Rastislav Metruk
Trnava University, Trnava, Slovak Republic
University of Ostrava, Ostrava, Czech Republic
ORCID ID 0000-0002-4786-0265
rastislav.metruk@truni.sk
rastislav.metruk@osu.cz

GENERATIVE AI IN EFL SPEAKING AND PRONUNCIATION PRACTICE: A SYSTEMATIC REVIEW OF CHALLENGES AND MITIGATIONS

Abstract. While the role of generative artificial intelligence (GenAI) in English as a foreign language (EFL) writing instruction has received growing research attention, its impact on speaking and pronunciation remains underexplored. This PRISMA-based systematic review examines the challenges associated with using GenAI tools, particularly large language model (LLM)-based chatbots, for EFL speaking and pronunciation practice and synthesizes evidence-based mitigation strategies. Twelve studies that met the eligibility criteria yielded 48 challenges. These were organized into six thematic categories, including technical, pedagogical, learner-related, and equity concerns. The studies also varied in research design and instructional context. Several recurrent patterns emerged across the categories. Output accuracy and reliability represented the most persistent challenge and were linked to concerns about learner trust and uncritical reliance. AI-generated feedback was also repeatedly described as mechanical, formulaic, and lacking contextual sensitivity, highlighting a tension between the accessibility and scalability of these tools and limitations in areas that are important to learners. The studies commonly identified a human-in-the-loop approach as a key mitigation principle, alongside technical improvements and pedagogical interventions, including prompt literacy training. However, important questions remain unresolved, particularly whether GenAI reduces or increases teacher workload and if these tools narrow or widen existing equity gaps. Overall, the findings suggest that effective integration of GenAI into EFL speaking instruction requires attention to technical design, pedagogy, and institutional support, while several aspects of its implementation remain insufficiently understood.

Keywords: generative artificial intelligence (GenAI); English as a foreign language (EFL) speaking and pronunciation practice; Large Language Models (LLMs); challenges and limitations; mitigation strategies

1. INTRODUCTION

Statement of the problem.

The recent surge in generative artificial intelligence (AI), particularly Large Language Models (LLMs) like ChatGPT, has opened new avenues for language learning. These tools can transform pedagogical approaches, enhance learner engagement, and create personalized learning experiences across diverse EFL settings, offering novel opportunities for tailored instruction that address inherent challenges in language acquisition, especially in EFL speaking and pronunciation [1], [2].

Recent research has examined GenAI-supported writing from several perspectives: intervention effects on academic writing [3], GenAI-assisted feedback and revision [4], ChatGPT feedback on EFL writing development [5], and student use and misuse of ChatGPT in academic writing [6]. Research on other language skills is less extensive. Existing work includes a systematic review of AI-powered chatbots for EFL speaking [7], a review of empirical GenAI research across language learning and teaching [8], a review of ChatGPT applications in ESL/EFL education [9], a study of teachers' and students' perceptions of AI in English learning [10], and an AI-driven intervention focused on EFL pronunciation [11].

© Rastislav Metruk, 2026.

Analysis of recent research and publications.

1.1. The Integration of Generative AI in Language Learning

Generative AI tools are increasingly recognized for enhancing language instruction through instant feedback, adaptive learning paths, and interactive environments [12], [13]. Jiang [14] notes that LLMs can significantly enhance students' speaking and writing skills, though they also have shortcomings, such as producing superficial or incorrect feedback.

Researchers have also explored AI's role in developing grammar, vocabulary, reading, and writing [15], in supporting creative production [16], and in shaping critical thinking in EFL education [17], though the question remains whether these tools can genuinely increase EFL learners' language levels.

1.2. GenAI tools in speaking and pronunciation development

AI-powered tools have been developed to provide learners with interactive and adaptive speaking practice. Studies have shown that AI applications, such as voice recognition chatbots, can significantly improve EFL learners' speaking abilities [18].

Fathi et al. [19] found that AI-mediated interactive speaking activities were more effective at enhancing EFL learners' speaking skills and willingness to communicate, with students maintaining favorable attitudes toward such instruction. Similarly, Tai and Chen [20] found that a GAI-powered chatbot significantly enhanced elementary EFL learners' speaking abilities in both individual and paired interactions.

1.3. Pronunciation development and practice

Pronunciation instruction has been an overlooked area in EFL settings, though emerging research shows promise. Choi [21] found that Gen-AI use correlates with better pronunciation, lower speaking anxiety, and higher motivation. Georgiou [22] reported that teachers generally perceive AI as useful for pronunciation instruction, especially those with more comprehensive training. Mompean [23] concluded that ChatGPT might be added to the toolkit of teachers and learners for teaching or learning L2 pronunciation.

1.4. Challenges of GenAI

Using GenAI tools also comes with challenges. Mompean [23] argues that such tools should remain supplementary to human interaction, not replace it, with a blended approach combining chatbot use and human guidance recommended to preserve pragmatic nuance [24]. Achieving this requires guided chatbot use and teacher monitoring of feedback quality [25], structured pedagogical integration of GenAI tools [26], and learner training in effective tool use and prompting [27].

The research goal and the research objectives.

Therefore, this article attempts to identify and classify the challenges associated with using generative AI tools to support L2 speaking and pronunciation practice, and to synthesize the proposed mitigation strategies to address them. It is hoped that examining the challenges and associated mitigation strategies will offer better insight into this underdeveloped yet crucial issue in the field of EFL instruction. Although a number of the challenges discussed below are shared with other language skills, the review keeps its analytical focus on oral practice, interpreting these issues and drawing implications specifically for speaking and pronunciation instruction. Beyond synthesizing the existing evidence, it contributes a set of concrete,

classroom-oriented recommendations that translate the reported mitigation strategies into actionable guidance for EFL teachers and materials developers.

2. RESEARCH METHODS

This systematic literature review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines [28]. The research objectives and research questions were formulated as follows.

2.1. Research objectives

1. Identify and classify the challenges, limitations, and risks reported in studies using generative AI tools to support L2 speaking and/or pronunciation practice (2022–2026).
2. Synthesize the mitigation strategies proposed or discussed to address these challenges, including their categorization, the instructional contexts in which they have been studied, and the methodological approaches employed to investigate them.

2.2. Research Questions

1. What challenges, limitations, and risks are reported when using generative AI tools for L2 speaking and/or pronunciation practice between 2022 and 2026?
2. What mitigation strategies are proposed, recommended, or discussed to address these challenges, and how can these strategies be categorized?
3. What GenAI tools and research designs characterize the existing literature on generative-AI challenges and mitigation strategies for L2 speaking instruction?

