Feedback in College English Writing from Teachers and Artificial Intelligence Platform：A Comparative Study

Liu Houtao[1*], Ma Wenjia[2*], Wu Tingting[3], Xin Chuanhua[4]

[1]Jiamusi University, Jiamusi, China

[2]Jiamusi University, Jiamusi, China.

[3]School of Foreign Languages, Jiamusi University,Jiamusi, China.

4School of Foreign Languages, Jiamusi University,Jiamusi, China.

Email: 47345984@qq.com, mawenjia8466@163.com, yuemeng123@163.com, xinchuanhua_2000@163.com

Abstract: With the development of modern teaching technology, intelligent writing correction platform has emerged and become an effective way to supplement traditional feedback from teachers under the new situation. The main theoretical basis of this study is dynamic evaluation theory. The experimental group and the controlled group were assigned three writing tasks. The statistical analysis software SPSS was used to analyze the data results. The literature summarizes the development status of teacher feedback and intelligent platform correction at home and abroad. The experimental method is to compare the similarities and differences between teacher feedback and intelligent platform correction in specific experiments. Through questionnaires and interviews, we can understand the attitudes of teachers and students towards the intelligent correction platform.

The research results show that the two feedback methods are consistent and have obvious differences. Teacher feedback has more significant advantages in the two dimensions of text structure and content, while the intelligent writing correction platform is too general and procedural in the feedback of these two dimensions. In comparison, the feedback of the intelligent writing correction platform on vocabulary and grammar is more detailed, giving students detailed information. The combination of the two ways of feedback and the application of them to the teaching of English writing in college will have a more ideal effect, which will have a positive effect on the improvement of college students' English writing level.

Keywords: Dynamic evaluation theory; Teacher feedback; Feedback from AI writing platform; contrastive study

Introduction

The College English Curriculum Standards (2017 Edition) issued by the Ministry of Education mentioned that "special attention should be paid to improving students' ability to think and express in English", "English curriculum evaluation system should be improved to promote the effective formation of core literacy". The standard requires paying attention to the role of evaluation in promoting teaching and learning. Teachers should guide students to gradually adjust and supervise themselves in learning objectives, learning methods and learning process. It can be seen that composition plays an important role in the distribution of weight values in college English teaching.

However, at the teacher level, due to the lack of actual classroom teaching time, limited energy, and heavy task of evaluation and correction, the focus of classroom teaching is to explain vocabulary independently, refine grammar and structure sentences, but it does not have the opportunity to combine multiple elements of English writing for detailed teaching. In

addition, each English teacher undertakes the teaching tasks of multiple teaching classes. If careful composition correction is carried out, the work intensity of teachers will be greatly increased, and teachers cannot spend more time and energy in the classroom to correct problems in writing one by one. In addition, there is almost no time for explaining and writing compositions alone in college English classes. The number of compositions assigned each week is small, so the internal motivation of students' writing becomes weak. More time is spent on learning objective problems. A single traditional teacher feedback model can not fully meet the students' desire for substantial improvement in writing performance, which requires a new type of writing feedback model, which is complementary to teacher feedback.

The College English Curriculum Standards (2017 Edition) requires "to attach importance to the application of modern information technology and enrich curriculum learning resources". English teaching in the new era requires corresponding adjustment in the mode. Online English writing evaluation system has been rising in recent years. The evaluation system of the intelligent platform provides feedback from multiple dimensions - such as "vocabulary, word collocation, sentence composition, text, and content relevance". To a certain extent, the feedback of the intelligent platform will reduce the work pressure of front-line English teachers. However, the validity of its impact on college English writing, the similarities and differences between the two feedback methods, and whether they can work together are all worth further studying.

Literature Review

In the process of globalization and the rapid development of science and technology today, modern teaching methods have been replaced rapidly[1] . Online and offline teaching has played a more coordinated role in actual teaching. The advantages of computer as an auxiliary teaching method are highlighted[2]. The research and development of intelligent assessment system for English writing is in the ascendant in China.

