



Responsible AI Integration in University EFL Writing Classes in Japan Grammar-Checking Tasks and Ethical Framing at Kobe University

Mullin, Bertram

(Citation)

神戸大学国際コミュニケーションセンター論集, 22:1-10

(Issue Date)

2026-03-31

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/0100502588>

(URL)

<https://hdl.handle.net/20.500.14094/0100502588>



Responsible AI Integration in University EFL Writing Classes in Japan Grammar-Checking Tasks and Ethical Framing at Kobe University

MULLIN, Bertram
Kobe University, IPHE (Part-time Lecturer)

概要

本研究は、日本の大学 EFL 授業において人工知能 (AI) を責任ある形で統合するための授業実践を検討するものである。本稿は現在査読中の大規模調査と同一の学生群を背景としつつ、分析の焦点を質問紙ではなく教室内実践に置く。神戸大学における Academic English Literacy および Academic English Communication 科目において、AI を文法確認ツールとして導入し、学生に AI 使用前の原稿作成、使用後の修正、および学習内容の省察を求めた。成績評価を伴う試験や直接的な解答生成での AI 使用は禁止し、懲罰的対応ではなく倫理的対話を中心とした指導を行った。その結果、明確な使用境界と信頼を重視した指導により、学生の文法意識、透明性、および AI 利用に対する不安の低減が確認された。本実践は、学問的誠実性を維持しながらデジタル・リテラシーを育成する一つの有効なモデルを提示する。

キーワード

人工知能, EFL ライティング, 学問的誠実性, デジタル・リテラシー, 日本

1. Introduction

The rapid adoption of artificial intelligence (AI) in education has created both opportunities and anxiety across universities in Japan. Although many teachers recognize its potential to assist students with writing and research, concerns remain about plagiarism, dependency, and fairness. These concerns are especially visible in English as a Foreign Language (EFL) classrooms, where AI tools such as ChatGPT and Grammarly can provide feedback beyond what class time allows. However, clear pedagogical frameworks for responsible AI use are still limited.

This paper discusses an approach implemented in Academic English Literacy B1 and Academic English Communication B1 at Kobe University. The goal was to incorporate AI as a grammar-checking support tool while protecting the principles of independent

learning and academic integrity. Students were taught to write emails, essays, and other short assignments in English without AI assistance, then to check their writing afterward using AI and reflect on the corrections. Through this process, students developed awareness of grammar patterns, identified recurring mistakes, and learned to treat AI as a tool for review rather than production.

Rather than penalizing inappropriate AI use, the classes emphasized open discussion about why misuse damages community trust and creates fear among teachers and parents. By reframing AI as a shared learning aid instead of a threat, students became more comfortable discussing their use of technology honestly. The following sections outline the theoretical background for ethical AI use, describe classroom design and implementation, and present reflections on outcomes and implications for foreign language instruction in Japan.

2. Literature Review

2.1 AI and Academic Integrity in Higher Education

The emergence of generative AI has reshaped understandings of academic integrity across universities. Balalle and Pannilage (2025) note that rapid AI adoption has blurred boundaries between acceptable support and academic misconduct, calling for ethics-based frameworks rather than punitive enforcement. Similarly, Lund et al. (2025) found that students distinguish between major and minor AI infractions: using AI to generate full papers is viewed as severe misconduct, whereas grammar or citation support is often tolerated. Colby (2024) reported a comparable split among instructors, many of whom recognize AI's pedagogical potential but remain uncertain about how to regulate it consistently across courses. Together, these studies suggest that responsible integration, explicitly taught and modeled, can maintain integrity while improving engagement.

2.2 AI in EFL Writing Pedagogy

Research on AI use in second-language writing highlights both opportunity and risk. Guendouz and Ouahiani (2025) showed that AI-assisted writing tasks can enhance critical thinking if students are trained to reflect on the reasoning behind revisions rather than accept corrections passively. Praphan and Praphan (2023) similarly describe how Thai university instructors permitted AI during the prewriting and grammarchecking stages but prohibited full text generation. Their students demonstrated stronger ownership of writing after teachers emphasized that AI feedback should inform and not replace their voices. These classroom studies support an “Alliteracy” framework built on understanding, prompting, corroborating, and incorporating AI responsibly. Within Japan, Irwin (2024) reported that university students are open to AI assistance but seek clear guidance from instructors on acceptable use. Crosby (2024) found that Japanese

learners often remain hesitant to disclose AI use due to fear of negative evaluation, underscoring the need for transparent classroom policies. An approach, requiring students to write first, then check grammar with AI and explain what they learned fits within this growing body of pedagogically guided, ethically bounded AI practice.

