

TEACHERS' AND STUDENTS' PERCEPTIONS OF AI-BASED WRITING ASSESSMENT

NHẬN THỨC CỦA GIẢNG VIÊN VÀ SINH VIÊN VỀ VIỆC ĐÁNH GIÁ KỸ NĂNG VIẾT DỰA TRÊN TRÍ TUỆ NHÂN TẠO (AI)

Tran Thi Thuy Oanh, Tran Thi Minh Ngoc*

The University of Danang - University of Foreign Language Studies, Vietnam

*Corresponding author: ttmngoc@ufl.udn.vn

(Received: January 05, 2026; Revised: February 25, 2026; Accepted: March 11, 2026)

DOI: 10.31130/ud-jst.2026.24(2).006E

Abstract - In assessing English writing skills, AI-assisted tools are considered a promising solution due to their efficiency, accuracy, and ability to provide personalized feedback. This study investigates the use of AI-based writing assessment tools in evaluating ESP students' writing skills at The University of Danang - University of Foreign Language Studies. A mixed-methods approach was employed, including questionnaires and interviews with both students and teachers to explore their perceptions of AI-based assessment. The findings indicate that AI tools simplify the marking process, provide prompt feedback on grammar and coherence, and support improvements in students' writing. However, several challenges remain, particularly in evaluating creativity, depth of content, and the need for human intervention. This study contributes to a better understanding of the role of AI in writing assessment.

Key words - AI assistance; evaluation; assessment; English education; writing skill; UFLS

1. Introduction

With the boom of advanced technology, adopting Artificial Intelligence (AI) in education has significantly reformed the conventional methods of teaching and learning, and assessing the process of language acquisition. In terms of English for Special Purposes (ESP), learners are expected to not only be fluent in English but also employ efficiently English skills for special purposes (such as business, tourism, ...), hence, the roles of AI tools have become an essential role in advanced instructions and assessment [1]. Since writing is one of the most complex skills that requires good cognitive abilities to be excellent to assess it as a second language, making use of Automatic Writing Evaluation (AWE) such as ChatGPT offers a very bright future for rapid and effective and individualized methods of writing assessment [2], [3].

While traditional approaches struggle to provide immediate responses, they also lack accuracy and individuality, especially in the large-scale classrooms; AI tools are capable of mitigating these shortcomings by assisting teachers to evaluate several students' writing at the same time with consistency and immediacy (Bai & Hu, 2023). However, under certain pedagogical circumstances, especially in Vietnam, the aid of AI assisting tools has not been significantly noticed yet. Although the adoption of AI assessment on writing has gained a lot of popularity around the world, empirical research and application of AI tools to

Tóm tắt – Trong đánh giá kỹ năng viết tiếng Anh, các công cụ hỗ trợ trí tuệ nhân tạo (AI) được xem là giải pháp tiềm năng nhờ nâng cao hiệu quả, độ chính xác và khả năng cung cấp phản hồi cá nhân hóa. Nghiên cứu này khảo sát việc ứng dụng các công cụ đánh giá viết dựa trên AI trong đánh giá kỹ năng viết của sinh viên Tiếng Anh chuyên ngành tại Trường Đại học Ngoại ngữ - Đại học Đà Nẵng. Phương pháp hỗn hợp được sử dụng, bao gồm bảng hỏi và phỏng vấn đối với sinh viên và giảng viên nhằm tìm hiểu nhận thức của họ về việc đánh giá bằng AI. Kết quả cho thấy AI giúp đơn giản hóa quá trình chấm, cung cấp phản hồi nhanh về ngữ pháp và tính mạch lạc, đồng thời hỗ trợ cải thiện bài viết. Tuy nhiên, vẫn tồn tại những hạn chế liên quan đến đánh giá tính sáng tạo, chiều sâu nội dung và sự cần thiết của can thiệp con người. Nghiên cứu góp phần làm rõ vai trò của AI trong đánh giá viết.

Từ khóa – Hỗ trợ trí tuệ nhân tạo AI; kiểm tra đánh giá; giảng dạy Tiếng Anh; kĩ năng viết; UFLS.

mark writing answers of university students have not been widely adopted yet in Southeast Asian contexts.