Articles were searched in Scopus and Web of Science using parallel search strings structured around five Boolean blocks: (1) generative AI technology terms (e.g., GenAI, LLM, ChatGPT, GPT-4, Claude, Gemini, conversational AI, AI chatbots); (2) EFL/ESL context terms (e.g., EFL, ESL, TESOL, SLA, CALL, second language acquisition); (3) speaking and pronunciation terms (e.g., speaking, oral communication, fluency, pronunciation, intelligibility, ASR); (4) challenge-related terms (e.g., limitations, risks, barriers, hallucinations, bias, ethics, feedback quality); and (5) mitigation-related terms (e.g., strategies, scaffolding, prompt engineering, human-in-the-loop, AI literacy, guidelines). Proximity operators (NEAR/W) were used where appropriate to ensure relevant term co-occurrence.

The inclusion criteria required studies to be: (1) peer-reviewed journal articles, including Early Access and Articles in Press; (2) published between 2022 and 2026; (3) empirical studies, design-based research, classroom interventions, or evaluations, with conceptual papers included only if they provide a clear mitigation framework tied to the speaking/pronunciation GenAI use case; (4) written in English; (5) using a generative AI/LLM-based system as a central element of the learning/teaching activity; (6) situated in EFL/ESL/ELT/TESOL or related SLA/CALL contexts with English as the target language; (7) targeting speaking/oral communication practice and/or pronunciation; and (8) explicitly identifying at least one challenge/limitation/risk/concern related to GenAI use for speaking/pronunciation practice.

The exclusion criteria included: (1) non-journal or non-peer-reviewed publications; (2) systematic, scoping, or narrative reviews and meta-analyses; (3) publications outside the 2022–2026 window or not in English; (4) studies using non-generative AI systems only; (5) studies not situated in English language learning/teaching contexts; (6) studies lacking a clear focus on speaking/oral practice or pronunciation; (7) studies mentioning challenges without specifying mitigations, or vice versa; and (8) purely technical engineering papers without pedagogical linkage.

Searches were conducted in Web of Science and Scopus. In Web of Science, 56 records were retrieved and filtered to 2022–2026; document type (Article and Early Access); WoS categories relevant to education, linguistics, and allied fields; and English language, yielding 35 records. In Scopus, 184 records were retrieved and filtered to 2022–2026; subject areas (Social Sciences, Computer Science, Arts and Humanities, Psychology, and Multidisciplinary); document type (Article); publication stage (Final and Article in Press); and English language, yielding 90 records.

An inclusive subject classification filter was applied in both databases (WoS Categories and Scopus Subject Areas) to exclude records from clearly irrelevant domains (e.g., physics and astronomy, chemical engineering, pharmacology/pharmacy, etc.) while retaining interdisciplinary studies relevant to GenAI-supported language learning. Because individual records can be indexed under multiple categories/subject areas, a broad set of education-, language-, and learning-adjacent classifications was selected to minimize the risk of inadvertently excluding pertinent studies.

After applying the above-mentioned filters, a total of 125 records were identified through database searching (Web of Science = 35; Scopus = 90). After removing 26 duplicates, 99 unique records remained for title/abstract/keyword screening, during which 62 records were excluded and 37 reports were advanced for full-text consideration (28 provisionally included + 9 unclear). Of these 37, one report could not be retrieved, leaving 36 full texts assessed for eligibility; at this stage, 24 reports were excluded, resulting in 12 studies included in the final synthesis. This screening flow is fully consistent with PRISMA 2020 reporting conventions [28]. Title/abstract/keyword screening, full-text eligibility decisions, and data extraction were conducted by the author. AI tools were used only as secondary consistency checks; all inclusion/exclusion decisions and retained data were determined and verified by the author against the original records and full texts. The screening flow is shown in Figure 1.

A final corpus of 12 included studies is methodologically sound and not unusually small given the deliberately high-specificity eligibility criteria (2022–2026 + English-target EFL/ESL + GenAI central + speaking/pronunciation + explicit challenge→mitigation). In tightly scoped reviews, a smaller, well-justified set is expected and preferable to broadening inclusion with tangential studies, because prespecified eligibility criteria are designed to maintain focus, minimize bias, and keep the synthesis aligned to the review questions [29].

