 3163. https://doi.org/10.59429/esp.v10i2.3163

Slamet, J., & Basthomi, Y. (2025). Examining the challenges and opportunities of ChatGPT in EFL education: A systematic review. *Journal of University Teaching & Learning Practice, 22*(2). https://doi.org/10.53761/deezkh88

Teng, M. F. (2024). A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies, 6*(3), 36–57. https://doi.org/10.58304/ijts.20240304

Toncelli, R., & Kostka, I. (2024). A love-hate relationship: Exploring faculty attitudes towards GenAI and its integration into teaching. *International Journal of TESOL Studies, 6*(3), 77–94. https://doi.org/10.58304/ijts.20240306

Wiesner, P., Krauss, S., Stegmüller, N., & Binder, K. (2026). Is flipped classroom really superior? Questioning the flip in K-12 teaching. *Frontiers in Education, 11*, Article 1741733. https://doi.org/10.3389/feduc.2026.1741733

Zakaria, N. Y. K., Hashim, H., & Jamaludin, K. A. (2025). Exploring the impact of AI on critical thinking development in ESL: A systematic literature review. *Arab World English Journal* (Special Issue on Artificial Intelligence), 330–347. https://doi.org/10.24093/awej/AI.19

Zhang, X., & Umeanowai, K. O. (2025). Exploring the transformative influence of artificial intelligence in EFL context: A comprehensive bibliometric analysis. *Education and Information Technologies, 30*, 3183–3198. https://doi.org/10.1007/s10639-024-12937-z

Zhou, B., Pek, L. S., Kabir, N., Yang, J., Bouteraa, M., & He, X. (2026). Artificial intelligence in English-speaking education: A WoS-based bibliometric analysis of technology, pedagogy, and ethics (2021–2025). *International Journal of Learning, Teaching and Educational Research, 25*(3), 853–873. https://doi.org/10.26803/ijlter.25.3.37

KAYNAKÇA

Aktaş, F. N., Argün, Z. (2020a). Görme engelli bireylerin matematiksel iletişim süreçlerinde matematiksel dil kullanımlarının incelenmesi: Kabartma yazının rolü. Türk Bilgisayar ve Matematik Eğitimi Dergisi, 11(1), 128-156.

https://dergipark.org.tr/tr/pub/turkbilmat/issue/54035/602095

Aktaş, F. N., Argün, Z. (2020b). Görme engelli öğrencilerin sayı ve şekil örüntülerini genelleme süreçleri. Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi, 21(2), 947- 981.

"COPE-Dergi Editörleri İçin Davranış Kuralları ve En İyi Uygulama İlkeleri" beyanları:

Etik Kurul Belgesi:	Bu çalışma için etik kurul onayı gerekmemektedir.
Çıkar Çatışması Beyanı:	Bu makalenin araştırması, yazarlığı veya yayınlanmasıyla ilgili olarak yazar/ların potansiyel bir çıkar çatışması yoktur.
Finansal Destek:	Bu çalışmanın araştırma ve yazım aşamasında herhangi kişi/kurum veya kuruluşlar tarafından finansal destek alınmadığı bildirilmiştir.
Katkı Oranı Beyanı:	%50-%50
Destek ve Teşekkür Beyanı:	
Sorumlu Yazar:	Mustafa DOLMACI
Çifte Kör Hakem Değerlendirmesi:	Dış-bağımsız

The following statements are made in the framework of "COPE-Code of Conduct and Best Practices Guidelines for Journal Editors":

Ethics Committee Approval:	Ethics committee approval is not required for this article.
Declaration of Conflicting Interests:	No conflicts of interest were reported for this article.
Financial Support:	It has been reported that this study did not receive financial support from any person/institution or organization during the research and writing phase.
Author Contributions:	%50-%50
Statement of Support and Acknowledgment:	
Corresponding Author:	Mustafa DOLMACI
Double-Blind Peer Review:	External-independent