INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT

OPEN ACCESS JOURNAL

Vol 15, Issue 2, (2026) E-ISSN: 2226-6348

The Role of Generative AI Tools in English Writing Teaching

Jiang Yuexi, Amelia Binti Alias & Nurfaradilla Mohamad Nasri

Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

Email:P145939@siswa.ukm.edu.my, amelia.alias@ukm.edu.my, nurfaradila@ukm.edu.my

DOI Link: http://dx.doi.org/10.6007/IJARPED/v15-i2/28081

Published Online: 18 April 2026

Abstract

The rapid developmenat of generative artificial intelligence (AI) has introduced new possibilities and challenges for English writing instruction in higher education. This study investigates the role of generative AI tools, particularly ChatGPT and Quillbot, in university-level English writing teaching by comparing them with Grammarly, teacher feedback, and peer review. Using a comparative case study design, the research combines horizontal analysis across different feedback sources with vertical analysis of the evolution of writing support technologies. Data were collected from undergraduate students' drafts, revisions, reflective journals, interviews, and feedback records. The findings indicate that generative AI tools provide immediate, accessible, and scalable support, helping students with idea generation, paraphrasing, discourse organization, and revision. However, their use also raises concerns about over-reliance, academic integrity, authorship, and the weakening of students' critical engagement with writing. Compared with AI-generated feedback, teacher feedback remains more effective in addressing higher-order issues such as argumentation, rhetorical development, and disciplinary conventions, while peer review offers collaborative but uneven support. The study argues that generative AI should not replace human feedback, but should be integrated responsibly within writing pedagogy as a supplementary scaffold. By situating AI within the frameworks of the Zone of Proximal Development, the process writing approach, and critical pedagogy, this study contributes to a more balanced understanding of how AI may reshape writing instruction, teacher roles, and student agency in higher education.

Keywords: Generative AI, AI in Education, English Writing Teaching, AI Writing Tools, Language Learning Technology, NLP, Writing Skill Development, Automated Feedback, Academic Writing Support, Personalized Learning

Introduction

Research Background

The rapid development of artificial intelligence (AI) has transformed many aspects of education, including the teaching of English writing in higher education. In second language writing instruction, feedback has long been regarded as a crucial element in helping learners improve linguistic accuracy, develop arguments, and strengthen awareness of academic conventions (Hyland & Hyland, 2006). Traditionally, this support has been provided through teacher feedback, peer review, and automated writing evaluation tools such as Grammarly.

While these approaches remain valuable, they often face limitations in timeliness, personalization, and scalability, especially in contexts where teachers are required to support large groups of students.

The emergence of generative AI tools, particularly ChatGPT and Quillbot, has introduced a new stage in the evolution of writing support technologies. Unlike earlier systems that mainly focused on error detection and surface-level correction, generative AI can assist learners with idea generation, paraphrasing, text expansion, and discourse organization (Kasneci et al., 2023). This shift suggests that AI is no longer used only as a corrective device, but increasingly as a participant in the writing process itself. For learners of English as a foreign language (EFL), such tools may offer immediate and individualized support, allowing students to experiment with language, receive continuous feedback, and revise their writing more independently.

At the same time, the integration of generative AI into writing instruction has raised important pedagogical and ethical concerns. Recent studies have highlighted risks related to academic integrity, over-reliance on machine-generated output, and the possible weakening of students' critical and creative engagement with writing (Trachtenberg, 2023; Lund et al., 2023). These concerns suggest that the role of generative AI in writing classrooms cannot be understood simply in terms of efficiency or convenience. Rather, it must be examined in relation to broader questions of teaching practice, learner agency, and the changing responsibilities of teachers in digitally mediated educational settings.

Research Gap and Significance

Although research on AI-assisted writing has expanded rapidly, several important gaps remain. First, much of the existing literature focuses on a single tool at a time, such as ChatGPT, Grammarly, or automated writing evaluation systems, rather than comparing multiple feedback sources within one analytical framework. Second, many studies emphasize surface-level outcomes, including grammar correction and lexical improvement, while paying less attention to higher-order concerns such as argumentation, rhetorical development, and academic integrity. Third, AI tools are often discussed as supplementary technologies, rather than being examined in relation to established pedagogical frameworks of writing instruction.

This study addresses these gaps by comparing generative AI tools, especially ChatGPT and Quillbot, with Grammarly, teacher feedback, and peer review in university-level English writing instruction. Its significance lies not only in evaluating the strengths and limitations of different feedback modes, but also in examining how generative AI may reshape writing pedagogy, teacher roles, and students' engagement with academic writing. In this way, the study contributes to current discussions on the responsible integration of AI into higher education and highlights the need to balance technological innovation with pedagogical and ethical reflection.

Research Aims and Questions

Against this background, the present study investigates the role of generative AI tools in tertiary-level English writing instruction. It explores how tools such as ChatGPT and Quillbot compare with Grammarly, teacher feedback, and peer review in supporting or hindering students' writing development. In addition, it examines the opportunities and risks associated

with integrating generative AI into English writing pedagogy, with particular attention to teaching models, teacher roles, and academic integrity.

These aims are articulated through the following research questions:

1. How do generative AI tools such as ChatGPT and Quillbot compare with Grammarly, teacher feedback, and peer review in supporting or hindering students' writing performance?
2. What opportunities and risks arise from integrating generative AI into English writing pedagogy in higher education?
3. What implications does the development of generative AI hold for teaching models, teacher roles, and academic integrity in university writing instruction?

