



AI-ASSISTED GRAMMAR INSTRUCTION: THE ROLE OF CHATGPT IN IMPROVING ENGLISH GRAMMAR COMPETENCE

Abdumajidova Sevinch Xurshidbek qizi

<https://orcid.org/0009-0002-4863-1488>

Qo'qon davlat universiteti

<https://doi.org/10.5281/zenodo.21837174>

Abstract. Grammar remains one of the most persistent difficulties for learners of English as a foreign language, and traditional grammar instruction, built around rule explanation, mechanical drilling, and delayed correction, does not always give individual students the attention needed to internalize accurate structures. Over the past three years, ChatGPT has become a widely accessible source of grammar support, capable of producing instant explanations, corrections, and example sentences on demand. This paper reviews recent experimental, quasi-experimental, and survey-based research examining the effect of ChatGPT on learners' grammatical competence. Drawing on studies conducted with EFL learners in Turkey, Saudi Arabia, Vietnam, Indonesia, Iran, and the Czech Republic, the review identifies a consistent pattern: guided ChatGPT use tends to produce measurable gains in grammatical accuracy, particularly in subject-verb agreement, article use, and sentence-level cohesion, while also raising recurring concerns about overreliance, contextual misjudgment, and the weakening of learners' own error-correction habits. The paper concludes that ChatGPT works best as a supplementary grammar tutor rather than a substitute for teacher-led instruction, and it proposes several practical conditions under which the tool can be integrated responsibly into EFL grammar pedagogy.

Keywords: ChatGPT, grammar competence, artificial intelligence, English as a foreign language, computer-assisted language learning

Introduction. Grammatical accuracy has long occupied an uneasy place in English language teaching. Communicative approaches pushed grammar instruction toward the margins of the classroom for much of the past three decades, yet learners themselves continue to report grammar as the skill they find hardest to master and the one that most visibly betrays their proficiency level in writing and speech. Subject-verb agreement, article choice, tense consistency, and preposition use remain sources of persistent error even among advanced learners, largely because these features rarely receive the sustained, individualized correction that acquisition seems to require.

The release of ChatGPT in November 2022 introduced a tool that could, in principle, supply exactly this kind of individualized attention: instant grammatical feedback, unlimited patience, and explanations tailored to a learner's own sentences rather than to a generic textbook example. Within two years, a substantial body of research had accumulated on the question of whether this promise translates into measurable learning gains. Fang and Han's [1] narrative review of 71 papers published between November 2022 and May 2024 found that grammar was in fact one of the least studied linguistic domains within this literature, overshadowed by a much larger interest in writing more broadly, yet the handful of studies that did isolate grammar produced some of the clearest experimental results in the entire corpus.

This paper synthesizes that emerging evidence base to answer three questions. First, does ChatGPT use produce measurable improvement in learners' grammatical accuracy compared with conventional instruction? Second, through which specific grammatical structures and instructional arrangements does this improvement tend to appear? Third, what limitations and risks accompany these gains, and how might they be managed in classroom practice? The review draws on nine primary studies conducted across seven countries, supplemented by findings reported in two large-scale literature reviews, to build a picture of where the research currently stands and where its gaps remain.

Materials and methods. This paper adopts a narrative literature review design rather than a systematic review with predefined statistical pooling, since the body of experimental research on ChatGPT and grammar instruction is still too small and methodologically heterogeneous for meta-analysis. Sources were located through Google Scholar, ResearchGate, ScienceDirect, and the Arab World English Journal database, using combinations of the terms "ChatGPT," "grammar," "EFL," "quasi-experimental," and "writing accuracy." Priority was given to studies published between 2023 and 2026 that reported an intervention involving ChatGPT and a quantifiable grammar or writing-accuracy outcome, whether through pretest-posttest comparison, error-count analysis, or standardized rating.

Nine primary empirical studies met these criteria and form the core evidence base of this review: an experimental study of Saudi ESL learners [2]; a Turkish grammar-guide intervention summarized within a broader review [3]; a Vietnamese quasi-experimental study on grammar precision in writing [4]; a Cogent Education study on ChatGPT feedback for Gen Z EFL writers in the Czech context [5]; an Indonesian study on essay quality development [6]; a Saudi study on the syntactic complexity of EFL writing [7]; a Chinese comparative study contrasting learner writing with ChatGPT-generated text [8]; a systematic review of ChatGPT practice in Vietnamese English teaching [9]; and an Iranian study of metacognitive awareness during ChatGPT-assisted error correction, again summarized within a larger review [10]. Two broader review articles [1, 9] were used to contextualize these individual findings within the wider literature and to identify studies not independently accessible for this review. Each primary study was read for its research design, sample, intervention length, and reported outcome measures, and findings were grouped thematically into accuracy gains, structure-specific effects, and reported limitations, which structure the Results section below.

