Empirical Study on Generative Artificial Intelligence Enabling Chinese Writing Teaching

Manzi Zhu*

School of International Education, Shandong University, Jinan, China

***Author to whom correspondence should be addressed.**

Abstract: The digital transformation of international Chinese language education is an important direction for enhancing the global dissemination effectiveness of the Chinese language. In teaching Chinese as a second language, writing, as a form of higher-order language output, has long faced challenges such as insufficient teacher resources, lagging feedback, and a lack of personalized support. Generative artificial intelligence (GAI) offers a new path for teaching Chinese writing. This study, with intermediate Chinese learners as the subjects, systematically explores the effectiveness of GAI-assisted Chinese writing instruction through a combination of controlled experiments and interviews based on a Chinese learning platform embedded with GAI. The study found that the GAI class was significantly superior to the traditional class in terms of the adequacy dimension of content quality and the grammatical and lexical dimensions of language quality; teachers and GAI play complementary roles in this process; most Chinese learners have a positive attitude towards this model. Based on the above findings, this study presents three suggestions: using GAI to empower Chinese writing instruction, giving full play to the complementary advantages of human and machine, and enhancing the AI literacy of teachers and students to achieve two-way empowerment of human and machine, with the aim of promoting the intelligent development of international Chinese education.

Keywords: Generative artificial intelligence; Chinese writing instruction; Quality of writing

1. Introduction

Digital development is a key direction for the advancement of international Chinese language education, as well as an important platform for multi-party participation and collaborative success [1]. Chinese writing instruction starts with discourse organization and paragraph expression, with the goal of improving discourse expression ability. It covers basic training in various aspects such as Chinese character writing, word formation, sentence construction, and discourse cohesion, and is a more complex form of language output. However, the traditional Chinese writing instruction model is limited by teacher resources, feedback lag, and method singleness, making it difficult to meet the individualized needs of learners. In recent years, the emergence of generative artificial intelligence (GAI), represented by ChatGPT, has brought great possibilities for educational transformation. Lu pointed out that to

ensure Chinese reaches the world faster and better, efforts should be made to enable more people in various countries to learn and use the written Chinese well, and called for exploring the digital path of written language teaching. After a series of exploratory trials in the field of international Chinese language education, it has been found that this intelligent means can contribute well to the transformation of learning models, the construction of learning resources, and the scientification of teaching assessment. Despite the broad application prospects of GAI, existing research has mostly focused on the English context, and empirical exploration of Chinese writing is still in its infancy [2]. This study takes intermediate Chinese learners as the research subjects, conducts teaching experiments based on a Chinese learning platform embedded with generative artificial intelligence, explores the effectiveness of GAI in empowering Chinese writing instruction, and puts forward relevant suggestions for promoting the intelligent development of international Chinese language education.

2. Literature review

2.1. Current status and limitations of research on Chinese writing instruction

Chinese writing instruction focuses on the development of discourse ability and builds a multi-level language output system through systematic training from Chinese character writing, vocabulary application, sentence structure, to discourse cohesion [3]. The existing teaching models have diversified characteristics, including traditional paradigms such as the "result-oriented" focus on the finished product, and the "process-oriented" focus on writing behavior, as well as the "learning and application" model of English as a second language instruction [4]. At the practical level, eight teaching methods have been formed, including the control method, the genre method, and the task method, among which the process method and the task method have attracted much attention from the academic community. It is notable that the academic community is exploring teaching breakthroughs through integrated innovation, such as the "process genre method" [5] and the "read-write integration" teaching method [6]. Wen Qiufang's "outcome-oriented approach" was later applied to the practice of writing Chinese as a second language. However, the actual effectiveness of written language teaching has long been out of balance with the status of the discipline, and only 12 percent of students can achieve the goal of expressing thinking in Chinese [7]. The predicament stems from three constraints: the multi-dimensional bias caused by the particularity of the Chinese character system and the comprehensiveness of writing, the difficulty and applicability limitations of teaching methods, and the practical predicament of dynamic monitoring of the writing process. These systemic challenges urgently need to be addressed through teaching innovation in order to enhance the effectiveness of cultivating written expression ability in Chinese.