In foreign countries, the development of the intelligent writing correction platform began in the 1960s. Now the most influential correction platforms are: Project Essay Grader, Intelligent Essay Assessor, IntelliMetric and Electronic Essay Rater. The researchers pointed out that the research and development of multiple intelligent correction platforms has played a rich role in the dimension of feedback[3]. This intelligent correction platform not only has the automatic scoring function, but also provides feedback on students' compositions from multiple dimensions, such as vocabulary, phrase collocation, sentence pattern, style, etc. The improvement of students' writing performance depends on the guidance of teachers. [4]With the help of the intelligent correction platform, teachers can not only reduce the workload, but also supplement the deficiencies of teachers' feedback [5].

In terms of research on the advantages and disadvantages of intelligent platform feedback, there are also two voices in foreign countries. Scholars who support the use of intelligent platform correction believe that this modern feedback method is more time-saving and labor-saving and more objective when compared with teachers' manual correction. They believe that the intelligent correction platform will play a more objective role in improving evaluation skills and feedback efficiency[6].However, some scholars still believe that there is no unified scientific evaluation standard for this kind of intelligent feedback, and some even suggest that this new type of feedback violates the principle of direct communication between people in teaching, and can not provide vivid and direct feedback on the theme of the composition as teachers do[7]. In the positive and negative attitudes, there is still a neutral view, and researchers focus on the advantages and disadvantages of intelligent correction at the same time. It is suggested that in order to play the role of the intelligent correction platform to the utmost, it must be modified by students themselves several times before being corrected by teachers[8].

In China, the research and development of intelligent writing correction platform is later than that of foreign countries. The earlier research started in the 21st century[9] . Professor Liang Maocheng's team tried to establish an automatic composition correction model in 2005, becoming a leading figure in this field in China. After the research and development of Professor Liang's team, intelligent platforms have sprung up like mushrooms[10] . Among them, there are several intelligent correction platforms that are widely used - www.pigai.org, i-write, TRP teaching resource platforms. In terms of theoretical research, the role of computer feedback in second language writing, but this research is not deep enough in terms of validity and impact[5]. After that, many scholars conducted comprehensive research, mainly focusing on the comparison between teacher feedback and intelligent platform feedback. The biggest advantage of AI technology is that there is no time and space limitation, which can be used as an aid to the reform of English teaching methods. We should expand multiple channels to apply modern teaching resources and information[11] . Researchers conducted a comprehensive research on three feedback methods and students' attitudes based on the dynamic evaluation theory[12] . The three methods are traditional teacher

feedback, AI feedback and "teacher+AI feedback"[13] . The research results show that teacher feedback and AI feedback under guidance work together, with better results.

The intelligent correction platform used in this paper takes "Pigai correction network" as an example, which is an online composition correction system designed and developed based on corpus and cloud technology, with diversified software design[11] . the reliability and validity of the AI Platform and found that it mainly comments on sentences, refining and expanding vocabulary and phrases, but the comments on the logic of discourse and content are not ideal[14] .

To sum up, the main research on two forms of writing feedback at home and abroad is to demonstrate their effectiveness. This paper also focuses on the improvement of college English writing by the differences and complementary between feedback from teachers and the feedback of the correction in students' English writing practice.

Description of the Study Area:

In this study, 70 students from two parallel classes of Jiamusi University in Heilongjiang Province were selected as the research objects, and the two parallel classes were selected as the experimental group and the control group, and the teachers were the same. The research period was the first semester of freshman year. During the one semester experiment, the experimental group adopted the form of feedback from the correction network, while the controlled group adopted the form of teacher feedback.This study analyze the data of the pre-test, the experimental process, the post test.

1Analysis of pre-test data

After the research objects were determined, the two classes were divided into experimental group and controlled group . The same English topic was used for the test. The composition scores of the experimental group and the parallel class were tested by independent sample t test. SPSS data analysis software was used for comparison and analysis. The analysis results were as follows:

	N	M	SD	t	MD	P
Experiment group		69.66	9.652			
	35			0.768	1.829	.445
Controlled group		67.83	10.254			

Table 1 Independent sample t-test of pre-test scores of controlled group and experimental group

The results of independent sample t-test in Table 1 show that there is no significant difference between the scores of students in the experimental group and the controlled group (t=0.78, df=68, p=. 445>0.05), indicating that there is no significant difference in the scores of the selected two classes.

