

ARTICLE



## Bringing Cinderella into spotlight: GenAI-assisted grammar acquisition in academic writing

Shu Zhou\*, *Surrey International Institute,  
DUFE-Surrey Joint Research Centre,*

*Dongbei University of Finance and Economics, Dalian, Liaoning, China*

Gerhardus D. Du Preez, *Surrey International Institute,*

*DUFE-Surrey Joint Research Centre,*

*Dongbei University of Finance and Economics, Dalian, Liaoning, China*

### Abstract

*This study explores students' learning effects, evolving perceptions, and experiences of using a generative artificial intelligence (GenAI) tool in tandem with pedagogical theory to facilitate grammar acquisition within an academic writing context. While grammar is critical, limited research exists on systematic grammar instruction in English for academic purposes (EAP), and prevailing grammar learning approaches lack strategy-based pedagogical instructions. Existing studies on GenAI-assisted L2 writing focus predominantly on grammar correction rather than acquisition, and on students' universal perceptions rather than their dynamic trajectories. We conducted a six-week mixed-methods experiment involving pre-/post-essay tests and reflective journals with 20 first-year Chinese undergraduates enrolled in a year-long EAP course at a Sino-British institute. Quantitative analysis revealed a statistically significant improvement in grammatical performance; qualitatively, students revealed four developmental stages: (a) Enthusiastic Discovery, (b) Critical Engagement, (c) Empowered Mastery, and (d) Reflective Personalisation. Theoretically, this study advances the Grammar Learning Strategies (GLS) framework by demonstrating how GenAI can operationalise cognitive and metacognitive strategies (Pawlak, 2013) through an interactive, sequenced workflow. Practically, it proposes a pedagogical model for integrating GenAI into time-constrained EAP curricula, enabling learners to transition from grammar correction to autonomous grammar acquisition.*

**Keywords:** GenAI-assisted L2 writing; grammar acquisition; Grammar Learning Strategies (GLS); English for Academic Purposes (EAP)

**Language(s) Learned in This Study:** English

**APA Citation:** Zhou, S., & Du Preez, G. D. (2025). Bringing Cinderella into spotlight: GenAI-assisted grammar acquisition in academic writing. *Language Learning & Technology*, 29(1), 1–25.  
<https://doi.org/10.64152/10125/73639>

### Introduction

Grammar has been dubbed a ‘Cinderella’ in English as a second language/foreign language (ESL/EFL) due to the lack of studies compared with other language skills (Oxford et al., 2007). This trend extends to English for Academic Purposes (EAP), a sub-branch of ESL/EFL, where courses are typically characterised by short, fixed timeframes (usually 4–12 weeks for pre-sessional or one year for foundation courses) and a focus on formal academic genres (Smith, 2022). This lack of focus is problematic because grammar plays a critical role in EAP, directly impacting clarity, formality, and precision—key academic writing requirements (Cullen, 2008; Evans & Green, 2007; Nguyen, 2010).

\* **Corresponding Author:** Shu Zhou, [s.zhou@surrey.ac.uk](mailto:s.zhou@surrey.ac.uk)

Compounding this issue, many EAP learners enter courses with only pre-intermediate proficiency such as IELTS 5.0 to 5.5 (Nguyen, 2010), a band with frequent grammatical errors (Müller & Han, 2022). Consequently, learners struggling with grammar face difficulty in meeting academic writing demands. As noted by Barnard and Scampton (2008) and Fedorova (2022), many EAP students explicitly express dissatisfaction and frustration with the lack of grammar instruction in writing courses, which ultimately hinders their academic success. To address this deficit, students have sought computer-mediated feedback or correction for their writing assignments, especially in higher education contexts where they generally have their own portable computing device and Internet connection (Taskiran & Goksel, 2022). While traditional automated writing evaluation (AWE) tools (e.g., Grammarly, QuillBot, Copy.ai, Wordtune) assist with error correction and surface-level revisions, they fail to engage learners in meaningful cognitive processing necessary for internalising grammatical rules and metacognitive awareness for strategic planning (John & Woll, 2020; Shi & Aryadoust, 2023). Thus, they fail to facilitate further grammar acquisition, transcend mere error reduction, or support the unconscious internalisation of rules through exposure and use (Ellis, 2015); they do not help learners automatise knowledge for communicative use.

The critical role of cognitive and metacognitive skills in grammar acquisition is central to Pawlak's (2013) Grammar Learning Strategies (GLS) framework, with cognitive and metacognitive strategies particularly crucial for developing grammatical competence. Cognitive strategies (e.g., deduction, practice through writing) directly facilitate explicit knowledge development and its gradual automatisisation into implicit knowledge. Meanwhile, metacognitive strategies (e.g., planning, self-monitoring, self-evaluation) enable learners to systematically regulate their grammar learning processes (Pawlak, 2018), justifying our focused investigation of these two strategies. However, existing GLS research (e.g., Alnufaie & Alzahrani, 2024; Cahyani et al., 2022) has predominantly measured students' use of these strategies, with limited attention to how teachers can effectively instruct students in employing them to achieve better grammatical performance. In EAP contexts in particular, as previously noted, even basic grammar instruction remains minimal, let alone systematic instruction on GLS implementation.

The emergence and rapid development of large language models and generative artificial intelligence (GenAI) technologies, such as ChatGPT, offer transformative potential for addressing persistent gaps in grammar instruction, particularly in time-constrained EAP classrooms. Compared with traditional AWE tools, GenAI tools go beyond simple error correction to provide more detailed explanations, tailored assistance, and, most importantly, dynamic interactions and personalised learning resources (Kohnke, 2024; Vo & Nguyen, 2024; Zadorozhnyy & Lai, 2024), which can all support cognitive strategy development and metacognitive regulation. However, two critical research gaps remain: (a) while existing studies demonstrate GenAI's effectiveness in improving writing accuracy, pedagogical frameworks to guide proper usage are lacking, potentially leading to over-reliance (e.g., Yan, 2024) and superficial learning rather than genuine grammar acquisition; and, (b) current research reports students' overall positive perceptions, but few studies have examined dynamic changes in learners' experiences throughout the AI adoption process, particularly how their understanding, skills, and attitudes evolve as they learn to use these tools more effectively with instructional guidance.

This study aims to bridge these gaps and 'bring Cinderella into the spotlight'; accordingly, we adopted a mixed-methods approach through a six-week experiment, consisting of four core experimental sessions (each 1.5–2 hours) and two writing tests (one at the beginning and one at the end of the experiment). We aimed to explore both learning effects and the developmental trajectory of students' perceptions and experiences when using a GenAI tool alongside existing pedagogical theory to facilitate grammar acquisition within an EAP writing context. Our research seeks to establish an innovative pedagogical model where GenAI serves not merely as a grammar correction tool but as a cognitive partner facilitating the internalisation of grammatical knowledge and the development of metacognitive learning strategies. More broadly, this study may advance traditional GLS in the emerging GenAI era.

## Literature Review

### Insufficient Grammar Instruction in EAP Despite Substantial Demand

Grammar has a neglected position in ESL; debate over whether and how grammar should be taught is ongoing, and research on GLS is still limited (Oxford, 2017; Oxford et al., 2007; Pawlak, 2012, 2018). Though linguistic accuracy plays an important role in the quality of written texts, whether explicit grammar teaching promotes students' grammar acquisition remains unclear (Zhou, 2009). Krashen (2003), who advocated for a natural approach, claimed that grammar instruction does not improve learners' language use and that explicit grammar input is insignificant and not indicative of long-term language production accuracy, concluding that grammar instruction is unnecessary for L2 acquisition. While Ellis (2006) contradicted this approach, citing one meta-analysis and concluding that explicit instruction contributes to significant 'target-oriented gains', he noted limited and mixed evidence supporting the effectiveness of grammar teaching. Given the weak support for grammar instruction and the lack of GLS research, Oxford et al.'s (2007) metaphor of grammar as a 'Cinderella' appears justified.

While the debate over grammar instruction persists in general ESL contexts, its role proves more crucial in EAP settings. Cullen (2008) described grammar as a 'liberating force', enabling learners to distinguish meaning, express formality/informality, and convey clarity or ambiguity—key academic writing aspects. Evans and Green (2007) emphasised the importance of accurate use of grammar for expressing complex ideas in academic writing and speaking. Hinkel (2004) also highlighted syntactic accuracy's vital role in EAP writing, and accordingly, suggested teaching explicit grammar structures in academic writing.