2.2. Opportunities and challenges for GAI to assist Chinese writing teaching

Generative artificial intelligence (GAI), as an emerging intelligent technology, provides an innovative path to break through the traditional predicament of Chinese writing teaching with its multimodal intelligent processing capabilities. This technology can be applied throughout the entire writing process. In the conception stage, it can expand understandable input through intelligent retrieval and demonstrate language transfer strategies with functions such as multilingual translation and grammar correction [8]. The formative assessment stage enables automated grading and personalized feedback. Empirical research shows that ChatGPT's immediate feedback in the field of English writing has both efficiency and quality advantages over traditional teacher feedback [9], and assisted modification significantly improves micro-language elements such as vocabulary and grammar [10]. However, GAI still faces technical bottlenecks and ethical dilemmas. Functionally, GAI has problems such as sometimes inaccurate semantic understanding, the lack of classification of vocabulary and grammar levels,

the need for enrichment of high-quality basic Chinese corpora, and the generation of inaccurate or fabricated information [11]. In terms of ethical risks, learners may fall into a crisis of mental inertia, leading to weakened writing innovation and academic integrity issues [12]. These dual effects suggest that AI tools need to form educational synergy with teacher-led instructional design rather than a simple substitution relationship.

To sum up, existing research has initially revealed the potential of GAI in assisting second language writing instruction, particularly in providing immediate feedback, enhancing language input, and stimulating learning motivation. However, there are still three deficiencies in the related research. First, the existing empirical research is mostly focused on the field of English writing, and the application research of GAI for Chinese writing teaching is still in the exploratory stage, especially lacking systematic controlled experiments to verify its effectiveness. Second, the research perspective is rather limited. Most studies mainly rely on the comparison of pre- and post-assessment scores to present the results, and few scholars have delved into specific writing evaluation indicators such as organizational structure, grammar, and vocabulary, making it difficult to fully reveal the micro impact of GAI on the development of second language writing. Third, the role of GAI in teaching is not clear, such as the lack of theoretical exploration of its role as an auxiliary tool, co-subject, or feedback mediator. In view of this, this study intends to explore the impact of GAI support on learners' Chinese writing ability by conducting a controlled teaching experiment and combining interview and writing scoring data, and on this basis, explore the role of GAI in assisting Chinese writing teaching, with a focus on answering the following questions:

(1) Can GAI-assisted writing instruction improve the quality of learners' writing?

(2) What roles do GAI and teachers play in the process of writing instruction?

(3) How do Chinese learners view GAI-assisted Chinese writing instruction?

3. Research design

3.1. Research subjects

This study employed an educational quasi-experimental method. The subjects were students from the intermediate Chinese Comprehensive course at a domestic university: 29 students from the GAI class and 29 students from the traditional class (as shown in **Table 1**). These students participated in the unified class placement examination, including listening, speaking, reading, and writing tests. Their Chinese proficiency was basically at HSK level 4, and their abilities were comparable. Their study duration was concentrated between 1 and 4 years. The GAI class had 21 students from developed countries, while the traditional class had 20 students from developed countries, a roughly equal ratio.

Table 1. Basic information of participants

Basic information	Options	GAI class ($n = 29$)		Traditional class ($n = 29$)	
		Number	Proportion (%)	Number	Proportion (%)
Gender	Male	9	31.03	1	3.45
	Female	20	68.97	28	96.55
Chinese language study duration	1 year or less	4	13.79	4	13.79
	1 to 4 years	21	72.41	23	79.31
	More than 4 years	4	13.79	2	6.90

Table 1 (Continued)

Basic information	Options	GAI class (*n* = 29)		Traditional class (*n* = 29)	
		Number	Proportion (%)	Number	Proportion (%)
Country	Developed countries	21	72.41	20	68.97
	Developing countries	8	27.59	9	31.03
HSK levels	Level 4	21	72.41	25	86.21
	Level 5	3	10.34	1	3.45
	Not taking the exam	5	17.24	3	10.34

3.2. Experiment process

This study conducted a two-month Chinese teaching experiment based on an online learning platform built by the research team, which enabled teachers to customize teaching processes, upload teaching resources, and set up teaching interventions (as shown in **Figure 1**). Students in both the GAI class and the traditional class went through three stages: pre-experiment preparation, Chinese-themed teaching activities, and post-experiment interviews. Before the experiment began, the teacher introduced the course content and the operation instructions of the learning platform to the students to ensure that each student was proficient in using it.