This study attempts to address this gap by investigating the integration of AI tools into the assessment of ESP students' writing ability at The University of Danang - University of Foreign Languages Studies (UD-UFLS). This study aims to explore how teachers and students welcome and perceive the application of AI in writing assessment, and furthermore, to help compare some assessment criteria obtained when using different AI tools, thereby pointing out the strengths and weaknesses of two popular AI websites, contributing to the growing literature on AI-enhanced language education. This study focuses specifically on two AI-based writing support tools, Grammarly and Lexibot, and does not aim to generalize the findings to all AI writing assessment systems.

2. Literature Review

Writing in ESP is a specialized skill that requires learners not only to master grammar and general vocabulary but also to demonstrate the ability to communicate appropriately within professional and technical contexts [4]. Therefore, the assessment of ESP writing needs to take into account both linguistic accuracy and genre-specific conventions. Traditional methods of assessing writing-such as universal scoring or analytic rubrics-are often time-consuming and subject to

subjectivity and variability. Besides, especially in large classes, individual feedback becomes impractical [5]. Furthermore, many ESP instructors have difficulty providing timely and formative feedback that can effectively guide student progress [6].

The use of AI in writing assessment has received considerable attention over the past decade. AWE systems such as Grammarly, Criterion, and Write & Improve provide immediate feedback on grammar, vocabulary, coherence, and overall writing quality [7], [8]. More recently, large language models (LLMs) like OpenAI's ChatGPT have been explored for their ability to generate personalized comments, suggest rephrases, and engage in iterative feedback with learners [12].

Studies have shown that AI tools can boost the self-study spirit of learners, reduce teacher workload, and improve the quality of feedback on students' answers [10]. However, there are still some concerns remaining about overreliance on AI, potential inaccuracies in evaluation, and ethical issues such as data privacy and academic integrity [11]. Moreover, while AI systems are increasingly accurate in detecting surface errors, their ability to evaluate other complex aspects like argument structure, audience perception, and genre conventions remains limited [12].

Although there is increasing research on the application of AI and digital technologies in language teaching and learning, previous studies have explored the integration of technological tools in foreign language education in Vietnamese universities [13]. However, there is still limited research on the use of AI specifically for assessing writing skills for ESP learners, especially in Southeast Asia and in Vietnam. Most studies focus on the use of technology in general English language teaching, while research on technology-assisted assessment of English writing remains relatively scarce [14].

Furthermore, there is limited understanding of how both students and teachers perceive the affordances and limitations of AI tools in actual classroom settings. This lack of contextualized research impedes evidence-based implementation of AI in ESP programs, especially in institutions like the UFLS, where specialized English writing plays a pivotal role in the curriculum.

This study seeks to fill this research gap by exploring how AI tools are currently being used-or could be used effectively-in assessing ESP students' writing performance, taking into account both pedagogical and technological dimensions in the Vietnamese higher education context.

3. Research Questions

In order to achieve these aims, the study is guided by the following research questions:

- How are AI-assisted tools integrated into the evaluation and assessment of ESP students' writing in the instructional context at UD-UFLS?
- How do ESP teachers and students at UD-UFLS perceive the use of AI tools in writing evaluation in terms of usefulness, limitations, and pedagogical value?

To address these research questions, the study employed a mixed-methods approach including content analysis of AI-generated feedback and survey data from students and teachers.

4. Methods

This study employed purposive convenience sampling, as the participants were selected from intact ESP writing classes at UD-UFLS, where the researchers were directly involved in teaching.

The sample included 20 second-year undergraduate students, with 10 students majoring in English for Tourism and 10 in English for Business. All participants were enrolled in a compulsory ESP writing course during the second semester of the 2024-2025 academic year.

Regarding English proficiency, the students shared similar language backgrounds, having completed General English courses prior to the ESP program. Their proficiency level was estimated to range from B1 to B2 according to the CEFR, based on institutional placement and course requirements. Both groups followed the same curriculum framework and assessment criteria, and received comparable instructional input.

These two ESP majors were selected because, while they differ in disciplinary focus, they place similar demands on professional writing skills. This allowed the study to explore perceptions of AI-assisted writing assessment across different ESP contexts within the same institutional setting.

In addition, five ESP instructors teaching writing-related courses at UD-UFLS voluntarily participated in the survey and interviews, offering supplementary perspectives on the pedagogical use of AI-based writing assessment tools.

5. Design of the Study

The data were analyzed through a combination of quantitative and qualitative techniques:

- Quantitative analysis: Responses from Likert-scale items in both surveys were analyzed using descriptive statistics (e.g., mean scores, percentage distributions) to summarize general perceptions of AI tools among students and teachers. Key metrics included satisfaction with AI feedback, perceived time-saving, and clarity of comments.

