

# The Impact of Automated Evaluation Feedback on Students' Writing Revision

Ximing Yuan

**Abstract**—To make essay scoring more manageable, Automated Writing Evaluation (AWE) programs have been developed and widely used in China. It makes immediate feedback possible and helps students do self-revision instantly. This study attempts to explore the influences exerted by automated scoring feedback on students writing revision and analyze their revision process. with Juku. Pigai AWE system as an example, the research involves 45 non-English major students in Northwestern Polytechnical University and adopts their writing samples to build a small corpus. With careful study of their modification process and detailed analysis of their writing, the results shows that apart from providing holistic ideas on the article, Pigai system provides a specific review from 6 aspects as well. Besides, students can actively respond to the feedbacks and constantly revise their writing according to the feedback points, but mostly in technical details, morphemes, words, and phrases, short sentences. Online modification expands knowledge input, reduces language errors, improves students' writing quality, and effectively promotes autonomous learning in students' interaction with machines. However, limitations existed as well, such as lack of effective feedback on the content. Finally, the researcher proposed a "four-step" writing and evaluation model for future improvement.

**Index Terms**—Automated writing evaluation, college English writing, feedback, revision

## I. INTRODUCTION

Writing is considered the most essential way to improve English learning and the most useful tool to assess learning outcomes. Writing essays can effectively assess students' vocabulary band, grammar, the skill to organize and construct ideas, and the ability to use English to express themselves. But in the meantime, writing is also a complex cognitive skill in language learning. It requires not only sustained, deliberate practice but frequent, appropriate, and effective feedback as well [1]. Immediate ratings and evaluation from the teacher provide students with essential information related to the revision of their essays, and thus achieve best effect of writing improvement. However, grading essays is often hindered by a lot of difficulties. It is rather time consuming, which is always a headache for teachers. To make essay scoring more manageable, Automated Writing Evaluation (AWE) programs within the range of those awarded by human raters have been developed as a way to meet this challenge. There have appeared many well-known AWE systems at home and abroad, such as Project Essay Grade (PEG), E-rater, which claim not only to provide total and sub-scores of essays, but also offered personalized feedback, including both macro (essay center, structural arrangement, content organization, etc.) and micro (grammar, spelling, idiomatic usage, etc.)

Manuscript received August 15, 2023; revised September 15, 2023; accepted October 4, 2023.

Ximing Yuan is with Northwestern Polytechnical University, Xi'an, Shaanxi, China. E-mail: yuanxm@nwpu.edu.cn

aspects [2]. With the popularity of AWE systems, much research has been conducted on its application in writing [3–5]. Some empirical studies have focused on the effects on students' writing ability [6–8]. Some scholars have investigated the reliability, validity and accuracy [9–11]. To this end, this study aims to explore how the feedback from an AWS system affect student writing revision, and find out what effective revision aspects the system offers.

## II. LITERATURE REVIEW

### A. Problems in Learning and Teaching L2 Writing in China

Focusing mainly on vocabulary and grammar, but lack of systematic training in English writing in high school, Chinese students seldom have chances to put English learning into practical use. After entering college, they are required to take the national English tests, College English Test Band 4 (CET-4) and Band 6 (CET-6), which have writing as the first testing item. But we always find that their essays are made up of ill-formed sentences. Owing to the bad writing environment and the simple way of assessment, the English writing quality of Chinese EFL learners at college allows no optimism. There exist a lot of problems which need immediate action: 1) Students' writings are full of high frequency words and syntactical mistakes. Their writings are always full of high frequency words and various kinds of mistakes in grammar, word choice, word forms and collocation. 2) Students' writings are full of Chinese-English expressions. 3) Sentence writing is lack of variety.

In the meantime, due to the large number of students a Chinese English teacher usually has to face, problems like consuming a great amount of time and energy on evaluating repeated drafts of student writing stand in the breach. Under this condition, students have very limited practices. The teacher can also hardly give them immediate and detailed diagnostic feedback, instead, they evaluate the students' writing with the general impression and score the writing. Their feedback contains little individual feedback information for every student, and can hardly achieve the purpose of helping the students promote their writing to a higher level. Considering all these special problems, the researcher proposes online AES system to deal with the situation and improve students writing.