2.3 Cultural Context and Japanese Educational Values

While international studies highlight ethics and literacy, successful AI integration in Japan also depends on alignment with its educational culture. Silence and deference, long recognized features of Japanese classroom interaction (Kawabata & Gastaldo 2015; Bao 2015), , further influence how students discuss or conceal their use of AI tools. In this context, transparent instruction that frames AI as a collective learning aid rather than a personal shortcut aligns more closely with Japanese norms of harmony and shared responsibility. Thus, by inviting students to reflect on the communal consequences of AI misuse, such as eroding trust among teachers, parents, and peers, educators can preserve these cultural values while advancing digital literacy.

2.4 Summary of Gaps

Existing research provides strong evidence for guided AI literacy and ethical awareness but seldom connects these pedagogical practices with the socio-cultural dimensions of Japanese higher education. This study addresses that gap by demonstrating how explicit boundary-setting and reflective grammarchecking tasks foster responsible AI use in a culturally compatible way within EFL writing courses at Kobe University. Therefore, the research questions are as follows:

2.5 Research Purposes and Questions

RQ1: How can AI tools be integrated into EFL writing instruction without undermining academic integrity?

RQ2: How do explicit boundaries and reflective activities shape student understanding of ethical AI use?

RQ3: What classroom-based evidence suggests that guided AI use improves grammar awareness and writing confidence?

3. Research Design

3.1 Course Context

The practice took place in three English communication courses at Kobe University: Academic English Literacy B1 and Academic English Communication B1. Each class met once per week for ninety minutes and included between twenty and thirty students. Most

students were second-year English majors or global studies majors with proficiency ranging from upper-B1 to lower-B2 on the CEFR scale.

The classes emphasized academic reading and writing, email correspondence, and oral presentations.

The goal was to integrate artificial intelligence as a supportive tool for grammar awareness and self-editing without reducing student responsibility for content. Students were encouraged to view AI as an aid comparable to a dictionary or reference tool rather than a substitute for composition.

3.2 Data

Data for the current study were drawn from classroom observations, student reflections, and anonymized writing samples. The broader survey of student attitudes toward AI, reported in Mullin (in review), served as background evidence. The present analysis focuses on pedagogical outcomes within the Kobe University context, emphasizing student behavior, classroom discourse, and instructor reflection.

3.3 Pedagogical Design

At the start of the semester, students were given explicit explanations of what constitutes appropriate AI use. The rules were summarized in class handouts and discussed collectively. Acceptable uses included checking grammar, confirming vocabulary accuracy, and comparing sentence clarity after completing an assignment. Prohibited uses included generating answers on graded tests, copying AI-written paragraphs, or submitting AI-produced text as one's own work.

The following sequence was applied in each major writing task:

1. Students produced an original draft without AI assistance.
2. After submitting the draft, they were instructed to use AI to check grammar and vocabulary.
3. Each student wrote a short reflection describing what they learned from the AI feedback and which corrections they accepted or rejected.
4. The instructor reviewed both the original and the revised versions to assess learning progress rather than mechanical accuracy alone.

This design followed the principle that AI should facilitate awareness and reflection. Activities also emphasized peer discussion about AI feedback, allowing students to explain why certain corrections were valid or misleading.

3.4 Classroom Management and Ethical Framing

Rather than employing a punitive approach to plagiarism or unauthorized AI use, the course highlighted the broader social impact of misuse. Students discussed how academic dishonesty contributes to community fear and leads institutions to restrict beneficial technologies. By connecting ethics to community responsibility, this framing aligned with Japanese values of group harmony and mutual trust.

When cases of misuse occurred, students were not penalized automatically. Instead, the incident became a topic for class dialogue about integrity, responsibility, and the shared role of technology in learning. The discussion aimed to transform mistakes into opportunities for growth.

3.5 Assessment Methods

Assessment combined language proficiency, reflection quality, and engagement with feedback. Grades for writing tasks considered three main elements: originality of ideas, improvement after AI review, and clarity of written reflections. A separate participation component rewarded students who actively discussed the role of AI in writing and demonstrated awareness of ethical guidelines.

Tests and in-class tasks were completed without AI to confirm independent ability. This clear division between formative and summative contexts reinforced that AI use is permissible only when learning, not when being evaluated.

4. Results and Discussion

4.1 Transparency and Trust

The first major finding concerns the effect of explicit boundaries on classroom trust. At the beginning of the term, several students expressed uncertainty about whether using AI tools would be considered cheating. By clarifying the distinction between acceptable and prohibited uses, students began to disclose their AI practices more openly in reflective writing. Their comments showed that transparency from the instructor reduced anxiety and confusion about academic expectations.

