US-China Education Review A, December 2025, Vol. 15, No. 12, 845-849
doi: 10.17265/2161-623X/2025.12.004

Generative Artificial Intelligence for L2 Writing Feedback: Potentials and Pitfalls

ZUO Xiuyuan, ZHU Rongrong
University of Shanghai for Science and Technology, Shanghai, China

There has been a burgeoning interest in the application of generative artificial intelligence (GenAI) in second language (L2) writing instruction and evaluation. Based on extant literature, this paper synthesizes and discusses the potential opportunities as well as the possible risks of using GenAI for L2 written feedback. Furthermore, this paper proposes strategies to optimize GenAI feedback by maximizing its benefits and minimizing its weaknesses, thereby shedding light on the development of novel feedback approaches in the digital era.

Keywords: generative artificial intelligence, second language writing, feedback

Introduction

While effective feedback is considered useful for facilitating second language (L2) learners' writing development, it is oftentimes challenging and difficult for teachers to provide such feedback due to the time-consuming nature of specific feedback provision as well as contextual constraints like large-class instruction. The rise of generative artificial intelligence (GenAI) offers new possibilities and opportunities for alleviating teachers' workload and enhancing the effectiveness of L2 writing feedback (Kim & Chon, 2025). Considering the great potential of GenAI for writing feedback, L2 teachers and teaching researchers across the globe have been attempting to utilize this powerful tool in feedback practice. An increasing number of studies exploring the pedagogical effects of integrating GenAI in L2 writing feedback could be identified (Derakhshan, 2025; Wu et al., 2025). It has been found that GenAI-assisted feedback could benefit L2 learners' writing development (Muñoz, Nassaji, & Carrillo, 2025). Despite the benefits, problems and concerns about GenAI's application in L2 writing feedback exist (Cengiz, Bilki, Ataş, & Celik, 2025; Tai, Lin, & Chen, 2025) and should not be ignored. Sorting out the upsides and downsides of GenAI-assisted feedback based on existing research and empirical evidence is helpful for its optimal utilization in L2 writing courses. Hence, this paper aims to present the merits and potential issues related to the use of GenAI in L2 writing feedback drawing on extant literature and the first author's own experiences with GenAI feedback.

Acknowledgement: This paper is funded by Research Project for Teacher Development at University of Shanghai for Science and Technology (Project No. CFTD2025YB15).

ZUO Xiuyuan, Ph.D., Associate Professor, School of Foreign Languages, University of Shanghai for Science and Technology, Shanghai, China.
ZHU Rongrong, School of Foreign Languages, University of Shanghai for Science and Technology, Shanghai, China.

Merits of GenAI for L2 Writing Feedback

Immediacy and Timeliness

One notable advantage for using GenAI in L2 writing feedback lies in its immediacy and timeliness in feedback provision, which has been widely recognized in academia (Hawkins, Taylor-Griffiths, & Lodge, 2025; Tai et al., 2025). Once the learners complete their first drafts of writing, they could upload the drafts, either directly or as an attachment, to GenAI tools like ChatGPT or Deepseek for feedback. The GenAI tools could offer feedback to the learners based on the prompts given within a short period of time, making immediate and timely feedback accessible to the learners. The immediacy and timeliness of such written feedback can make up for the deficiency of traditional delayed written feedback provided by teachers or peers. Compared with delayed feedback, immediate and timely feedback can better facilitate learners' prompt identification of problems existing in their writing and help them clarify directions for writing improvement. As such, learners could possibly better sustain their motivation and investment in L2 writing, which is oftentimes more challenging than L1 writing.

Customization and Personalization

Another merit of using GenAI tools in L2 writing, as documented in previous literature (Derakhshan, 2025; Hawkins et al., 2025), is that they can provide learners with customized and personalized written corrective feedback, which is tailored to learners' specific needs. Unlike the often unidirectional written feedback from peers or teachers, GenAI tools offer interactive experiences to L2 learners when they seek feedback. Specifically, after receiving the first round of feedback from GenAI tools, learners could raise questions to these tools regarding the feedback or ask for more feedback on certain areas of concern. The interaction between learners and GenAI tools can continue iteratively until learners have resolved all their questions. The iterative, dialogic process enables GenAI tools to capture learners' individual needs, such as weaknesses in lexicogrammatical accuracy or logical flow of ideas, thus providing customized feedback that addresses each learner's unique problems in L2 writing and better helping them enhance their writing proficiency.