However, the neglect of systematic grammar teaching in EAP stems from both pedagogical and practical constraints. Most EAP programmes' time-bound nature results in prioritising genre conventions over grammatical instruction (Smith, 2022). This institutional limitation is compounded by prevailing teaching methodologies in EAP classrooms: content- and task-based instruction (Fedorova, 2022), which relate grammar to incidental learning (Barnard & Scampton, 2008). Evans and Green (2007) argued that this immediate communication fails to develop sustained competence.

Paradoxically, transnational education's rapid expansion has brought increasing numbers of students to EAP programmes—many with over 10 years of prior English study yet unprepared for the grammatical demands of academic discourse (Zhou et al., 2022). Hyland's mid-1990s study (cited in Evans & Green, 2007) found academic writing a major concern for university students, confirming their perceived shortcomings with notable grammatical limitations. Similarly, Nguyen (2010) reported that, while students' speaking and listening skills were adequate, writing remained their greatest hurdle, with many appearing 'academically [i]lliterate'. A recent IELTS Official Report (Müller & Han, 2022) revealed that grammar errors do not decline steadily with score gains; regression occurs in key categories (e.g., nouns, adjectives, and conjunctions at 6.5–7.0), demonstrating that even high-scoring test-takers (7.0+) may exhibit accuracy fluctuations due to cognitive shifts or skill attrition. Consequently, there is a clear need for enhanced, systematic grammar instruction in EAP contexts, or teachers must actively guide students to develop achievable grammar awareness and target deliberate acquisition through explicit correction and contextualised practice (Fedorova, 2022; Zhou, 2009).

### Grammar Learning Strategies: The Path to Acquisition

GLS refers to 'actions and thoughts that learners consciously employ to make language learning and/or language use easier, more effective, more efficient and enjoyable' (Oxford et al., 2007, p.117). Cohen and Pinilla-Herrera (2010) further defined GLS as 'deliberate thoughts and actions that students consciously employ for learning and getting better control over the use of grammar structures' (p. 64). Oxford (2017) emphasised that GLS are 'teachable, dynamic' strategies in 'specific contexts' that enhance 'self-regulated, autonomous L2 grammar development and long-term efficiency' (p. 244).

Three key principles emerge from these definitions, guide the present study. First, although GLS are often instinctively adopted by learners, teachers' instruction and guidance are indispensable. As Alnufaie and Alzahrani (2024) suggest, GLS use cannot be isolated from grammar learning and instruction. Second, these strategies are most effective when integrated within specific learning contexts and tailored to purposeful learning tasks; the importance of context in different cultural and educational backgrounds is also highlighted by Pawlak (2018). Third, learning strategies fostering long-term grammatical proficiency should evolve continuously.

To further examine learners' processes when learning grammar constructs, this study adopts Pawlak's (2013) framework due to its learner-centred design (Pawlak, 2012) and relevance to university-level learners (Pawlak, 2019). In particular, this study focuses on two core strategies: (a) Metacognitive strategies (META) for planning, monitoring, and evaluating grammar learning; and, (b) cognitive strategies, which engage learners in understanding and retaining grammar, with four subcategories: communication-based strategies (COG 1), explicit grammar knowledge (COG 2), implicit grammar knowledge (COG 3), and corrective feedback processing (COG 4). These two core strategies are prioritised as they directly govern grammar learning processes: cognitive strategies, which directly influence grammar learning, are central to the process, while metacognitive strategies play a supervisory role (Pawlak, 2018)—a pattern observed in high-achieving higher education students (Cahyani et al., 2022; Nakachi, 2021).

These strategies collectively address both grammar learning, involving the conscious study of grammatical rules and resulting in explicit knowledge, and grammar acquisition, broadly defined as the internalisation of grammatical knowledge, enabling its use in comprehension and production (Ellis, 2015). The measurement of grammar acquisition is not synonymous with error-free performance, since learners may demonstrate a knowledge of rules in controlled tasks (explicit knowledge) yet struggle to apply them fluently in writing (implicit knowledge); or vice versa, development in complexity might mask changes in accuracy (Larsen-Freeman, 2006). While accuracy is reflected by 'the ability to be free from errors while using language', complexity is shown in syntactic sophistication (e.g., clause embedding) or diversity (Polio & Shea, 2014). Thus, to operationalise acquisition in writing, studies increasingly combine accuracy and complexity measures (Larsen-Freeman, 2006; Norris & Ortega, 2009; Polio & Shea, 2014).

Therefore, to respond to Pawlak's (2019) call for strategy-based interventions leading to autonomous learning, this study investigates how a GLS-based intervention fosters acquisition, measured through both accuracy and complexity gains. This is critical in EAP contexts, where grammar instruction is often marginalised.

## **GenAI-assisted L2 Writing**

Since emerging in November 2022, ChatGPT, a generative large language model, has demonstrated unprecedented potential in assisting with and enhancing L2 learners' writing skills. Compared to previous chatbots or AWE systems such as Grammarly, AI-powered tools such as ChatGPT provide enhanced feedback and more dynamic interactions with learners; they can offer tailored assistance and personalised learning strategies and resources (Zadorozhnyy & Lai, 2024). Studies from the past three years have consistently highlighted GenAI's ability to generate high-quality corrective feedback and students' overall positive perceptions. For example, Mahapatra (2024) revealed that ChatGPT, as a formative feedback tool, significantly enhanced undergraduate ESL students' academic writing skills. Similarly, Chen et al. (2022) used a chatbot to support learners' writing, finding that the automated grammar, vocabulary, and sentence structure feedback led to improved writing performance over time. Notably, learners generally perceive these AI tools positively and see AI as a 'valuable learner partner' (Xiao & Zhi, 2023) and a 'dialogic and reliable writing tool' (Mahapatra, 2024). Likewise, students universally rate ChatGPT higher than human teachers for its more detailed, interactive, and engaging learning process (Escalante et al., 2023; Kohnke et al., 2023).

However, key issues in adopting GenAI tools in writing include learners' potential misuse, academic integrity violations (Kasneci et al., 2023), AI's own limitations such as inaccurate and misleading feedback (Barrot, 2023), and an over-reliance that reduces autonomous learning (Yan, 2023). To mitigate these risks, scholars call for the development of regulatory policies and pedagogical guidance—informing students to utilise AI legitimately and productively (Xiao & Zhi, 2023), guiding students to use AI purposefully and systematically in specific writing contexts, and exploring collaborative strategies to harness their full potential (Barrot, 2023). More importantly, students should be trained to evaluate AI's output critically and modify their prompts to maximise learning benefits and thus develop lifelong learning skills (Escalante et al., 2023). Only in this way can GenAI act as a more knowledgeable scaffold rather than a substitute.

Despite these insights, two critical gaps persist in current research. First, while AI's general feedback-provider role is confirmed, little attention has been paid to its role in systematic grammar acquisition—a foundational skill requiring targeted interventions (e.g., metalinguistic explanations, repeated practice) aligned with grammar learning theories (e.g., Pawlak, 2013). No pedagogical framework exists to demonstrate how ChatGPT might scaffold this process, from error detection to internalisation and then production. Second, despite students' positive perceptions of GenAI tools, the dynamic learning process remains overlooked: how do students' perceptions and behaviours evolve during learning processes with AI? These developmental perceptions are essential as students' perceived usefulness and ease of use are key motivators for technology adoption (Davis, 1989).

Addressing these gaps, our study proposes a theory-informed GenAI-assisted workflow for grammar acquisition, combining grammar learning strategies with GenAI's capabilities. We investigate this workflow's effectiveness while tracking students' behavioural and emotional responses across the learning process. Additionally, to respond to the 'Cinderella' role of grammar in EAP, we set this study in a particular academic writing context. Accordingly, we present the following research questions (RQs):

**RQ1:** To what extent does a GenAI-assisted strategic intervention enhance students' grammar acquisition in academic writing tasks?

**RQ2:** How do students' perceptions and experiences evolve throughout the GenAI-assisted grammar acquisition process?

## Methodology

### Research Context and Participants

Participants were first-year Chinese undergraduates at a Sino-British institute in a Chinese university. Students take EAP courses for a year before progressing to their main degrees; approximately 340 students enrol in each of these year-long courses. The EAP curriculum includes ten modules over two semesters, covering Academic Listening, Reading, Speaking, Writing, and English for Specific Purposes. All lectures and assessments are delivered in English. The students were all native Chinese speakers. They typically had 10 years of English learning experience before university and had to score at least 115 out of 150 on the Gaokao (Chinese national matriculation to college) English section to gain admission. Based on the institution's English Language Admission Test, students were evaluated to possess at least B1-level English proficiency, according to the Common European Framework of Reference.