This openness extended to peer discussions. Students compared their experiences using AI grammar checkers and occasionally shared which corrections seemed inaccurate. These conversations encouraged critical thinking about machine feedback. Students learned that AI suggestions could be useful but not infallible. As a result, their reliance on instructor judgment shifted toward a more balanced interaction between human and digital feedback.

4.2 Metacognitive Grammar Awareness

The second finding relates to improved grammatical awareness through reflective use of AI. Students were required to identify three to five examples of grammar corrections

provided by AI and explain what they learned from each one. The reflections revealed a pattern of metacognitive growth.

Many students reported noticing repeated errors in article usage, prepositions, or verb tense consistency. Some explicitly stated that they started predicting which parts of their writing the AI might flag before even running the check.

This awareness was also visible in later assignments. The number of basic grammatical errors in final drafts declined compared with the first writing task of the semester. More importantly, students demonstrated stronger control over sentence rhythm and coherence. The focus on learning from AI feedback, rather than submitting polished but unearned writing, appeared to strengthen independent editing skills.

4.3 Ethical Framing and Community Responsibility

A third finding centers on ethical reflection. Class discussions about AI misuse produced consistent insights: students recognized that using AI dishonestly could damage not only their reputation but also the credibility of the entire program. When presented with the argument that community fear of AI often results from unethical use, students responded that the best solution was not restriction but education.

Over time, students began to use language reflecting collective awareness rather than individual blame. Phrases such as “we should show teachers we can use AI correctly” appeared in reflection notes. This shared sense of responsibility aligns with prior research on Japanese collectivist learning culture, where ethical reasoning is framed in relation to group harmony rather than personal consequence. By linking classroom ethics to broader community trust, the instructor’s approach transformed what might have been a disciplinary issue into a learning opportunity about integrity and accountability.

4.4 Instructor Reflection

From the instructor’s perspective, these classes demonstrated that ethical boundaries must be accompanied by consistent modeling. Simply stating the rules was not enough. Demonstrating AI checks live in class and discussing their limitations proved essential for student understanding. When students saw that the instructor used AI critically, accepting some corrections while rejecting others, they developed similar evaluative habits.

This reflective practice also revealed how instructor transparency reduces institutional resistance. Sharing these methods with full-time colleagues generated interest rather than opposition, whereas part-time instructors who preferred traditional methods often remained skeptical. The experience confirmed that institutional change depends on demonstrating responsible outcomes rather than promoting technology abstractly.

4.5 Summary of Pedagogical Outcomes Three outcomes summarize the findings.

1. Clear explanation of acceptable AI use increased student honesty and reduced anxiety.
2. Guided reflection after AI grammar checking improved grammatical awareness and self-editing ability.
3. Ethical framing based on community trust reinforced both digital literacy and cultural values of collective responsibility.
4. These results indicate that well-structured AI integration can coexist with Japanese educational norms. The data supports university-level language learning without diminishing academic integrity.

5. Implications for EFL Education

5.1 Integration of AI and Ethical Instruction

The results from these courses suggest that successful AI integration in EFL instruction depends on pairing technology use with ethical education. When students understand that AI tools are part of a structured learning process rather than a replacement for effort, they become more confident and less likely to misuse them. Introducing AI through classroom demonstrations and follow-up discussions encourages students to view technology as a shared learning aid. This aligns with findings by Busso and Sanchez (2024) that ethical framing enhances learner engagement while preserving academic integrity.

For EFL teachers, especially in contexts where institutional guidelines remain unclear, explicit modeling of responsible AI use can serve as a safeguard against misunderstanding. Teachers who provide transparent explanations about when and how AI may be used create an environment where students feel both supported and accountable.

5.2 Pedagogical Design and Learner Autonomy

The structured approach used in these classes, which required students to produce original writing followed by AI grammar checking and reflection, proved effective in balancing autonomy with oversight. Students developed editing skills and learned to question AI feedback critically. This suggests that AI can be integrated into process-oriented writing pedagogy without weakening human judgment.

Teachers may adapt this model by emphasizing reflection logs or comparison exercises between AI suggestions and teacher feedback. The focus should remain on the reasoning behind revisions, not the volume of corrections. Such activities train students to act as informed decision-makers rather than passive recipients of automated feedback.

5.3 Cultural Relevance in Japanese Higher Education

In Japanese universities, where deference and silence often characterize classroom interaction, the reflective use of AI can help students express individual perspectives in a controlled setting. Encouraging reflection after AI use provides a bridge between collectivist classroom norms and individual critical thinking. This approach respects cultural expectations of harmony while still fostering self-expression.