Figure 1. Experimental platform function module

In the second phase, the teacher organized a total of five language knowledge and cultural inquiry learning activities, with themes including high-speed rail travel, online shopping, self-service pick-up, mobile payment, and low-carbon living. Each learning topic was composed of sections such as an introduction to learning objectives, prior knowledge pretest, learning of new words and grammar knowledge, cultural content experience and interactive communication, opinion summary and expression, learning effect assessment and feedback, self-reflection and improvement, among which the students' output in the "Opinion summary and expression" section was used as their writing text, and the researchers scored and analyzed it. After the experiment, the researcher conducted in-depth interviews with the students to gain a more detailed understanding of their experience and evaluation of the use of human-machine collaborative teaching.

The difference between the two groups was that the GAI class platform was embedded with the Zhipu Qingyan large model, and the role and function of the large model were pre-trained to limit the number of

words, language level, content topic, etc., so that it could act as a Chinese learning and communication partner and writing intelligent mentor to have real-time conversations with learners. The traditional class, however, does not have the support of this model.

3.3. Research tools

3.3.1. Language quality and content quality scoring sheets

The assessment of writing quality is divided into two sections: language quality and content quality. The scoring criteria for language quality are based on the composition scoring criteria of the New Chinese Proficiency Test, combined with the writing quality evaluation indicators developed by Wu *et al.* and Cheng [13,14], mainly including three sub-dimensions: grammar, vocabulary, and Chinese characters. The scoring criteria for content quality draw on the second language writing function sufficiency scale by Kuiken and Vedder [15], which specifically includes four sub-dimensions: sufficiency of viewpoints, task requirements, coherence and cohesion, and comprehensibility. The above scale is divided into five grades, with a full score of 5 and a minimum score of 1.

To make the essay scoring fairer, our scoring is divided into two steps: one round of trial evaluation and one round of formal scoring. In the first round, the two raters will conduct a trial evaluation of 58 written texts, which account for about 20% of the total, based on the scoring criteria. After the scoring is completed, the Pearson correlation coefficient of the two raters' scores will be calculated; When the coefficient is below 0.85, the two raters discuss the scores of the essays with greater differences and re-score them until the correlation coefficient reaches 0.85 or above. In the second round, two raters complete the scoring of the remaining written texts. The final score for each essay is the average of the scores given by the two examiners. The quality of each learner's writing was rated by averaging the scores of the five active essays.

3.3.2. Outline of the semi-structured interview

Referring to Liu *et al.*'s [16] questions on the students' experiences and feelings regarding intelligent composition feedback, and interview questions on human-computer collaborative instructional design for graduate students, an interview outline on the effectiveness of GAI-assisted Chinese writing teaching was developed. The specific questions include: (1) What roles did GAI play in the stages of conception, formal writing, and revision and reflection? (2) What do you think are the advantages of the teacher and GAI in Chinese writing instruction? (3) What roles do you think GAI and the teacher play respectively in the writing process? (4) What difficulties did you encounter in the process of using GAI? Are there any aspects of GAI that you are not satisfied with? In what aspects do you expect future AI-assisted Chinese writing tools to improve further? (5) Would you like to continue using GAI in your Chinese writing process? Explain the question in more detail during the interview and ask questions at the right time to gain rich, deep, and detailed insights and understanding of the question from the natural context.

4. Results

4.1. Comparison of the differences in Chinese writing ability and emotion among learners

The study conducted an independent sample *t*-test using SPSS software to statistically test the learning outcomes of the two primary dimensions of content quality and language quality and their corresponding secondary indicators of the students in the two classes. The results are shown in **Table 2**. In the content quality dimension, the GAI class was significantly better than the traditional class in terms of viewpoints

sufficiency (t = -2.393, P = 0.020), but no significant difference was achieved in terms of task requirements, coherence and cohesion, and comprehensibility. In the language quality dimension, the GAI class significantly outperformed the traditional class in grammar (t = -3.057, P = 0.003) and vocabulary (t = -3.612, P < 0.001), while there was no significant difference in Chinese characters between the two groups.