Novelty and Contribution of the Study

The novelty of this study lies in its comparative examination of multiple feedback sources within a single framework, including ChatGPT, Quillbot, Grammarly, teacher feedback, and peer review. Rather than treating generative AI as an isolated technological tool, the study situates it within broader pedagogical practices of English writing instruction and analyzes both its instructional potential and its ethical implications. Its contribution to the field of social sciences lies in showing how emerging AI technologies may reshape educational relationships, teacher authority, student agency, and academic integrity in higher education. In doing so, the study offers a more balanced understanding of AI in writing pedagogy by connecting technological change with social, pedagogical, and institutional dimensions.

Structure of the Essay

This essay is organized into six chapters. Following this introduction, Chapter 2 reviews existing literature on English writing pedagogy, automated writing tools, and the emergence of generative AI in education. Chapter 3 presents the theoretical framework, drawing on Vygotsky's Zone of Proximal Development, the process writing approach, and perspectives from critical pedagogy. Chapter 4 outlines the methodology, combining horizontal comparisons among feedback modes and vertical comparisons across the historical development of writing technologies. Chapter 5 presents and discusses the findings in relation to the research questions, highlighting both opportunities and risks. Finally, Chapter 6 concludes the study, offering pedagogical implications, reflections on limitations, and directions for future research.

By situating generative AI tools within the continuum of writing feedback practices, this study contributes to the ongoing scholarly conversation on the role of technology in language education. It aims to balance optimism about technological innovation with critical reflection on ethical and pedagogical challenges, offering insights for the informed and responsible integration of AI into higher education writing instruction.

Literature Review

Current Study of English Writing Instruction in Higher Education

The development of English writing skills has a rich heritage, especially in higher education. In the context of English as a Second or Foreign Language writing instruction, it is difficult to provide accurate language, good summaries, and well-developed arguments. Despite many years of studies on English teaching, a lot of university students have difficulties with academic writing due to the lack of credible models, weak feedback, and high cognitive

load in a foreign language discussion (Ferris, 2003). The problems of this magnitude are particularly serious in the lexicon of science in which English is used as the lingua franca; thus, international students, who may have urgent tasks to publish their research articles in academic journals or to achieve academic progress, will be under pressure (Flowerdew, 2015).

In the list of influential university writing approaches, the process writing one, which suggests recursive stages of brainstorming, first drafting, multiple revising, and last editing, according to Matsuda (2003), is the most popular. This method is consistent with learning principles, which view the use of metacognition as a way to help students understand writing better as a developmental process. Studies have proven that a process-focused method is more helpful in raising students' ability to self-regulate as well as make more substantive amendments to the writing (Seow, 2002). However, this method has challenges in large class sizes.

Regular assessment was considered as a keystone of successful writing by numerous scholars. Teacher feedback, for example, charts the way for students in academic practices and language proficiency (Lee, 2017). However, research also suggests that students may view the feedback from teachers as imperative or perceive it as being tough to grasp, without which they may fail to fully utilize it (Hyland & Hyland, 2001). They can do this by incorporating a system of peer review, where they assess the other's content critically and reflect on their own (Rollinson, 2005). Co-operation has many benefits. Among them are the creation of new ideas and achievements in critical thinking, but all these advantages significantly depend on students' language proficiency level and the training of constructive comments by teachers (Liu & Edwards, 2018). In the past, scholars often had varying opinions on the value of using peer feedback tools due to issues of institutional culture and the willingness of the learners to practice meaning negotiation together. Presently conducted research also underpins the observation that, to some extent, the success of peer feedback operation can largely be determined by such factors.

Although these classical ways of communication usually take the lion's share, technology-enhanced feedback seems to become more demonstrative to higher learning. Criterion and Grammarly are both long-standing and well-used automation sighting AWE (Automated Writing Evaluation) tools, which give instant surface-level feedback on grammar, spelling, punctuation and style to the highest extent. Numerous surveys have found that Automated Writing Evaluation Systems (and they are believed to perform this function) can be valuable for improving students' awareness about writing errors and guiding language practice (Li et al., 2015; Ranalli et al., 2017). Yet one usually finds criticism focused on their shortfalls in regard to minimum standards of concern that mainly point out the ability of the same to address higher sentence grammar and (Dikli & Bleyle, 2014) coherence. Learning deficiencies can also occur due to the learners' excessive reliance on automatic grammar checkers, which makes their interest in applying additional strategies for editing not only more challenging but also gets impeded during the process (Link et al., 2014). Hence, technology cannot fully replace the teachers in the writing pedagogy, such as at the college level, which focuses on critical engagement and the development of discipline-specific conventions.

Moreover, researchers have claimed that writing pedagogy is more and more regarded from the viewpoint of academic literacy and socialization through disciplines (Lea & Street, 1998; Hyland, 2004). Students also lose access to some niches for academic literacy and discipline socialization (Lea & Street, 1998; Hyland, 2004). Therefore, according to them, on the college/university level, students do more sophisticated academic writing than before, and therefore preparing students to handle these complicated steerers is a crucial anti-plagiarism strategy. The change in approach also emphasizes the various features of language discourse and advocates for an all-inclusive approach towards developing the ability to participate meaningfully within a disciplinary community through language. The need for incorporating linguistic competence with higher-order thinking along with knowledge of academic norms heightened by the emergence of new technologies endorses the idea of making transformative instruction. On the other hand, such process-oriented practices of education are very demanding both with regard to the preparation of teachers and regarding the extended support that can hardly be ensured for everyone (Wingate, 2012).