Results: Overall gains in grammatical accuracy. Across the reviewed studies, the experimental groups that used ChatGPT for grammar practice or feedback consistently outperformed control groups that relied on textbook instruction or teacher feedback alone. Ali et al. [2] found that Saudi ESL students who used ChatGPT as a grammar-comprehension aid showed significantly higher post-intervention scores than students taught through conventional classroom methods, a pattern the authors attributed to the tool's capacity to answer follow-up questions on the spot rather than waiting for the next scheduled lesson. In the Turkish study summarized by Fang and Han [1, 3], a seven-week intervention comparing a ChatGPT-supported experimental group with a textbook-only control group produced no significant difference at pretest but a significant gap at posttest, with the experimental group's grammar knowledge improving considerably more.

A similar pattern held in writing-based interventions, where grammar was measured as one component of overall text quality rather than in isolation. Ivenz [5] reported that a twelve-

week structured program of ChatGPT feedback led to statistically significant gains in grammar, vocabulary, and organization for beginner-level EFL writers, with inter-rater reliability checks confirming that the improvements were not an artifact of a single evaluator's judgment. Puteri, Cahyono, Widiati, and Suryati [6] likewise found that Indonesian university students who used ChatGPT as a writing assistant produced essays of measurably higher quality after the intervention than a control group receiving no AI-based feedback, although the sample in that study was modest at thirty-five participants.

Structure-specific effects Several studies moved beyond aggregate accuracy scores to identify which grammatical structures responded best to ChatGPT-assisted instruction. Subject-verb agreement, article use, and preposition choice appeared repeatedly as the structures showing the sharpest error reduction, likely because these are the categories ChatGPT flags most reliably and explains most consistently across repeated queries. Vo and Ho [4] found that Vietnamese first-year English majors who received real-time ChatGPT feedback on writing assignments over a semester achieved significantly better grammatical precision than a control group taught conventionally, with the largest gap appearing in sentence-level accuracy rather than paragraph-level organization. Alkhalifah and Almutlaq [7], examining Saudi undergraduates' academic writing, found that combining ChatGPT feedback with teacher feedback produced greater gains in syntactic complexity than teacher feedback alone, suggesting that ChatGPT's contribution is additive to, rather than a substitute for, expert human input.

Not every comparison favored ChatGPT unconditionally, however. Zhou, Cao, Zhou, Zhang, and He [8] found that Chinese intermediate English learners actually outperformed ChatGPT-generated text on measures of narrative cohesion, a reminder that the tool's grammatical fluency does not automatically translate into superior discourse-level coherence, and that human writers retain an advantage in structuring extended narrative that current models do not fully replicate.

Reported limitations and risks Alongside these gains, nearly every study in the review flagged limitations that qualify an otherwise favorable picture. The Iranian metacognition study summarized by Fang and Han [1, 10] found that a week of ChatGPT-assisted error correction improved learners' accuracy in spotting linguistic errors but produced no corresponding improvement in metacognitive awareness, meaning students got better at the task without necessarily understanding why their corrections were right. Pham and Cao's [9] systematic review of ChatGPT use in Vietnamese English education similarly noted that students tended to treat the tool as a source of instant answers rather than a resource for developing independent grammatical judgment, and the authors called explicitly for further research into safeguards against this kind of passive dependence.

A further limitation concerns contextual sensitivity. Several of the reviewed studies observed that ChatGPT's corrections, while grammatically sound in isolation, sometimes altered a learner's intended register or stylistic voice, correcting errors the student had not actually made or imposing a more formal tone than the assignment required. This pattern appeared most clearly in studies involving longer compositions rather than single-sentence exercises, where ChatGPT's suggestions had more room to diverge from the writer's original intent.

Discussion. Taken together, the reviewed evidence supports a cautiously positive answer to the first research question: ChatGPT use is associated with measurable gains in grammatical accuracy across a range of EFL contexts, sample sizes, and intervention lengths. The consistency of this finding across studies conducted in Saudi Arabia, Turkey, Vietnam, Indonesia, and the Czech Republic is notable given how much these educational contexts otherwise differ, and it suggests that the mechanism at work, namely immediate and repeatable feedback, operates fairly independently of the broader instructional culture in which it is embedded.

