ISSN 2663-8169 Vol. 8, Issue 6: 62-67, DOI: 10.25236/IJNDE.2026.080610

Research on AI-driven Personalized English Writing Feedback Mechanism

Ying Zhai

Guangdong University of Science & Technology, Dongguan, 523083, Guangdong, China

Abstract: *This study focuses on the construction of personalized English writing feedback mechanism driven by artificial intelligence, which breaks through the limitation that existing automatic writing evaluation tools only focus on surface language form error correction, and proposes a four-layer system architecture (data layer, processing layer, strategy engine layer and interaction layer). In this study, a three-level differentiated feedback model based on students' level (L1 basic level, L2 development level and L3 proficiency level) is designed, which combines the wrong "fingerprint" strengthening strategy and metacognitive excitation mechanism to guide students to correct themselves instead of directly providing answers. At the same time, a three-stage teacher-AI collaborative process is constructed to retain the core value of teachers in complex content evaluation and emotional support. Through an 8-week quasi-experimental study of two parallel classes (60 students in total) in Grade Two of a middle school, the results show that the post-test writing performance of the experimental group is significantly higher than that of the control group ($t=5.67$, $p<0.01$), with an average increase of 4.34 points. The recurrence rate of high-frequency errors in the experimental group was significantly lower than that in the control group ($P < 0.05$). The feedback viewing rate of students reached 94.2%, and the revised adoption rate reached 76.8%, and the students with weak foundation (L1 layer) made the most obvious progress. The research shows that the AI-driven personalized feedback mechanism can effectively improve students' writing performance, reduce the error recurrence rate and enhance students' active revision behavior, which provides a feasible technical path and practical paradigm for the digital transformation of foreign language education.*

Keywords: *AI; Personalized feedback; English writing*

1. Introduction

The importance of English as an international common language is increasingly prominent. As the core skill of language output, writing is not only the embodiment of students' comprehensive language application ability, but also the key link of cultivating their logical thinking and intercultural communication ability. At present, English writing teaching in China is still generally faced with the dilemma of "teaching in large classes, heavy burden of teachers' correction, lagging feedback and homogenization" [1]. The traditional "teacher's correction-student's correction" model is often difficult to take into account the individualized learning needs of students of different levels, which leads to students' low interest in writing, insufficient motivation for revision, and bottlenecks in improving teaching efficiency [2].

The new generation of information technologies, represented by natural language processing (NLP), deep learning, and generative AI, is reshaping the education ecosystem at an unprecedented speed. Especially with the rise of Large Language Models (LLMs), machines are able to understand, analyze, and even generate human language more accurately. These technologies provide infinite possibilities for realizing the organic combination of large-scale education and personalized training [3]. Introducing AI technology into English writing teaching and building an intelligent and personalized instant feedback mechanism have become an important breakthrough to solve the current teaching problems and promote the digital transformation of foreign language education.

Although automatic writing evaluation tools such as Grammarly have appeared in the market, the existing systems mostly focus on surface language form error correction, and lack deep diagnosis and adaptive guidance on article content logic, text structure, rhetorical devices and ideological depth. How to effectively integrate the efficiency of AI with teachers' humanistic care and establish the optimal feedback path of "man-machine collaboration" is still a weak link in current research. This study focuses on the AI-driven personalized English writing feedback mechanism. Breaking through the

ISSN 2663-8169 Vol. 8, Issue 6: 62-67, DOI: 10.25236/IJNDE.2026.080610

limitation that traditional AI-assisted writing is limited to "right and wrong judgment", we are committed to building an intelligent model that can provide differentiated and developmental feedback according to individual differences of students.

2. AI-driven personalized feedback mechanism construction

2.1. System architecture design

In order to realize an efficient and scalable personalized feedback system, a four-layer architecture (see Figure 1) is adopted, which is from bottom to top: data layer, processing layer, policy engine layer and interaction layer.