In conclusion, writing in English curricula of universities has moved from focusing on the final product to process-oriented methods and literacy-based approaches controlled by many feedback techniques. Even with that, predicaments occur in attaining effectiveness and comprehensiveness at the same time, and in settling the issues of syntactic and writing comprehension. Such problems serve as a pedagogical background where AI tools have been developed just recently, contributing as developing tools to give feedback for the course, but also creating a threat of making academic dependency and damaging academic integrity. Subsequently, it appears unassailable for identifying the capacities and limitations of AI-assisted machine grading if the status of the existing writing instruction is not correctly assessed.

The Evolution of Automated Writing Assistance Tools

Enormous advancement has been noted in computer design supporting writing as evident through a series of transformations in the past four decades, starting from basic spell checkers to cutting-edge generative artificial intelligence systems. The pre-existing assistive tools of the 1980s and 1990s only managed with such basic problems as spelling or relatively simple (compared to grammar tracking) grammar checking. Applications developed to manage to fix surface errors fail to incorporate any additional compound of pedagogy as the weaknesses of the writing such as rhetoric, stylistic, and discourse levels were criticized (Ranalli, 2018). While early automation tools, in brief, enable to ease off the process, they, however, leave higher-order issues of writing instruction, and that is a bugbear in the face of the product-oriented writing instruction frameworks that gained prevalence in those years.

Recently added facilities such as an advanced version of Grammarly introduce new opportunities, where users can choose among a context-specific grammar checker, suggest vocabulary enhancements, or even give stylistic recommendations. Grammarly implements practices of machine learning and natural language processing (NLP) to catch missing grammar, point out stylistic errors, unclear writing, or prolixity. Besides, other studies have also shown that Grammarly was often used to check writing texts (Li et al., 2015). Though Grammarly is able to evoke a critical habit while correcting, it still lacks improvement in advanced writing skills, such as argumentation, genre awareness, and rhetorical organization.

Another AI-based tool widely used in writing is Quillbot. Unlike Grammarly, it aims at producing genuine and unique writing and compels users to replace and rewrite sentences creatively, which can lead to the improvement of vocabulary and stylistic complexity. Quillbot was also noted for building scholarly authorities. The extensive application of paraphrases can lead to the simplification and loss of semantic ties in students' own work to the extent that students use the automated reformulation more, thus reducing the creativity of students and their ability to grasp the notion behind the writing to the fullest extent.

Unlike other AI innovations, ChatGPT has the unique characteristic of aiding in the automated production of texts. In particular, this technology provides an automated conversation feature that engages users by generating lengthy and naturally occurring content that is relevant. Generative AI will be a new type of corrective tool because users do not need to submit drafts or outlines. Although this makes it possible to provide tailored comments, it also raises issues of originality and true authorship that must be addressed to uphold academic integrity (Cotton et al., 2024).

It might seem a little sluggish to go from spell-checkers to Grammarly, QuillBot, and ChatGPT. But as each tool is improved, it offers more features, such as the ability to teach a course that may eventually be entirely automated and localised language correction. This procedure reflects the overall divergence in the development of AI and NLP. In addition, it presents new opportunities and problems for academic writers working in higher education. Teachers need to think particularly about the formative actions teachers perform as part of feedback - and how to help students understand the ethical, social, and cognitive dimensions of writing.

Generative AI in Language Learning and Writing

The use of generative AI in language acquisition has been heavily deliberated. In the lines that follow, the views influenced by applied generative AI, and those inspired by criticism of some of its aspects, come to the fore. Some experts say these tools could significantly broaden the scope of writing help for ordinary students, while others argue that they carry risks of overreliance, ethical concerns, and increased cognitive burden.

Supportive Perspectives

The amount of research discussing the teaching strategies of generative AI programs, such as ChatGPT, is increasing. They have been continually demonstrated as a means to not only help students become more engaged but to reduce their writing anxiety, while also allowing for it to be personalised for the needs of the student. For instance, it was demonstrated by Zhai (2022) that ChatGPT could be successful in teaching students how to produce language in a stress-free environment. Learning in this case, therefore, is an iterative process where immediate feedback was provided without the kind of social pressure that students may experience while they are under the care of their teacher or their peers. Similarly, Lo et al. (2024) took a similar position, holding that through the use of avenues like AI-aided systems, the teacher's role changes from giving knowledge to a facilitator as the students are able with language and make many choices.

Generative AI has been positively reviewed for helping learners to overcome the fear of academic writing and the confidence to make their arguments. It is postulated in the Master

Plan that students' accomplishments in elaborating complex writing tasks will greatly increase if these students are supported with adequate feedback (El Weber & Windeatt, 2010). Notably, ChatGPT and related technology do just that by extending this principle to the entire community by offering large-scale, thorough, and realistic discussions by answering the learner's questions with complete instructions in the natural language syntax and even customers' quizes to mimic the tutoring system (Kasneci et al., 2023). Then, the AI system learns and adapts from learners' interactions to craft scaffolding materials for each personalized lesson. This scaffolding support may be related to Vygotsky's (1978) concept of Zone of Proximal Development (ZPD).

Critical Perspectives

Despite its benefits, there is a note of warning in the critical literature about the consequences of uncritically embracing generative AI in academia. An underlying issue is of intellectual integrity. ChatGPT, and other such platforms, can produce essays and assignments engineered to look slick and authentic—leaving teachers unable to tell the difference between student work and that of a computer (Trachtenberg, 2023). This muddies the waters of authorship, and there are ethical questions around plagiarism and whose work deserves fair credit.