The study was conducted in March and April 2024. 23 students volunteered to participate. Two withdrew due to scheduling conflicts and one due to technical issues with the AI software, leaving 20 complete data sets. All participants had experience using GenAI and AWE tools such as Grammarly in their English studies. The invitation included sufficient information about the project to enable students to provide informed consent. Wenxin Yiyan (Baidu ERNIE Bot: <https://yiyan.baidu.com>), a popular LLM in China, was utilised in all experimental sessions. Officially launched in March 2023, this tool was selected because of its credibility as a product of Baidu, a leading Chinese technology company. We employed the

basic (free) version as it is capable of interacting with college students to facilitate grammar learning without requiring additional cost.

## Research Design and Procedure

This mixed-methods study employed a quasi-experimental design paired with qualitative enquiry to address its dual research aims. To evaluate the extent to which the GenAI-assisted intervention enhances grammar acquisition (RQ1), we compared students' essay scores in a pre-test (baseline) and post-test (after four sessions with the AI tool). Additionally, to trace the evolution of students' perceptions and experiences during each stage of the learning process (RQ2), we collected qualitative data in the form of written reflections after each of the experimental sessions, analysing thematic patterns over time.

The overall data collection followed three sequential stages: (a) quantitative pre-test essay scores; (b) four experimental intervention sessions with embedded qualitative reflection points; and (c) quantitative post-test essay scores.

## Data Collection

To assess RQ1 and gauge students' progress, we implemented a diagnostic writing test administered in the first (pre-test) and last week (post-test). Both tests used argument-type prompts (see [Appendix A](#)), requiring students to write a 350-word essay including an introduction, body paragraphs, and a conclusion. The task was selected because argumentative writing is both a common mode of academic discourse for the target student group and an integral component of their regular syllabus.

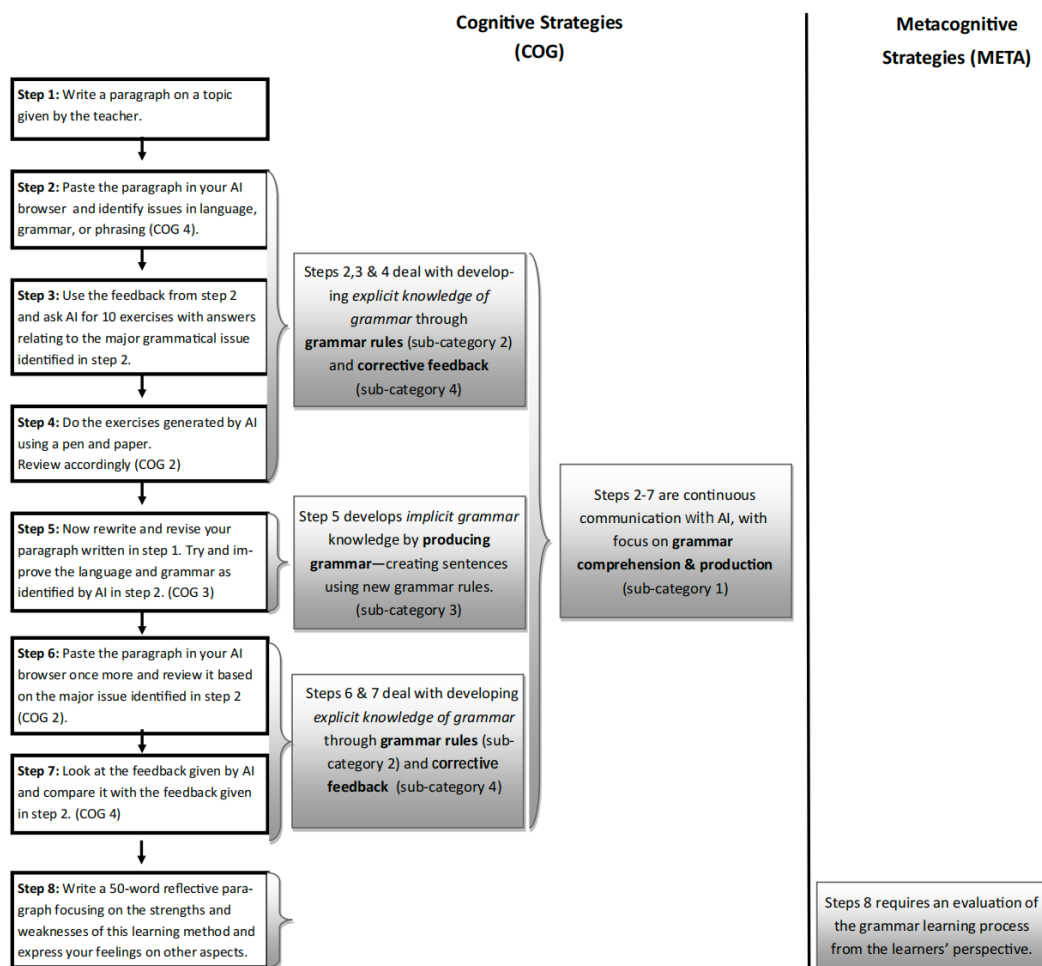
Pre- and post-tests were independently marked (total score of 100) by the two authors, who are both experienced EAP instructors. Grading criteria were derived from existing course writing rubrics developed by language specialists from the UK partner university. As this study specifically examined grammar acquisition, the researchers focused exclusively on grammar-related assessment criteria extracted from these comprehensive rubrics. Additionally, building upon participants' engagement in multiple grammar-focused training sessions, our GenAI-assisted GLS intervention aimed to promote grammar acquisition, operationalised as learners' ability to deploy grammatical rules in writing production. Following Polio and Shea's (2014) recommendation, as mentioned earlier, we employed a holistic evaluation approach that simultaneously assessed grammatical accuracy (demonstrating controlled application) and complexity (reflecting rule internalisation and syntactic sophistication, e.g., subordinate clauses, passive voice). We added quantitative thresholds ([Appendix B](#), e.g.,  $\leq 1$  major error and  $\geq 3$  error-free complex structures per 100 words for 'Excellent'), as 'major errors' were defined as meaning-disrupting violations (e.g., tense shifts, agreement errors), based on researchers' consensus during piloting.

To address RQ2, participants were asked to engage in four interactive learning sessions with GenAI, each lasting between 1.5 and 2 hours. Prior to each session's commencement, a 5–10-minute training or warm-up session was conducted to familiarise students with the process and address their queries, aligning with Mathur and Mahapatra's (2022) recommendation that learners receive training before using digital tools in the classroom. Students were directed to strictly adhere to the steps outlined on the instruction sheet in Sessions 1–3 (see [Appendix C](#)). We meticulously designed the initial instruction sheet, incorporating Pawlak's (2013) GLS (mainly cognitive and metacognitive) at every step. Considering feedback gathered from students during Session 1, modifications were made to Instruction Sheet 2 to better align with students' needs. Similarly, Instruction Sheet 3 underwent minor changes based on feedback received during Session 2. To evaluate students' mastery of the learning method and promote autonomous learning, no instructions were given for Session 4. The interactive learning process is summarised in [Figure 1](#).

Data collected from Stage 2 included reflective paragraphs elaborating on each session's learning perceptions and experiences, and the entire chat history with the GenAI (from Steps 2–7). Ultimately, 80 reflective paragraphs (written mainly in English) and 80 chat histories were gathered from the 20 students. These were submitted via a WeChat group following each session and stored in a virtual shared folder.

**Figure 1**

*GenAI-Assisted Grammar Acquisition Process*



As outlined in Step 8 of the worksheets, students were instructed to evaluate their GenAI-assisted learning experiences mainly through two prescribed lenses: (a) Strengths: beneficial aspects of the learning process (e.g., AI's immediate feedback helped me identify grammatical errors); and, (b) Weaknesses: limitations of the learning process (e.g., AI's explanation is confusing). This dual focus was theoretically grounded in prior studies tracking students' perceptions on ChatGPT-facilitated language learning, where learners consistently demonstrated two dimensions: affordances and challenges (e.g., Huang et al., 2022; Kasneci et al., 2023; Yan, 2023). Chat histories were archived to triangulate reflection claims (e.g., verifying 'confusing explanations' with actual AI responses in Session 2 logs).

## Data Analysis

As mentioned previously, all pre-test and post-test essays were evaluated using grammar-focused rubrics (Appendix B) that holistically integrated accuracy and complexity through a six-level performance band. The final holistic score (0–100) balanced both dimensions to reflect grammar acquisition—prioritising accuracy as a prerequisite while rewarding complexity when achieved without compromising correctness. This mirrored real-world writing contexts where fluent grammatical control precedes stylistic elaboration (Ellis, 2015).

