

Original Paper

Application of Artificial Intelligence Feedback in College English Writing Assessment

Jun Chen

Faculty of Foreign Languages, Huai'an University, Huai'an, China

Received: August 20, 2026 Accepted: September 22, 2026 Online Published: October 4, 2026

doi:10.22158/selt.v14n4p1

URL: <http://dx.doi.org/10.22158/selt.v14n4p1>

Abstract

In the fast-changing digital education and the rapid development of artificial intelligence, AI-generated feedback has been widely used in college English writing assessment as a valuable teaching support. The traditional manual marking is typically characterized by a heavy workload, slow feedback, and subjective judgment, making it difficult to meet students' individual needs for writing improvement. AI feedback can detect surface-level language errors and improve the efficiency of assessment because it can provide immediate feedback, follow standard correction procedures, and offer detailed diagnostic guidance, thereby helping students to revise their writing autonomously. In this study, the analysis of student essays titled "The Impact of Influencer Culture on Youth Consumption Patterns" will be done to evaluate the utility of feedback provided by artificial intelligence. Additionally, rule-based AI correction and a more detailed review by teachers considering the practical value and importance of the learning process will be applied. It is proved that AI is quite efficient when dealing with language mistakes, yet it is insufficient in the logical development of an essay, in-depth interpretation, and critical thinking that are still the responsibility of teachers. From the author's teaching experience, a new model that combines the efforts of AI and teachers can be proposed, according to which AI makes surface-level corrections and teachers concentrate on developing the higher-order skills of students.

Keywords

artificial intelligence feedback, college English writing, formative assessment, automated scoring system, human-machine collaboration

1. Background of Digital Foreign Language Teaching and Research Purpose

1.1 Existing Background of Digital Foreign Language Teaching

Under the background of the digital wave and deep integration of "Internet Plus" education, the paradigm of foreign language teaching is witnessing a transformation from the previous

experience-based approach to a new data and intelligence-driven approach. The development of technologies including natural language processing and deep learning makes artificial intelligence no longer just a tool but an essential part of the processes of language acquisition, which changes the interaction environment. Intelligent automated writing evaluation and feedback systems can solve a lot of problems during writing classes (Shen, 2026). As for higher education, college English writing plays an important role in language development, logical thinking, and academic expression. Therefore, the system of college English writing assessment faces unprecedented reform pressure and challenges.

Currently, teachers in universities have to deal with the contradiction of large-class teaching and big individual differences of students. In particular, traditional manual grading has the limitations of time and energy, leading to delayed, subjective and incomplete feedback. One single form of assessment cannot be sufficient to satisfy the needs of the students for instant and individualized learning aid. In the meantime, the generative artificial intelligence, which is equipped with characteristics like instant reaction, affordable repetition, and scalability, has soon developed into one of the most effective supports in English writing education (Cui et al., 2026). The intelligent revolution in second language writing education has turned into one of the most unstoppable issues in the college English education reform (Yang & Wu, 2026).

Nevertheless, the integration of technological elements faces several controversies. University foreign language teachers should take into account numerous challenges and implications that are imposed by artificial intelligence on English writing training, including the possible deterioration of students' capacity for independent reasoning (Liu, 2026). In this context, simply integrating additional technology has ceased to be sufficient to cope with emerging teaching demands, and creating a human-machine collaboration approach with multi-dimensional feedback has become necessary. The purpose of the study is to conduct a critical analysis of the benefits of using AI feedback in terms of increasing the efficiency and accuracy of the process of assessment, and at the same time, explore the structural weaknesses of the existing algorithms in processing high-level semantic, emotional, and cultural aspects. Within the framework of this work, the empirical case connected with the topic "The Impact of Influencer Culture on Youth Consumption Patterns" will be explored, and the low-score traps associated with the violation of formatting guidelines by students will be analyzed in detail. Moreover, the mechanical error correction of rule-based AI will be contrasted with the pragmatic evaluation conducted by teachers. On the basis of this comparison, an attempt to develop the dynamic human-machine collaboration. In this way, it seeks to move from reliance on a single technology toward a more balanced, multidimensional educational ecology, offering theoretical support and practical pathways for improving the quality of college English writing instruction.