There is a related critical view that brings to question the possibility of AI augmenting the thinking ability of humans and not obstructing or mitigating such ability. Once again, the danger is to abolish thinking itself. If instead of students doing logical work of creating ideas and structure of their writing using AI, they just crush the buttons and think the work is done, they'll acquire little academic writing skills except for the one to make sense of the source, organizing their ideas and constructing logical reasoning (Lund et al., 2023), (Bender et al., 2021). Moreover, the information to which an AI-trained language model will co-emit will reflect actual biases and prejudices that relate to this domain and thus might end up continuing to support the reproduction of stereotypes or preference to some dominant discourses. This concern are the most important in global higher education, where students of very different linguistic and cultural backgrounds may face normative bias that tend to be encoded within AI-generated English.

There are also pedagogical dangers of relying on AI tools. Overuse of paraphrasing and text generation tools might hinder students from participating in the recursive processes of drafting and revising, which is fundamental to process writing approach (Hyland, 2019). Rather than learning that writing is an iterative process of discovery, they will learn that writing is a chore spent tweaking AI output—dulling their development as language users.

In sum, the literature on generative AI in language learning is characterized by a productive yet uneasy tension between its promises of enhanced fluency, confidence, and accessibility, and the challenges it poses towards authenticity, integrity, and critical engagement. The Janus face of generative AI emphasises the importance of pedagogical frameworks that aim to optimise the benefits of AI, but at the same time to mitigate its concerns.

Research Gaps

This study of AI-assisted writing process outlines the features and limitations of each element for the support it offers, but it also shows that several areas remain insufficiently addressed. A major limitation is that most prior research investigates a single tool at a time, rather than comparing multiple sources of feedback in combination. This single-mode approach fails to reflect the reality of higher education, where students frequently shift between teacher comments, peer review, and various AI systems, often blending them in unpredictable ways.

Another gap lies in the focus of measurement. Much scholarship prioritizes surface-level outcomes such as grammatical accuracy and lexical range, while overlooking higher-order writing abilities, including argumentation, rhetorical awareness, and disciplinary identity. These dimensions are critical in academic contexts, where writing is not only a skill but also a form of socialization into scholarly communities.

Furthermore, existing work is disproportionately situated in Western, Anglophone settings. Comparatively little research has examined how students in non-Anglophone contexts adopt and interpret AI tools, even though differences in cultural perceptions of feedback, authorship, and plagiarism may strongly influence learning outcomes.

Finally, there is a conceptual gap in how AI is positioned within established pedagogical frameworks. Rather than being studied as an integrated element of the process writing approach or academic literacies model, AI tools are often treated as supplementary add-ons. Without theorizing their role more systematically, research risks underestimating both their transformative potential and the challenges they pose to critical engagement, authorship, and academic integrity.

Theoretical Framework

The Zone of Proximal Development (ZPD)

Lev Vygotsky's (1978) concept of the Zone of Proximal Development (ZPD) continues to be another revolutionary idea for the mediation of learning interactions. According to the ZPD, learners exhibit a set difference between their present capabilities and those they can achieve with the help of educated others. The model underscores the notion that effective learning is not just about the right answer from the learners but through a collaboration of events leading to the eventual appropriation of the responsibility by the learners due to the reduction of the scaffolds (Lantolf et al., 2014).

ZPD can be described by how generative AI systems fall within university-level English writing, and they can be identified as helping students expand the English learning zone of mediation conditions. ChatGPT is, therefore, a tool to help students through brainstorming, discourse structuring, and linguistic elaboration that can be beyond their independent capability (Lo et al., 2024). Quillbot similarly serves this art of rewriting and the improved manner with the same reasonability, assisting them in moving beyond the simplest of sentence structuring to the most advanced academic registers. On the other side, AI models these specific functions; therefore, it extends the learners' capacities for development.

Furthermore, ZPD makes clear the possible abuses of the theory of over-scaffolding. As Godwin-Jones (2020) put it, the generative AI approach is “distributed agency”, meaning that it can give or discourage student development, depending on how it is applied. Over-reliance on AI is associated with the fact that students often bypass the hard slog of internalization, since they are not made to push the envelope of what is temporarily acceptable to learn how they are fully capable of. At the teaching levels, this brings to light issues around the proportion or combination of AI as support tentacles with the cognitive or mental efforts that should still be intact. This implies that this generative AI in the ZPD model should not be looked at as a lone teaching aid but as a media socializing force whose success depends on how well students can simultaneously involve themselves in its feedback.

The Process of Writing Approach

The writer-based process writing paradigm, which was championed in the late twentieth century, shifted the sole focus from the end product to the recursive steps formulated in prewriting, drafting, revision, and editing (Hyland, 2019). The reflective workshops for all learners most probably play a central role here, as writing is traced by them to be longitudinal and cyclical, with constant feedback and inter-learning throughout the process. In this approach, the teachers and fellow students are given principal roles to play as they help pupils to develop ideas, ask questions, design possible solutions, encourage feedback, and ensure that they revise.

New technology introduces other functions to assist mediation across these stages. While overtone of prewriting, aid, such as ChatGPT, lower entrance levels of thought development to recommend instances of ideas, arguments, or even an outline (Lo et al., 2024). During the working out of the draft, smart AI could suggest the syntax rules but also the right word replacements and the logical connections between the thoughts, lowering the burden to compose the content and thus, accordingly, granting the student to prioritize conceptual understanding. And on revising, Quillbot will recommend similar tweaks to fill gaps, main words and maintain connectivity, while a grammar checker is important for editing with regard to grammar and awkwardness.