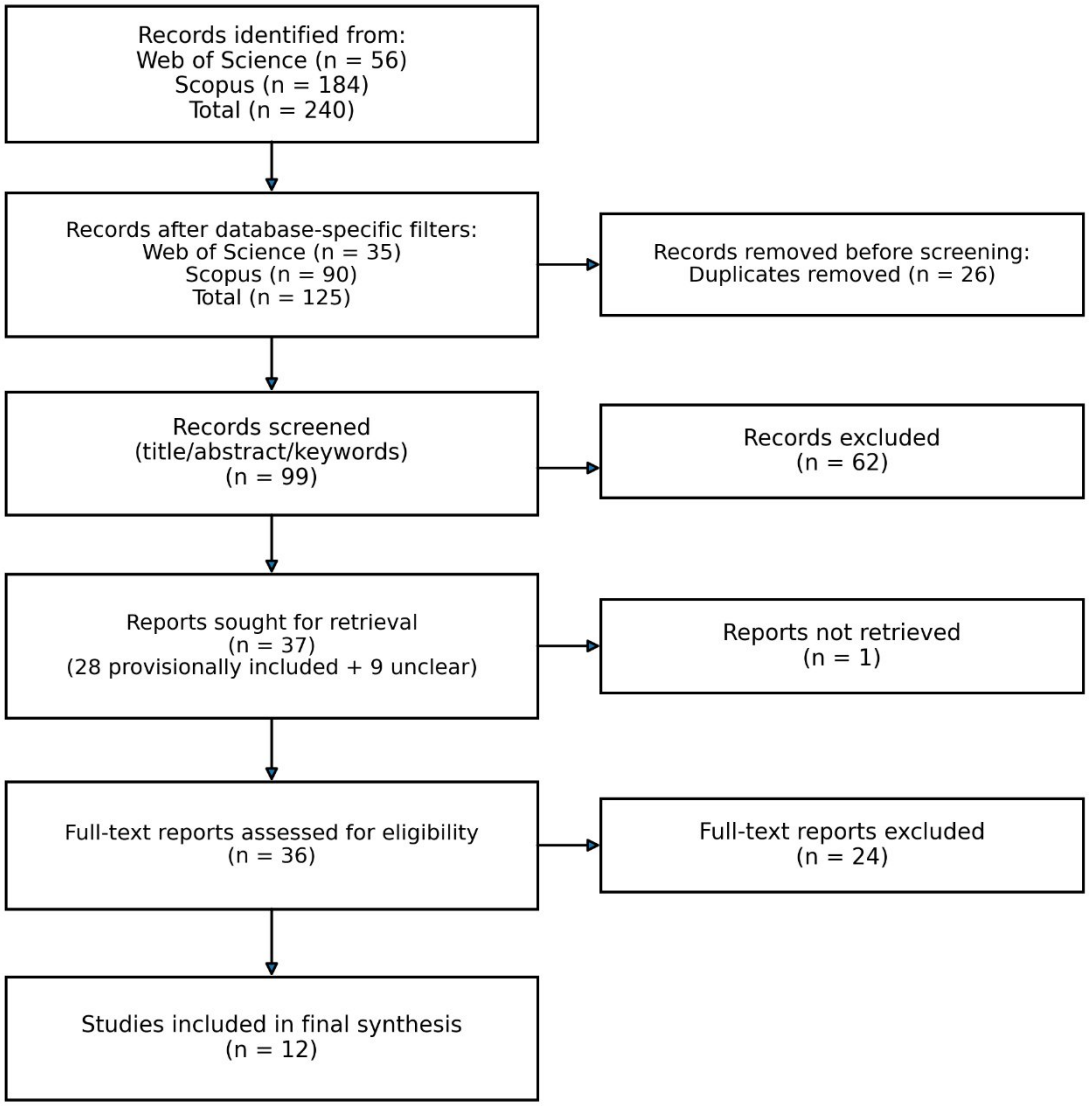

Figure 1. PRISMA Flow Chart

3. FINDINGS

3.1. Challenges, Limitations, and Risks (RQ1)

3.1.1. Overview of Reported Challenges

This section presents the findings for Research Question 1, which inquired about the challenges, limitations, and risks reported when using generative AI tools for L2 speaking and/or pronunciation practice from 2022 to 2026. Among the reviewed studies, a total of 48 challenges were identified, with an average of four challenges per study.

Using thematic synthesis, these challenges were categorized into six groups: technical and infrastructure issues, concerns about output quality and accuracy, problems with feedback and interaction, pedagogical and instructional difficulties, learner-specific factors, and issues of equity and accessibility.

The most frequently mentioned issue was output accuracy and reliability, cited in eight studies. This was followed by technical challenges related to ASR performance and system

reliability, mentioned in six studies, and nearly half of the studies emphasized the risk of over-reliance on AI as a significant pedagogical concern. Overall, these findings indicate that challenges in the technological, pedagogical, and learner-related areas of GenAI-supported speaking practice demand coordinated attention across all these areas.

3.1.2. Technical and Infrastructure Limitations

Technical and infrastructure issues were prevalent across the reviewed studies. Muniandy and Selvanathan [27] observed server hangs, latency, and disruptions in system availability, while Raimkulova et al. [30] highlighted connectivity and device management barriers in EFL primary classrooms.

ASR limitations were often regarded as a major technical issue. Li [31] reported lower G-ASR performance on non-native than native-speaker data and identified classroom noise as an environmental factor requiring adaptive speech-recognition sensitivity. Raimkulova et al. [30] observed that some students spoke too softly for voice recognition and also noted speech-to-text conversion limitations that could lead to inaccurate feedback.

Technical limitations were also evident: Zhang et al. [32] described failures to comprehend intended meanings, limited memory, and difficulties with complex prompts and pronouns, while Mohamed [33] reported limited sensory customization in existing AI tools.

3.1.3. Output Quality and Accuracy Concerns

Output quality and accuracy were raised in eight studies, making this the most frequently cited challenge. Huang et al. [34] reported occasional inaccuracies and doubts about GenAI reliability, and Muniandy and Selvanathan [27] documented inaccurate information, including fabricated or nonexistent references, which learners identified as a major reliability concern. In addition to outright fabrication, Ling and Jan [35] noted that AI outputs occasionally appeared vague, generic, or inaccurate, and occasionally offered misleading suggestions, such as those that were contextually inappropriate idioms.

These limitations extended beyond factual accuracy: Muniandy and Selvanathan [27] documented grammar mistakes, repetitive responses, and outdated content; Darmawansah et al. [36] identified prompt sensitivity as a significant constraint; and Nguyen [37] observed limited linguistic realizations for some communication strategies and minimal variation in strategy range across prompts. Taken together, these limitations reinforce the need for human oversight and critical evaluation of AI output.

3.1.4. Feedback and Interaction Quality Issues

Multiple studies identified feedback quality limitations. Huang et al. [34] noted learners described AI feedback as mechanical or robotic, while Liu et al. [38] found non-authentic multimodal cues reduced perceived naturalness in human–chatbot interactions.

Zhang et al. [32] reported that even previously successful prompts could sometimes elicit irrelevant chatbot feedback, while Liu et al. [38] described formulaic, template-like feedback and unidirectional interaction lacking reciprocity. Li [31] incorporated teacher-based human validation of feedback quality and found balanced AI–teacher involvement most effective, reinforcing the value of continued human involvement.

Interaction flow and multi-turn continuity remain persistent challenges. Liu et al. [38] identified response delays and learner formulation difficulties that were exacerbated by limited chatbot scaffolding, while Zhang et al. [32] and Darmawansah et al. [36] highlighted prompt sensitivity – with Zhang et al. [32] specifically noting brittle handling of complex prompts and pronouns – indicating that chatbot effectiveness depends heavily on prompt quality. These issues are compounded by limitations in memory, context handling, and personalization.

3.1.5. Pedagogical and Instructional Challenges

Several pedagogical and instructional challenges were identified, with over-reliance on AI identified in five studies. Almineeai et al. [39], drawing on prior literature, likewise cautioned against overreliance on ChatGPT because of its potential to hinder social communication skills. Huang et al. [34] observed student concerns about impeded independent thinking, confirmed by Ling and Jan [35] who documented over-reliance on AI outputs. Li [31] found AI-dominant feedback less effective than balanced AI–teacher feedback, with motivation decreasing over time in the AI-dominant condition, while Juan et al. [40] highlighted the need for further research on pronunciation and pragmatic development and suggested combining ChatGPT with complementary instructional strategies for more comprehensive oral language support.

Teacher roles and workload were also affected. Ling and Jan [35] described how instructional unpredictability increased teachers' cognitive load and stress. Nguyen [37] emphasized preserving teacher agency in AI integration, and Mohamed [33] expressed concerns about teacher preparedness and lack of resources for effective AI integration. Without a doubt, this area needs further exploration, as teacher-related issues remain one of the key concerns in using GenAI tools in language education.

Limitations in pronunciation and suprasegmental instruction were also specifically highlighted. Juan et al. [40] described pronunciation gains as comparatively more modest in their discussion and linked remaining limitations to suprasegmental features such as intonation and stress, and Mohamed [33] reported limited support for social-pragmatic skills. Raimkulova et al. [30] expressed concerns over assessment complexities in ChatGPT-mediated voice-interaction games.

Finally, Darmawansah et al. [36] raised ethical issues of originality and academic integrity in AI-assisted student work, a topic that merits further investigation.

3.1.6. Learner-Related Factors

Several studies highlighted learner-related challenges. Huang et al. [34] noted student comprehension and tool-navigation difficulties, while Muniandy and Selvanathan [27] observed that creating accurate prompts required significant extra effort. Juan et al. [40] found that the progress-tracking feature received the lowest ease-of-use score, indicating persistent usability difficulties.