- Qualitative analysis: Open-ended responses from students and teachers were thematically coded using an inductive approach. Recurring themes were identified, such as perceived benefits (e.g., immediacy, objectivity), concerns (e.g., superficial feedback, lack of content understanding), and suggestions (e.g., combining AI with teacher feedback). Representative quotes were selected to support each theme.

- AI Feedback Comparison Analysis: A comparative content analysis was conducted between students' original texts and AI-revised versions. The analysis was based on a writing assessment rubric currently used in ESP writing courses at UD-UFLS, which was developed by the teaching team with reference to commonly adopted writing

assessment frameworks in ESP and EAP contexts.

Corrections identified in the AI-generated feedback were classified into four main categories:

- + Grammar (e.g., verb tense, subject–verb agreement);
- + Lexical choice (e.g., word appropriacy, word repetition);
- + Coherence and cohesion (e.g., sentence structure, transitions);
- + Mechanics (e.g., punctuation, capitalization).

The frequency of each type of correction was recorded to determine which areas were most commonly targeted by AI tools. Both quantitative and qualitative responses were collected.

A small number of samples were manually analyzed to assess the depth and pedagogical quality of feedback provided by Grammarly and Lexibot, offering insights into the tools' alignment with instructional goals in ESP contexts.

6. Data collection & analysis

Data collection instruments and procedure: The data collection process involved three key sources:

- Student Writing Samples: Students were assigned one ESP-related writing task relevant to their major, such as a formal email, a short report, or a tourism-related text. Submissions were collected via Google Drive. Each student's original text was then processed using two AI-based tools, Grammarly and Lexibot, to generate automated feedback, including direct corrections and comments. Both the original and AI-revised versions were saved for comparative analysis. Each student contributed one ESP writing task, which was used as a reference point for generating AI-based feedback and for prompting students' reflections in the questionnaire. The purpose of collecting these writing samples was not to examine writing improvement over time, but rather to support the exploration of students' and teachers' perception of AI-assisted writing assessment. Given the exploratory and perception-focused nature of the study, the number of writing tasks was considered appropriate for addressing the research objectives. Detailed descriptions of the writing tasks are provided in Appendix A.

- Student Survey: All 20 students completed an online questionnaire designed to explore their experiences with AI-assisted writing feedback. The survey included both closed-ended and open-ended questions. Closed-ended items used a five-point Likert scale and focused on perceived accuracy, usefulness, clarity, feedback speed, and overall satisfaction with AI feedback. Open-ended questions invited students to share their opinions, expectations, and perceived limitations of AI-assisted feedback. The full student questionnaire is presented in Appendix B.

- Teacher Survey: A separate online survey was distributed to five ESP instructors to gather their perceptions of using AI tools such as Grammarly and Lexibot for writing assessment. The survey focused on pedagogical effectiveness, efficiency, reliability, and

limitations of AI-assisted feedback in comparison with traditional human-based assessment methods. Details of the teacher survey are provided in Appendix C.

In addition to the questionnaire, semi-structured interviews were conducted with five ESP teachers to explore their perceptions of AI-assisted feedback in greater depth. The interviews focused on their experiences with AI tools in writing evaluation, particularly in terms of pedagogical relevance, contextual limitations, and the tools' potential role in ESP instruction. Each interview lasted approximately 15–20 minutes and was audio-recorded with consent. Responses were transcribed and thematically analyzed to identify recurring concerns and insights.

7. Findings

This section presents the results derived from student and teacher surveys, as well as the analysis of AI-generated corrections on student writing. The findings are organized into three sub-sections: student perceptions, teacher perceptions, and the comparative analysis of AI feedback content.

Students' perception of AI feedback

Survey responses from 20 ESP students revealed overall positive attitudes toward AI-assisted feedback. As illustrated in (Figure 1 - Student Perceptions of AI Feedback), the majority of students (80%) agreed that AI tools provided feedback in a timely manner, while 70% found the suggestions clear and easy to understand. Around 60% believed that the feedback helped them identify common grammar and vocabulary errors.

However, less than half of students (45%) of students felt that Grammarly and Lexibot could help with content organization and idea development. This suggests that while students appreciated the speed and clarity of AI feedback, they remained cautious about its capacity to support higher-level writing skills.