### B. Application of AWE in L2 Writing

The development of AWE systems began in the 1980s to offer instantaneous diagnostic feedback on essays [12]. Most systems have adopted the latest achievements in statistics, natural language processing and artificial intelligence to offer students multidimensional, concrete and timely feedback. After 40 years of hard work, the automated assessment has progressed greatly in many countries. The current existing

approaches, such as Project Essay Grade (PEG), Intelligent Essay Assessor (IEA), Educational Testing Service (ETS), Electronic Essay Rater (E-Rater), Criterion, My Access!, Pigai, and Iwrite, have their own strengths in different aspects and their scoring quality is even comparable to human raters. The feasibility and advantages of scoring essays with such computer-based techniques can be concluded in the following perspectives: 1) practicality: grading writing drafts is no longer a burden and time-consuming task for English teachers; 2) consistency: assessment consistency in the scores could be achieved; 3) immediate feedback. Experts and language teachers regard it as a revolution. However, on the other hand, many researchers also pointed out some limitations. For example, they argued that AWE systems cannot provide effective higher-level revisions such as ideation, organization, and style [8–13].

### *C. Revision and Feedback*

According to cognitive model of writing, improving and enhancing the quality of writing depend largely on writing revision. It refers to any change made during the writing process, which involves observing the gap between the expected and the original writing, deciding on appropriate and feasible changes, and taking action to achieve the desired effect of the revision [14]. The changes can be minor lexical revisions to even significant idea modification of the first draft.

In revision process, feedback is the critical element, which provides necessary references for students to make changes to their writings, such as inadequate content, poor logical linking, errors in vocabulary and grammar, etc. It includes both corrective feedback at the surface level and feedback at the discourse level. With the help of the feedback, students learn about the negative evidence in their writing and start the second writing process of revision. According to Interactionist Hypothesis, the process of continuous modification based on automatic feedback is actually the process of interaction between students and the machine.

## III. THE RESEARCH STUDY

### *A. Research Questions*

A large number of previous research studies have been conducted to prove the effectiveness of AWE systems in essay writing but seldom analyze the revision impact. In order to find the influences exerted by automated scoring feedback on students writing improvement, this paper attempts to answer the following questions:

- 1) What feedback points can Automated Evaluation System offer?
- 2) How do students utilize the feedback points to revise their writing?
- 3) How do students respond and feel about using Juku in their writing revision?

### *B. Research Participants*

The research involved 45 non-English major graduate students in Northwestern Polytechnical University and adopted their writing samples to build a small corpus. These students used AWE system for writing practice for the first time.

### *C. Research Instruments*

The AWE system adopted in this study is Juku-pigai system ([www.pigai.org](http://www.pigai.org)), which is based on corpus linguistics, statistics, and cloud computing technology. Its core algorithm is to calculate the distance between students' essays and the standard corpus, and then use a mapping to convert the distance into essay scores and comments. Each submitted text essay will be cut into sentences, and compared with the standard learner corpus from 192 measurable dimensions. After a detailed analysis from perspective as discourse, syntax, collocation, and structure, the system will automatically count out the distance of the essay from the standard corpus, which will finally generate the score, comment and diagnostic feedback. Taking into account of the practical needs of Chinese students, Juku sets up different optional scoring criteria, such as criteria for CET-4, CET-6, English test for postgraduate entrance examination, and even IELTS and TOEFL. Students can revise and submit their essays multiple times, and the platform can give them grades for each submission and save all draft versions. The system offers the information of each student, such as the name, score of writing, and all the drafts submitted by the student, which serves the basic materials for the study.

In order to see how the students feel about using the system and have a further study of their perceptions to the effect of Pigai system on their writing revision process, an interview was conducted after several writing practices with Pigai. The interview questions include: How do you feel of Pigai AWE? Is it helpful? Will you revise your writing exactly according to the recommendation of the system? In what aspects do you think the AWE system offered you the most useful revision feedback?