Regarding cognitive and affective engagement, Huang et al. [34] reported that while affective and behavioral engagement increased with AI use, cognitive engagement did not. Almineeai et al. [39] documented substantial pre-intervention speaking anxiety, low motivation, embarrassment, and reluctance to participate; after ChatGPT-supported practice, participants reported reduced anxiety and greater motivation, confidence, and willingness to speak.

Response clarity also seems to influence learner experiences. In this connection, Muniandy and Selvanathan [27] noted that unclear and vague AI answers added obstacles to effective learning.

3.1.7. Equity and Accessibility Concerns

Equity and accessibility emerged as important challenges. Darmawansah et al. [36] noted that differences in learners' language skills, technology access, and familiarity may affect learning outcomes, raising equity concerns. Nguyen [37], drawing on prior communication-strategy research, noted that learners at different proficiency levels tend to use strategies differently, supporting proficiency-sensitive tailoring of AI-generated materials.

Mohamed [33] focused on autistic ESL learners and identified challenges including the lack of AI tools specifically designed for autistic learners and limited sensory customization in

existing applications. These findings underscore the need for more inclusive, flexible, and adaptable AI tools tailored to the communicative and sensory needs of autistic L2 learners.

3.1.8. Summary

The synthesis across the 12 reviewed studies shows that challenges are not evenly distributed – technical limitations and output quality issues dominate the literature, reported in six and eight studies, respectively. Yet these categories are closely intertwined: ASR limitations, for instance, directly feed into the feedback quality concerns documented across multiple studies, while output inaccuracy compounds the over-reliance risk noted in five studies, as learners who cannot reliably judge AI output may default to uncritical acceptance. Equity and accessibility, though less frequently discussed, may amplify all other challenge categories for underserved learner populations. These interconnections suggest that addressing challenges in isolation is unlikely to be sufficient – a point examined in greater detail in the following Discussion section.

3.2. Mitigation Strategies

All 12 studies proposed, recommended, or discussed mitigation strategies, varying from classroom-level pedagogical adjustments to broader institutional and technical recommendations, with most proposing between three and five distinct approaches.

3.2.1. Thematic Categories of Mitigations

Using thematic synthesis, the mitigation strategies were classified into six categories: Pedagogical; Technical; Assessment & Usability; Governance; Process; and Data. These categories were not mutually exclusive; most studies reported strategies covering multiple categories.

Pedagogical. All 12 studies incorporated pedagogical strategies such as instructional scaffolding, feedback design, learner training, and combining human instruction with AI-mediated practice. Common approaches involved using GenAI as a supplementary tool alongside peer and teacher interactions [34], creating tasks that require critical evaluation and verification of AI suggestions [34], [27], and providing clear prompt guidance through scaffolded prompt structures [27], [32], [36].

Raimkulova et al. [30] recommended structured pre-activity technology checks. Nguyen [37] recommended that learners reflect on their own strategic performance and compare it with that of proficient speakers, with AI potentially providing additional automated feedback. Mohamed [33] described multisensory social stories for autistic learners, and Almineeai et al. [39] described personalized, low-pressure AI-mediated practice as beneficial for reducing speaking anxiety.

Governance, Process, and Data. Governance and process strategies are discussed in ten studies, while data-related strategies – such as cross-checking procedures, validation of AI outputs, and the use of reliable data sources – are covered in eight of them. These strategies encompass institutional oversight, stakeholder collaboration, ethical considerations, and implementation procedures.

Recommendations support human-in-the-loop approaches combining AI with educator feedback [34], [40], [38], the use of more sophisticated, adaptive AI tools [32], and AI–teacher collaboration with teachers as curators and validators [37].

The studies also highlight monitoring and clear guidelines for ethical AI use [36], redefining the teacher's role as mediator [35], establishing institutional governance [31], adding validation steps [27], creating adaptable AI systems [33], and encouraging collaborative learning with both humans and chatbots [38].

Technical. Seven studies addressed technical mitigations. Muniandy and Selvanathan [27] proposed enhancing platform infrastructure; Li [31] recommended noise-resistant processing for ASR accuracy; and Zhang et al. [32] called for more sophisticated systems capable of adapting and learning from learner interactions.

Additional technical measures included sensory-adjustable interfaces [33], simplified dashboards [31], portable hotspots [30], and taxonomy-informed prompts to elicit specified communication strategies [37].

Assessment & Usability. Six studies proposed assessment and usability mitigations, including targeted support for progress-tracking functions [40], developing assessment methods for GenAI contexts [30], sensory-adjustable interfaces [33], promoting equitable access [36], standardized rubric templates [31], and proficiency-based AI prompt customization [37].

3.2.2. Evidence and Trends

Detailed frequency and co-occurrence analyses of individual mitigation strategies were not performed because multiple strategies could be coded within the same study. The study-level category counts reported above therefore indicate broad thematic coverage rather than frequencies of individual mitigation techniques.

Most studies proposed strategies spanning multiple categories, with pedagogical strategies consistently highlighted alongside governance, technical, and assessment measures. This strongly suggests that successfully integrating GenAI into educational environments requires coordinated, multidimensional efforts.

3.2.3. Summary of mitigations

The mitigation strategies identified across the 12 studies show a clear pattern: no single category of intervention is treated as sufficient on its own. Pedagogical strategies appeared universally but were almost always paired with technical, governance, or assessment measures – suggesting that authors recognize the interconnected nature of the challenges outlined in the previous section. Notably, while the literature is rich in proposed mitigations, it remains difficult to determine which specific strategies are most effective, under what conditions, and for which learner populations.

Future research would therefore benefit from comparative studies that directly examine the efficacy of different mitigation approaches, enabling practitioners and curriculum designers to make evidence-informed decisions when integrating GenAI into L2 speaking instruction. These implications are explored further in the Discussion section.

3.3. AI tools used and research design involved

Among the studies reviewed, LLM-based chatbots, especially ChatGPT, were the most frequently used generative-AI tools for L2 speaking instruction. Across the corpus, GenAI tools were examined in higher-education, teacher-education, primary-school, and special-education/ESL contexts. ChatGPT functioned either as an AI conversation partner or as instructional support for oral-language tasks. Some reports explicitly mentioned using ChatGPT-4 or ChatGPT-4o mini, while others simply referred to ChatGPT without specifying the version. Overall, these chatbots supported oral-language learning through direct learner interaction or through the generation of instructional materials for speaking-related tasks.

Additional studies utilized avatar-based conversational agents (e.g., Call Annie), ASR-driven speaking systems (e.g., G-ASR), and commercial GenAI chatbots such as iFLYTEK Spark; Zhang et al. [32] also included FiF Speaking as an intelligent tutoring system comparator.

Research designs were heterogeneous. The corpus included quasi-experimental and mixed-methods intervention studies, some with pre- and post-test comparisons, qualitative

studies based on interviews and classroom observations, and AI-output analyses without human participants. Outcomes and focal variables included speaking performance, fluency, pronunciation, anxiety, willingness to communicate, perceptions, and implementation experiences.