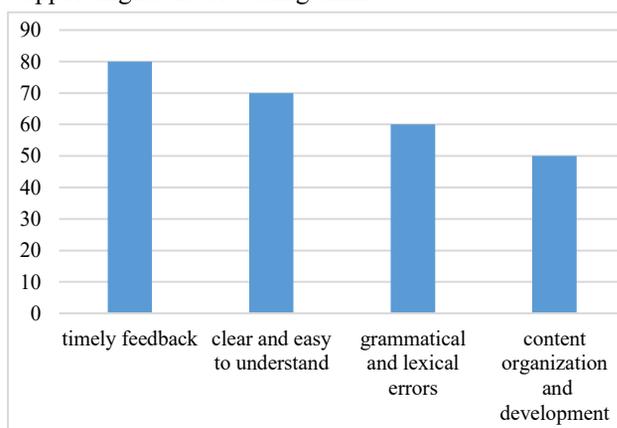


Figure 1. *Students' perception of AI feedback*

“Grammarly helped me spot grammar mistakes quickly, but I still wasn't sure if my ideas were logical.” - Student S4.

“Lexibot gave me better word choices, but sometimes they didn't fit the tourism topic.” - Student S12.

Additionally, most students expressed interest in

continuing to use AI tools in future writing tasks but emphasized the need for complementary teacher feedback to guide structure and content development.

Teacher's perceptions of AI-Assisted Assessment

Among the 05 ESP teachers surveyed, 60% acknowledged that AI tools offered practical support in reducing marking time and providing objective, consistent feedback on surface-level errors. However, only 33% believed that these tools could be trusted to perform comprehensive assessments, particularly in specialized writing contexts.

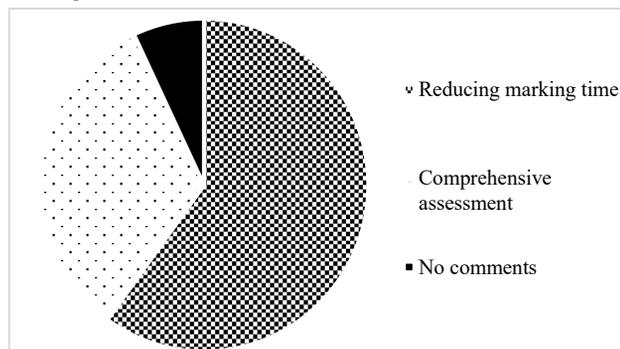


Figure 2. Teachers' perceptions of AI-assisted assessment

Figure 2 show that teachers consistently highlighted the inability of AI to handle genre-specific, contextual, and content-based feedback, which is crucial in ESP courses such as business correspondence or tourism writing. Their qualitative responses revealed concerns about overreliance on AI, lack of personalization, and the absence of pedagogical reasoning behind corrections.

"AI can correct grammar, but it cannot evaluate tone or appropriateness in ESP writing." - Teacher T3.

"Students need guidance on ideas and structure - AI doesn't teach, it corrects." - Teacher T5.

In fact, teachers generally agreed that AI tools should be viewed as supportive, rather than substitutive, especially in formative assessment.

These survey findings were supported by qualitative insights from the teacher interviews. Several teachers emphasized that while AI tools were useful for identifying surface-level errors, they lacked the pedagogical nuance and contextual awareness required for ESP genres. For example, one teacher noted, "AI can spot grammar issues quickly, but it doesn't know the difference between a formal business proposal and a casual tourism blog post." Another added, "Students might rely too much on AI and skip the critical thinking process we try to teach."

Comparison of correction types by Grammarly and Lexibot

A comparative analysis was conducted on the original student texts and the AI-revised versions. Corrections made by Grammarly and Lexibot were categorized into four types: grammar, vocabulary, coherence/cohesion, and mechanics.

Table 1 is the table which shows that Grammarly provided more accurate grammar-based corrections, while Lexibot demonstrated strength in offering lexical and stylistic suggestions, including paraphrasing. However,

both tools made limited contributions to global coherence and thematic relevance.

Table 1. Types of Correction AI

Type of Correction	Grammarly (%)	Lexibot (%)
Grammar	42	37
Vocabulary	28	36
Coherence/Cohesion	15	18
Mechanics (punctuation)	15	9

As shown in Table 1, Grammarly and Lexibot demonstrated different strengths across the four categories of writing correction. Grammarly generated a higher proportion of grammar-related feedback, particularly in areas such as verb tense consistency, subject-verb agreement, and article use. For example, Grammarly frequently identified tense shifts within a paragraph and suggested more accurate verb forms, which were common issues in students' drafts.