Comprehensiveness and Specificity

The third notable benefit brought by GenAI to L2 writing feedback is the comprehensiveness and specificity of the feedback. It has been pointed out in previous studies that in large-class instruction context, it is challenging for teachers to provide comprehensive, detailed and specific feedback to every student (Lee, 2014). Yet, this challenging undertaking is not so difficult and complex for generative artificial intelligence tools, which are able to rapidly analyze and process vast amounts of information. When provided with learners' compositions and a comprehensive assessment rubric, GenAI tools can generate detailed responses to learners' writing quality based on the rubric within seconds, pinpointing learners' strengths and weaknesses regarding each writing dimension listed in the rubric and giving suggestions on possible ways to modify the compositions from varied facets. Such feedback enables learners to form a full picture of their current L2 writing level, identify the specific areas for improvement in their writing and figure out actionable measures to achieve progress, thereby promoting their self-efficacy in future L2 writing.

Effectiveness and Sustainability

The fourth merit of GenAI-supported L2 written feedback is its effectiveness and sustainability. Several empirical studies have confirmed the effectiveness of Gen-AI supported feedback in helping L2 learners identify and correct lexical and grammatical errors (Deng & Lin, 2023; Pfau, Polio, & Xu, 2023), increase lexical diversity

and syntactic complexity of writing (Kim & Chon, 2025), and maintain confidence in writing (Wu et al., 2025). Compared with traditional automatic writing evaluation (AWE) systems, GenAI, with its deep thinking process transparent, has the potential to provide more effective feedback that could stimulate L2 learners' reflection on their language use, idea transmission and arrangement of structure, which in turn promotes their progress in writing. Moreover, GenAI-assisted feedback is possibly more sustainable compared with teacher feedback, considering that the learners can seek suggestions and responses from GenAI anytime and anywhere. While teacher feedback is often inaccessible to learners after the completion of a course, GenAI-supported feedback is readily available to learners, providing continuous scaffolding for their L2 writing development.

Potential Issues When Using GenAI for L2 Writing Feedback

Concerns About Feedback Quality

Since the quality of artificial intelligence generated content is largely contingent upon its underlying datasets and the prompts given, using GenAI for L2 written feedback might raise concerns about the accuracy of the feedback it provides. According to research, learners reported concerns about the quality of GenAI feedback on their argumentative essay writing (Tai et al., 2025). Specifically, learners were unsure of whether GenAI-supported feedback in the form of model texts was relevant to the specific writing context and whether the generated language was natural in style (Lu & Zeng, 2025). Furthermore, when GenAI is requested to provide feedback based on an assessment rubric which encompasses several dimensions (e.g., cohesion, grammatical accuracy) with detailed descriptions, it might misunderstand certain descriptions, classify feedback into incorrect dimensions and thus offer misleading guidance to learners. Confronted with such inaccurate feedback, learners will probably be puzzled, reluctant to deeply engage with the feedback, and therefore unable to gain the desired benefits from it. The experience with dissatisfactory feedback from GenAI may have long-term consequences, constraining learners' belief in the quality of GenAI-supported feedback in the future.

Limited Trust and Acceptance

Given concerns about the quality of GenAI-supported feedback and traditional reliance on teacher feedback, L2 learners' trust and acceptance of GenAI feedback is limited. As demonstrated in previous studies exploring students' perceptions about the written feedback provided by GenAI (Thomas, Yildirim-Erbasli, & Hariharan, 2025), second language or foreign language learners exhibited insufficient trust in and acceptance of GenAI feedback, which could be attributed to their prior GenAI-assisted learning experiences. Generally speaking, feedback from teachers, who have been playing a central and authoritative role in instruction and feedback, is considered more trustworthy compared with GenAI feedback. Learners' lack of trust in GenAI-supported feedback is likely to prevent them from in-depth processing and reflecting on such feedback, thereby restraining their potential L2 writing development in the process. Likewise, learners' low acceptance of GenAI feedback is possibly detrimental to their agentic engagement with such feedback, which in turn hinders learners' significant improvement of their writing proficiency.