3.4. Discussion

This article aimed to identify and categorize the challenges of using generative AI tools for L2 speaking and pronunciation practice and to synthesize mitigation strategies. Based on 12 studies meeting high-specificity eligibility criteria, the findings are discussed below.

The 48 challenges, organized into six categories (technical/infrastructure, output quality/accuracy, feedback/interaction quality, pedagogical/instructional, learner-related, and equity/accessibility), yielded a broad range of findings.

Accuracy and reliability were the most commonly mentioned disadvantages. Errors, inaccuracies, and misleading suggestions were associated with concerns about learner trust and reliability. Trust itself was not systematically measured across the included studies; however, Song [41] suggests that trust may be closely linked to successful adoption.

Outdated content and inconsistent output quality further compounded trust concerns and co-occurred with warnings about over-reliance; Nguyen [37] additionally identified limited linguistic realizations and restricted variation in some communication-strategy patterns. While generative AI can positively influence learning, its ethical implications and risk of over-dependence require careful consideration [42]. Human-in-the-loop remains a key notion, as human guidance appears essential for effective GenAI use in language development [35], aligning with the present review's mitigation findings that emphasize critical evaluation of AI outputs alongside peer, teacher, and textbook-based interaction.

Technical and infrastructure limitations – including network connectivity issues, classroom noise, chatbot comprehension and memory limitations, and limited sensory customization – negatively influence the learning process. Both instructors and developers need to address these challenges, with mitigations including portable hotspots, enhanced platform infrastructure, and improved ASR accuracy in noisy environments to increase output reliability and facilitate personalized learning, including for autistic learners.

Mechanical, unclear, or template-like feedback poses a serious threat to GenAI use in language education, potentially demotivating and frustrating learners. Shahini [43] recommends that developers move urgently from simple error correction to emotionally responsive designs with adaptive, supportive language. Human-in-the-loop validation remains an imperative complement, as learners value both AI convenience and the contextual understanding and emotional support provided by teachers [44]. While AI can offer more balanced feedback, it often remains impersonal and surface-level; human feedback delivers deeper insights and more nuanced critiques [45]. Incorporating collaborative learning modes with both human–human and human–chatbot interactions may also help address AI feedback limitations.

Teacher preparedness was another key challenge. Instructional unpredictability raised teachers' cognitive load and stress. This somewhat contradicts Lee and Moore [46], who suggest GenAI can reduce workload through automated grading and feedback. Achieving this ideal state requires AI training programs, improved infrastructure, ready-made materials, and ethical standards to promote responsible AI use in EFL practices [47]. Such a solution involves learners, teachers, and developers.

Prompting issues affect both learners and instructors, as prompt literacy directly influences the relevance, accuracy, and creativity of AI-generated results [48].

Affective barriers also matter: learners may enter GenAI speaking practice with anxiety, low motivation, embarrassment, or hesitation; in Almineeai et al. [39], these barriers decreased after ChatGPT-supported practice. Providing proper GenAI training for both learners and teachers, including proficiency-based prompt customization, can help address these concerns.

Regarding equity and accessibility, students' varying language skills, technology access, and familiarity with technology may contribute to unequal learning outcomes [36]. AI tools may reinforce existing inequalities, particularly where technology access and digital literacy are uneven [36], [49]. Nguyen [37] further supports proficiency-sensitive tailoring of AI-generated materials. It remains unclear whether generative AI tools narrow or widen existing equity gaps in language education. Autistic learners also require AI tools tailored to their sensory and communicative needs [33].

In practice, several concrete activities can help EFL teachers integrate generative AI into speaking and pronunciation work while containing the risks identified above. For controlled pronunciation practice, chatbot- and ASR-based tools can support shadowing and minimal-pair drills in which learners imitate model utterances and receive immediate feedback, with the teacher confirming or correcting the tool's judgments given the lower G-ASR performance reported for non-native-speaker data [31]. Structured pre-activity technology checks help ensure that connectivity or device problems do not disrupt these tasks [30]. For fluency development, low-stakes scripted role-plays with an AI interlocutor can precede freer, unscripted exchanges, letting learners rehearse before performing and reducing pre-existing speaking anxiety. Prompt literacy can be developed through scaffolded, proficiency-sensitive prompt design that makes the task, context, and desired output explicit [37], [48]. AI feedback should be treated as a first draft to be validated. Pairing human-human with human-chatbot interaction and having learners compare AI suggestions against textbook or teacher input keeps human judgment central to the activity [34], [38], [44], [26].

4. CONCLUSIONS AND PROSPECTS FOR FURTHER RESEARCH

This PRISMA-based systematic review examined the challenges of using generative AI tools for L2 speaking and pronunciation practice (2022–2026) and synthesized proposed mitigation strategies. Analyzing 12 studies meeting strict eligibility criteria, the review provides an evidence-based overview of a rapidly expanding yet fragmented field. The small number of studies suggests that rigorous empirical research in this domain is still in its early phase.

Several common patterns appeared across the six challenge categories. First, accuracy and reliability stand out as the primary weaknesses of generative AI in oral language practice, not only in isolated cases but as a persistent issue across various tool types, study settings, and learner populations. Second, concerns about learner trust were associated with perceived inaccuracies and inconsistent AI output, although trust was not systematically measured across the included studies. Third, feedback produced by AI was often described as mechanical, superficial, and lacking the contextual awareness that learners associate with meaningful teaching. These patterns indicate that although generative AI tools are valued for their accessibility and scalability, they frequently fall short in essential areas, such as accuracy, trustworthiness, and depth of feedback, which are vital for making technology-supported practice truly effective in enhancing authentic language development.

The review also identified areas of genuine debate and uncertainty. The impact of generative AI on teacher workload is still unresolved: while some evidence suggests these tools could lighten instructional tasks, the results here imply that increased instructional unpredictability might actually heighten rather than reduce teacher stress. It is still unclear whether this reflects a temporary transition or a deeper structural issue. Likewise, whether generative AI tools help close or widen equity gaps remains uncertain. The studies reviewed

pointed out the risk of reinforcing existing inequalities, but comparative evidence on whether GenAI narrows or widens broader equity gaps remains limited.

Central to this review is the essential role of human agency in AI-supported language learning. The human-in-the-loop approach emerged as a recurrent and prominent theme across the literature, consistently positioning teachers or peers as validators, mediators, or facilitators working alongside AI. In addition to this core idea, the review identified three broad strands within the wider mitigation framework: technical solutions (such as improving ASR accuracy, platform stability, and infrastructure), pedagogical interventions (including structured AI and prompt literacy training, and collaborative learning methods), and design-level changes (shifting from error correction to emotionally responsive, learner-sensitive AI feedback). These strategies are interconnected; their success probably relies on coordinated implementation, although evidence for such integration is limited. In practical terms, this means treating generative AI as a supplementary, teacher-mediated resource for oral practice: valuable for low-stakes rehearsal, controlled pronunciation work, and additional speaking exposure, with teacher validation remaining important for evaluating its feedback and accuracy.