In contrast, Lexibot provided a greater number of vocabulary-related suggestions, including word choice refinement and paraphrasing. In several cases, Lexibot suggested alternative lexical items or rephrased sentences to reduce repetition and improve stylistic variation, even when the original sentence was grammatically acceptable.

With regard to coherence and cohesion, both tools offered limited support. Suggestions in this category were mostly restricted to sentence-level adjustments, such as reordering clauses or adding basic connectors, rather than addressing global organization or logical flow across paragraphs.

Similarly, feedback related to mechanics, such as punctuation and capitalization, accounted for a relatively small proportion of the total corrections, particularly in the case of Lexibot. Overall, while Grammarly appeared more effective in addressing form-related issues, Lexibot showed relative strength in lexical enhancement. However, neither tool provided substantial feedback on higher-level discourse features, such as thematic development or overall coherence.

Comparison between AI and traditional assessment

The table below compares the capabilities of AI-assisted tools with traditional teacher feedback in writing assessment, based on teacher responses and literature.

As can be seen in Table 2, the findings underscore that while AI tools can offer fast and efficient feedback on language form, they fall short in semantic, pedagogical, and disciplinary aspects, reaffirming the need for human-AI collaboration in writing evaluation.

The practical distinctions between these two feedback modes become particularly evident when considering the specialized needs of ESP students in Tourism and Business. In terms of Grammar and Mechanics, while AI can efficiently rectify generic syntax errors, it often lacks an understanding of professional tone. For instance, in a business email, an AI might suggest a formal correction that unintentionally makes the message sound too stiff or robotic, whereas a teacher can advise on how to maintain a professional yet approachable tone to build rapport with a client.

Table 2. Comparison between AI vs Traditional Assessment

Assessment Criteria	AI Tools (Grammarly, Lexibot)	Traditional Teacher Feedback
Grammar & Mechanics	✓ Fast and accurate	✓ Accurate, with pedagogical explanation
Vocabulary Use	✓ Suggests alternatives	✓ Contextual appropriateness considered
Content Relevance & Logic	X Not evaluated	✓ Fully assessed
Organization & Coherence	X Limited capacity	✓ Structured guidance
Genre-specific Conventions (ESP)	X Often inaccurate or irrelevant	✓ Expert-based, aligned with course goals
Personalization	X Generic, rule-based	✓ Tailored to student ability & needs
Pedagogical Intent	X None (automated rule application)	✓ Supports learning and revision strategies
Efficiency	✓ Immediate feedback	X Time-consuming
Student Engagement	✓ Interactive (in-app suggestions)	✓ Motivational through teacher interaction

Regarding Vocabulary Use, the limitations of AI are more pronounced when dealing with industry-specific terminology. An AI tool might suggest replacing "itinerary" with "schedule" for simplicity, failing to realize that "itinerary" is the standard professional term in the tourism sector. A teacher, conversely, can guide students toward genre-specific precision, ensuring they use the exact terminology required in a business contract or a travel brochure to meet industry expectations.

The human element is most critical in evaluating Content Relevance and Logic within these fields. A student might write a response to a customer complaint that is grammatically perfect but fails to address the customer's emotional concerns or provide a logical solution. While AI focuses on the "form," a teacher evaluates the "substance," identifying whether the student has effectively employed persuasion or empathy-key components in business communication. Similarly, for Organization and Coherence, AI tools often treat coherence as a mere presence of transition words. However, a human evaluator can provide nuanced guidance on the "move-structure" of a sales pitch or a promotional flyer, ensuring that the information flows in a way that successfully captures a potential traveler's interest and leads to a call to action.

8. Discussion

The integration of AI tools into writing assessment for ESP learners offers both opportunities and challenges, particularly in contexts like Vietnam where class sizes are large and formative feedback is often delayed. This section discusses the key findings of the study, makes comparisons

with relevant literature and identifies implications for future practice.