The two authors independently rated all essays after establishing inter-rater reliability through standardised practice scoring and discussion. Initial discrepancies exceeding five points (occurring in 17% of cases) were resolved through consensus meetings with reference to the rubric descriptors. Paired-samples *t*-tests were analyzed using SPSS (Version 28.0), and were employed to compare pre- and post-test scores, with normality assumption confirmed via Shapiro-Wilk tests for all score distributions (pre-test, post-test, and difference scores; all  $p > .05$ ). Given the small sample size ( $n = 20$ ), Cohen's *d* was calculated to complement null-hypothesis testing by quantifying the intervention's practical effect size.

The qualitative data analysis followed Strauss and Corbin's (1990) three-step coding process (open, axial, and selective) to capture the dynamic trajectory across the four sessions. Two additional senior EAP lecturers collaborated in coding and theme extraction, and a joint discussion session was organised where discrepancies and disagreements amongst the coders were addressed and resolved.

During the open coding phase, line-by-line coding identified 37 recurrent themes across the sessions, mainly categorised into perceived strengths (S, e.g., 'spotting subtle mistakes I overlooked'; 'making my essay more professional/academic') and weaknesses (W, e.g., 'ambiguous explanations'; 'exercises are too simple'). Special attention was paid to affective markers (e.g., 'surprised', 'frustrated') to investigate emotional responses.

In the axial coding phase, these 37 themes were consolidated into core categories. For example, 'Quality of explanation' merged four open codes criticising the AI's unsatisfying feedback (e.g., 'It points out errors but doesn't explain why').

In the final selective coding phase, axial categories were theorised into a four-stage model aligning with grounded theory's aim to derive processual explanations (Charmaz, 2006), in which axial categories are synthesised into a central narrative. For instance, Enthusiastic Exploration summarised the initial excitement on this learning approach's novelty, a tentative yet motivated exploration.

## Results

### Research Question 1: Students' Progress in Grammar Acquisition

As shown in Table 1, paired-samples *t*-tests revealed a statistically significant improvement from pre-test ( $M = 76.52$ ,  $SD = 5.97$ ) to post-test ( $M = 80.38$ ,  $SD = 5.81$ ):  $t(19) = -3.573$ ,  $p = .002$ , with a mean gain of 3.86 points ( $SD = 5.81$ ). The effect size reached a medium-to-large magnitude (Cohen's  $d = 0.780$ ), indicating a substantive pedagogical impact. While individual gains varied considerably with a range between -6.5 to 11 points, the Winsorised analysis (5% truncation) confirmed result robustness,  $t(19) = -3.665$ ,  $p = .002$ . These findings demonstrate that the 4-week GenAI intervention effectively elevated students' integrated grammar acquisition, balancing accuracy and complexity as theorised by Polio and Shea (2014), enhancing students' scores by approximately one proficiency band from Satisfactory (70–80) to lower Good (80–90).

**Table 1***Comparison Between Pre-test and Post-Test (N = 20)*

Measure	Pretest	Posttest	Mean Difference	<i>t</i> (19)	<i>p</i>	Cohen's <i>d</i>
M (SD)	76.52 (5.97)	80.38 (5.81)	3.86 (5.81)	-3.573	.002	0.780
Score Range	[66.5, 87.5]	[70.0, 89.5]	[-6.5, 11.0]			

*Note.* Results remained significant after 5% Winsorisation ( $t(19) = -3.665, p = .002$ ).

These quantitative findings are supported by teachers' feedback on the essays. A representative case is Student 15 (see two excerpts from Student 15's essays below), whose pre-test introduction paragraph exhibited multiple major errors according to the grading criteria. For example, frequent errors in verb forms ('made' vs 'make'), word order ('medicine's made need'), and sentence fragments resulted in  $\geq 5$  major errors per 100 words, alongside limited complexity, evidenced by predominant SVO structures and flawed attempts at subordination (e.g., 'which can be used...'), collectively resulting in a Soft Fail tier (60–70). After the Gen-AI intervention sessions, this student's work demonstrated marked improvement: minor prepositional errors ('on their studies') with no meaning obstruction ( $\leq 3$  errors per 100 words), as well as the effective use of relative clauses ('things that people can hardly do'), transitional phrases ('What is more...'), meeting Good tier standards (80–90). This progression demonstrates the significant effectiveness of GenAI-assisted strategic intervention on students' grammar acquisition, not just improved error control but genuine internalisation and production of grammatical patterns.

Nowadays, more and more kinds of illnesses are found in the population. This requires the doctors to do a lot of reseaches to made new kinds of medicines which can be used to fight these illnesses. However, a new kind of medicine's made need a long time. The researchers have to do a lot of tests and then the medicine can be used on people. Medicine should be tested on animals will be discussed in this essay. (Pre-test essay, Student 15)

Currently, AI is becoming more and more important in people's daily life. AI can finish a lot of things that people can hardly do and it can help people in many different ways. What is more, AI can also do a lot of things for students on their studies. Some people may think students should not be allowed to use AI in their studies. However, the reasons why students should be allowed to use AI in their studies will be discussed in this essay. (Post-test essay, Student 15)

## Research Question 2: Students' Perceptions and Experiences

To examine students' evolving perceptions and experiences, we analysed reflective paragraphs in each intervention session and revised the Instruction sheet according to students' feedback, eventually revealing four developmental stages: (a) Enthusiastic Discovery, (b) Critical Engagement, (c) Empowered Mastery, and (d) Reflective Personalisation. [Table 2](#) describes the evolving process.

### Stage 1: Enthusiastic Discovery

During the first session, students explored this new GenAI-assisted learning method, following the steps outlined in Instruction Sheet 1 under the researchers' guidance. Regarding the positive reviews, most students felt amazed by AI's ability to identify subtle errors in their writing, with many describing the experience as 'shocking' or 'surprising'. They also expressed excitement about this instructed learning method compared with their previous approach of simply asking AI to proofread their essays, though they still seemed reliant on AI's corrective feedback function during their first session.

I think AI is a relatively useful tool that helps me to find some small mistakes that are difficult to detect in ordinary times and refine my articles to make them more professional. (Student 1)

In contrast, the most common criticism was that the AI-generated exercises were overly simple. Additionally, two students noted that the exercises did not address their specific grammatical issues, two reported inaccuracies or insufficient explanations in the AI's feedback, and three stated that the limited grammar training did not transfer into improved writing skills.

I don't think the 10 practice questions are very useful because they are too simplistic, and we pay special attention to our grammar when doing these questions, but we make mistakes again unconsciously when we write. (Student 7, Session 1)

**Table 2**

*Four Stages of Students' Perceptions and Experiences*

Session	Theme	Description
1 Instruction Sheet v.1	Enthusiastic Discovery	Students were initially captivated by this learning workflow's novelty, expressing awe at its ability to detect subtle grammatical errors, though students rated grammar correction higher than grammar learning since many perceived AI-generated exercises to be overly simplistic.
2 Instruction Sheet v.2	Critical Engagement	With modified prompts, students valued AI-generated exercises and actively engaged in understanding explicit grammar rules, yet growing frustration emerged over AI's ambiguous explanations.
3 Instruction Sheet v.3	Empowered Mastery	Through continuous experimentation and crafting precise prompts, students' overall confidence grew in navigating the learning process with the AI and in developing empowerment to extract finer AI responses.
4 No Instruction Sheet	Reflective Personalisation	Without any teacher guidance, students transitioned from passive tool users to active strategic planners, as critical evaluation and personalised methods emerged. Students also transferred strategies to other language domains to accommodate their needs.

To further analyse the student feedback, we reviewed chat logs and identified several issues. First, AI primarily provided fill-in-the-blank and error-correction exercises, which, though tailored to identified issues, were overly basic in vocabulary and sentence complexity, failing to match students' proficiency levels. Second, nearly half the students did not follow instructions strictly. For example, in Step 3, some omitted the two specific issues from Step 2 when requesting exercises, resulting in less relevant tasks. Third, when AI gave inaccurate explanations, students often failed to ask follow-up questions. To address these issues, three improvements were planned for the second experiment: (a) A specification was added to the Step 3 prompt (Instruction sheet, V.2): 'the level of language must be suitable for IELTS 6.5'; (b) strict adherence to the process was emphasised; and (c) students were encouraged to ask AI for clarification when needed (e.g., requesting specific examples) or to seek help from researchers.