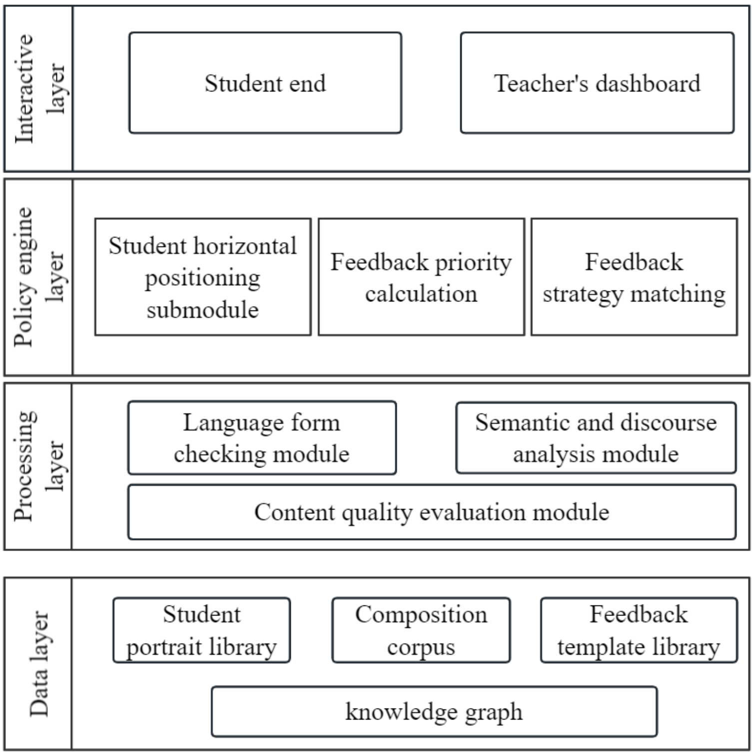

Figure 1 Personalized feedback system architecture

The data layer is used as multi-dimensional writing data collection and storage, and the student portrait library stores students' historical writing scores, common error types, weakness labels and modification behavior data. The composition corpus stores students' original texts and revised versions by stage, genre, and topic classification. Feedback template inventory includes feedback statement templates tailored to different types of errors and skill levels. Constructing an English writing knowledge point network using a knowledge graph for association diagnosis.

The processing layer is the AI diagnostic module, and the language form checking module detects spelling, grammar and punctuation errors based on open source tools, and marks the error level. The semantic and discourse analysis module uses the pre-training model to score coherence and identify arguments. Using dependency syntax analysis to detect sentence monotonicity. The content quality evaluation module calls the large language model to evaluate logical clarity, argument sufficiency and reader awareness, and outputs the structural analysis results.

The strategy engine layer is the core of personalized decision-making, and the student horizontal positioning sub-module divides students into three levels: L1 (basic), L2 (developing) and L3 (proficient) according to their historical error frequency, revision success rate and current composition complexity. The feedback priority calculation adopts weighting algorithm to determine the feedback order. For example: Priority score = 0.4 × error severity+0.3 × matching degree between the error and students' weaknesses+0.2 × importance of the error in the current genre-0.1 × student history correction rate. The strategy engine selects the feedback style from the template library or the generation model according to the student level and the error type.

ISSN 2663-8169 Vol. 8, Issue 6: 62-67, DOI: 10.25236/IJNDE.2026.080610

The interactive layer is the front-end and collaborative interface, and the student side displays feedback in the form of annotation and sidebar, and supports the buttons of "one-click modification", "ask AI reasons" and "view examples". The teacher's bulletin board displays the overall error distribution of the class, essays with low system confidence, and records of student AI interactions, supporting the teacher to cover or supplement feedback. The interactive layer uses Python+FastAPI to build the back end, and the front end uses React; AI module handles conventional grammar with fine-tuning small model, and calls commercial LLM API for complex content evaluation and sets cache to reduce costs.