Notwithstanding, the scholars advise against total usage of AI. According to Hyland (2019), the main potential of the process approach lies in the students’ active role in eliciting versions in their revision and self-checking. The tendency of students to depend on automated tools, yet they are obliged to go through the recursive cycle of concerns, ideas, judgment, observations, and judicious correction, to experience the education that is truly developmental, is compromised by AI. Studies based on the evidence suggest that at times students unknowingly accept changes in texts made AI-based by missing the need for a detailed and careful look, thereby limiting their metalinguistic awareness (Godwin-Jones, 2020). Respectively, the process writing method emphasizes both the positive educational potential and the danger of integrating generative AI into teaching writing, thus making it no less important a lens of the research.

Critical Pedagogy

Freire (2020), whose pedagogy originates from educating as a practice of freedom, illuminates the educational process as a dialogue that transforms learners’ consciousness to a critical level. When writing becomes one of the areas of concern in the classroom,

Pennycook (2004) approaches the subject by emphasizing that language is a socially situated tool for action, which is inseparable from ideology and relations of power. Emphasizing critical pedagogy in AI-mediated writing brings to the forefront fundamental issues of power, editorial line, and morality.

Conversely, AI can mitigate the trend of students' intellectual abilities worsening with social diversity and help students better understand academic contents and break the language barrier (Lo et al., 2024). It may represent a lighthouse to the shore of the non-English speakers as AI has become the linguistic scaffolding that could help them taking part to the academic communities which are about to exclude them. A concern raised by critical pedagogy is this: that AI systems will turn students into passive consumers of machine-generated language and deny them the ground to build their own knowledge.

Moreover, ethical concerns demand scrutiny. AI and education Plagiarism, authorship, and algorithmic bias are central considerations in discourses around AI in education (Zawacki-Richter et al., 2019). A critical pedagogy approach would also posit that these are not simply side issues, but topics for discussion in the classroom. Students' thinking should be challenged over when AI improves learning and when the integrity of academic qualifications is put at risk. Kasneci et al. (2023) also assert that combining efficiency with a critical reflection is crucial, especially when it comes to the responsible implementation of AI in higher education. As a result, critical pedagogy not only views AI as an instrument, but as an object of study as well, prompting students to question its impact on authorship, equity, and scholarly standards.

Integrative Framework

The three theoretical approaches—ZPD process writing, and critical pedagogy—combine to provide a multi-faceted framework for this paper. ZPD theory clarifies AI's scaffolding power, showing how a tool like ChatGPT could function to augment learners' abilities in their developmental zones. The process writing perspective locates AI interventions throughout the recursive stages of writing, making evident their pedagogical interventions and potential interruptions. Critical pedagogy reframes the conversation by attending to questions of agency, equity, and ethics in questions about AI and higher education.

This integrated point of view has three main implications for the research questions. By putting AI in the context of ZPD, this work can assess whether generative AI supports or impedes student development along with each teacher and peer feedback. Overlaying AI functionality to the process writing cycle maps onto stage-specific opportunities and threats. Enacting a critical pedagogy position allows the research to remain sensitive to more institutional and ethical issues such as academic integrity, teacher subjectivity and student agency.

By synthesizing these perspectives, the study does not present generative AI as a neutral tool but as a socially situated artifact, a resource whose effects on writing pedagogy are contingent on its mediations, appropriations, and critical interrogations. This theoretical basis means that the analysis has relevance for immediate pedagogical effectiveness as well as much bigger-picture values in higher education in the age of AI.

Research Methodology

Research Design

The research employs a comparative case study research design and examines the use of generative AI tools in university level English writing instruction. A case study design is appropriate for complex educational phenomena since it allows access to processes not available through only quantification (Yin, 2017). For large-scale surveys or experiments can yield broad generalizations, but in the process they tend to miss the contextual dimensions of students' engagement with new technologies, their understandings of feedback, and how they enter into and grow out of writing practices. A comparative case study employs what Bartlett and Vavrus (2016) call a horizontal and vertical strategy: it compares between cases (horizontal) and between cases and the historical and technological context in which they are embedded (vertical).

The horizontal dimension of this study focuses on how different feedback sources—ChatGPT, Grammarly, Quillbot, teacher feedback, and peer review—shape students' writing revisions and perceptions. The vertical dimension examines the evolution of writing support tools, tracing how students experience the shift from error-correction software to generative AI and how this evolution reconfigures teaching practices. By combining these two levels of analysis, the study provides a holistic account of how AI intersects with existing pedagogical traditions.

Unlike experimental designs, which require random assignment and controlled variables, the case study approach accepts the complexity of real classroom interactions. It acknowledges that writing development is influenced not only by the feedback itself but also by students' prior knowledge, motivation, and social context. This interpretive orientation is consistent with applied linguistics and educational research traditions, where the goal is to generate "naturalistic generalizations" (Stake, 1995) rather than statistically predictive claims.

Participants and Sampling

The sample included nine undergraduates in an English for Academic Purposes course. This required class is for second-year English majors to develop argumentative, investigative, and academic style skills. Purposive sampling ensured diversity in proficiency. Students were sorted into three levels: low group struggled with fluency, vocabulary, grammar; middle group created storylines but weak in reasoning; high group produced organized texts despite style issues. Each group of three had four tests. Participation was voluntary with written consent. The study observed how proficiency levels use feedback differently, considering prior AI experience, motivation, and acceptance.