Student Perceptions: Autonomy and Surface-Level Gains

The high level of student satisfaction with Grammarly and Lexibot, especially regarding speed and clarity of feedback, confirms previous research suggesting that timely, accessible feedback enhances learner engagement [15]. In particular, students appreciated the instant error detection and vocabulary suggestions, which fostered greater autonomy and confidence in editing their own writing—an outcome also supported by Wilson and Roscoe, who found that AI-assisted feedback promotes iterative learning behaviors [16].

However, the limited effectiveness of AI tools in evaluating coherence and idea development reflects a gap between form-focused corrections and meaning-focused writing development. Nneka similarly argues that while AI tools can efficiently polish surface-level language, they do not scaffold higher-order thinking, argumentation, or structure [17]. In this study, students acknowledged this shortcoming, especially when asked to write within professional or field-specific genres, such as promotional tourism materials or workplace emails.

Moreover, students reported some lexical suggestions being too generic or inappropriate for ESP contexts. This supports findings by Li et al., who observed that while AI can offer advanced vocabulary options, it lacks contextual intelligence and domain-specific nuance [18].

Teacher Perceptions: Trust, Pedagogy, and Human Judgement

While some teachers appreciated the efficiency of AI tools in reducing their marking burden, the majority remained skeptical about AI's ability to provide meaningful, pedagogically sound feedback. This caution was mentioned by Sadler, who emphasized that for feedback to be effective; it must go beyond error correction and help students understand the qualities of good work. In other words, AI cannot help to develop students' mindset and thinking [19]. Similarly, Lee stressed that teacher feedback fosters cognitive awareness, which cannot be achieved through automated corrections alone [20].

Teachers also emphasized that AI feedback lacked individual features, (sometimes the feedbacks are almost identical) and genre awareness which are critical components in ESP writing. Unlike general English writing, ESP requires register sensitivity, discourse conventions, and professional tone, which are often misinterpreted or overlooked by AI systems. This aligns with Hyland, who emphasized that writing in specific disciplines must be understood as a social and communicative practice, not merely a technical skill [21].

AI vs. Traditional Assessment: Tension or Complementarity?

Rather than viewing AI tools as a replacement for teacher evaluation, the findings suggest that a hybrid model of assessment may be more effective. AI tools such as Grammarly and Lexibot can serve as first-line evaluators

for grammar, punctuation, and vocabulary issues, freeing up teachers to focus on higher-order concerns such as idea development, audience awareness, and content quality. This blended approach supports the call for assessment for learning, where feedback is formative, continuous, and supportive of learner development [22].

Furthermore, in resource-constrained environments like Vietnam, where student-teacher ratios are high, AI feedback can reduce teacher workload and ensure students receive at least basic correction in a timely manner. However, this requires training both students and teachers to interpret AI feedback critically, and not accept it unconditionally. As argued by Nneka, AI tools must be seen not as autonomous evaluators, but as co-pilots in the learning process [17].

Towards a Context-Aware Use of AI in ESP

The results underscore the need for contextualized integration of AI tools. In ESP courses, where language use is highly goal-oriented and audience-specific, AI systems should ideally be customizable to domain language-something current commercial tools are not yet capable of. In the future, more research should focus on training AI models on ESP corpora, or creating hybrid AI-teacher systems where machine output is moderated by expert judgement.

Moreover, students need explicit instruction on how to analyze, filter, and apply AI feedback, to avoid mechanical acceptance and foster deeper revision strategies. Teachers, likewise, need guidelines on blending AI feedback into existing assessment rubrics, particularly in formative and peer review stages.

9. Conclusion

This study explored the integration of AI-assisted tools-specifically Grammarly and Lexibot-in evaluating and assessing the writing performance of ESP students at UFLS. The findings indicate that while AI tools are effective in detecting and correcting surface-level errors such as grammar and vocabulary, they remain limited in addressing content development, coherence, and genre-specific conventions-particularly critical in ESP writing.

Students generally responded positively to the speed and clarity of AI feedback, finding it helpful for self-editing and revision. Teachers, on the other hand, recognized the efficiency of AI tools but emphasized their pedagogical and contextual shortcomings. Both groups expressed support for a blended assessment approach, in which AI provides preliminary feedback while teachers offer deeper, personalized, and content-aware evaluation.

Implications for Practice

For ESP teachers: AI tools can mitigate the burden of marking a large volume of writing answers and offer teachers lots of benefits such as spare time for doing other tasks. Under some circumstances, students can use AI tools by themselves, so teachers should provide in-depth understanding of the use so that students can take advantage of these tools effectively.