The answer to the second question, concerning which structures benefit most, points toward a practical implication for teachers: ChatGPT appears best suited to the kind of rule-governed, high-frequency error that admits a clear right answer, such as subject-verb agreement or article choice, and less reliable for judgment calls involving register, cohesion, or authorial voice. This distinction matters for how the tool should be introduced in the classroom. Used to drill and explain discrete grammatical points, ChatGPT appears to function well as an on-demand tutor; used as a wholesale substitute for feedback on extended writing, its limitations become more visible, as the Chinese cohesion study [8] and the register-shifting problems noted above both illustrate.

The third question, concerning risks, is where the reviewed literature converges most strongly on a single concern: overreliance. Several studies [1, 9, 10] independently raised the possibility that easy access to correct answers may erode the slower, more effortful process of noticing and self-correcting errors, a process long considered central to second-language acquisition. This concern does not appear to be fully resolved by any study reviewed here, since none tracked learners over a period long enough to establish whether early gains in accuracy persist once ChatGPT support is withdrawn. This is arguably the most important open question for future research: whether ChatGPT-assisted gains reflect durable learning or a temporary scaffolding effect that disappears once the tool is removed.

Based on these findings, three conditions seem important for integrating ChatGPT responsibly into grammar instruction. First, the tool should supplement rather than replace teacher feedback, consistent with the additive effect found by Alkhalifah and Almutlaq [7]. Second, its use should be structured around specific, bounded tasks, such as reviewing a paragraph for a named category of error, rather than left open-ended, since bounded tasks appear less prone to the register-shifting problems noted in longer compositions. Third, some component of the intervention should require learners to explain, in their own words, why a ChatGPT-suggested correction is right, a step aimed directly at the metacognitive gap identified in the Iranian study [1, 10].

Conclusion. The research reviewed here indicates that ChatGPT can meaningfully support the development of English grammatical competence when used as a structured supplement to classroom instruction, with the clearest gains appearing in rule-governed categories such as subject-verb agreement and article use. At the same time, the same body of research consistently flags overreliance, weakened metacognitive engagement, and occasional loss of register as risks that unstructured use can bring. Neither the enthusiastic embrace nor the outright rejection of ChatGPT in grammar teaching is well supported by current evidence; what the studies converge on instead is a conditional endorsement, contingent on how deliberately the tool is folded into pedagogy. Future research would benefit from longer intervention periods, delayed posttests to assess retention after ChatGPT support is withdrawn,

and closer attention to how learners at different proficiency levels respond to the same intervention, since most existing samples remain concentrated among university-level learners in higher education settings.

References:

1. Fang, S., & Han, Z. (2025). On the nascency of ChatGPT in foreign language teaching and learning. *Annual Review of Applied Linguistics*, 45, 148–178. <https://doi.org/10.1017/S026719052510010X>
2. Ali, Y. A., Ali, M. M., Alharbi, W., Hamid, S., & Abbas, S. (2025). The influence of ChatGPT on English grammar among Saudi ESL learners. *Arab World English Journal (AWEJ)*, Special Issue on Artificial Intelligence, 358–371.
3. Kucuk, S. (2024). [Grammar-guide intervention study], as cited in Fang, S., & Han, Z. (2025). On the nascency of ChatGPT in foreign language teaching and learning. *Annual Review of Applied Linguistics*, 45, 148–178.
4. Vo, T. L., & Ho, T. M. P. (2026). A quasi-experimental study on enhancing EFL learners' grammar precision in writing skills under supported-ChatGPT approach at a university in a remote region of Vietnam. *Arab World English Journal (AWEJ)*.
5. Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1), Article 2410101. <https://doi.org/10.1080/2331186X.2024.2410101>
6. Puteri, C. G., Cahyono, B. Y., Widiati, U., & Suryati, N. (2024). Using ChatGPT as a writing assistant: A study on essay quality development among Indonesian EFL university students. *Indonesian Journal of Applied Linguistics*.
7. Alkhalifah, E., & Almutlaq, S. (2026). The combined impact of ChatGPT and teacher feedback on the syntactic complexity of EFL learners' writing. *SAGE Open*. <https://doi.org/10.1177/21582440261453413>
8. Zhou, T., Cao, S., Zhou, S., Zhang, Y., & He, A. (2023). Chinese intermediate English learners outdid ChatGPT in deep cohesion: Evidence from English narrative writing. *System*, 118, Article 103141. <https://doi.org/10.1016/j.system.2023.103141>
9. Pham, M. T., & Cao, T. X. T. (2025). The practice of ChatGPT in English teaching and learning in Vietnam: A systematic review. *International Journal of TESOL & Education*, 5(1).
10. Mousazadeh, R. (2023). [Metacognitive awareness and ChatGPT-assisted error correction study], as cited in Fang, S., & Han, Z. (2025). On the nascency of ChatGPT in foreign language teaching and learning. *Annual Review of Applied Linguistics*, 45, 148–178.