### *D. Data Collection*

To achieve the purpose of examining the revision process of the students, the researcher assigned two writing tasks. Before writing for the first assignment, the instructor explained in detail about the functions and rules of using Pigai system. Students were required to submit the first draft within one week. They were encouraged to learn more about the feedback from the platform and revise their essays as much as possible until they were satisfied. The basic data of this research were extracted from the second assignment after they were familiar enough with the Pigai system. Students were asked to write an English business letter of 150-300 words with unlimited revision times from the automated feedback. The rating criteria followed the rules of general letter writing scoring standard with 100 points as the full mark.

### *E. Data Analysis*

In this study, both quantitative and qualitative research methods were employed.

To study how the students revise their writing and in what aspects they revise their writing, revision type was analyzed. The revision coding scheme was based on Roscoe [15], which was categorized into four types: substitution, deletion, addition, and reorganization. Substitution is to use one expression to substitute another. Deletion refers to the removal of words or expressions. Addition means the students revise the writing by adding some words, sentences or even paragraphs. Reorganization is changing the order of words, phrases, or sentences in the original writing.

To have a quantitative and textual analysis of the revision process, students' submitted drafts for the writing task was recorded and calculated. 45 students submitted their compositions, with 405 revised submissions, an average of 8 times as shown in Table I. With the feedback from AWE system, 89% students actively revised their writing and increased their score from draft 1 to the final draft (see in Fig. 1). Obviously, repeated submissions give students the opportunity to revise and increase interest in writing and revising [16].

TABLE I: NUMBER OF MODIFICATIONS

Student number	Total modifications	Unmodified students	Maximum of modifications	Minimum of modifications	Average number of modifications
45	405	5	49	2	8



Fig. 1. An example of the score-change of a student in Pigai.

To answer the research questions, AWE feedback on students' drafts were classified based upon Pigai system. Table II clearly displays the error feedback in students' first draft. Most corrective feedback was concerned with spelling, punctuations, word forms and collocations. As to sentence expression, Pigai system not only pointed out the errors, but also offers suggestions and learning tips for corrections in several versions (see in Fig. 2). After the revision process, students' errors reduced dramatically. Revision types occurred in all types, such as substitutions (52%), deletions (26%), additions (13%), reorganizations (9%). Most of the modifications were superficial revisions, while substantive content revisions occurred much less.

TABLE II: AWE ERROR FEEDBACK

	feedback error focus (draft 1) Total 443	Feedback error focus (final draft) Total 145
spelling	103	20
sentence expression	79	43
punctuation	76	20
Noun/article	48	10
verb	42	16
capitalization	33	13
collocation	21	7
Preposition/adjective/ other word use	18	6
	23	10

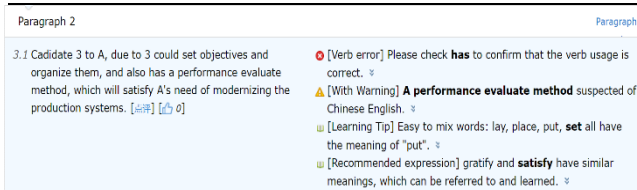


Fig. 2. An example of a sentence corrective feedback in Pigai.

Apart from the detailed specific feedbacks, the Pigai AWE system also provides general positive comment, frequency

information, and extended learning point, which can help students identify their strengths and lead to more substantial revisions.

For further qualitative study, interview was recorded and transcribed for data analysis. From the interview, the researcher finds that students think of it as an effective teaching method to improve their writing.

- 1) Around 80% students acknowledge that this net-based writing assessing mode create a positive learning environment. They can set their own writing time and writing speed. And more important, the system provides them enough chances to practice themselves.
- 2) Students think the quick individual feedback Jukupigai offers is beneficial in solving most of the problems, especially grammar mistakes and collocation problems which they sometimes cannot detect themselves. The recommended revision suggestions are of great help to improve the essay quality and satisfy their needs of lifting their scores.
- 3) 60% to 70% of the students can repeatedly correct their writings according to the feedback. They get not only encouragement but also great fun from the writing process. Furthermore, the key feature of plagiarism detection of Juku also helps prevent and cultivate self-learning responsibility, and increase motivation.