2.2. Personalized feedback strategy design

Traditional tools provide unified error correction sentences, and this study emphasizes differentiation, development and meta-cognitive stimulation [4-5]. The three-layer feedback mode based on students' level is shown in Table 1.

Table 1 Three-layer feedback model based on students' level

Student level	Feedback focus	Feedback language features	Example (for the grammatical error "She go to school")
L1 (basic)	Language accuracy, simple structure	Direct prompt+correct demonstration, avoiding terminology; Encourage trying	"The subject here is She, and the verb should be goes. Try changing it: She ___ to school.”
L2 (developing)	Sentence richness and paragraph cohesion	Point out mistakes and compare correct/wrong sentences, and provide 2~3 synonymous rewriting options; Arouse thinking	“She goes to school. But when the last sentence is used in the past, does it need to be consistent? "
L3 (proficient)	Logical coherence, rhetorical style, argument depth	Ask enlightening questions; Recommend academic writing strategies	"Your tense is correct. Can you try to reorganize this sentence with' not only ... but also' to strengthen the contrast? "

Dynamic personalization is based on the reinforcement strategy of error "fingerprint", which records the top three types of errors in each student's first five writings. When these types appear again in the new text, the system automatically increases the feedback intensity. The radar chart is generated by the change of error rate in each writing and the success rate of rewriting after adopting feedback, so that students can see their four-dimensional ability evolution of "grammar-vocabulary-coherence-logic".

Meta-cognitive stimulation guides self-correction rather than giving answers directly. For students with above-average level, they don't directly say "there is no coherence here" when they encounter logical fracture, but ask: "There is a transitional sentence between the second paragraph and the third paragraph. Which sentence do you think can be used to connect' problem description' and' solution'?" Personalized feedback strategy can show the similar wrong sentences that the student successfully modified last time as a reference, and strengthen the positive transfer. Feedback generation combines rule template with LLM dynamic generation. Common errors of rule template used in L1/L2; LLM is called for L3 and complex logic problems, and constraint style is prompted by Few-shot.

2.3. Teacher -AI collaborative model

Personalized feedback can not be fully automated, and the core value of teachers lies in calibration, position compensation and strategy optimization. This paper designs a three-stage collaboration process (Figure 2).

ISSN 2663-8169 Vol. 8, Issue 6: 62-67, DOI: 10.25236/IJNDE.2026.080610

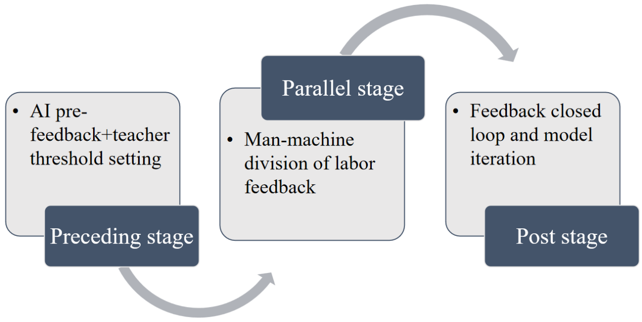

Figure 2 Three-stage collaborative process

1) Preceding stage

After students submit their compositions, AI generates initial feedback within 30 seconds and does not display it for the time being, and calculates the confidence of each feedback. Teachers can set in the "intervention rules" of the system that when the argument recognition confidence of an article is less than 0.7, or the logical score is lower than the threshold, the composition will be automatically pushed to the teacher's priority manual correction queue. Therefore, it prevents AI from giving false guidance on complex content and reserves the space for teachers to deeply intervene.

2) Parallel stage

AI is responsible for low-order errors, high-frequency error identification, model matching suggestions, and real-time data statistics. Teachers are responsible for high-level thinking evaluation, complex emotional support, and handling innovative expressions that AI cannot recognize [6]. Every composition that teachers see in the system has been pre-marked with AI. Teachers can adopt, modify or delete AI feedback with one click, and add their own comments. The system records every intervention of the teacher for fine-tuning the subsequent model.