### **Stage 2: Critical Engagement**

Session 2 demonstrated a significant improvement in students' investment in the grammar acquisition process. Unlike Session 1, where the remarks mainly covered 'detecting grammar errors' and 'providing corrective feedback', students in Session 2 consistently highlighted 'strengthening target grammatical knowledge' (13 out of 20). Additionally, thanks to the upgraded instruction sheet, students expressed that the AI-generated targeted exercises effectively enhanced their understanding of specific grammar concepts. Many reported learning new grammar rules or reviewing previously learnt concepts. Excerpts from their reflections reveal their motivation and inspiration:

The questions provided by the AI have a certain level of difficulty, which helps us practice grammar effectively. After completing 20 questions, I have gained a better grasp of grammar knowledge that I didn't know before or had forgotten. I am also able to correct mistakes in the original text on my own. (Student 6, Session 2)

The practices are more specific and useful than before! Today's class was so useful for me! Thank you, teachers! (Student 9, Session 2)

However, more students (compared with Session 1) expressed dissatisfaction with the AI's explanations of grammatical rules, citing inaccuracy or ambiguity. This may reflect students' growing need for more detailed instructions as their understanding of explicit grammar rules deepened.

Furthermore, we reviewed chat logs from students expressing dissatisfaction (e.g., 'not helpful', 'frustrating'), finding a few issues. Student 21 persistently requested logical feedback despite evident grammatical issues (e.g., article usage), limiting AI's effectiveness for his actual needs. Student 4 called the method 'confusing', but the analysis revealed he began his AI interaction without completing his initial writing (Step 1). Additionally, he followed instructions mechanically without seeking clarification. To improve the subsequent sessions, we made additional minor revisions to Instruction Sheet V.3: (a) Modified Step 2's prompt so it identified one rather than two grammar issues to elicit more focused AI outputs; (b) lowered the word limit for the two essay-writing tasks (Steps 1 and 5) from 100 to 70 words, encouraging students to allocate more time to explicit learning; and, (c) allocated additional attention to Students 4, 5, 16, and 21 to enhance their learning experience.

### **Stage 3: Empowered Mastery**

In this stage, students exhibited a marked shift in their approach to learning, demonstrating greater autonomy and strategic thinking. They no longer passively accepted AI's suggestions but actively questioned and refined them, developing personalised strategies to optimise the tool's utility. Their growing confidence was evident in their ability to craft precise prompts, which yielded more targeted and useful AI responses. The emotional tone of their reflections also evolved, reflecting satisfaction derived from their newfound competence and mastery over the learning process.

I think AI indeed make a favour in my grammar. Without AI, I could not be aware of these grammar mistakes I made. AI reminds me and I will correct the mistakes consciously in future writing. I think this is the most useful help for me. (Student 11, Session 3)

As noted earlier, special attention was given to Students 4, 5, 16, and 21. Students 4 and 5 completed the grammar exercises under a researcher's supervision and found the session more helpful than previous ones. Student 16, a more proficient writer, had no grammatical issues identified in her original essay. Following our suggestion, she used the AI tool for advice on improving her expressions and ultimately learnt advanced vocabulary, which enriched her learning experience. Student 21, however, remained focused on his self-perceived issue with logic rather than grammar. Despite having significant problems with article usage, he did not invest time in learning grammar.

#### **Stage 4: Reflective Personalisation**

The final session revealed students' ability to critically evaluate and adapt the AI-assisted learning method to their individual needs. Without structured instructions, half adhered to the trained steps, while the other half innovated. Some adopted a hybrid approach, combining AI feedback with traditional resources like grammar books, achieving what one described as 'the best results yet'. Others experimented with unconventional yet insightful methods. A student used prompts in Chinese to interact with the AI, finding that phrasing questions in their native language yielded clearer explanations of complex grammar rules. Another employed a 'question ladder' technique, starting with broad AI queries (e.g., 'Explain past perfect tense') and progressively narrowing them (e.g., 'Compare past perfect and simple past with examples from my essay'), deepening their understanding. These adaptations highlighted learners transitioning from passive tool users to active, strategic learners, as they thoughtfully evaluated what worked best for their individual learning processes.

This time I made some innovations: First, I gave the original paragraph to the AI to analyse it from the perspective of language (vocabulary, grammar); then, I directly asked it to rewrite the paragraph to make it more natural and clear; afterwards, I asked it why it made such changes; finally, based on the reasons for the changes, I asked it to create several exercises for practice (the exercises it provided were all very good: Vocabulary Expansion, Conciseness Challenge, Sentence Structure Variation, etc.). I used to rely entirely on AI, but now I know how to guide it. It feels like I've learnt how to learn. (Student 6, Session 4)

## **Discussion**

### **A Theory-GenAI Integrated Approach to EAP Grammar Instruction**

This study addresses three persistent gaps in current research and practice. First, while grammar remains fundamental to academic writing (e.g., Cullen, 2008; Evans & Green, 2007), systematic instruction in EAP contexts has long been constrained by time limitations (Smith, 2022) and pedagogical priorities (Fedorova, 2022), creating a pedagogical vacuum for learners struggling with foundational skills. Second, existing GLS and related frameworks (Oxford, 2017; Pawlak, 2013), though theoretically robust, lack concrete strategy-based interventions to foster autonomous learning (Pawlak, 2019). Third, while GenAI tools show promise for L2 writing (e.g., Vo & Nguyen, 2024; Zadorozhnyy & Lai, 2024), their application has predominantly focused on surface-level error correction rather than deeper grammar acquisition processes (Kohnke, 2024). Our study bridges these gaps by integrating Pawlak's (2013) GLS framework with generative AI technology, creating a novel workflow that moves beyond current reliance on corrective feedback to scaffold grammar acquisition, providing the 'glass slipper' to elevate grammar from its Cinderella status in EAP contexts.

Theoretically, this integrated workflow demonstrates how cognitive strategies and metacognitive regulation, two core GLS strategies, can be systematically scaffolded through AI's interactive

capabilities—a previously unexplored synergy in GLS research. Additionally, this integration helps transform GenAI from a reliable learning tool to a cognitive partner through sequenced interventions. Practically, this study proposes a new pedagogical model for time-constrained EAP classes that addresses grammar deficits and enables learners to progress from initial reliance on error identification to strategic and critical adaptation with minimal teacher intervention.

### **Efficacy of the Integrated Approach**

Our experimental results demonstrate the effectiveness of this integrated approach, with students showing significant improvement in grammar acquisition, measured by both grammatical accuracy and complexity. The overall cohort achieved a mean score increase of 3.86 points (76.52 to 80.38), from the high end of the Satisfactory band (70–80) to the lower Good tier (80–90) on our assessment rubric. This progress is remarkable given merely four weeks of intervention.

These gains resulted from the alignment of the integrated approach with established GLS principles (Pawlak, 2013), particularly cognitive and metacognitive strategies, and were operationalised and facilitated by GenAI's unique capabilities. The following dual mechanisms explain the outcomes.

On the one hand, this GenAI-assisted intervention incorporates and activates all four subcategories of cognitive strategies (Pawlak, 2013), facilitating an acquisition process from superficial error identification to rule internalisation and eventual production (Ellis, 2015). Communication-based learning (COG 1) laid the groundwork for the entire acquisition process from grammar comprehension to production, and AI's responses not only provided immediate corrective feedback (COG 4) but also enabled students to generate targeted exercises (COG 2) and apply rules in rewriting (COG 3). Beginning with COG 4, 'corrective feedback processing', described by Pawlak (2013) as a process of 'listening carefully to the teacher', traditional classrooms, in reality, struggle to provide individualised feedback (Mahapatra, 2024). This gap is now addressed by GenAI, which functions as a personal tutor, offering personalised, adaptive (Xiao & Zhi, 2023), real-time, need-based (Guo et al., 2022) feedback. Subsequently, COG 2 stresses developing explicit grammar knowledge by understanding grammar rules and analysing examples. AI-generated exercises and relevant explanations not only reinforce explicit grammar knowledge but also target individual specific errors. This also echoes previous claims that ChatGPT can enhance language abilities by offering tailored practice materials and detailed explanations (Kasneci et al., 2023), assisting in the acquisition of specific knowledge (vocabulary, writing, listening, and grammar) (Huang et al., 2022), though these studies did not demonstrate 'how'. Finally, COG 3 involves developing implicit grammar knowledge through producing sentences in meaningful contexts. When students rewrite their essays, AI shifts from a 'feedback provider' to an 'error-awareness trigger', reminding learners to actively recall and apply rules they have just learnt, with 'attention and consciousness' (Ellis, 1995). Such applications help students identify and rectify common errors, ultimately fostering grammar proficiency (Arora & Mehta, 2023) and enabling immersive grammar acquisition (Loncar et al., 2021).