Writing Tasks and Feedback Conditions

The study was built around three writing assignments completed over the course of the semester, all three designed to draw out unique aspects of writing. Firstly, Composition 1 asked students to write an argumentative essay of approximately 500 words for the general education topic. Next, Composition 2 required a 600-word cause-and-effect essay, which included at least one source from the paper. Composition 3, compared to Composition 1 and Composition 2, demands that students write a 1100-word argumentative essay, which certainly equates to writing a mini research paper with more in-depth and critical analysis.

Every task was composed and revised under the guidance of different feedback mechanisms. For Task 1, the feedback relied on AI tools such as Grammarly, Quillbot, and ChatGPT; for Task 2, the source of feedback was the teacher or a peer-review platform; and in Task 3, the feedback came from both AI tools and teachers.

This rotation ensured that each student experienced all five feedback modes across the semester. Drafts and revisions were submitted digitally, and students were asked to save feedback texts (from AI, teachers, or peers) for analysis.

To capture the process dimension of writing, students were also encouraged to keep a short reflective journal after each feedback experience, noting what they found useful, confusing, or limiting. These reflections provide valuable insights into how feedback is internalized, aligning with process writing principles (Hyland, 2019).

Data Collection

Communicated data incorporated text and reference assessments together with students' writings. Drafts and compositional works were collected, and variations were discovered as a result of three different types of feedback. These were coded as surface-level corrections or higher-order revisions. AI-generated feedback was downloaded, teacher feedbacks were reviewed, and peer review comments were recorded from the platform. Students' journals and interviews provided their opinions of acceptance. Students reflected in a 200–300 word essay, and after completion of the term, semi-structured interviews of 30–45 minutes were conducted. Conversations aimed at understanding feedback incorporation into the self-directed learning process. Context information came from syllabuses, facilitator's guidelines, and policy documentation, showing the broader pedagogical framework. The integration of data sources ensured credibility.

Comparative Case Study Procedure

Research used within-case and cross-case analysis. Drafts, revisions, feedback texts, and reflections were included. Feedback shaped learners' writing: weak learners focused on surface corrections from Grammarly, while peers offered deeper meaning-related feedback. Cross-case analysis examined similarities and divergences. Skilled students used ChatGPT to explore alternatives; less skilled used it only for ideas. The horizontal dimension was AI-generated (ChatGPT, Grammarly, Quillbot) and personalized (teacher, peer) feedback. The vertical dimension compared older tools (Grammarly, Quillbot) with newer systems (ChatGPT), highlighting development of writing support technology.

Evaluation Metrics and Operationalization

To operationalize the evaluation of feedback practices, the study adopted four analytical dimensions: feedback speed, depth, scalability, and transferability of learning. Feedback speed was defined as the interval between draft submission and the reception of comments, thereby capturing differences between the instantaneous responses of AI tools and the comparatively delayed feedback provided by teachers or peers. Depth of feedback was measured using a coding scheme adapted from Hyland and Hyland (2006), which distinguished between surface-level corrections (e.g., grammar and vocabulary) and higher-order concerns (e.g., organization, criticality, and evidential support). Scalability was conceptualized as the extent to which a feedback method could be feasibly applied in larger

classroom contexts, acknowledging that teacher feedback, while often comprehensive, is constrained by time, whereas AI tools can be scaled more efficiently. Finally, transferability of learning was operationalized through evidence of whether students were able to apply feedback principles in subsequent tasks, drawing on both textual analysis of later drafts and students' reflective accounts of internalizing lessons from feedback.

Revisions were categorized using Faigley and Witte's (1981) taxonomy of surface and meaning-changing revisions, supplemented by Ferris's (2010) framework on corrective feedback in L2 writing. Coding was conducted by two researchers independently, and discrepancies were discussed until consensus was reached, ensuring reliability.

Ethical Considerations

Informed consent was obtained, and pseudonyms were assigned to participants. Given concerns about AI and academic integrity, students were assured that their use of AI tools for this study would not be penalized. Data were stored securely in password-protected files, and participants had the right to withdraw at any time.

Findings and Discussion

Horizontal Comparison: AI Tools versus Human Feedback

The horizontal comparison revealed important contrasts between AI-generated feedback (ChatGPT, Grammarly, Quillbot) and human feedback (teachers, peers).

Students consistently highlighted feedback speed as the most salient difference. AI tools offered instant feedback, which several participants found "motivating" and "convenient," letting them revise when their thoughts were still fresh. Feedback coming from Teachers and peers on the contrary took several days, reflecting the institutional limitations with respect to workload and schedule. Although immediacy was appreciated, AI tools were also noted by certain students to generate generic comments, particularly Grammarly and Quillbot, which concentrated primarily on surface errors.

The richness of feedback varied significantly between sources. Teacher feedback was described as "extensive," not limited to grammar but involving argument and evidence use and critical position taking. Comments from reviewers were mixed; some reviewers were helpful in their comments to clarify and organize the structure of the review, while others made brief comments. Among AI tools, ChatGPT got closest to human feedback in facetiousness, giving me advice on paragraph construction and logical progression. Though, some students were skeptical of ChatGPT's advice and did not know if the recommendations made were in line with academic standards. This is consistent with previous research that warns against blindly trusting AI (Kasneci et al., 2023).