Table 2. Results of the independent sample *t*-test for learning outcomes between the GAI class and the traditional class (SD: standard deviation)

First-level dimension	Second-level dimension	GAI class		Traditional class		*t*	*P*
		Mean	SD	Mean	SD		
Content quality	Sufficiency of viewpoints	2.88	0.99	2.31	0.78	5.728	0.020
	Task requirements	4.18	0.72	3.73	0.98	3.930	0.052
	Coherence and cohesion	3.67	0.85	3.34	1.03	1.799	0.185
	Comprehensibility	4.17	0.41	4.26	0.37	0.630	0.431
Language quality	Grammar	3.57	0.52	3.11	0.62	9.346	0.003
	Vocabulary	3.63	0.50	3.13	0.56	13.050	<0.001
	Chinese characters	4.79	0.27	4.84	0.19	0.696	0.408

4.2. Learners' views and evaluations of GAI-assisted Chinese writing instruction

Based on semi-structured interview data, this study uses category coding and content analysis (as shown in **Table 3**) to explore the role division between teachers and students and GAI in human-computer collaborative writing instruction and students' evaluation of GAI-assisted design.

In terms of role division, learners consider the teacher to play the following roles: teaching leader and common resource provider, responsible for topic selection, section design, and guidance on questioning methods; Teaching organizer, adjusting the pace of teaching based on system feedback and student status; Professional instructors, relying on their teaching experience and humanistic care, provide personalized text feedback. In contrast, the role of GAI is to provide personalized resource providers with material for writing through machine learning and online resources; Companion communicators simulate real conversation environments to regulate expression habits; Language error correctors identify grammatical, lexical, and Chinese character errors to improve the accuracy of expression. In summary, teachers and GAI complement each other in Chinese writing instruction by playing to their strengths.

The interview text shows that the majority of Chinese learners hold a positive attitude towards GAI-assisted Chinese writing teaching and believe that it has many advantages, including resource expansion, supplementation of cultural background and examples; Contextualized training to enhance language application skills through immersive conversations; Interest stimulation to boost confidence and motivation in writing; Immediate feedback to assist in text revision and standardizing expression; Learning situation monitoring, precisely identifying weak points to adjust learning strategies; Multi-dimensional cultivation, integrating language habits, cultural cognition, and thinking innovation to promote the achievement of higher goals.

However, feedback was also received from students during the interviews that the more functions of the platform led to an increase in initial cognitive load; GAI's feedback focused on error identification, lacked in-depth analysis of vocabulary usage, and students still relied on teacher guidance; Expect to improve GAI's questioning skills and humanistic care.

Table 3. Interview text data categories, primary coding, and typical categories

Study questions	Categories	Elementary coding	Typical entries (examples)
What roles did the teacher and GAI play respectively in the writing process?	The role of teacher	Instructional design leader, common resource provider, instructional activity organizer, key opinion provider	The teacher combined cultural knowledge to teach us how to optimize expressions with GAI.
	The role of GAI	Personality resource provider, companion communicator, language error corrector	Robots accompany me to practice conversations and correct grammatical errors.
How do learners view GAI-assisted Chinese writing instruction?	Advantages	Enrich teaching resources, simulate real contexts, enhance writing confidence, provide real-time feedback for revision, support self-positioning, and cultivate advanced abilities	GAI recommends a lot of examples, and there are more materials for writing compositions.
	Shortcomings	Increased cognitive burden, insufficient humanization, inaccurate understanding of the question, broad and nonspecific answers	Robots respond too broadly and sometimes fail to understand complex questions.
	Expectations	Learn GAI questioning skills and enhance agent humanistic care	Hopefully the AI can imitate the teacher's tone and increase encouraging feedback.

5. Discussion

5.1. GAI can effectively improve the quality of students' writing content and language

In terms of content quality, GAI classes performed significantly better than traditional classes in the dimension of viewpoint sufficiency, indicating that generative AI can enrich students' comprehensible input and further supports scholars' functional positioning of GAI as providing input corpora and online translation and question-and-answer. However, the effects of GAI's support on other aspects of writing content quality have not yet emerged, possibly due to the limited duration of the experiment and other factors such as individual AI literacy and learning style. In terms of language quality, GAI classes have significant advantages in grammar and vocabulary, consistent with the findings of Boudouaia *et al.* [10] and Tsai *et al.* [17]. The interview texts of the learners showed that the intelligent chatbot was able to help students correct grammatical errors and expand vocabulary through diverse resource inputs and real-time interactive feedback. However, there was no significant difference in Chinese characters between the two groups, and it is speculated that both groups maintained high accuracy because the online input method reduced writing errors.