On the other hand, the GenAI-assisted reflective component (e.g., evaluating each session and finally adjusting prompts in the last session) technologically fostered metacognitive strategies (META), which Pawlak (2013) defines as planning, monitoring, and evaluating learning. During each session, students' metacognitive reflection (e.g., 'exercises are now more helpful') directly informed their use of cognitive strategies (e.g., prompt refinement), demonstrating metacognitive strategies' 'supervisory role' (Pawlak, 2018). This finding is in tandem with recent studies confirming the enhancement of metacognitive strategies through interacting with GenAI in ESL writing (e.g., Teng, 2024).

This approach's efficacy may provide some pedagogical implications. First, results support Alnufaie and Alzahrani's (2024) argument that GLS cannot be isolated from instruction. While GLS were traditionally learner-driven, the adoption of GenAI now systematises them within pedagogy, operationalising them as deliberate, structured processes transforming abstract strategies into actionable, scaffolded learning steps (Oxford, 2017). Crucially, this does not diminish the teacher's role; instead, it emphasises teachers' role

in strategy-based interventions (Pawlak, 2019) and also redefines their focus from repetitive correction to strategic mentoring.

More specifically, these strategies are more effective when teachers integrate them within specific contexts and purposeful learning tasks (Pawlak, 2018; Zhou, 2009); rather than leaving students to independently develop metacognitive and cognitive strategies, teachers could design GenAI-embedded, step-by-step, actionable learning routines, such as submitting drafts to AI for feedback, reflecting on errors using guided prompts, and revising drafts using both AI and teacher feedback. These instructive learning strategies may reinforce long-term acquisition and may also address current concerns about ‘over-reliance’ and ‘academic integrity’ (e.g., Kasneci et al., 2023; Yan, 2023).

## The Learner’s Journey

Our qualitative findings reveal a four-stage developmental trajectory in students’ engagement with the AI-assisted learning process. Initially characterised by ‘enthusiastic exploration’, students then progress through phases of ‘critical engagement’ and ‘empowered proficiency’ to eventual ‘reflective personalisation’. This finding extends prior research on general AI acceptance (e.g., Xiao & Zhi, 2023) or students’ static perceptions of AI tools (e.g., Guo et al., 2022) in ESL contexts by capturing the dynamic, non-linear nature of learner-AI interaction.

This stage model reflects the effectiveness of structured prompt-engineering training as well as iterative refinement based on learners’ evolving needs. Initially, students’ excitement stemmed from their unfamiliarity with step-by-step, enquiry-based interactions with AI; most had previously used AI tools only for direct essay corrections rather than guided learning. These positive emotional responses include enjoyment (Kohnke, 2023), motivation, increased confidence, and willingness to communicate (Zadorozhnyy & Lai, 2024). While mixed perceptions emerged in the second stage due to unsatisfying AI outputs, this was mitigated through instructor-supported prompt adjustments, echoing findings that ChatGPT performs more effectively when given specific instructions (Xiao & Zhi, 2023) and that the efficacy of GenAI relies partly on skilled prompt engineering (Strobelt et al., 2023). By the third stage, students began proficiently applying learnt strategies with more tailored prompts, and by the final stage, without further instructor guidance, many had developed personalised adaptations, demonstrating metacognitive regulation and autonomous learning (Pawlak, 2018). This process shows that technology fosters learner motivation and autonomy, enabling students to take control of their learning. This also accomplished Krashen’s (1982, cited in Fedorova, 2022) assertion that teachers should facilitate natural language acquisition by gradually transferring learning responsibility to students.

These findings yield actionable implications for pedagogy. Early intervention stages require systematic scaffolding, particularly in prompt engineering, such as explicit training in formulating targeted queries (Xiao & Zhi, 2023). These prompts should also be continuously refined based on learners’ needs. While later stages should emphasise learner autonomy, as Pawlak (2019) and Oxford (2017) claim, strategy-based interventions need to foster self-regulated, autonomous acquisition and long-term efficiency. This also holds particular significance in light of the diminishing emphasis on grammar instruction within EAP curricula across various contexts. The potential of GenAI in facilitating students’ self-learning beyond the classroom can in turn help decrease writing teachers’ workload within traditional classroom settings (Farrokhnia et al., 2023).

## Conclusion

This study resolves the persistent ‘Cinderella status’ of EAP grammar instruction by developing a theoretically grounded GenAI integration workflow. Our six-week intervention demonstrates that scaffolded AI interactions can enhance grammar acquisition, and qualitative findings reveal a four-stage developmental trajectory in technology adoption from following step-by-step prompts provided by teachers to greater learner autonomy. These outcomes advance GLS theory by operationalising Pawlak’s

(2013) strategies through AI's capabilities, contributing to ongoing investigations into how AI can support language learning and providing valuable insights for EAP teachers.

Nonetheless, several limitations must be acknowledged. First, the four-week intervention period may have been insufficient to fully assess grammar acquisition progress and its sustainability over time. Second, varying English proficiency levels amongst participants introduced heterogeneity in learning needs that the current study design could not address systematically. For future research, longitudinal studies tracking sustained grammar improvement beyond short-term interventions are recommended. Additionally, investigating the impact of GenAI tools on diverse learner profiles could provide further insights, allowing for the development of tailored interventions that maximise individual learning benefits.

## Acknowledgements

We would like to give special thanks to Professor Xiaojun Zhang who contributed a number of inspiring ideas to this research project. We are grateful to our students for their contributions as well as the comments from anonymous reviewers.

## Funding

This research has been supported by:

1. 2024 Fundamental Research Project for Universities of the Education Department of Liaoning Province, China. Grant No.: LJ112410173035.
2. 2022 Project under the 14th Five-Year Plan for Educational Science in Liaoning Province, China. Grant No.: JG22DB236.

## References

- Alnufaie, M. R., & Alzahrani, I. H. (2024). EFL grammar learning strategy use: Utilizing grammar learning strategy inventory in an Arabic context. *The Electronic Journal for English as a Second Language*, 27(4), 1–15. <https://doi.org/10.55593/ej.27108a6>
- Arora, A., & Mehta, N. K. (2023). Enhancing writing skills through ChatGPT: An experimental study in the context of ergonomics. *Speech and Context*, 1(16), 93–103. <https://zenodo.org/records/10452223>
- Barnard, R., & Scampton, D. (2008). Teaching grammar: A survey of EAP teachers in New Zealand. *New Zealand Studies in Applied Linguistics*, 14(2), 59–82.
- Barrot, J. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, Article 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Cahyani, R., Abdullah, M., & Komara, C. (2022). The investigation of English grammar learning strategies on high, middle, and low achievers' students in Indonesia. *ELLTER Journal*, 3(2), 54–63. <https://doi.org/10.22236/ellter.v3i2.10063>
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. SAGE Publications.
- Chen, Z., Chen, W., Jia, J., & Le, H. (2022). Exploring AWE-supported writing process: An activity theory perspective. *Language Learning & Technology*, 26(2), 129–148. <https://doi.org/10.125/73482>

- Cohen, A. D., & Pinilla-Herrera, A. (2010). Communicating grammatically: Constructing a learner strategies website for Spanish. In T. Kao & Y. Lin (Eds.), *A new look at language teaching and testing: English as subject and vehicle* (pp. 63–83). The Language Training and Testing Center.
- Cullen, R. (2008). Teaching grammar as a liberating force. *ELT Journal*, 62(3), 221–230. <https://doi.org/10.1093/elt/ccm042>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Ellis, R. (1995). Interpretation tasks for grammar teaching. *TESOL Quarterly*, 29(1), 87–105. <https://doi.org/10.2307/3587806>
- Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40(1), 83–107. <https://doi.org/10.2307/40264512>
- Ellis, R. (2015). *Understanding second language acquisition* (2nd ed.). Oxford University Press.
- Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20, Article 57. <https://doi.org/10.1186/s41239-023-00425-2>
- Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, 6(1), 3–17. <https://doi.org/10.1016/j.jeap.2006.11.005>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 460–474. <https://doi.org/10.1080/14703297.2023.2195846>
- Fedorova, N. (2022). EAP learners' attitudes towards grammar instruction. *International Journal of English for Academic Purposes*, 2022(Autumn), 109–137. <https://doi.org/10.3828/ijeap.2021.13>
- Guo, K., Wang, J., & Chu, S. K. W. (2022). Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing*, 54, Article 100666. <https://doi.org/10.1016/j.asw.2022.100666>
- Hinkel, E. (2004). *Teaching academic ESL writing: Practical techniques in vocabulary and grammar*. Lawrence Erlbaum Associates.
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38, 237–257. <https://doi.org/10.1111/jcal.12610>
- John, P., & Woll, N. (2020). Using grammar checkers in an ESL context. *CALICO Journal*, 37(2), 193–196. <https://doi.org/10.1558/cj.36523>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kohnke, L. (2023). L2 learners' perceptions of a chatbot as a potential independent language learning tool. *International Journal of Mobile Learning and Organisation*, 17(1–2), 214–226. <https://doi.org/10.1504/IJMLO.2023.128339>
- Kohnke, L. (2024). Exploring EAP students' perceptions of GenAI and traditional grammar-checking tools for language learning. *Computers and Education: Artificial Intelligence*, 7, Article 100279. <https://doi.org/10.1016/j.caeai.2024.100279>
- Krashen, S. D. (2003). *Explorations in language acquisition and use: The Taipei lectures*. Heinemann.

- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners. *Applied Linguistics*, 27, 590–619. <https://doi.org/10.1093/applin/aml029>
- Loncar, M., Schams, W., & Liang, J. S. (2021). Multiple technologies, multiple sources: Trends and analyses of the literature on technology-mediated feedback for L2 English writing published from 2015–2019. *Computer Assisted Language Learning*, 36(4), 722–784. <https://doi-org.eproxy.lib.hku.hk/10.1080/09588221.2021.1943452>
- Mathur, M., & Mahapatra, S. (2022). Impact of ePortfolio assessment as an instructional strategy on students' academic speaking skills: An experimental study. *CALL-EJ*, 23(3), 1–23.
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11, Article 9. <https://doi.org/10.1186/s40561-024-00295-9>
- Müller, A., & Han, W. (2022). *IELTS Writing band scores 5.5–7.5: Grammatical error rates, stakeholder perceptions, and risk. IELTS Research Reports Online Series, No. 1/22*. British Council.
- Nakachi, K. (2021). Grammar learning strategy use by English-major university students: An investigation with grammar learning strategy inventory (GLSI). *Nagoya JALT Journal*, 2(1), 53–70. <https://doi.org/10.37546/jaltchap.nagoya2.1-3>
- Nguyen, T. M. P. (2010). EAP and LLNP students: Common problems and strategies. *English Language Teaching*, 3(2), 17–21. <https://doi.org/10.5539/elt.v3n2p17>
- Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578. <https://doi.org/10.1093/applin/amp044>
- Oxford, R. L., Lee, K. R., & Park, G. (2007). L2 grammar strategies: The second Cinderella and beyond. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies* (pp. 117–139). Oxford University Press.
- Oxford, R. L. (2017). *Teaching and researching language learning strategies: Self-regulation in context*. Routledge.
- Pawlak, M. (2012). Instructional mode and the use of grammar learning strategies. In M. Pawlak (Ed.), *New perspectives on individual differences in language learning and teaching* (pp. 263–287). Springer. [https://doi.org/10.1007/978-3-642-20850-8\\_17](https://doi.org/10.1007/978-3-642-20850-8_17)
- Pawlak, M. (2013). Researching grammar learning strategies: Combining the macro- and micro-perspective. In L. Salski, W. Szubko-Sitarek, & J. Majer (Eds.), *Perspectives on foreign language learning* (pp. 191–220). University of Łódź Press.
- Pawlak, M. (2018). Grammar Learning Strategy Inventory (GLSI): Another look. *Studies in Second Language Learning and Teaching*, 8(2), 351–379. <https://doi.org/10.14746/ssllt.2018.8.2.8>
- Pawlak, M. (2019). Grammar learning strategies as a key to mastering second language grammar: A research agenda. *Language Teaching*, 53, 358–370. <https://doi.org/10.1017/S0261444819000314>
- Polio, C., & Shea, M. C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, 26, 10–27. <https://doi.org/10.1016/j.jslw.2014.09.003>
- Shi, H., & Aryadoust, V. (2023). A systematic review of automated writing evaluation systems. *Education and Information Technologies*, 28(1), 771–795. <https://doi.org/10.1007/s10639-022-11200-7>

- Smith, S. C. H. (2022) *What is EAP?* EAP Foundation.
- Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research*. SAGE Publications.
- Strobelt, H., Webson, A., Sanh, V., Hoover, B., Beyer, J., Pfister, H., & Rush, A. M. (2023). Interactive and visual prompt engineering for ad-hoc task adaptation with large language models. *IEEE Transactions on Visualization and Computer Graphics*, 29(1), 1146–1156. <https://doi.ieeecomputersociety.org/10.1109/TVCG.2022.3209479>
- Taskiran, A., & Goksel, N. (2022). Automated feedback and teacher feedback: Writing achievement in learning English as a foreign language at a distance. *Turkish Online Journal of Distance Education*, 23(2), 120–139. <https://doi.org/10.17718/tojde.1096260>
- Teng, M. F. (2024). Metacognitive awareness and EFL learners' perceptions and experiences in utilising ChatGPT for writing feedback. *European Journal of Education*, 60(1), Article e12811. <https://doi.org/10.1111/ejed.12811>
- Vo, A., & Nguyen, H. (2024). Generative artificial intelligence and ChatGPT in language learning: EFL students' perceptions of technology acceptance. *Journal of University Teaching and Learning Practice*, 21(6), 1–19. <https://doi.org/10.53761/fr1rkj58>
- Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, 8(3), Article 212. <https://doi.org/10.3390/languages8030212>
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28, 13943–13967. <https://doi.org/10.1007/s10639-023-11742-4>
- Yan, D. (2024). Comparing individual vs. collaborative processing of ChatGPT-generated feedback: Effects on L2 writing task improvement and learning. *Language Learning & Technology*, 28(1), 1–19. <https://doi.org/10125/73597>
- Zadorozhnyy, A. & Lai, W. Y. (2024). ChatGPT and L2 written communication: A game-changer or just another tool? *Languages*, 9(1), Article 5. <https://doi.org/10.3390/languages9010005>
- Zhou, A. A. (2009). What adult ESL learners say about improving grammar and vocabulary in their writing for academic purposes. *Language Awareness*, 18(1), 31–46. <https://doi-org.eproxy.lib.hku.hk/10.1080/09658410802307923>
- Zhou, S., Zhao, S., & Groves, M. (2022). Towards a digital bilingualism? Students' use of machine translation in international higher education. *Journal of English for Academic Purposes*, 60, Article 101193. <https://doi.org/10.1016/j.jeap.2022.101193>

### Appendix A. Pretest and Posttest Worksheet

Pre-Test

Student Number \_\_\_\_\_

Task: Write a short essay to the following essay question. Your essay should be about 350 words long consisting of an introduction, body paragraph, and conclusion. Do not make use of any assistance like computers or mobile phones.

Should medicine be tested on animals?

---

---

---

---

---

---

Post-Test

Student Number \_\_\_\_\_

Task: Write a short essay to the following essay question. Your essay should be about 350 words long consisting of an introduction, body paragraph, and conclusion. Do not make use of any assistance like computers or mobile phones.

Should students be allowed to use AI in their studies?

---

---

---

---

---

---

## Appendix B. Grading Rubrics (Grammar Focus)

Level	Score range	Description	Accuracy Criteria	Complexity Criteria
<b>Excellent</b>	90–100	Grammar is consistently accurate with precise use of tenses, subject-verb agreement, and sentence structure. Demonstrates advanced use of complex sentences and minimal errors that are rare and do not affect meaning.	<ul style="list-style-type: none"> <li>• <math>\leq 1</math> <b>major error</b>/100 words</li> <li>• No meaning obstruction</li> <li>• Consistent control of advanced forms (e.g., perfect tenses, mood)</li> </ul>	<ul style="list-style-type: none"> <li>• <math>\geq 3</math> <b>error-free</b> complex structures/100 words</li> <li>• Diverse clause types (e.g., relative, adverbial)</li> <li>• Effective use of passive/conditional constructions</li> </ul>
<b>Good</b>	80–90	Grammar is mostly accurate with occasional errors in complex sentences, such as incorrect use of conditionals or articles. Good variety of sentence structures with effective use of subordinate clauses; errors rarely obscure meaning.	<ul style="list-style-type: none"> <li>• <math>\leq 3</math> <b>major errors</b>/100 words</li> <li>• Occasional meaning ambiguities (localized)</li> <li>• Minor errors in complex forms</li> </ul>	<ul style="list-style-type: none"> <li>• 2-3 <b>error-free</b> complex structures/100 words</li> <li>• Moderate clause variety</li> <li>• Attempts at advanced forms with partial success</li> </ul>
<b>Satisfactory</b>	70–80	Grammar shows control of basic structures (e.g., simple and compound sentences), but errors in complex forms like passive voice or relative clauses are evident. Errors in prepositions or word forms occasionally affect meaning.	<ul style="list-style-type: none"> <li>• <math>\leq 5</math> <b>major errors</b>/100 words</li> <li>• Noticeable but manageable meaning issues</li> <li>• Errors in basic structures rare</li> </ul>	<ul style="list-style-type: none"> <li>• 1-2 <b>error-free</b> complex structures/100 words</li> <li>• Limited clause variety (e.g., simple subordination)</li> <li>• Occasional advanced forms with errors</li> </ul>
<b>Soft Fail</b>	60–70	Grammar is limited to simple structures (e.g., subject-verb-object); frequent errors in tense consistency, plurals, or word order. Attempts at complex sentences often result in significant errors that sometimes obscure meaning.	<ul style="list-style-type: none"> <li>• Frequent errors (<math>\leq 8</math>/100 words)</li> <li>• Meaning occasionally obscured</li> <li>• Reliance on simple structures</li> </ul>	<ul style="list-style-type: none"> <li>• <math>\leq 1</math> <b>error-free</b> complex structure/100 words</li> <li>• Predominantly simple sentences</li> <li>• Attempts at complexity often erroneous</li> </ul>