Scalability also differed significantly. Teachers reported that it was challenging for them to give personalised feedback to large classes. This problem was somewhat mitigated by peer review, but its quality was unimproved. AI-based tools, by comparison, had virtually infinite scalability, delivering feedback to many students at once without any added labor. As argued by Zou and Thomas (2019), the potential for expansion could revolutionize writing instruction, but not without human oversight that maintains pedagogical depth.

Lastly, in terms of generalization of learning, teacher feedback seemed most potent in encouraging students to apply lessons to future writing. Students said that teacher explanations often made issues of argument and style clearer. According to the students, teachers' explanations frequently helped to distill principles of argument and style. Peer feedback had low transfer value because it mainly addressed immediate concerns, not generalizable writing processes. English Model Feedback was more hit or miss: Grammarly and Quillbot led to revisions but rarely to deeper learning, and ChatGPT prompted reflection about structure and coherence at times, although students noted the challenges of knowing when to trust it.

Horizontal Comparison: Contrasting AI Tools

Within AI helpers, Grammarly, Quillbot and ChatGPT stood out in relative strengths and weaknesses.

Grammarly performed well in detecting errors in grammar and lexis, presenting succinct feedback 1. Students liked that it was clear, but also said it was "narrow." Hardly if at all did it attend to larger issues of, say, argument coherence or stylistic range.

Quillbot was used primarily for paraphrasing purposes and to add lexical variety. Low-level students found it useful for creating varied sentence forms; however, the teachers were wary of their depending on it, as in a few cases it gave awkward or incorrect sentences.

ChatGPT excelled at offering discourse-level feedback. However, both students and instructors cautioned that ChatGPT occasionally produced "overconfident but misleading" suggestions, aligning with Godwin-Jones's (2022) observation that generative AI risks promoting fluency without accuracy.

Taken together, these findings suggest that AI tools should not be treated interchangeably. Rather, they occupy different niches: Grammarly as a surface-level corrector, Quillbot as a paraphrasing assistant, and ChatGPT as a generative scaffold for idea exploration.

Vertical Comparison: Tool Evolution and Writing Pedagogy

The vertical comparison traced students' experiences across the evolution of writing support tools, from error-focused systems to generative AI.

Several students noted that Grammarly felt like a "digital grammar checker," similar to older spell-check systems, reinforcing a form-focused orientation. Quillbot introduced variation but remained primarily at the sentence level. ChatGPT, however, was perceived as a "writing partner" capable of engaging with ideas and structure. This transition mirrors broader technological shifts in CALL (Computer-Assisted Language Learning), where tools have evolved from corrective to collaborative (Chapelle & Sauro, 2017).

Pedagogical Opportunities and Risks

This study uncovered the potentials and pitfalls of adopting generative AI in argumentative writing instruction.

Opportunities

AI tools supported personalized and instant feedback, allowing students to iteratively revise at their own pace. Some respondents found motivation in having feedback available "instantly and anytime. The second element of AI that fostered learner autonomy was the chance to experiment via multiple drafts without having to rely on external feedback. Third,

generative AI fostered student engagement, as students "can simply brainstorm counter-arguments or improve the thesis statement" using ChatGPT, consistent with the process writing approach of drafting and redrafting (Hyland, 2019).

Risks

Yet students also recognized risks of over-reliance. Others conceded to submitting revisions created by AI without critically reading them, raising fears of surface learning. One of the biggest concerns cited by teachers was the issue of academic integrity, with teachers stating that generative AI might make it difficult to differentiate between support and cheating (Cotton et al., 2023). Finally, the proliferation of AI feedback risked weakening the teacher's role, potentially shifting authority from human educators to algorithmic systems. From a critical pedagogy perspective (Freire, 1970), such a shift could reproduce new power asymmetries, where students defer uncritically to machine-generated authority.

Conclusion and Implications

Summary of Findings

This study investigated the role of generative AI tools in university-level English writing instruction, using a comparative case study design to analyze both horizontal (AI tools versus teacher/peer feedback) and vertical (tool evolution over time) dimensions. The findings addressed three guiding research questions.

First, in comparing generative AI tools with other feedback sources, the study found that AI feedback was consistently faster and more scalable, but often less trustworthy in depth and transferability. Teacher feedback provided the most comprehensive guidance, especially for rhetorical and critical dimensions of writing, while peer review offered variable but occasionally insightful perspectives. Of the AI tools, Grammarly worked best for surface-level focus, Quillbot for paraphrasing, and ChatGPT for discourse-level revision. Nevertheless, students approached AI feedback with caution, consistent with concerns about academic norms and veracity.

Second, higher education's adoption of generative AI exposed both opportunities and dangers. Opportunities were greater personalization, immediacy, and independent learning, which promoted motivation, and iterative writing. Risks were overuse, confounding academic integrity, and creation of a teacher less powerful in scaffolding student progress.

Third, the evolution of AI writing instruments has disrupted teaching models, the status and responsibilities of teachers, and the norms of academic honesty. The transition from corrective towards generative systems represents the move to tools that reason about meaning and structure rather than simply about form. But this transformation also forces teachers to imagine how they can winnow the wheat of writing and judgment and agency and critical consciousness in the classroom.

Pedagogical Implications

There are a number of implications in the findings for higher education writing instruction.

Generative AI offers new openings for rewriting pedagogy. Educators may choose to incorporate AI feedback in a nuanced way into the process writing cycle, for example with

students using Grammarly for editing, Quillbot for practicing written paraphrasing, and ChatGPT for ad hoc brainstorming or structural revision. This differentiation recognises that each of these tools occupies its own pedagogical niche, thereby supplementing rather than replacing human intervention.