1.2 Purpose and Importance of Research

The goal of this study is to critically analyze the two-fold impact of AI-based feedback systems in the process of college English writing assessment. Although the benefits that AI brings into this process are significant in terms of increasing assessment efficiency and accuracy of language form analysis, the

study is going to take a closer look at the shortcomings of algorithmic models, when it comes to their ability to deal with higher-level semantics, intentions of the writer, and cultural complexity. Given the change in digital foreign language teaching paradigm, the traditional practice of manual grading is restricted by the lack of time and energy, which leads to the delay of feedback and its incompleteness, and, thus, cannot satisfy students' needs in terms of timely and personalized assistance (Shen, 2026). Therefore, the development of the dynamic approach to human-machine collaboration in the process of evaluation becomes a necessity. AI is supposed to be used as an instrument for checking basic language forms, owing to the high level of its language processing skills. At the same time, teacher feedback should focus on stimulating critical thinking, deepening the exploration of ideas, and guiding rhetorical variety, thereby moving from reliance on a single technology toward a more balanced, multidimensional educational ecology.

In presenting a concrete case on the topic "The Impact of Influencer Culture on Youth Consumption Patterns" for empirical discussion, this essay aims to shed light on the problem of low scores resulting from students' neglect of basic writing formatting rules. It is clear from the analysis of the case that, despite the fact that AI can easily fix various mistakes like word spelling errors, inappropriate mixing of Chinese and English punctuation, and sentence grammar, it does not show great results in encouraging students to move to logical reasoning and proper academic writing. It demonstrates the drawbacks of using purely AI writing evaluation systems and proves the indispensable role of human teachers in conducting thorough evaluation from a practical point of view. In addition, this work touches upon the issue of the matching of the moment of AI generation with writing cognitive processes of learners and concludes that an appropriate strategy of intervention can help alleviate students' concerns about over-reliance on technology and develop their independent writing skills (Wu & Ji, 2026; Yang & Wu, 2026).

The theoretical significance of this study lies in enriching the theoretical framework of human-machine collaborative evaluation in second language writing instruction and offering a new perspective for understanding the integration of technology and humanistic concerns in educational assessment. Its practical significance is reflected in providing university foreign language teachers with an actionable pathway for instructional improvement: by optimizing the division of feedback responsibilities, AI can be used to increase the efficiency of basic language error correction, while teachers can play a leading role in cultivating higher-order thinking. In this way, the quality of college English writing instruction can be comprehensively improved, helping students develop sound digital academic literacy.

2. Benefits of AI Feedback on English Writing Evaluation

2.1 Systems of Instant and Tailored Feedback

Long (2014) suggests an interaction hypothesis, which claims that interaction and negotiation of feedback may contribute to noticing the gap between learners' own output and target language forms and may facilitate language acquisition. The main advantage of AI feedback systems in the assessment

of college English writing is the deep integration of immediate and personalized feedback, first of all. Traditional manual grading is constrained by the cost of teachers' energy and time, which often results in delayed feedback that does not meet students' need for timely correction (Shen, 2026). However, automatic systems can use NLP technologies to recognize and correct any surface errors such as grammar, spelling, and punctuation mistakes immediately after students submit their work. Not only does such an approach help reduce the cycle time from writing to getting feedback, but it also makes the whole assessment process a part of the writing process itself, helping students to do self-revision frequently.

Consider the following examples of student essays "The Influence of Influencer Culture on the Consumption Behavior of the Youth." Despite the presence of rather decent insights, these papers suffer from a lot of mechanical mistakes in terms of formatting. Specifically, these include lack of spacing between words ("youngpeople"), the usage of different languages for punctuation marks (full-width commas in Chinese), and other similar issues. The AI system is able to identify these spots easily and to list all of them one-by-one, suggesting appropriate corrections (the usage of English commas and proper spacing rules). The task may seem insignificant, but it solves the problem of the low-score trap because of inconsistent formatting. This proves the efficiency of using technology to make the assessment more efficient.

It becomes possible due to the implementation of a personal feedback mechanism that allows for analyzing students' learning trajectory via diagnostic capabilities. Contrary to human teachers, it is hard to keep track of each student's progress individually. That is why the AI system is able to analyze frequent errors in terms of language acquisition and offer personalized strategies. For instance, in case when a student's writing lacks lexical variety. For instance, in case a student does not use enough different words, AI can recommend synonym substitutions to enrich expression.