For students: AI feedback fosters autonomy and

confidence. However, students should be trained to recognize the limitations of automated feedback and seek human input for idea development and genre conventions.

For institutions: Adopting AI-supported assessment tools can promote formative feedback practices and address time constraints in large classes. However, professional development for teachers is essential to ensure that AI integration aligns with pedagogical goals and ESP objectives. In some situations, institutions should invest money on apps for teachers to use it more efficiently.

Limitations of the Study

While this study offers meaningful insights into the integration of AI tools in ESP writing assessment, several limitations should be noted.

First, the sample size was relatively small, comprising only 20 ESP students and 05 teachers from a single institution. As such, the findings may not be generalizable to other contexts, especially institutions with different technological access, curriculum goals, or student profiles. In addition, each student contributed one writing task, which was used to support the investigation of participants' perceptions of AI-assisted writing assessment rather than to examine changes in writing performance. This approach was considered suitable for the exploratory nature of the study. However, future research could include a larger number of writing tasks or a longitudinal design to provide further insights into the impact of AI feedback on students' writing development.

Second, the AI tools analyzed - Grammarly and Lexibot- are general-purpose writing aids not specifically designed for ESP writing. Their limited ability to handle discipline-specific genres and register-sensitive expressions may have influenced the perceived effectiveness by both students and teachers. Future studies could explore the performance of AI tools trained on ESP corpora or designed for field-specific communication.

Third, the study focused mainly on perceptions and surface-level text analysis. While survey responses and correction categorization provide useful data, longitudinal or experimental designs would be necessary to determine the actual learning impact of AI feedback on student writing development over time.

Finally, the research relied on self-reported satisfaction and beliefs, which are inherently subjective and may be influenced by novelty effects or lack of familiarity with AI tools. Including in-depth interviews, classroom observations, or writing portfolios in future research could yield a more comprehensive understanding of how students interact with and benefit from AI-supported feedback.

Future Research

Further research could investigate the long-term impact of AI-assisted feedback on students' writing proficiency, as well as explore how customized AI tools trained on ESP corpora might improve genre-specific writing performance. Comparative studies across institutions or disciplines would also provide broader insights into the generalizability of these findings.