Overreliance and Ethical Concerns

Another noteworthy issue with using GenAI for L2 written feedback is learners' overreliance on AI tools and other potential ethical problems arising in the feedback receiving and uptake process. As indicated in previous studies (Shi, Chai, Zhou, & Aubrey, 2025), learners may encounter overreliance problems when using GenAI feedback for modifying their writing, which is possibly detrimental to learners' agency and creativity. Since

GenAI is inclined to provide direct written feedback (i.e., explicit error correction and text revision) without specific instructions, learners are deprived of the opportunity to analyze how to correct the error and modify the text by themselves. Moreover, some learners, fascinated by the revision power of GenAI, tend to reduce investment in formulating their first drafts in the following writing tasks. Additionally, if learners unquestioningly accept all the corrections and revisions generated, an ethical dilemma might emerge: who should be credited as the author of the final draft? This dilemma is particular evident when a substantial number of corrections and revisions were made by GenAI. Under such situations, learners' L2 writing self-efficacy and authorial identity may be undermined, which is harmful to their writing development in the long run.

Ways to Maximize the Benefits of GenAI for L2 Writing Feedback

As shown above, using GenAI for L2 writing feedback presents opportunities as well as challenges. To maximize the benefits and mitigate the potential problems of GenAI feedback, enhancing the feedback quality is crucial. One strategy to ensure GenAI feedback quality is to adopt the latest version of GenAI tools, considering that version updates are found to significantly increase their effectiveness as feedback instruments (Yang & Chen, 2025). Using appropriate prompts in the GenAI feedback seeking process is another strategy to guarantee feedback quality. Wu et al. (2025) suggested that task-specific prompting was helpful in driving GenAI to produce preferable feedback. Hence, both L2 teachers and learners should raise their awareness of the importance of proper prompting in eliciting high-quality GenAI feedback.

In order to realize the potential benefits of GenAI feedback, promoting learners' acceptance and trust cannot be neglected. As learners' knowledge about and exposure to GenAI affect their acceptance of GenAI-powered feedback (Thomas et al., 2025), it is critical to develop learners' AI literacy (Hossain, Çelik, & Hınız, 2024) to enable them to impartially examine the strengths and weakness of GenAI feedback and skillfully operate GenAI tools to obtain feedback that could facilitate their improvement of L2 writing skills. To foster learners' trust in GenAI feedback, teachers should provide guidance before, during and after learners interact with GenAI for responses, particularly when learners are not so familiar with using GenAI for L2 written feedback. Furthermore, teachers' comments on GenAI feedback could help learners better aware of the trustworthy aspects of the feedback.

To alleviate learners' overreliance on GenAI feedback and avoid academic integrity issues, teachers can guide learners to seek indirect feedback (i.e., just pinpointing error types and locations without providing corrections) from GenAI tools since such feedback is considered to be advantageous for encouraging learners to engage in self-directed reflection, enhancing their problem-solving skills and helping them develop into independent learners (Ferris, 2010). In addition, teachers can instruct learners to ask GenAI tools for detailed explanations of the underlying reasons for the errors so that leaners are less likely to make similar errors when engaging in L2 writing on their own in the future. For complex and abstract GenAI feedback, teacher can advise learners to seek exemplification, which helps learners gain a better understanding about their L2 writing problems and solutions.

Conclusion

This paper discusses the potential opportunities and challenges of employing generative artificial intelligence tools in L2 writing feedback, with a view to providing implications for the appropriate and effective use of GenAI technology in L2 writing courses. On one hand, L2 writing teachers should fully acknowledge the

transformative potential of GenAI, with its remarkable aforementioned strengths in feedback provision, to writing instruction. On the other hand, teachers should be well aware of the possible technical and ethical risks of GenAI feedback and take proactive actions to navigate the risks, thereby fostering learners' authentic writing development.

To maximize the benefits of generative artificial intelligence in second language writing feedback, a synergetic model integrating AI empowerment and teacher guidance needs to be established. In terms of AI empowerment, feedback quality needs to enhanced through technological optimization. As for teacher guidance, teachers should provide scaffolding for students to facilitate their rational and proper use of GenAI feedback. Future research could explore different ways of combining GenAI feedback and teacher guidance to identify the optimal feedback approach.