Lyu and Zhang (2026) note that artificial intelligence can help build thinking frameworks, and they argue for a human-machine collaborative writing model in which AI handles basic skill training and immediate feedback. This division allows teachers to have time for more important activities that will develop higher thinking skills in students and will help to form their values.

But immediacy and personalization do not mean that artificial intelligence can completely replace human judgment. Previous research has shown that AI still relies on teachers' professional judgment to interpret deeper argumentation or complex linguistic and cultural phenomena, and that the quality of its feedback is very sensitive to prompt design (Cui et al., 2026). The ideal feedback process, then, would be the use of AI as a tool for the correction of basic language forms, with teacher feedback focused on the development of critical thinking and academic expression. This human-machine collaborative evaluation strategy maintains the efficiency of technology while enhancing the depth of humanistic engagement, making it possible to move from dependence on a single technology to a more balanced, multidimensional educational ecology.

2.2 Establishing Objectivity Within the Multidimensional Scoring System

Establishing the objectivity within a multidimensional scoring system is one of the fundamental ways in which AI feedback systems address the subjectivity issue in traditional human scoring systems. In the case of college English writing assessment, the teachers' scores tend to be influenced by individual experience, their own emotion and even fatigue. On the other hand, an algorithm-driven automated assessment tool converts abstract linguistic skills into measurable data dimensions by means of quantitative metrics.

Such an objective depends firstly on structuring the breakdown of writing elements. Thus, taking into consideration the student's essays devoted to "The Impact of Influencer Culture on Youth Consumption Patterns," the AI system did not just give some total score. Instead, the process of evaluation was decomposed into several independently weighted modules like the content relevance, language accuracy, discourse consistency, and viewpoint clarity. In this particular case, in spite of numerous formatting mistakes in the student's essay, like the lack of spacing between words (for example, "youngpeople") and mixing Chinese and English punctuation, which made the language accuracy score equal to 75.0, the student got rather high scores for viewpoint clarity (85.0) and structural organization (100.0). Such a multifaceted evaluation helps students to realize their performance in terms of different writing qualities and prevents the situation when a general impression score hides some specific advantages or disadvantages. As pointed out by Wang and Wang (2026), the advent of generative artificial intelligence makes knowledge production paradigm shift, and research literacy becomes a crucial factor in the training of innovative English majors. Similarly, a multidimensional and objective scoring system in writing instruction helps students to develop a precise awareness of their own writing abilities, which enables them to improve their academic expression in a more targeted way.

The multidimensional scoring system increases the scientific basis of the evaluation procedure by using the weighted calculation procedure. Scores of different dimensions are not simply summed up; on the contrary, different weights are set to the scores depending on the significance of the dimensions in a particular genre. For instance, in the case of argumentative writing, content relevance and logical consistency are usually more important than mere lexical richness. Such weighted calculation, based on linguistic theory, allows for making sure that the results of the evaluation procedure reflect students' real writing skills rather than mere statistics. However, one should keep in mind that algorithmic objectivity does not guarantee absolute accuracy. The AI program is still capable of misunderstanding higher-level semantic comprehension as well as cultural context. Thus, the goal of creating an objective multidimensional scoring system is not to make machine evaluate students independently but to provide teachers with an objective reference point. With the help of such cooperation between humans and machines, the AI program can be used for objective correction of basic language structures, while teachers' professional experience will compensate for the shortcomings of the algorithm.

3. Problems of AI-powered Writing Evaluation Systems and Evidence from Cases

3.1 Drawbacks of AI in Evaluating Deep Logic and Intentions

There are many examples of how effective AI evaluation systems have been in improving the efficiency of marking college writing tasks by focusing on language forms. However, recently, the issue of structural flaws in such systems regarding problems of deep logic and intentions has arisen. Today, the mainstream natural language processing is mostly statistical and rule-based. That is, they can detect grammatical mistakes, but not the practical meaning of the text and the cultural context.