Of course, Juku-Pigai AES system is not perfect. Students feel the system cannot recognize all the mistakes and problems of students' writing. Sometimes, it even offers unnecessary or wrong feedback.

#### IV. DISCUSSION AND CONCLUSION

##### A. Findings

This paper focuses on the specific feedback points from the Pigai AWE system and studies students' revision process based on the feedback. 45 students were required to write a business letter with Pigai system, which will not only offer them the instant scores but also provide positive error feedbacks and recommendations in different aspects. Based on their perceptions of the diagnostic feedback, students made revisions to improve their writing. With careful study of the modification process and detailed analysis of their writing records, the results showed that apart from commenting on the article in general, Pigai system provides a specific review from six aspects, which are collocation statistics, learning tips, correction tips, low-frequency alerts, verb errors, and collocation errors. Out of the errors in the 45 sample writings, most of them focused on wrong use of sentence components, nouns, articles, collocation, and sentence structure. Besides, the study also found out that students actively responded to the feedbacks and constantly revised their writing according to the feedback points, but mostly stayed in surface and lower-level revisions, such as technical details, morphemes, words, and phrases, short sentences. The total revision is up to 405 times and 8 times per person. Thirdly, online modification expands knowledge input, reduces language errors, improves students' writing quality, and effectively promotes autonomous learning in students' interaction with machines. Most students regarded it a good way to improve their writing proficiency. The

number of students who have engaged in writing revisions displayed that this AWE system does stimulate their interests in further revisions, and even serves as a prompt for students to do some other modifications without being pointed out by the system like adding more paragraphs or even changing ideas. Although most revisions are substitutions, reorganizations still cover 9%.

However, defects existed as well, such as lack of effective feedback on the content, coherence and general structure, and the incompetence of recognizing sentence errors from long language units. For some complicated sentence expressions, the system even failed to recognize the errors, or provided wrong suggestions. Anyway, AWE system is computer-based, which lacks the flexible thinking ability of human raters.

### B. Limitations

There also exist some limitations in this study. The sample size of 45 is too small to generalize the research findings. Secondly, the data adopted in this study is only from one writing task sample which is not adequate enough to have an all-rounded exploration into the revision process of students' writing. Students still need more practices to get familiar with using the Pigai system. Besides, different writing tasks may also lead to different response of students' revision.

### C. Implications for Future Teaching

In L2 writing, simply relying on the engine-score is not enough. To avoid the limitations and the defects of AWE system, mixed and flexible teaching method can be utilized in future teaching to improve students' revision and writing. Teachers can set the topics and the criteria tailored to their own requirements and needs of the students. To achieve an ideal teaching effect, the traditional classroom instructing and online rating can be integrated. The suggested 4-step method is: Prewriting instructing + unlimited online rating + one human rating + classroom feedback

#### 1) Pre-writing preparation

Before writing, the writer asks the students to read as many as possible similar articles to make preparation for the final writing.

#### 2) Peer-evaluation

After finishing the draft, students are given enough time for peer-evaluation. Suggestions may range from essay content and structure to grammar and spelling.

#### 3) Online submission

Submit the text essay within the time limit, and repeatedly correct the essay according to the feedback and suggestions to make it more perfect.

#### 4) Classroom feedback

In the classroom time, the teacher can give face-to-face feedback aiming at the problems in general and especially those the system fails to reflect. This after-writing feedback tends to be very effective.