Level	Score range	Description	Accuracy Criteria	Complexity Criteria
<b>Hard Fail</b>	50–60	Frequent and noticeable errors in basic structures, such as missing auxiliaries, incorrect verb forms, or word endings. Meaning is often unclear, and reliance on fragmented or run-on sentences is evident.	<ul style="list-style-type: none"> <li>• Persistent errors (&gt;8/100 words)</li> <li>• Meaning frequently unclear</li> <li>• Basic grammar inconsistencies</li> </ul>	<ul style="list-style-type: none"> <li>• No sustained complex structures</li> <li>• Fragmented or run-on sentences</li> <li>• Minimal subordination</li> </ul>
<b>Unacceptable</b>	< 50	Severe and persistent grammatical errors in even simple forms, such as subject-verb agreement or sentence fragments. Structures are disorganized, and errors consistently impede comprehension	<ul style="list-style-type: none"> <li>• Severe errors in basic syntax</li> <li>• Meaning consistently impeded</li> <li>• No control of core grammar</li> </ul>	<ul style="list-style-type: none"> <li>• No evidence of complex structures</li> <li>• Disorganized sentence boundaries</li> </ul>

## Appendix C. Instruction Sheets in Each Experimental Session (with Changes Highlighted)

### AI Grammar Learning Workshop

#### Student Instruction Sheet –Session 1

---

Step 1)-Write a paragraph (100 words) on the following topic:

*Discuss the impact of technology on education.*

Step 2)-Copy and paste the paragraph into your AI browser with following prompt:

*“Look at the following paragraph below and identify two major issues (only two) in terms of language, grammar, and phrasing.”*

Step 3)-Use the feedback given in Step 2) and insert the following prompt:

*“Give me 10 exercises (followed by an answer key) focusing on (first issue) and (second issue) respectively. The level of language must be suitable for a first-year university student.”*

Step 4)-Do the exercises generated by the AI using a pen and paper. Review accordingly.

Step 5)-Now rewrite and revise your paragraph written in Step 1). Try and improve the language and grammar as identified by the AI in Step 2).

Step 6)- Copy and paste the revised paragraph into your AI browser and use the following prompt:

*“Look at the following paragraph below and review it in terms of (first issue) & (second issue).”*

Step 7)-Look at the feedback given by AI and compare with the feedback given in Step 2).

Step 8)-Write a 50-word reflective paragraph to reflect on your learning experience today (focus on the strengths and weakness of this learning method, and feel free to express your feelings on other aspects).

\*Thanks for your cooperation today.

At the end of this session, please copy and paste your entire chat history with AI, along with your reflection into one Word Document, name it as follows: <Session 1\_Student No.>.

# AI Grammar Learning Workshop

## Student Instruction Sheet –Session 2

---

Step 1)-Write a paragraph (100 words) on the following topic:

*How does technological development affect the relationship between people?*

Step 2)-Copy and paste the paragraph into your AI browser with following prompt:

*“Look at the following paragraph below and identify two major issues (only two) in terms of language, grammar, and phrasing.”*

**Attention : If the issues are not specific, ask for specific examples in terms of grammar using the following prompt:**

*“Can you give me specific examples which can be used to generate language exercises?”*

Step 3)-Use the feedback given in Step 2) and insert the following prompt:

*“Give me 10 exercises (followed by an answer key) focusing on (first issue) and (second issue) respectively. The level of language must be suitable for IELTS level.”*

Step 4)-Do the exercises generated by the AI using a pen and paper. Review accordingly.

Step 5)-Now rewrite and revise your paragraph written in Step 1). Try and improve the language and grammar as identified by the AI in Step 2).

Step 6)- Copy and paste the revised paragraph into your AI browser and use the following prompt:

*“Look at the following paragraph below and review it in terms of (first issue) & (second issue).”*

Step 7)-Look at the feedback given by AI and compare with the feedback given in Step 2).

Step 8)-Write a 50-word reflective paragraph to reflect on your learning experience today (focus on the strengths and weakness of this learning method, and feel free to express your feelings on other aspects).

\*Thanks for your cooperation today.

At the end of this session, please copy and paste your entire chat history with AI, along with your reflection into one Word Document, name it as follows: <Session 2\_Student No.>.

# AI Grammar Learning Workshop

## Student Instruction Sheet –Session 3

---

Step 1)-Write a paragraph (**70 words**) on the following topic:

*What are the environmental impacts of not taking public transport?*

Step 2)-Copy and paste the paragraph into your AI browser with following prompt:

*“Look at the following paragraph below and identify **one major issue(only one)** in terms of language, grammar, and phrasing.”*

**Attention : If the issues are not specific, ask for specific examples in terms of grammar using the following prompt:**

*“Can you give me specific examples which can be used to generate language exercises?”*

Step 3)-Use the feedback given in Step 2) and insert the following prompt:

*“Give me 10 exercises (followed by an answer key) focusing on **(main issue)**. The level of language must be suitable for IELTS level.”*

Step 4)-Do the exercises generated by the AI using a pen and paper. Review accordingly.

Step 5)-Now rewrite and revise your paragraph written in Step 1). Try and improve the language and grammar as identified by the AI in Step 2).

Step 6)- Copy and paste the revised paragraph into your AI browser and use the following prompt:

*“Look at the following paragraph below and review it in terms of **(main issue)**. .”*

Step 7)-Look at the feedback given by AI and compare with the feedback given in Step 2).

Step 8)-Write a 50-word reflective paragraph to reflect on your learning experience today (focus on the strengths and weakness of this learning method, and feel free to express your feelings on other aspects).

\*Thanks for your cooperation today.  
At the end of this session, please copy and paste your entire chat history with AI, along with your reflection into one Word Document, name it as follows: <Session 3\_Student No.>.

# AI Grammar Learning Workshop

## Student Instruction Sheet –Session 4

---

Step 1)-Write a paragraph (**100 words**) on the following topic:

*What are the environmental impacts of not taking public transport?*

NOW feel free to work with AI with a focus on improving your grammar.

After learning, rewrite and revise your paragraph written in Step 1). Try and improve the language and grammar .

Write a 50-word reflective paragraph to reflect on your learning experience today (focus on the strengths and weakness of this learning method, and feel free to express your feelings on other aspects).

\*Thanks for your cooperation today.  
At the end of this session, please copy and paste your entire chat history with AI, along with your reflection into one Word Document, name it as follows: <Session 4\_Student No.>.

## About the Authors

**Shu Zhou**, Director and Senior Lecturer at Surrey International Institute (DUFE), holds an MA in TESOL from Durham University and is pursuing a PhD at the University of Liverpool. With over 10 years of EAP teaching experience, her research focuses on EAP pedagogy and AI-assisted language learning.

**E-mail:** [s.zhou@surrey.ac.uk](mailto:s.zhou@surrey.ac.uk)

**ORCID:** <https://orcid.org/0009-0004-6371-8025>)

**Gerhardus D. du Preez**, Namibian/South African EAP Senior Lecturer at Surrey International Institute (DUFE). He holds an MA in Medieval English Literature (Stellenbosch University), and is currently pursuing a PhD. His research focuses on EAP writing pedagogy, while also contributing as a columnist for a Namibian Afrikaans newspaper.