Enhancing CSL Learners' Story Writing through Metacognitive Instruction and AI Support

Sarah H.J. Liu
Kainan University
sarah.hjliu@gmail.com

Abstract: *This pilot study examines whether metacognitive instruction combined with ChatGPT can support Chinese as a Second Language (CSL) learners' story writing. Eight undergraduate learners completed three narrative writing tasks and a closed-ended questionnaire. Overall writing scores improved significantly ($p < .05$) across the three assignments; organization received the highest score, whereas word choice/conventions obtained the lowest. The learners also reported generally positive attitudes toward ChatGPT, especially for ease of use, but comparatively weaker usefulness and metacognitive engagement. The findings suggest that AI-supported writing can be beneficial when accompanied by explicit strategy instruction, although the extent of improvement depends on how actively learners evaluate and apply AI feedback during revision.*

Keywords: Artificial Intelligence (AI), ChatGPT, Metacognitive Instruction, Chinese as a Second language

1. Introduction

Writing in a second language (L2) is a demanding process because it requires the integration of linguistic knowledge, idea development, and self-regulation. In addition to vocabulary and grammar, writers must plan what to write, monitor whether ideas are coherent, and evaluate whether a text meets task expectations (Alamri, 2019; Godwin-Jones, 2022). Recent studies in L2 have suggested that artificial intelligence (AI) tools such as ChatGPT can partially address this challenge by providing instant and personalized feedback (Pitychoutis, 2024). AI tools have been reported to enhance opportunities for interaction and support writing development (Kohnke, Moorhouse, & Zou, 2023). However, AI tools do not automatically lead to better learning outcomes. Their pedagogical value depends on whether learners use them critically and strategically. From a metacognitive perspective (Flavell, 1976; Yeh, 2015), particularly in L2/FL contexts, the effectiveness of AI feedback should be understood in relation to learners' planning, monitoring, and evaluation. Learners who actively regulate their writing are more likely to benefit from AI-generated suggestions than those who simply copy corrections without reflection.

2. Methodology

The participants were eight international CSL learners recruited from a university setting. Two were Japanese (25%) and six were Vietnamese (75%). All participants were at the B1 or B2 level according to the Test of Chinese as a Foreign Language (TOCFL), representing intermediate to upper-intermediate proficiency. The study lasted five weeks. In Week 1, students were introduced to the text-chat function of ChatGPT and received basic technical training. In Weeks 2 to 4, they completed three story-writing tasks based on different 10-minute Mr. Bean video clips. For each task, students first wrote their own approximately 150-character stories without AI assistance and then revised them with ChatGPT. The design was intended to ensure that the initial content originated from learners rather than from the AI tool.

Metacognitive instruction was explicitly embedded in the writing process. In the planning stage, students predicted what might happen in the story before watching the video. In the monitoring stage, they identified the main character and key events while viewing the clip. In the evaluation stage, they summarized and revised their writing, using ChatGPT feedback to improve expression and accuracy. Students were also asked to submit screenshots of AI interactions and final revised drafts. Writing performance was evaluated across four categories: content, organization, grammar, and word

choice/conventions; each category corresponded with the scores from 0 to 10, the higher scores that represented more accuracy (Hsiung, Lee, & Sung, 2015). A questionnaire was used to measure learners' attitudes toward ChatGPT, perceived ease of use, perceived usefulness, and metacognitive processes on a 5-point Likert scale (from 1 = strongly disagree to 5 = strongly agree). The revised drafts were analyzed, and the researcher established 92% intra-rater agreement through repeated coding.

3. Results, Discussions, & Conclusion

The results revealed a significant improvement in learners' overall story-writing performance across the three assignments, as indicated by the Friedman test ($p < .05$). The total mean score increased from 19.88 ($SD = 3.19$) in the first assignment to 22.31 ($SD = 2.15$) in the second and 24.38 ($SD = 2.25$) in the third. This pattern indicates that students gradually produced better revised texts over time. A closer look at the four writing categories revealed that organization received the highest overall mean score ($M = 17.38$, $SD = 2.08$), followed by content ($M = 16.81$, $SD = 2.00$), grammar ($M = 16.56$, $SD = 1.64$), and word choice/conventions ($M = 15.81$, $SD = 1.28$). The learners improved most in the overall organization of their stories, but their lexical precision and orthographic control remained weaker. In addition, text length increased across tasks, with the third assignment reaching an average of 190 words, suggesting that students became more able or willing to elaborate their ideas.

Questionnaire results showed generally positive perceptions of ChatGPT. Ease of use received the highest percentage of agreement (84.38%), followed by attitudes toward use (71.87%). Usefulness and metacognitive processes both received lower but still positive levels of agreement (68.75%). Within the metacognitive dimension, self-evaluation was the weakest item, with a mean of 3.1 on the 5-point scale, whereas planning and monitoring received higher means of 4.5 and 4.3, respectively. In other words, learners felt more confident using the tool than critically evaluating their own writing.

Pedagogically, the study suggests that teachers should not treat AI as a stand-alone writing assistant. Instead, they should explicitly train learners to plan before writing, monitor meaning while drafting, and evaluate AI feedback critically during revision. This is particularly important in CSL contexts, where orthographic accuracy poses a persistent challenge. Teachers may also need to give direct support on character form and radical awareness because general-purpose AI tools are not specifically designed for Chinese orthography.

Acknowledgements

This study was funded by Taiwan's Science and Technology Council under grant no. NSTC 114-2410-H-424-002.

References

Alamri, B. (2019). Exploring metacognitive strategies employed by ESL writers: Uses and awareness. *International Journal of English Linguistics*, *9*(1), 159–168.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–235). Hillsdale, NJ: Lawrence Erlbaum.

Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, *26*(2), 5–24.

Hsiung, Y. W., Lee, H. H., & Sung, Y. T. (2014). Examining the ACTFL writing assessment rating scale for L2 Chinese learners. *Journal of Chinese Language Teaching*, *11*(4), 111–139

Kohnke, L., Moorhouse, B.L., & Zou, D. (2023b). ChatGPT for language teaching and learning. *RELC Journal*, 1–14. DOI: 10.1177/00336882231162868

Pitychoutis, K.M. (2024). Harnessing AI chatbots for EFL essay writing: A paradigm shift in language pedagogy. *Arab World English Journal* Special Issue on ChatGPT, *1*, 197–209.

Yeh, H. C. (2015). Facilitating metacognitive processes of academic genre-based writing using an online writing system.

Computer Assisted Language Learning, *28*(6), 479–498.