Prospects for further research

Research should now examine the specific conditions under which generative AI can support L2 speaking practice effectively, appropriately, and fairly. Three key areas deserve attention. First, future research needs to create and validate comprehensive mitigation frameworks that combine technical, pedagogical, and design interventions, instead of testing isolated solutions. Second, longitudinal studies are essential to determine whether the issues observed – such as trust concerns, inadequate feedback, and over-reliance – are temporary challenges or long-term structural limitations. Finally, research on equity should go beyond just identifying risks, aiming instead to gather comparative data on how these tools affect learners differently depending on their proficiency, socioeconomic background, and educational needs. Beyond these directions, and consistent with the review's focus on oral skills, a further priority is fine-grained research on pronunciation itself: future studies should isolate segmental and suprasegmental outcomes using validated speaking and pronunciation assessments, because the corpus provided limited fine-grained evidence on these features.

This review has limitations. It was limited to 12 studies, reflecting the field’s early stage. The rapid pace of AI development means some documented challenges may be partially addressed by newer iterations, while new ones may emerge. Variability in study designs, tools, and learner groups restricts generalizability. Additionally, several mitigation strategies were analytically derived rather than empirically tested, meaning their effectiveness remains an open question. Screening and extraction were conducted by a single reviewer; although AI-based secondary checks supported consistency, the process did not include independent human verification. No formal risk-of-bias instrument or certainty-of-evidence framework was applied, given the heterogeneity of the included study designs and the descriptive aim of the review. The synthesis therefore maps the challenges and mitigation strategies reported in the literature rather than grading their comparative effectiveness.

In conclusion, generative AI shows promise for L2 speaking and pronunciation practice, but realizing this potential depends on key conditions. The challenges extend beyond technical issues to pedagogical approaches, relationships, and ethics. For optimal results, generative AI in language learning should always prioritize human judgment, empathy, and expertise over automation. The focus should be on creating conditions that ensure responsible, fair, and truly effective integration of generative AI in language learning. Beyond mapping these challenges, the review offers EFL practitioners concrete, evidence-informed guidance for using generative AI responsibly and effectively to support speaking fluency and pronunciation.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

Claude Opus 4.6 [50] was used to edit the language and style of the manuscript and to assist with citation formatting, and GPT-5.2 [51] assisted in generating Figure 1 (PRISMA flow diagram). Both tools also served only as secondary consistency checks, flagging potentially ambiguous, inconsistent, or provisional items for further review. Study selection, data extraction, and all inclusion, exclusion, and coding decisions were carried out and verified by the author against the original records and full texts; AI outputs were not treated as evidentiary sources. The author manually checked each citation for accuracy, reviewed and verified the final figure, and takes full responsibility for the content of the publication.

ACKNOWLEDGEMENT

The study is from the project „Implementation of the Reform of Teacher Education at the University of Ostrava: Competence-based learning founded on practice, research, internationalisation and cooperation with the clinical school CZ.02.02.XX/00/23_019/0008272“ which is co-financed by the European Union.

CONFLICT OF INTEREST

The author declares that there is no conflict of interest.

REFERENCES (TRANSLATED AND TRANSLITERATED)

[1] N. Abbas, I. Ali, R. Manzoor, T. Hussain, and M. H. AL Hussaini, "Role of artificial intelligence tools in enhancing students' educational performance at higher levels," *J. Artif. Intell., Mach. Learn. Neural Netw.*, vol. 3, no. 5, pp. 36-49, 2023, doi: 10.55529/jaimlnn.35.36.49. (in English).

[2] H. Alisoy, "The impact of AI-powered personalized learning on ESL student outcomes," *EuroGlobal J. Linguistics Lang. Educ.*, vol. 2, no. 2, pp. 89-98, 2025, doi: 10.69760/egjlle.2500187. (in English).

[3] S. Mahapatra, "Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study," *Smart Learn. Environ.*, vol. 11, 2024, Art. no. 9, doi: 10.1186/s40561-024-00295-9. (in English).

[4] M. Mekheimer, "Generative AI-assisted feedback and EFL writing: A study on proficiency, revision frequency and writing quality," *Discover Educ.*, vol. 4, 2025, Art. no. 170, doi: 10.1007/s44217-025-00602-7. (in English).

[5] P. Polakova and P. Ivenz, "The impact of ChatGPT feedback on the development of EFL students' writing skills," *Cogent Educ.*, vol. 11, no. 1, 2024, Art. no. 2410101, doi: 10.1080/2331186X.2024.2410101. (in English).

[6] A. Jankovic and D. Kulic, "Use and misuse of ChatGPT in academic writing among the English language students," *Inf. Technol. Learn. Tools*, vol. 105, no. 1, pp. 178-188, 2025, doi: 10.33407/itlt.v105i1.5955. (in English).

[7] J. Du and B. K. Daniel, "Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice," *Comput. Educ. Artif. Intell.*, vol. 6, 2024, Art. no. 100230, doi: 10.1016/j.caeai.2024.100230. (in English).

[8] B. Li, Y. L. Tan, C. Wang, and V. Lowell, "Two years of innovation: A systematic review of empirical generative AI research in language learning and teaching," *Comput. Educ. Artif. Intell.*, vol. 9, 2025, Art. no. 100445, doi: 10.1016/j.caeai.2025.100445. (in English).

[9] C. K. Lo, P. L. H. Yu, S. Xu, D. T. K. Ng, and M. S. Jong, "Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies," *Smart Learn. Environ.*, vol. 11, 2024, Art. no. 50, doi: 10.1186/s40561-024-00342-5. (in English).