Consider the empirical analysis of the essays about “The Impact of Influencer Culture on Consumption Behavior among Youth” as an example. Despite the presence of proper structuring and the development of the definite point of view as well as satisfying the minimal requirements to be met for a good CET-4 argumentative essay, the text is rife with various issues of formatting, including lack of spaces between some words (e.g., “youngpeople”) and using both Chinese and English punctuation marks. The system is able to pinpoint these issues and give recommendations regarding the necessary adjustments, solving the low-grade problem caused by formatting inconsistencies. However, when relying only on the machine-made feedback, students might tend to pay attention only to mechanical aspects and overlook the issue of developing proper logic and academic language proficiency.

Furthermore, there is a lack of the necessary flexibility of the AI assessment systems when dealing with sophisticated semantic networks. For example, consider the terms “blind consumption” and “impulsive buying” used by the student in the essay. The artificial intelligence can interpret them only as a collocational problem but will not be able to understand the youth subculture psychology and the society anxiety hidden behind the terms. According to Cui et al. (2026), task-oriented composite prompting can significantly raise the percentage of higher order cognitive feedback provided. However, even with the optimal prompt engineering, the AI will not be able to mimic the quality of the feedback that human teachers give from the pragmatic point of view. This lack of evaluation prevents students from creative thinking in solving writing tasks that require deep analysis and development of rhetorical diversity.

The trend towards standardized feedback generated by AI may also hinder students’ unique academic expression. Since the algorithmic models usually rely on massive corpora to generate recommendations, those will be more universal rather than specific and would be inappropriate for the needs of particular disciplines and writing styles. As observed by Wang & Wang (2026), the generic nature of recommendations generated by AI makes it less suitable for academic tasks; moreover, the problem of “hallucinations” only increases the awareness of risks for the students without helping them develop creativity. Consequently, when developing a collaborative human–AI approach to assessing students’ performance, it is important to consider AI as an instrument of simple language correction, leaving the responsibility for critical thinking, ideas’ elaboration, and rhetorical diversity to teachers. In this way, it is possible to transition from one-sided use of technologies to a more comprehensive system.

3.2 Case Analysis Based on Student Writing

Drawing on the empirical case “The Impact of Influencer Culture on Youth Consumption Patterns,” it is particularly instructive to examine the structural limitations of AI feedback systems in college English writing assessment. The case centers on an argumentative essay discussing how social media influencers promote products through short videos and live streaming, thereby shaping adolescent consumption habits. Although the essay’s overall structure conforms to CET-4 conventions, with a clear thesis and logical coherence, it contains numerous mechanical errors at the level of basic language form—such as run-on words (e.g., “youngpeople”), mixed use of Chinese and English punctuation (e.g., full-width commas), and missing spaces.

From this perspective, the AI-assessment system reveals its effectiveness as an error correction device. It not only identifies but also categorizes all eleven mistakes that need to be corrected, making clear suggestions on how it should be done, such as using English commas rather than Chinese commas and correcting spacing norms. The use of the rule engine in this case helps to correct low score mistakes associated with formatting issues and shows the evident benefit of the technology from the perspective of increased efficiency of assessment. However, if students start relying only on this type of feedback, they will definitely get too fixated on form and forget about the need to work on their logical arguments and academic writing skills. According to Wu and Ji (2026), the level of students’ dependency on generative AI can be significantly affected by intrinsic motivation and writing anxiety.

Moreover, AI analysis is highly ineffective in coping with higher-level semantics and culture. In the context of student’s paper, such expressions as “blind consumption” and “impulsive buying” might be considered by AI only as collocational phenomena without being able to recognize the psychology of youth subculture and social anxieties associated with them. From the existing literature, it can be concluded that although AI is able to detect mistakes on the surface level, it lacks in helping students with rhetorical diversity and critical thinking. According to Yang and Wu (2026), completely embedded intelligent assistance enables iterative ideation and acquisition of higher-order writing skills, while post-hoc feedback groups tend to experience poor revision outcomes because of their “black box” thinking. In the current situation, although the AI detected repetitive connectives used, it did not help the student to think about the ethical problems associated with influencer culture and find alternative solutions.