2. Comparison of teacher feedback and online feedback on feedback content

In terms of the content to be corrected, the concerns of the two feedback are different. In the first experiment, feedback from teachers and feedback from Artificial Intelligence Platform were conducted in five dimensions: vocabulary, grammar, structure, writing norms, and ideological content. Vocabulary includes part of speech selection and phrases; grammar includes tense, number, person, voice, etc.; writing specifications include the correct use of letters, capital and lowercase, punctuation marks, whether the word spacing is reasonable, whether the handwriting is neat and neat, whether the structure includes whether it is divided into three paragraphs, whether it has cohesion and coherence, whether the theme idea includes accurate conception, and whether the required information points are presented completely. After the feedback, the students' errors in these five aspects were counted. (See Table 2)

	Feedback from teacher	Feedback from AI Platform	t	MD	SD	P 值
Vocabulary	149	158	1.759	.257	.146	.083
Grammar	139	143	.642	.114	.178	.523
Norms	155	152	-1.413	-.257	.182	.162
Structure	60	38	-6.896*	-.629	.091	.000
theme	49	23	-6.350*	-.743	.117	.000

* p＜0.05

Table 2 Statistics of errors after feedback from teachers and feedback from AI platform

3. Comparison of scores of two feedback methods in correcting compositions

Any of the feedback methods in this experiment will ultimately serve English teaching, and strive to improve the teaching effect of teachers and students' academic performance. Therefore, the score must be the most critical indicator point. Therefore, this experiment is to find out whether the two kinds of feedback are consistent in the composition scores given. In order to make the experiment more scientific, students in the experimental group and the controlled group need to produce corresponding scores in their compositions. The experimental results in Table 4-3 below are obtained by analyzing and sorting out the scores of the two classes after three times of writing. The statistical tools are SPSS and Excel.

	Cotrolled group (n=35)	Experiment group (n=35)	MD	SD	T	P
	M	M				
First Test	67.829	69.657	1.829	2.380	.768	.445
Second Test	67.829	69.914	2.086	2.330	.895	.374
Third Test	70.257	72.829	-2.571	2.129	-1.208	.231

Table 3 Comparison of Three Composition Scores between the experimental group and the controlled group

According to the data of independent sample t-test in Table 3, there is no significant difference in the three composition scores between the experimental group and the controlled group (the first time: t=0.768, df=68, p>0.05); (The second time: t=0.895, df=68 p>0.05); (The third time: t=-1.208, df=68 p>0.05,). The scores given by the feedback from teachers are consistent with those given by the AI Platform, and the scores obtained by the students in the three composition exercises are distributed normally. According to the significant difference of the third writing performance in the independent sample t-test results, the difference between the controlled group and the experimental group is large. After finding out the actual reasons, the author found that two students did not complete the required writing within the specified time due to the unskilled operation of the platform, and two students only wrote more than 80 words, which did not reach the required number of words, feedback from teachers deducted a lot of marks for the unfinished information points of students, while the AI platform gave higher marks than the teachers. One of the students had an improper learning attitude and wrote an article in the textbook that had nothing to do with the theme of the composition. It was completely off the topic, and the teacher gave it 0 points. However, the AI platform suggested that it was "suspected off the topic", but it still gave it 10 points. Therefore, in the third exercise, the average scores of the two classes are relatively different from the other two.

Effect comparison of two feedback methods

After three composition exercises, an experimental post test was conducted. Compare the average values of the five dimensions in the first composition exercise of the experimental group with the average values of the five dimensions in the post test of the experiment, and also compare the average values of the five dimensions in the first composition exercise and the post test composition of the controlled group (see Tables 4 and 5).