5.2. Teachers and GAI play complementary roles in writing instruction

In the human-machine collaborative writing instruction designed in this study, the teacher played the roles of instructional design leader, instructional activity organizer, key opinion provider, and common resource provider. Students were guided to use GAI reasonably and accumulate cultural knowledge through task chain construction, schedule adjustment, and emotional interaction. GAI, for personality resource providers, language error correctors, and companion communicators, uses natural language processing and machine learning techniques to dynamically generate cross-cultural case libraries, analyze learning data in real time, and precisely diagnose language errors. The learner said in the interview that the intelligent robot can provide more abundant resource materials according to their needs than teachers, and can provide real-time feedback and modification

of their writing. Teachers, on the other hand, have the advantage of emotional communication and humanistic care, and their rich teaching experience makes their feedback suggestions more feasible. This functional positioning is in line with the findings of Cai [18] and Gu [19]. This study confirms that teachers and GAI complement each other, forming a synergistic ecological pattern that significantly improves the effectiveness of writing teaching.

5.3. The majority of Chinese learners have a positive attitude towards GAI-assisted Chinese writing

Learners generally use GAI to complete core tasks such as language resource acquisition, real-time error correction, and writing process support, which are divided into three stages: the conception stage, building a framework through dialogue and expanding thinking with the help of the question guidance mechanism; in the output stage, materials are obtained by using multilingual translation functions and cross-cultural case libraries, vocabulary and grammar examples are sought by relying on the standard language expression of the agent, and language expression is optimized; the revision phase relies on error diagnosis to improve accuracy. Learners' functional descriptions of GAI further support the results of the data analysis, which are consistent with the findings of Xu and Zhao [20]. Meanwhile, the model significantly boosts writing confidence, reduces language anxiety, and stimulates students' willingness to continue using it. It further supports the findings of Boudouaia *et al.* However, some learners also pointed out that the current system has three limitations: a heavy initial cognitive load due to its complex functions, an insufficient in-depth analysis of GAI feedback vocabulary, and the need to improve the ability to guide questions and human interaction design.

6. Conclusion

This study, through empirical analysis and theoretical exploration, reveals the multi-dimensional value of generative artificial intelligence (GAI) in empowering Chinese writing teaching. Based on this, the following three inspirations are drawn. Firstly, GAI technology can construct student portraits through multi-dimensional data modeling, supporting teachers in precisely formulating hierarchical goals and matching resources. Combined with the "practice-feedback-correction" closed-loop mechanism to achieve dynamic monitoring and personalized error correction, ultimately forming a "teaching-learning-evaluation integration" framework to promote the coordinated development of language ability and cultural literacy. Secondly, taking advantage of the complementary strengths of humans and machines, based on the "teacher-machine-student" ternary structure, teachers need to focus on the cultivation of higher-level abilities and promote the transformation of the teaching paradigm towards "holistic development" through emotional support and in-depth cultural interpretation; GAI, relying on multimodal data analysis and language generation technology, provides real-time feedback and resource support, and the two form a symbiosis of "value guidance-technology support." Finally, the key to the transformation of the human-machine collaborative education paradigm lies in the mutual empowerment of AI literacy between teachers and students. Teachers need to develop a composite ability of "intelligent diagnosis-human intervention," taking into account technology adaptation and cultural sensitivity. Students need to enhance their digital competence and regulate human-computer interaction behavior. Ultimately, through technological tools, we elevate from "auxiliary means" to "relationship reconstructors" to achieve a new form of education in human-machine intelligence symbiosis.

To sum up, this study, with intermediate Chinese learners as the subjects, systematically explores the effectiveness of GAI-assisted Chinese writing teaching through a combination of controlled experiments and

interviews. The results showed that the GAI class was significantly superior to the traditional class in terms of the dimension of the sufficiency of viewpoints in content quality and the dimension of grammar and vocabulary in language quality. In the interviews, learners indicated that in GAI-assisted Chinese writing instruction, teachers and GAI played complementary strengths and held a positive attitude towards the model, but also pointed out problems such as insufficient depth of feedback and initial cognitive load. Based on this, this study further proposes an optimized path for human-machine collaborative Chinese writing teaching at the theoretical level, with the aim of providing a more practical reference for the digital transformation of international Chinese language education.