In this context, evaluative feedback provided by human instructors is indispensable. They focus not only on the correctness of linguistic form but, more importantly, on pragmatic-oriented assessment, including the logic of argumentation, relevance of evidence and affective intention. Comparing the two feedback types, we see that AI feedback is stereotyped and formulaic, while teacher feedback is customized and heuristic. Thus, it is crucial to build up an interactive assessment method between humans and AI. The former will play the role of the tool for correcting the linguistic form due to its ability to detect errors at a quick pace, while the latter will aim to encourage critical thinking, explore the ideas more deeply, and cultivate rhetorical diversity. According to Liu (2026), university foreign

language teachers must help students accept and adopt the new normal of technology-assisted writing. This is how the change from one-dimensional technological addiction to the multi-dimensional educational ecology can be achieved.

4. Practical Implementation Strategies and Concluding Remarks for College English Instruction

The development of an integrated human-AI collaborative evaluation scheme is the only way to address the current contradiction of technology dependence vs. loss of humanistic depth in the context of college English writing teaching. From the point of view of this research, it is necessary to create a two-channel feedback model that includes “AI-based fundamental error correction + teacher-led high-level guidance.” In such a model, the role of AI technology can be understood as the use of algorithms for rapid identification of errors of linguistic form, which will help greatly in freeing teachers from the task of mechanical proofreading. In addition, the teacher’s role will move away from micro-corrections related to linguistic forms to the macro-construction of logical thinking and critical thinking development through the deep exploration of ideas and rhetorical variety. This will not only reduce the fears of students who rely too much on automatic tools but will also ensure the correct transfer of affective intention and cultural context in academic writing.

In terms of the practice of instruction implementation, great efforts need to be made to develop the feedback literacy of college students. With regard to the cognitive problems generated due to the involvement of AI in writing tasks, teachers need to encourage the students to critically evaluate the strengths and weaknesses of AI feedback to avoid mechanically following the feedback of AI and weakening the critical thinking of the students. Teachers can choose topics related to practical life and help students analyze the value conflicts involved in those problems based on AI feedback in their text revisions and thus make English writing an integrated process of linguistic-cognitive production, critical thinking, and value investigation. Lyu and Zhang (2026) argue that the future writing instruction will have to focus more on speculative writing than technical writing in order to foster innovation in intelligent and personalized teaching. In this way, instruction needs to view AI immediate feedback as the beginning rather than the end of self-revision of the students.

The use of AI feedback in college English writing assessment does not mean replacing the role of the teachers, but reconstructing the interactive relation between teaching and learning through the empowerment of technology. While current algorithmic models may remain limited by structural factors in handling higher-order semantics, human-AI collaboration in assessment is capable of using the strengths of AI in increasing the efficiency and accuracy of the evaluation process while maintaining teachers' dominance in promoting higher-order thinking skills among students. Such a solution provides an effective way to ensure the quality of college English writing education in order to develop high-quality talents with sound digital academic literacy and an independent critical spirit. With technological advancement, more attention should be given to the integration of technology and humanities in the future development of assessment systems to meet the challenges of the new

paradigm in digital foreign language education.

References

- Cui, Y., Schunn, C. D., Liu, H. S., & Yu, X. C. (2026). A study on model differences and prompt-regulating effects of generative-AI feedback for English writing. *Foreign Language Research*, (01), 96-102.
- Liu, T. T. (2026). Research on practice and reform of artificial intelligence-empowered English writing teaching. *Journal of Shandong Institute of Commerce and Technology*, (02), 60-67.
- Long, M. (2014). *Second Language Acquisition and Task-Based Language Teaching*. Hoboken: Wiley Blackwell.
- Lv, M. L., & Zhang, Y. (2026). Artificial intelligence and the innovation of English writing teaching: Impacts and prospects. *Journal of Western*, (16), 22-25.
- Shen, X. M. (2026). Research on AI-driven automated evaluation and feedback system for English writing. *Information & Computer*, (12), 233-237.
- Wang, Y. L., & Wang, C. H. (2026). Generative-AI-enabled cultivation of research literacy for Chinese English majors: A case study of the course “Research Methods and Academic Writing”. *Foreign Languages and Their Teaching*, (01), 53-65+93+147.
- Wu, Y., & Ji, J. J. (2026). Research on influencing factors of undergraduates’ reliance on generative artificial intelligence in English writing. *Foreign Languages Research*, (03), 78-83.
- Yang, L., & Wu, H. W. (2026). The intervention timing of generative artificial intelligence and college students’ English argumentative writing quality: A quasi-experimental study. *Foreign Languages and Their Teaching*, (02), 41-53+145.