	First Test	Post-Test	D-value
Vocabulary	13.628	14.200	.572
Grammar	14.325	14.771	.346
Writing Norms	13.800	13.971	.171
Structure	13.571	13.857	.286
theme	13.886	13.686	-0.2

Table 4 Comparison of average scores of five dimensions in the first composition and post test composition of the experimental group

	First Test	Post-Test	D-Value
Vocabulary	13.542	14.029	.487
Grammar	14.613	14.743	.013
Writing Norms	13.529	13.6000	.071
Structure	13.329	15.012	1.683
Theme	13.486	15.228	1.899

Table 5 Comparison of average scores of five dimensions in the first composition and post test composition of the controlled group

From Table 4 and Table 5, it can be found that the performance of the controlled group using teacher feedback in the five dimensions has improved to varying degrees. For the experimental group feedback by the AI Platform, the average scores of each item tested after the experiment were not improved except for the theme dimension. Compared with the first exercise, the scores of the other four dimensions were improved. This shows that the two feedback methods are effective in terms of vocabulary, grammar, norm, structure and theme.

After completing the independent comparative analysis of the two classes, the average scores of post test items in the controlled group and the experimental group shall be compared and analyzed, and the software SPSS shall be used for data analysis. See the following table (Table 6) for the data analysis results

	Controlled Group(n=35)	Experimental Group(n=35)	MD	SD	t	P
	M	M				
Vocabulary	14.029	14.2000	.171	.782	.219	.827
Grammar	14.743	14.771	.286	.451	.063	.950
Writing Norms	13.6000	13.971	.371	.775	.479	.633
Structure	15.012	13.857	-1.155	.645	-2.128*	.037
Theme	15.228	13.686	-1.543	.640	-2.410*	.019

*p＜0.05

Table 6 Comparison of the average scores of the five dimensions in the post test compositions of the controlled group and the experimental group

The post test results of independent sample t-test in Table 6 show that there is no significant difference in vocabulary, grammar and norm (vocabulary; t=0.219, df=68, p>0.05); (grammar: t=0.063, df=68, p>0.05); (Norms: t=0.479, df=68, p>0.05). There was significant difference between the two dimensions of structure and theme (structure: t=-2.128, df=68, p<0.05); (theme: t=-2.410, df=68, p＜0.05) It can be found that the feedback of this AI platform used in the experimental group will have a positive effect. Especially in the aspects of word hints, recommendation and expansion; The average score of the controlled group is lower than that of the experimental group. Therefore, in the future writing teaching and practice, it is suggested that we can use AI Platform to complete vocabulary, grammar and normative feedback, which can greatly reduce the workload of teachers and improve the accuracy of students' words. However, in the two discourse categories of grammar and structure, the average scores of the two classes are quite different, and the significance value is far lower than 0.05, so it shows that the teacher feedback used in the controlled group is much better than the AI platform in these two aspects. The two feedback methods are obviously different in these two aspects. Therefore, in the future writing correction, teacher feedback can mainly focus on the structure and theme.

Results and Discussion
Based on the results of the experiment and data analysis, this chapter will discuss the main findings of the experiment, the significance of the experiment to the future teaching of English writing, the development prospects and the existing limitations. The main findings are as follows:

The similarities and differences between teachers' feedback and the feedback of the AI Platform in terms of the content, score and method of evaluation are as follows: First, teachers' feedback and the feedback of the AI Platform will correct errors in grammar and vocabulary in students' compositions. Secondly, both feedback methods can give close scores. Thirdly, the feedback of teachers and the feedback of the AI Platform of correction have the feedback of the text structure to varying degrees.

The differences are as follows: First, although the two feedback methods can correct errors in vocabulary and grammar, the external forms are different. After pointing out the problems, the feedback of the AI Platform will give suggestions in written form, such as replaceable words, structures, especially vocabulary. The suggestive content is more than the teacher's feedback, and students can choose a suitable revision content from multiple reasonable plans. Second, even though the two will give relatively similar composition scores, according to the survey results, the scores given by the AI Platform are generally higher than the scores given by teachers. Third, teachers' feedback will give detailed tips on the overall logic, structural arrangement,

clear theme, and accurate conception of the composition. When teachers correct the composition, they will give icon and text feedback on the structure of the text, and will detail the shortcomings and brief suggestions on the paper. However, the AI Platform will not give too many suggestions in this regard. In terms of text structure and subject content, the feedback of the AI Platform is relatively formatted, and there is little or no personalized analysis.