[10] A. Alenezi and A. Alenezi, "The use of artificial intelligence in English learning: Teachers' and students' perceptions," *Inf. Technol. Learn. Tools*, vol. 108, no. 4, pp. 37-52, 2025, doi: 10.33407/itlt.v108i4.6127. (in English).
[11] I. Xodabande, S. Shiri, and M. Zohrabi, "Exploring the impacts of an AI-driven instructional intervention on Iranian EFL learners' pronunciation skill development," *Discover Educ.*, vol. 4, 2025, Art. no. 307, doi: 10.1007/s44217-025-00782-2. (in English).
[12] FX. Baskara, "Navigating the complexities and potentials of language learning machines in EFL contexts: A multidimensional analysis," in *Proc. 4th Int. Conf. Language, Literary, and Cultural Studies (ICON LATERALS)*, 2023, doi: 10.4108/eai.11-7-2023.2340636. (in English).
[13] M. B. Ulla and M. F. Teng, "Generative artificial intelligence (AI) applications in TESOL: Opportunities, issues, and perspectives [Editorial]," *Int. J. TESOL Stud.*, vol. 6, no. 3, pp. 1-3, 2024, doi: 10.58304/ijts.20240301. (in English).
[14] H. Jiang, "Applications and research gaps of LLM-based English-as-a-foreign-language education," *J. Educ. Humanit. Social Sci.*, vol. 39, pp. 429-434, 2024, doi: 10.54097/jkhv4m38. (in English).
[15] L. Wei, "Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning," *Front. Psychol.*, vol. 14, 2023, Art. no. 1261955, doi: 10.3389/fpsyg.2023.1261955. (in English).
[16] A. R. Doshi and O. P. Hauser, "Generative AI enhances individual creativity but reduces the collective diversity of novel content," *Sci. Adv.*, vol. 10, no. 28, 2024, doi: 10.1126/sciadv.adn5290. (in English).
[17] J. Liu, A. J. B. Sihes, and Y. Lu, "How do generative artificial intelligence (AI) tools and large language models (LLMs) influence language learners' critical thinking in EFL education? A systematic review," *Smart Learn. Environ.*, vol. 12, 2025, Art. no. 48, doi: 10.1186/s40561-025-00406-0. (in English).
[18] T. Duong and S. Suppasetseree, "The effects of an artificial intelligence voice chatbot on improving Vietnamese undergraduate students' English speaking skills," *Int. J. Learn., Teaching Educ. Res.*, vol. 23, no. 3, pp. 293-321, 2024, doi: 10.26803/ijlter.23.3.15. (in English).
[19] J. Fathi, M. Rahimi, and A. Derakhshan, "Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions," *System*, vol. 121, 2024, Art. no. 103254, doi: 10.1016/j.system.2024.103254. (in English).
[20] T.-Y. Tai and H. H.-J. Chen, "Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions," *Comput. Educ.*, vol. 220, 2024, Art. no. 105112, doi: 10.1016/j.compedu.2024.105112. (in English).
[21] W. Choi, "The role of changes in pronunciation ability and anxiety through Gen-AI on EFL learners' self-directed speaking motivation and social interaction confidence: A CHAT perspective," *System*, vol. 133, 2025, Art. no. 103788, doi: 10.1016/j.system.2025.103788. (in English).
[22] G. P. Georgiou, "Artificial intelligence in pronunciation teaching: Use and beliefs of foreign language teachers," *J. Educ. Teach.*, vol. 52, no. 4, pp. 821-840, 2026, doi: 10.1080/02607476.2026.2676279. (in English).
[23] J. A. Mompean, "ChatGPT for L2 pronunciation teaching and learning," *ELT J.*, vol. 78, no. 4, pp. 423-434, 2024, doi: 10.1093/elt/ccae050. (in English).
[24] A. Matiienko-Silnytska, N. Mikava, I. Savranchuk, N. Tkhor, and H. Poliakova, "Conversational analysis of learner-AI chatbot interactions in developing spoken fluency," *Arab World English J.*, vol. 16, no. 4, pp. 224-238, 2025, doi: 10.24093/awej/vol16no4.13. (in English).
[25] J. Taeza, "The role of AI-powered chatbots in enhancing second language acquisition: An empirical investigation of conversational AI assistants," *Edelweiss Appl. Sci. Technol.*, vol. 9, no. 3, pp. 2616-2629, 2025, doi: 10.55214/25768484.v9i3.5853. (in English).
[26] K. L. R. Chan and J. Liu, "Using multiple GenAI tools in pronunciation teaching: An AI-assisted teaching flowchart in blended learning and flipped learning," *J. Educators Online*, vol. 22, no. 3, 2025, doi: 10.9743/JEO.2025.22.3.6. (in English).
[27] J. Muniandy and M. Selvanathan, "ChatGPT, a partnering tool to improve ESL learners' speaking skills: Case study in a public university, Malaysia," *Teach. Public Admin.*, vol. 43, no. 1, pp. 4-20, 2025, doi: 10.1177/01447394241230152. (in English).
[28] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, 2021, Art. no. n71, doi: 10.1136/bmj.n71. (in English).
[29] J. P. T. Higgins et al., Eds., *Cochrane handbook for systematic reviews of interventions, version 6.5, updated Aug. 2024*. Cochrane, 2024. [Online]. Available: https://www.cochrane.org/authors/handbooks-and-manuals/handbook (in English).