3) Post stage

The system records whether students view, modify, ignore or ask each AI feedback. If some kind of feedback is ignored by more than 80% students, it will be marked as inefficient strategy and submitted to teachers for review. Every weekend, the teacher's revised feedback will be used as high-quality training data to fine-tune the AI strategy model and make the system closer to the teaching style and standards of specific classes. The system generates "Top3, the weak point of writing this week" and "Distribution of common logical fallacies" to assist teachers in adjusting next week's teaching plans.

3. Experimental implementation and data analysis

3.1. Experimental purpose

Compared with traditional teachers' unified feedback, the AI-driven personalized English writing feedback mechanism proposed in this study is verified to be effective in improving students' writing performance, reducing error recurrence rate and enhancing students' correction behavior.

3.2. Experimental subjects and groups

The subjects are 60 students in two parallel classes of Grade Two in a middle school, and there is no significant difference in English proficiency pre-test. The experimental group (30 people) used the personalized feedback system of this study, including hierarchical strategy, metacognitive prompt and teacher -AI collaboration. The control group (30 people) only received written feedback from regular teachers, without AI assistance and personalized stratification.

3.3. Experimental process

The experimental process lasts for 8 weeks, and the specific experimental operations are shown in Table 2.

ISSN 2663-8169 Vol. 8, Issue 6: 62-67, DOI: 10.25236/IJNDE.2026.080610

Table 2 Experimental process

Stage	Time	Experimental group operation	Control group operation
Pre-test	Week 1	Unified proposition writing (argumentative essay), collecting pre-test scores and error baselines	Same left
Intervention implementation	2~7 weeks	Once a week, the writing task, the system immediately generates personalized feedback, students modify; Teachers check kanban every week and supplement it selectively	Once a week for the same writing task, the teacher will grade and send it back, and the students will make revisions
Post-test	Week 8	Another proposition writing (quite difficult), collecting post-test results	Same left

3.4. Data collection and indicators

The total writing score is based on the English writing scoring standard of NMET, with a score of 25 out of four dimensions: content, language, structure and standardization. Error recurrence rate, the type of TOP3 error that the same student appeared in the pre-test, and the proportion of recurrence in the post-test. In the study, the activity of behavior was modified, and the experimental group recorded it automatically, feeding back the viewing rate and the adoption rate of modification; The control group was obtained by manually counting and modifying traces.

3.5. Data analysis

The difference of post-test scores between the two groups was compared by independent sample T test. Paired sample t test was used to compare the improvement of pre-and post-test scores in each group. Chi-square test was used to compare the difference of error recurrence rate between the two groups. Descriptive statistics are used to modify behavior data.

3.6. Result

There was no significant difference between the two groups ($P > 0.05$). The post-test experimental group was significantly higher than the control group ($t=5.67$, $p<0.01$), and the improvement range was even greater (See Table 3).

Table 4 shows that the error recurrence rate of the experimental group is significantly lower than that of the control group, indicating that personalized feedback is helpful for long-term error correction. The high adoption rate and questioning behavior show that students' participation and metacognition activation are effective (see Table 5).

Table 3 Comparison of pre-and post-test writing scores

Group	Pre test mean ± standard deviation	Post test mean ± standard deviation	Average increase	T-value	P-value
Experimental group (n=30)	17.23±2.14	21.57±1.86	+4.34	6.23	<0.01
Control group (n=30)	17.01±2.33	18.45±2.07	+1.44	2.14	0.038

Table 4 Error reproduction rate (using high-frequency error types as an example in previous measurements)

Type of error	Reproduction rate of experimental group	Recurrence rate of control group	Chi square value
Subject-predicate agreement	23.3%	56.7%	7.03
Missing article	30.0%	63.3%	6.53
Weak sentence cohesion	36.7%	70.0%	6.79

Table 5 Analysis of internal modification behavior of experimental group (system log)