Moreover, the ethical framing of AI as a community issue rather than an individual infraction resonates with broader Japanese educational values. By teaching that responsible AI use protects group trust, educators can integrate new technologies without undermining traditional classroom relationships. This model demonstrates that technological literacy and cultural sensitivity can coexist productively.

5.4 Institutional Policy and Teacher Development

Institutions should recognize that effective AI education depends on teacher preparedness. As the studies by Irwin (2024) and Balalle and Pannilage (2025) emphasize, faculty confidence and consistency in policy implementation are essential for maintaining integrity. Workshops or faculty discussions on AI use could help reduce polarization between instructors who support or resist the technology.

Providing a unified framework, such as designating which tasks permit AI use, can ensure fairness across courses. For example, instructors might adopt a shared statement clarifying that AI is acceptable for grammar correction and research preparation but prohibited for graded assessments. Such policies would reduce confusion among students and teachers alike.

5.5 Long-Term Educational Value

Finally, the responsible use of AI contributes to broader educational goals beyond English instruction. Students who learn to evaluate automated feedback, verify accuracy, and recognize ethical boundaries gain transferable critical thinking skills. These skills are increasingly relevant in an environment where technology permeates communication and research.

EFL instruction can therefore serve as a foundation for digital literacy education more broadly. By embedding AI awareness into writing pedagogy, universities can prepare students for professional environments that demand both language proficiency and ethical technology use.

6. Conclusion

This paper explored practical ways to integrate artificial intelligence into EFL writing courses at Kobe University while maintaining academic integrity. The classroom evidence showed that when students are given clear boundaries and guided opportunities to reflect on AI feedback, they learn to use technology purposefully and take greater responsibility for their writing. Structured reflection also encouraged awareness of recurring grammar patterns and fostered confidence in independent editing.

The teaching approach highlighted the importance of linking ethics with everyday classroom activity. By focusing on shared responsibility instead of punishment, students and teachers created a learning environment built on mutual trust. This perspective fits well within Japanese higher-education values that emphasize cooperation and accountability within the group.

Although the results represent a limited set of courses, they point toward a sustainable direction for AI use in language education. Carefully guided interaction with AI can support digital literacy, reinforce honesty in learning, and contribute to the broader educational aims of Japanese universities. Further studies could expand this work by comparing instructor training models or by observing similar approaches in larger programs.

Bibliography

- Balalle, H., & Pannilage, S. (2025). Reassessing academic integrity in the age of AI: A systematic literature review on AI and academic integrity. *Social Sciences & Humanities Open*, 11, 101299. <https://doi.org/10.1016/j.ssaho.2025.101299>
- Bao, D. (2015). *Understanding silence and reticence: Ways of participating in second language acquisition*. Bloomsbury.
- Busso, A., & Sanchez, B. (2024). Advancing communicative competence in the digital age: A case for AI tools in Japanese EFL education. *Technology in Language Teaching & Learning*, 6(3), 1–17. <https://doi.org/10.29140/tltl.v6n3.1211>
- Colby, J. (2024, July 29). AI has hurt academic integrity in college courses but can also enhance learning, say instructors, students. *Wiley*.
<https://newsroom.wiley.com/press-releases/press-release-details/2024/AI-HasHurt-Academic-Integrity-in-College-Courses-but-Can-Also-Enhance-LearningSay-Instructors-Students/default.aspx>
- Crosby, A. (2024). The complexity of silence in Japanese classrooms: Teachers' perspectives. *Journal of Silence Studies in Education*, 4(1), 13–23.
<https://doi.org/10.31763/jsse.v4i1.100>
- Guendouz, A., & Ouahiani, A. B. (2025). Using AI-powered tools to develop critical thinking skills among students in EFL writing: Perceptions and challenges. 441–426, (2) 10, *تمجل آفاق موللعل*. <https://doi.org/10.37167/1677-010-002-032>

- Irwin, B. (2024). Exploring generative AI tools for English language learning: Student adoption and perceptions. *Sophia University Junior College Division Faculty Journal*, 45, 127–137.
- Kawabata, M., & Gastaldo, D. (2015). The less said, the better: Interpreting silence in qualitative research. *International Journal of Qualitative Methods*, 14(1), 1-10.
<https://doi.org/10.1177/1609406915618123>
- Lund, B. D., Lee, T. H., Mannuru, N. R., & Arutla, N. (2025). AI and academic integrity: Exploring student perceptions and implications for higher education. *Journal of Academic Ethics*, 23, 1545–1565. <https://doi.org/10.1007/s10805-025-09613-3>
- Prapahan, P. W., & Prapahan, K. (2023). AI technologies in the ESL/EFL writing classroom: The villain or the champion? *Journal of Second Language Writing*, 62, 101072.
<https://doi.org/10.1016/j.jslw.2023.101072>