Instead of being displaced, teachers can be repositioned as mediators that support students to assess, interpret, and question feedback that is produced by AI. This similar to Vygotsky's idea of the ZPD scaffolding where support should be faded to promote independent competence. Teachers could prompt students to question AI suggestions, to compare them with educational standards, and to implement selective revisions, thus maintaining agency.

Finally, students require explicit training about trade-offs between the benefits and risks of AI utilisation. This involves building habits of critically engaging with AI-produced feedback, becoming savvy to academic integrity norms, and learning how to distinguish between legitimate support and illegitimate replacement with machined-produced text. Integrating ethics and authorship discussions in to writing curricula can reduce harm while reinforcing the academy.

Limitations and Directions for Future Research

The study, which analyzes the roles of generative AI in education, both contributes to and is limited by the existing conversation on the topic. The exploitation of a comparative case study approach led to in-depth analysis; however, it had limitations on the scope of effort to a single form of universities and academic discipline. Results, for instance, might not be adopted throughout different fields, cultures, and levels of English competence. The aspect of time that provides us with whether the information is past, present, or future is also being taken into consideration. The study described how AI technologies had been applied in a moment of AI technologies' quick proliferation, and this implies that further development of ChatGPT systems may lead to different complexities or shifts in what is pedagogically possible. A longitudinal research study will be conducted to learn how the use of AI technology affects the writing skills of college students throughout the semester or even the whole college academic program.

Finally, in the women's and gender studies seminar courses being taught, I largely focused on the English Language and Literature instruction area. Nevertheless, in future work, the interdisciplinary study is to be adopted, which also takes infection prevention into the areas of STEM, consultants' services business, social science, as well as in languages and writing, with a lot of different religious and cultural backgrounds, which have their own traditions and values. Similar research in the spheres of culture and the system of educational institutions' connections with AI could also show the picture of cultural shifts that are shaped by technological advances.

References

Bartlett, L., & Vavrus, F. (2016). *Rethinking case study research: A comparative approach*. Routledge.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in education and teaching international*, *61*(2), 228-239.

Dikli, S., & Bleyle, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback?. *Assessing writing*, *22*, 1-17.

El Ebyary, K., & Windeatt, S. (2010). The impact of computer-based feedback on students' written work. *International Journal of English Studies*, *10*(2), 121-142.

Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition & Communication*, *32*(4), 400-414.

Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Routledge.

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, *32*(2), 181-201.

Flowerdew, J. (2015). John Swales's approach to pedagogy in Genre Analysis: A perspective from 25 years on. *Journal of English for Academic Purposes*, *19*, 102-112.

Freire, P. (2020). Pedagogy of the oppressed. In *Toward a sociology of education* (pp. 374-386). Routledge.

Godwin-Jones, R. (2020). Emerging Technologies: Artificial Intelligence and Language Learning. *Language Learning & Technology*, *24*(2), 1-17.

Hyland, K. (2019). *Second language writing*. Cambridge university press.

Hyland, K. (2004). *Disciplinary discourses, Michigan classics ed.: Social interactions in academic writing*. University of Michigan Press.

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language teaching*, *39*(2), 83-101.

Hyland, F., & Hyland, K. (2001). Sugaring the pill: Praise and criticism in written feedback. *Journal of second language writing*, *10*(3), 185-212.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274.

Lantolf, J. P., Thorne, S. L., & Poehner, M. E. (2014). Sociocultural theory and second language development. In *Theories in second language acquisition* (pp. 221-240). Routledge.

Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in higher education*, *23*(2), 157-172.

Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Singapore: Springer Singapore.

Leki, I., Cumming, A., & Silva, T. (2010). *A synthesis of research on second language writing in English*. Routledge.

Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of second language writing*, *27*, 1-18.

Link, S., Dursun, A., Karakaya, K., & Hegelheimer, V. (2014). Towards better ESL practices for implementing automated writing evaluation. *Calico Journal*, *31*(3), 323-344.

Liu, J., & Edwards, J. G. H. (2018). *Peer response in second language writing classrooms*. University of Michigan Press.

Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. Y. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, *11*(1), 50.

Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, *74*(5), 570-581.

Matsuda, P. K. (2003). Process and post-process: A discursive history. *Journal of second language writing*, *12*(1), 65-83.

Pennycook, A. (2004). Critical applied linguistics. *The handbook of applied linguistics*, 784-807.

Ranalli, J. (2018). Automated written corrective feedback: How well can students make use of it?. *Computer Assisted Language Learning*, *31*(7), 653-674.

Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, *37*(1), 8-25.

Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT journal*, *59*(1), 23-30.

Seow, A. (2002). The writing process and process writing. *Methodology in language teaching: An anthology of current practice*, *315*(320), 67.

Stake, R. (1995). Case study research. *Cham: Springer*.

Trachtenberg, B. (2023). ChatGPT, artificial intelligence, and academic integrity. *Office of Academic*.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard university press.

Wingate, U. (2012). Using Academic Literacies and genre-based models for academic writing instruction: A'literacy'journey. *Journal of English for academic purposes*, *11*(1), 26-37.

Yin, R. K. (2017). Case study research and applications: Design and methods. *(No Title)*.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators?. *International journal of educational technology in higher education*, *16*(1), 1-27.

Zhai, X. (2022). ChatGPT user experience: Implications for education. *Available at SSRN 4312418*.