REFERENCES

- [1] Y. Li, H. Zhang, and X. Liu, "Limitations of AI feedback in academic writing: A genre-based perspective", *Language Learning & Technology*, vol. 26, no. 1, pp. 45-61, 2022.
- [2] J. Ranalli, S. Link, and E. Chukharev-Hudilainen, "Automated writing evaluation for formative assessment of second language writing: Investigating the potential of Criterion®", *Journal of Second Language Writing*, vol. 37, pp. 1-17, 2017.
- [3] S. Wang and C. Vasquez, "Web 2.0 and second language learning: What does the research tell us?", *CALICO Journal*, vol. 29, no. 3, pp. 412-430, 2012. <https://doi.org/10.11139/cj.29.3.412-430>.
- [4] K. Hyland, "Genre pedagogy: Language, literacy and L2 writing instruction", *Journal of Second Language Writing*, vol. 16, no. 3, pp. 148-164, 2007. <https://doi.org/10.1016/j.jslw.2007.07.005>.
- [5] S. C. Weigle, *Assessing writing*, Cambridge University Press, 2022.
- [6] L. Hamp-Lyons, "Feedback in portfolio-based writing assessment" In G. T. Topping (Ed.), *Assessing Writing in Context*, pp. 21-37, Routledge, 2016.
- [7] M. Warschauer and D. Grimes, "Automated writing assessment in the classroom", *Pedagogies: An International Journal*, vol. 3, no. 1, pp. 22-36, 2008.
- [8] S. Link, A. Dursun, T. Karakaya, and V. Hegelheimer, "Automated writing evaluation feedback and revision: Evidence from L2 students' writing", *Language Learning & Technology*, vol. 24, no. 1, pp. 26-45, 2020.
- [9] Y. K. Dwivedi *et al.*, "So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy", *International Journal of Information Management*, vol. 71, Art. no. 102642, 2023. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>.
- [10] B. Bai and G. Hu, "Using automated writing evaluation in EFL writing instruction: A meta-analysis", *Language Learning & Technology*, vol. 27, no. 1, pp. 1-25, 2023.
- [11] J. Burstein, M. Chodorow, and C. Leacock, "Automated essay evaluation: The criterion online writing service", *AI Magazine*, vol. 34, no. 1, pp. 27-36, 2013.
- [12] S. Tian, Z. Wang, and M. Li, "Artificial intelligence and higher-order thinking in second language writing", *System*, vol. 105, Art. no. 102735, 2022.
- [13] T. T. Ngoc, N. T. T. Thuy, and N. M. Tam, "A study on the application model of AI ChatGPT to improve writing skills for Korean language-majored students," *The University of Danang - Journal of Science and Technology*, vol. 23, no. 8A, pp. 123-128, 2025, doi: 10.31130/ud-jst.2025.23(8A).056.
- [14] T. T. H. Pham, and T. T. Nguyen, Technology-enhanced language learning in Vietnam: Practice and perspectives, *Asia CALL Online Journal*, vol. 12, no. 3, pp. 32-43, 2021.
- [15] D. Boud and E. Molloy, "Rethinking models of feedback for learning: The challenge of design", *Assessment & Evaluation in Higher Education*, vol. 38, no. 6, pp. 698-712, 2013. <https://doi.org/10.1080/02602938.2012.691462>.
- [16] J. Wilson and R. Roscoe, "Automated writing evaluation and the development of self-regulated learning", *Computers and Composition*, vol. 56, Art. No. 102572, 2020. <https://doi.org/10.1016/j.compcom.2020.102572>.
- [17] E. Nneka, "AI feedback in L2 writing: A mixed-methods study on Grammarly use in university ESL classrooms", *TESOL Quarterly*, vol. 57, no. 1, pp. 75-100, 2023. <https://doi.org/10.1002/tesq.325>.
- [18] J. Li, H. Zhang and Y. Wang, "Artificial intelligence in English language education: A systematic review", *Computer Assisted Language Learning*, vol. 35, no. 3, pp. 520-546, 2022. <https://doi.org/10.1080/09588221.2021.1888755>.
- [19] D. R. Sadler, "Formative assessment and the design of instructional systems", *Instructional Science*, vol. 18, no. 2, pp. 119-144, 1989.
- [20] I. Lee, *Classroom writing assessment and feedback in L2 school contexts*, Springer, 2017.
- [21] K. Hyland, *English for specific purposes*, Cambridge University Press, 2007.
- [22] P. Black and D. Wiliam, "Assessment and classroom learning", *Assessment in Education: Principles, Policy & Practice*, vol. 5, no. 1, pp. 7-74, 1998. <https://doi.org/10.1080/0969595980050102>.

APPENDIXES

Appendix A. ESP Writing Tasks

The ESP writing tasks were designed to reflect common writing genres used in tourism and business-related professional contexts. Students completed one main writing task relevant to their major.

Task description: Students were required to write one ESP-related text of approximately **250–300 words**, based on one of the following task types:

Formal Email (English for Business) Write a formal email to a business partner or client to request information, respond to a complaint, or propose a business arrangement.

Short Report (English for Business) Write a short report describing a business situation or problem and providing a brief recommendation.

Tourism-related Text (English for Tourism) Write a promotional or informative text (e.g., a destination description, tour introduction, or service overview) for international visitors.

Students submitted an initial draft via Google Drive. The draft was then processed using AI-based tools to generate automated feedback. The original and AI-revised versions were collected for analysis.

Appendix B. Student Survey Questionnaire

The student survey consisted of **12 items** designed to explore students' experiences with AI-assisted writing feedback.

Section 1: Closed-ended questions (8 items) Students responded using a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). The questions focused on:

- Accuracy of AI feedback
- Usefulness for improving writing
- Clarity of feedback
- Feedback speed
- Support for grammar correction
- Support for vocabulary use
- Support for coherence and organization
- Overall satisfaction with AI-assisted feedback

Section 2: Open-ended questions (4 items) Students were asked to respond to the following prompts:

- Which aspects of AI feedback were most helpful?
- What limitations of AI feedback did you notice?
- How does AI feedback compare with teacher feedback?
- What suggestions do you have for improving the use of AI tools in writing courses?

Appendix C. Teacher Survey Focus Areas

The teacher survey included **10 items**, combining Likert-scale and open-ended questions. The survey focused on:

- Pedagogical effectiveness of AI-assisted feedback
- Efficiency and workload reduction
- Reliability of AI-generated feedback
- Limitations and concerns in using AI tools
- Comparison between AI-assisted and human-based writing assessment.