Behavior index	Numerical value
Average feedback viewing rate	94.2%
Average modification adoption rate (click one-click modification or manual modification)	76.8%
The average number of times each composition asks AI	1.3 times
Students take the initiative to check the scale of growth radar map	68%

The average progress of L1 (basic) students is +5.2 points, L2 (developing) students is +4.1 points, and L3 (skilled) students is +2.8 points. It shows that students with weak foundation benefit most obviously, which is in line with the design expectation of personalized feedback. Personalized feedback mechanism has significantly improved students' writing performance, especially for L1 students. It effectively reduces the recurrence rate of similar errors, indicating that feedback has long-term migration. Students' active modification behavior is active, and their metacognitive participation is higher than that of traditional feedback mode.

4. Conclusions and suggestions

This study shows that the AI-driven personalized English writing feedback mechanism can significantly improve students' writing ability, especially for students with weak foundation (L1 level), and their writing scores are improved by an average of 5.2 points, far exceeding the traditional feedback mode. Through the differentiated hierarchical strategy (L1/L2/L3), the system realizes both accurate error correction and ability development: L1 students focus on language accuracy, L2 students strengthen sentence cohesion, and L3 students focus on logical depth. In addition, the dynamic reinforcement strategy effectively reduces the error recurrence rate, which proves that feedback has long-term migration. Students' active participation is high, and the feedback rate is 94.2%. Questioning stimulates metacognitive ability and promotes independent correction.

It is suggested that man-machine collaboration should be deepened, teachers should lead complex logic evaluation and emotional support, and AI should undertake high-frequency and low-order error analysis to form a complementary mechanism; It is suggested that the feedback intensity should be dynamically adjusted based on students' wrong fingerprints, and personalized suggestions should be generated by combining LLM. It is suggested that priority should be given to covering groups with weak foundations and gradually expanding to middle and above levels; Incorporate the teacher's correction records into the fine-tuning of the model to form an intelligent teaching ecology of "feedback-learning-iteration".

Acknowledgements

The authors acknowledge the Research project (Teaching Research Program) of Guangdong University of Science and Technology in 2025: “Innovative application and ethical norm exploration of generative AI in foreign language writing teaching” (Project Number: GKZLGC2025194).

References

[1] Wang Xueli. Three-dimensional Personalized Teaching Reform of Business English Writing Course Based on Learner Profiles [J]. Western Quality Education, 2026, 12 (04): 40-44. DOI:10.16681/j.cnki.wcqe.202604009.

[2] Li Zhoukang, Yan An, Xie Kaiyun, et al. Exploration and Practice of English Teaching in Grassland Science under the Background of Artificial Intelligence [J]. Guide to Smart Agriculture, 2025, 5 (19): 172-175. DOI:10.20028/j.zhnydk.2025.19.041.

[3] Zhang Yongwen. Research on the Integrated Model of Teaching Learning Evaluation for Regional Adaptation of English Listening and Speaking in Vocational Colleges Empowered by Artificial Intelligence: A Case Study of Guangxi Vocational Colleges [J]. Journal of Guangxi Vocational and Technical College, 2026, 19 (01): 15-20. DOI:10.26996/j.cnki.1674-3083.26.1.003.

[4] Chen Yuping. Application and Exploration of AI in English Personalized Teaching [J]. Journal of Huaiyin Normal University (Natural Science Edition), 2025, 24 (04): 365-368. DOI:10.16119/j.cnki.issn1671-6876.2025.04.015.

[5] Mo Qiulan. Research on English Assignment Design in the Era of Artificial Intelligence: A Case Study of Writing Design [J]. Language and Cultural Studies, 2025, 33 (05): 60-63. DOI:10.19954/j.cnki.lcr.2025.05.040.

[6] Wang Qi. Research on Personalized Learning Path for College English Based on Artificial Intelligence [J]. Journal of Hubei Adult Education College, 2025, 31 (04): 63-68. DOI:10.16019/j.cnki.cn42-1578/g4.2025.04.004.