References

Cengiz, B. C., Bilki, Z., Ataş, A. H., & Celik, B. (2025). Exploring second language writers' engagement with ChatGPT feedback: Revision behaviors and perceptions. *System*, *134*, 103837.

Deng, J., & Lin, Y. (2023). The benefits and challenges of ChatGPT: An overview. *Frontiers in Computing and Intelligent Systems*, *2*(2), 81–83. https://doi.org/10.54097/fcis.v2i2.4465

Derakhshan, A. (2025). Tracking the effects of Gemini as a GenAI tool on L2 learners' writing proficiency and anxiety: Latent growth curve modeling approach. The *Asia-Pacific Education Researcher*. Advance online publication. https://doi.org/10.1007/s40299-025-01042-5

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. https://doi.org/10.1017/S0272263109990490

Hawkins, B., Taylor-Griffiths, D., & Lodge, J. M. (2025). Summarise, elaborate, try again: Exploring the effect of feedback literacy on AI-enhanced essay writing. *Assessment & Evaluation in Higher Education*, 1–13. https://doi.org/10.1080/02602938.2025.2492070

Hossain, Z., Çelik, Ö., & Hınız, G. (2024). Exploring EFL students' AI literacy in academic writing: Insights into familiarity, knowledge and ethical perceptions. *Journal of Theoretical Educational Science*, *18*(1), 157–181. https://doi.org/10.30831/akukeg.1538011

Kim, M., & Chon, Y. V. (2025). The impact of self-revision, machine translation, and ChatGPT on L2 writing: Raters' assessments, linguistic complexity, and error correction. *Assessing Writing*, *65*, 100950. https://doi.org/10.1016/j.asw.2025.100950

Lee, I. (2014). Feedback in writing: Issues and challenges. *Assessing Writing*, *19*, 1–5. https://doi.org/10.1016/j.asw.2013.11.009

Lu, D., & Zeng, Y. (2025). Exploring the use of ChatGPT-generated model texts as a feedback instrument: EFL students' text quality and perceptions. *Innovation in Language Learning and Teaching*, 1–21. https://doi.org/10.1080/17501229.2025.2525341

Muñoz, B. C. M., Nassaji, H., & Carrillo, F. I. B. (2025). ChatGPT-generated versus human direct corrective feedback on L2 writing. *System*, *134*, 103805. https://doi.org/10.1016/j.system.2025.103805

Pfau, A., Polio, C., & Xu, Y. (2023). Exploring the potential of ChatGPT in assessing L2 writing accuracy for research purposes. *Research Methods in Applied Linguistics*, *2*(3), 100083. https://doi.org/10.1016/j.rmal.2023.100083

Shi, H., Chai, C. S., Zhou, S., & Aubrey, S. (2025). Comparing the effects of ChatGPT and automated writing evaluation on students' writing and ideal L2 writing self. *Computer Assisted Language Learning*, 1–28. https://doi.org/10.1080/09588221.2025.2454541

Tai, H. Y., Lin, M. F., & Chen, Y. S. (2025). Incorporating ChatGPT into genre-based instruction for argumentative writing among EFL college students. *International Journal of Applied Linguistics*, e12777. https://doi.org/10.1111/ijal.12777

Thomas, M. L., Yildirim-Erbasli, S. N., & Hariharan, S. (2025). Exploring undergraduate students' perceptions of AI vs. human scoring and feedback. *The Internet and Higher Education*, 101052. https://doi.org/10.1016/j.iheduc.2025.101052

Wu, J., Li, J., Ge, Z., Xu, M., Lin, L., & Zhang, R. (2025). Effectiveness of generative AI in automated written corrective feedback with prompting. *Journal of Educational Computing Research*, *63*(6), 1493–1527. https://doi.org/10.1177/07356331251359430

Yang, C. T. Y., & Chen, H. H. J. (2025). ChatGPT and L2 Chinese writing: Evaluating the impact of model version and prompt language on automated corrective feedback. *Computer Assisted Language Learning*, 1–29. https://doi.org/10.1080/09588221.2025.2453205