[30] A. Raimkulova, K. Ybyraimzhanov, M. Halmatov, G. Mailybayeva, and Y. Khaimuldanov, "Pre-service EFL primary teachers adopting GenAI-powered game-based instruction: A practicum intervention," *Educ. Sci.*, vol. 15, no. 10, 2025, Art. no. 1326, doi: 10.3390/educsci15101326. (in English).
[31] Z. Li, "Bridging pedagogy and technology: A generative AI and IoT approach to transformative English language education," *Humanit. Social Sci. Commun.*, vol. 12, 2025, Art. no. 1879, doi: 10.1057/s41599-025-06151-6. (in English).
[32] D. Zhang, J. G. Wu, and Z. Fu, "From shy to fly: Facilitating EFL learners' willingness to communicate with an AI chatbot and an intelligent tutoring system," *System*, vol. 127, 2024, Art. no. 103501, doi: 10.1016/j.system.2024.103501. (in English).
[33] A. M. Mohamed, "Using AI to teach English to multilingual autistic children: A qualitative study," *Int. J. Inclusive Educ.*, vol. 30, no. 9, pp. 2469-2489, 2026, doi: 10.1080/13603116.2025.2564336. (in English).
[34] J. Huang, A. Mizumoto, and Y. L. Huang, "Comparing peer feedback and generative artificial intelligence feedback in Japanese English as a foreign language speaking context: Impacts on motivation, engagement, and writing self-efficacy," *Int. J. Appl. Linguistics*, vol. 36, no. 2, pp. 1696-1707, 2026, doi: 10.1111/ijal.70020. (in English).
[35] Y. Ling and J. M. Jan, "Voices from the flip: Teacher perspectives on integrating AI chatbots in flipped English classrooms," *Educ. Sci.*, vol. 15, no. 9, 2025, Art. no. 1219, doi: 10.3390/educsci15091219. (in English).
[36] D. Darmawansah, D. Rachman, F. Febiyani, and G.-J. Hwang, "ChatGPT-supported collaborative argumentation: Integrating collaboration script and argument mapping to enhance EFL students' argumentation skills," *Educ. Inf. Technol.*, vol. 30, pp. 3803-3827, 2025, doi: 10.1007/s10639-024-12986-4. (in English).
[37] P.-A. Nguyen, "Evaluating AI-generated language as models for strategic competence in English language teaching," *IAFOR J. Educ.: Stud. Educ.*, vol. 12, no. 3, pp. 325-349, 2024, doi: 10.22492/ije.12.3.13. (in English).
[38] H. Liu, C. Lv, and J. Chen, "Willingness to communicate and anxiety in human-human and human-chatbot interaction contexts: An idiodynamic investigation," *Innov. Lang. Learn. Teaching*, to be published, doi: 10.1080/17501229.2025.2560101. (in English).
[39] A. M. Almineeai, M. A. Alrashed, T. A. Hezam, A. F. Alharthi, and S. F. Alharthi, "The impact of ChatGPT on EFL medical track students' affective filter," *J. Lang. Teaching Res.*, vol. 16, no. 1, pp. 158-167, 2025, doi: 10.17507/jltr.1601.17. (in English).
[40] W. Juan, H. H. Binti Ismail, and A. Z. Mansor, "Enhancing Chinese EFL learners' speaking proficiency through AI-integrated POA," *Educ. Sci.: Theory Pract.*, vol. 25, no. 2, pp. 44-57, 2025, doi: 10.12738/jestp.2024.2.018. (in English).
[41] D. Song, "How learners' trust changes in generative AI over a semester of undergraduate courses," *Int. J. Artif. Intell. Educ.*, vol. 35, pp. 1445-1464, 2025, doi: 10.1007/s40593-024-00446-6. (in English).
[42] I. P. Y. Laksana, P. D. Hudiananingsih, K. Suciani, P. T. Virginiya, and L. G. E. Wahyuni, "Exploring the possible use of generative artificial intelligence in supporting students' speaking performance," *Prosiding Seminar Nasional Riset Bahasa dan Pengajaran Bahasa*, vol. 6, no. 1, pp. 208-218, 2024. [Online]. Available: https://ojs2.pnb.ac.id/index.php/SENARILIP/article/view/2219 (in English).
[43] A. Shahini, "Emotional dimensions of feedback: How AI and human responses shape ESL learning outcomes," *Ampersand*, vol. 15, 2025, Art. no. 100235, doi: 10.1016/j.amper.2025.100235. (in English).
[44] M. Huang and F. Hu, "From comparison to integration: A hybrid feedback model based on student perceptions of AI and teacher support in EFL speaking," *Accents Asia*, vol. 20, no. 2, pp. 12-35, 2025. [Online]. Available: https://accentsasia.org/issues/20-2/Huang.pdf (in English).
[45] K. Hyland, "EFL teachers and feedback fatigue: AI to the rescue?" *Lang. Teaching*, vol. 59, no. 3, pp. 310-326, 2026, doi: 10.1017/S0261444825101018. (in English).
[46] S. S. Lee and R. L. Moore, "Harnessing generative AI (GenAI) for automated feedback in higher education: A systematic review," *Online Learn.*, vol. 28, no. 3, pp. 82-104, 2024, doi: 10.24059/olj.v28i3.4593. (in English).
[47] A. Ilma and Z. Rohmah, "AI in EFL education: Teachers' competence and the roadblocks to teaching material development," *Cogent Educ.*, vol. 12, no. 1, 2025, Art. no. 2588471, doi: 10.1080/2331186X.2025.2588471. (in English).

[48] E. Tour and A. Zadorozhnyy, "Conceptualizing and operationalizing prompt literacy for English language learners," *J. Adolesc. Adult Literacy*, vol. 69, no. 3, 2025, Art. no. e70020, doi: 10.1002/jaal.70020. (in English).
[49] L. Vesna, P. S. Sawale, P. Kaul, S. Pal, and S. N. V. R. M. Banda, "Digital divide in AI-powered education: Challenges and solutions for equitable learning," *J. Inf. Syst. Eng. Manage.*, vol. 10, no. 21s, pp. 300–308, 2025, doi: 10.52783/jisem.v10i21s.3327. (in English).
[50] Anthropic, "Claude Opus 4.6," 2026. [Online]. Available: https://www.anthropic.com/ (in English).
[51] OpenAI, "GPT-5.2," 2026. [Online]. Available: https://openai.com/ (in English).

Text of the article was accepted by Editorial Team 26.03.2026

ГЕНЕРАТИВНИЙ ШТУЧНИЙ ІНТЕЛЕКТ У ПРАКТИЦІ НАВЧАННЯ СТУДЕНТІВ РОЗМОВНОЇ МОВИ ТА ВИМОВИ АНГЛІЙСЬКОЇ ЯК ІНОЗЕМНОЇ: СИСТЕМАТИЧНИЙ ОГЛЯД ПРОБЛЕМ ТА ЗАХОДІВ З ЇХНЬОГО ПОДОЛАННЯ

Растіслав Метрук
Трнавський університет, м. Трнава, Словацька Республіка
Остравський університет, м. Острава, Чеська Республіка
ORCID ID 0000-0002-4786-0265
rastislav.metruk@truni.sk
rastislav.metruk@osu.cz

Анотація. Хоча роль генеративного штучного інтелекту (GenAI) у навчанні письма англійської мови як іноземної (EFL) привертає дедалі більшу увагу дослідників, його вплив на мовлення та вимову залишається недостатньо вивченим. Цей систематичний огляд, проведений відповідно до принципів PRISMA, аналізує проблеми, пов'язані з використанням інструментів GenAI, зокрема чат-ботів на основі великих мовних моделей (LLM), для відпрацювання мовлення та вимови англійської як іноземної, а також узагальнює науково обґрунтовані стратегії їхнього подолання. У дванадцяти дослідженнях, що відповідали критеріям відбору, було виявлено 48 проблем. Їх було розділено на шість тематичних категорій, зокрема технічні, педагогічні, пов'язані з учнями та питання рівності. Дослідження також відрізнялись за дизайном та навчальним контекстом. У різних категоріях простежувались кілька повторюваних закономірностей. Точність і надійність результатів були постійною проблемою і пов'язувались з побоюваннями щодо довіри учнів та їхньої безкритичної залежності. Зворотний зв'язок, генерований ШІ, також неодноразово описували як механічний, шаблонний та такий, що не враховує контексту, що підкреслює суперечність між доступністю та масштабованістю цих інструментів і їхніми обмеженнями в сферах, важливих для учнів. У дослідженнях підхід із залученням людини (human-in-the-loop) зазвичай визнавався ключовим принципом пом'якшення ризиків поряд із технічними вдосконаленнями та педагогічними заходами, зокрема навчанням навичкам роботи з підказками. Однак важливі питання залишаються невирішеними, зокрема щодо того, чи зменшує чи збільшує GenAI навантаження на викладачів і чи звужують чи розширюють ці інструменти існуючі розриви в рівності. Загалом отримані результати свідчать про те, що ефективна інтеграція GenAI у викладання розмовної англійської як іноземної (EFL) вимагає уваги до технічного проєктування, педагогіки та інституційної підтримки, тоді як деякі аспекти її впровадження залишаються недостатньо вивченими.

Ключові слова: генеративний штучний інтелект (GenAI); практика розмовної та вимови англійської як іноземної (EFL); великі мовні моделі (LLM); виклики та обмеження; стратегії пом'якшення наслідків.

(cc) BY-NC-SA

This work is licensed under Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.