Disclosure statement

The author declares no conflict of interest.

References

[1] Center for Language Education and Cooperation, 2022, The Forum on Digital Development of International Chinese Language Education Successfully Concluded: Technology Empowerment and Innovation-Driven Advancement Propel High-Quality Development of Online International Chinese Education.

[2] Yang Y, Jian, JR, Zhang YT, 2025, Second Language Writing Research in the Context of Generative AI: Current Status, Issues, and Prospects. Language and Culture Studies, 33(2): 135–138.

[3] Sun L, 2011, Principles and Models of Teaching Chinese as a Foreign Language Writing. Education and Vocation, (11): 104–105.

[4] Wang CM, 2016, The Learning-Promoting Function of the "Continuation Task" in College Entrance English Examinations. English Learning, (11): 10–11.

[5] Wu S, 2008, Application of the Process-Genre Approach in Teaching Chinese as a Foreign Language Writing. Modern Chinese (Teaching Research Edition), (3): 21–24.

[6] Zong SH, Han XP, 2015, Analysis of Unit Structure in Malaysian Chinese Independent Secondary School Textbooks Based on the "Integration of Reading and Writing" Concept. TCSOL Studies, (3): 59–70.

[7] Zhu XY, 2007, Research on Multimedia-Assisted Teaching of Chinese as a Foreign Language Writing. Journal of University of International Relations, (4): 69–74.

[8] Liu L, Zhou XB, Gao XS, et al., 2023, Panel Discussion: "ChatGPT Is Here: New Opportunities and Challenges for International Chinese Education" (Part 1). Language Teaching and Linguistic Studies, (3): 1–14.

[9] Guo K, Wang DL, 2024, To Resist It or to Embrace It? Examining ChatGPT's Potential to Support Teacher Feedback in EFL Writing. Education and Information Technologies, 29(7): 8435–8463.

[10] Boudouaia A, Mouas S, Kouider B, 2024, A Study on ChatGPT-4 as an Innovative Approach to Enhancing English as a Foreign Language Writing Learning. Journal of Educational Computing Research, 62(6): 1509–1537.

[11] Wang M, Xu X, et al., 2023, Unleashing ChatGPT's Power: A Case Study on Optimizing Information Retrieval in Flipped Classrooms via Prompt Engineering. IEEE Transactions on Learning Technologies, 2023(8): 1–13.

[12] Wang YM, Wang D, Liang WY, et al., 2023, "Aladdin's Lamp" or "Pandora's Box": Potential and Risks of ChatGPT in Education. Modern Distance Education Research, (2): 48–56.

[13] Wu JF, Zhou W, Lu DW, 2019, Assessing Writing Quality in Chinese as a Second Language by Korean Native Speakers: Linguistic Features and Content Quality. Chinese Teaching in the World, 33(1): 130–144.

[14] Cheng Y, 2022, Research on Chinese L2 Writing Quality Based on Grammatical Complexity. Language Teaching

and Linguistic Studies, (5): 10–22.

[15] Kuiken F, Vedder I, 2017, Functional Adequacy in L2 Writing: Towards a New Rating Scale. Language Testing 34(3): 321–336.

[16] Liu SJ, Li Y, He YW, et al., 2022, Can AI-Generated Feedback Improve Secondary School Writing Instruction? Journal of East China Normal University (Educational Sciences), 40(9): 90–104.

[17] Tsai C, Lin Y, Brown IK, 2024, Impacts of ChatGPT-Assisted Writing for EFL English Majors: Feasibility and Challenges. Education and Information Technologies, 2024(29): 22427–22445.

[18] Cai W, 2023, Chinese Learning and Teaching in the ChatGPT Era. Language Teaching and Linguistic Studies, (4): 13–23.

[19] Gu L, 2023, The Impact of ChatGPT on Training Teachers of Chinese as a Foreign Language and Countermeasures. Journal of Yunnan Normal University (Teaching and Research on Chinese as a Foreign Language Edition), 21(3): 63–70.

[20] Xu JJ, Zhao C, 2024, The Role of Large Language Models in English Language Teaching. Foreign Language Education in China, 7(1): 3–10 + 90.