Teacher feedback plays an irreplaceable role in all teaching, including writing teaching. The effective feedback information output by teachers is helpful to the improvement of students' learning and revision ability, and the effect of this help is huge. However, due to the existence of large classes, the proportion of teachers and students does not match, and the teaching task of teachers is very heavy, so it is impossible to feed back all aspects of the composition. All the major elements of the composition correction requirements cannot be fed back to every student, and every composition can be fed back without omission. In addition, as required by the domestic teaching trend, the application of modern auxiliary teaching means is also a requirement to improve the core quality in the "New Curriculum Standard". Therefore, in the actual English writing teaching, in order to achieve the best teaching effect for teachers, an auxiliary intelligent platform is needed. First of all, due to the lack of feedback from the intelligent platform on the text and theme content of the composition, teachers are suggested to make intervention corrections in these two aspects, so as to make students' mistakes in the text and theme more detailed, and gradually cultivate students' awareness of text writing. At the same time, the intelligent platform has strong advantages in terms of vocabulary expansion, high reliability of standardization, and efficient and comprehensive statistics of grammar error rate.

In general, the advantages of AI platform are obvious to all. Schools can use AI platform to correct students' English compositions to reduce the burden of teachers' correction. When a large number of basic errors are checked, AI platform can provide feedback efficiently and quickly, which can save a lot of human resources. However, it is unrealistic to rely solely on AI platform to improve students' English composition level. AI platform has great limitations in correcting and commenting, and it still has great room for progress in promoting students to write truly high-quality English compositions. Therefore, teachers can apply the AI platform according to the actual teaching situation, but they can't rely on it too much. On the basis of comprehensive understanding of students writing level, every teacher still needs to guide and comment on writing after refining their English compositions. Only in this way can English writing level of students be truly improved.

REFERENCES

[1] Berg,E. (1999)The effects of trained peer response on ESL students' revision types and writing quality[J]. *Journal of Second Language Writing*, (3): 215-241.

[2] Bull, J. & Mc Kenna, C. (2004) *A Blueprint for Computer-assisted Assessmen*[M]. London: Routledge.

[3] Chandler, J. (2003) The Efficacy of Various Kinds of Error Feedback for Improvement in the Accuracy and Fluency of L2 Student Writing [J]. *Journal of Second Language Writing,* (12):267-329.

[4] Feuerstein, R,Y.Rand & M.B. Hoffman.*The Dynamic Assessment of RetardedPerformers: The Learning Potential Assessment Device, Theory, Instruments, and Techniques*[M]. Baltimore: University Park Press, 1979.

[5] Gue'nette,D.(2007) *Is Feedback Pedagogically Correct? Research Design Issues in Studiesof Feedback on Writing* [J]. Journal of Second Language Writing, (16):40-53.

[6] Halliday, M. A. K.(1984). *Textual cohesion and Chinese Speaker of English* . Language Learning and Communication. (10):12-18

[7] Keh,C. *Feedback in the Writing Process: A Model and Methods for Implementation*. ELT Journal, 1990(44): 294-305

[8] Lantolf, J. P.(2004) *Dynamic Assessment of L2 Development: Bringing the Past into the Future*. Journal of Applied Linguistics, (1): 123-129.

[9] Larsen-Freeman,D (2005).*Teaching language: from grammar to grammaring*.Beijing: Foreign Language Teaching and Research Press,

[10] Lids, C. S. (1987). *Dynamic Assessment: An International Approach to assessment of Learning Potential*[M]. New York: The Guilford Press,

[11] Lidz, C.S.(2003). *Dynamic assessment (Learning potential testing, testing the limits)* [A].London:SAGE, , 337-343.

[12] Yin Xiaobei Lin Fuxian (2019) *Practice of Multi-Dimensional Feedback Model of English Writing from the Perspective of Dynamic Evaluation* . Journal of Fujian Medical University, (03): 32-38

[13] Zhang Baoguo, Guo Xiaoyu (2018) *The Application and Research of Dynamic Assessment Theory in Middle School English Writing*. Teaching Overseas English,(07): 64-65

[14] Zhu Qiujuan (2010) *Feedback Mechanism and Writing Teaching - Domestic and Foreign Research and Its Implications for College English Writing* Teaching Journal of Changchun University of Science and Technology (Social Science Edition), (3): 160-162