### D. Conclusion

This paper aims to explore the impact of AWE system on students' English writing revision process from the perspective of diagnostic suggestions Pigai system provides. It is obvious that the automated feedback leads to a positive reinforcement of the correct use of advanced words or phrases, correct grammar, and collocations, which helps improve the

accuracy and quality of student's writing. Though the system offers little feedback in essay structure and content, most students still regard it beneficial. Anyway, AWE system cannot take the place of human raters. If automated rating and human rating can be integrated, online feedback and classroom assessment can be combined, that could be an ideal strategy for improving writing quality of L2 learners.

### CONFLICT OF INTEREST

The author declares no conflict of interest.

### REFERENCES

- [1] C. Palermo and M. M. Thomson, "Teacher implementation of self-regulated strategy development with an automated writing evaluation system: Effects on the argumentative 24 Journal of Educational Computing Research 0(0) writing performance of middle school students," *Contemporary Educational Psychology*, vol. 54, no. 1, pp. 255–270, 2018.
- [2] L. F. Bai and J. Wang, "A critical review of the effectiveness of English AWE feedback over the past twenty years," *Foreign Languages Research*, vol. 173, no. 1, pp. 65–71, 2019.
- [3] C. S. Chew, W. C. V. Wu, N. Idris, E. F. Loh, and Y. P. Chua, "Enhancing summary writing of ESL learners via a theory-based online tool: System development and evaluation," *Journal of Educational Computing Research*, vol. 58, no. 2, pp. 398–432, 2019.
- [4] Y. Han, S. Zhao, and L. L. Ng, "How technology tools impact writing performance, lexical complexity, and perceived self-regulated learning strategies in EFL academic writing: A comparative study," *Frontiers in Psychology*, vol. 12, no. 1, pp. 752–793, 2021.
- [5] M. Liu, Y. Li, W. Xu, and L. Liu, "Automated essay feedback generation and its impact on revision," *IEEE Transactions on Learning Technologies*, vol. 10, no. 4, pp. 502–512, 2017.
- [6] B. Sun and T. Fan, "The effects of an AWE-aided assessment approach on business English writing performance and writing anxiety: A contextual consideration," *Studies in Educational Evaluation*, vol. 72, no. 1, pp. 101123, 2022.
- [7] Z. Wang, "Computer-assisted EFL writing and evaluations based on artificial intelligence: A case from a college reading and writing course," *Library Hi Tech*, vol. 40, no. 1, pp. 80–97, 2020.
- [8] J. Wilson and A. Cziki, "Automated essay evaluation software in English Language Arts classrooms: Effects on teacher feedback, student motivation, and writing quality," *Computers & Education*, vol. 100, no. 1, pp. 94–109, 2016.
- [9] X. L. He, "Reliability and validity of the assessment by the Pigaiwang on college students' writings," *Modern Educational Technology*, vol. 23, no. 5, pp. 64–67, 2013.
- [10] M. D. Shermis and J. Burstein, *Automated Essay Scoring: Across-Disciplinary Perspective*, Eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 2003.
- [11] J. Wilson, "Universal screening with automated essay scoring: Evaluating classification accuracy in grades 3 and 4," *Journal of School Psychology*, vol. 68, pp. 19–37, 2018.
- [12] J. H. Wang and M. S. Brown, "Automated Essay Scoring versus Human Scoring: A Comparative Study," *Journal of Technology, Learning, and Assessment*, vol. 6, no. 2, pp. 2–27, 2007.
- [13] S. Dikli and S. Bleyle, "Automated essay scoring feedback for second language writers: How does it compare to instructor feedback?" *Assessing Writing*, vol. 22, pp. 1–17, 2014.
- [14] J. Fitzgerald, "Research on revision in writing," *Reviews of Educational Research*, vol. 57, pp. 481–506, 1987.
- [15] R. D. Roscoe, J. Wilson, A. C. Johnson, and C. R. Mayra, "Presentation, expectations, and experience: Sources of student perceptions of automated writing evaluation," *Computers in Human Behavior*, vol. 70, pp. 207–221, 2017.
- [16] M. Warshcauer and D. Grimes, "Automated writing assessment in the classroom," *Pedagogies: An International Journal*, no. 3, pp. 32–36, 2008.

Copyright © 2023 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).