English Language Teaching and Linguistics Studies
ISSN 2640-9836 (Print) ISSN 2640-9844 (Online)
Vol. 5, No. 5, 2023
www.scholink.org/ojs/index.php/eltls

*Original Paper*

# A Study on Teacher Feedback and AES Feedback in Chinese College students' English Writings

Tong Guo[1]

[1] Dalian Maritime University, Dalian, China

Received: September 29, 2023 Accepted: November 22, 202 Online Published: December 05, 2023

doi:10.22158/eltls.v5n5p81 URL: http://dx.doi.org/10.22158/eltls.v5n5p81

***Abstract***

*Feedback holds significant importance in second language writing instruction, as evidenced by numerous studies indicating its direct or indirect impact on the quality of learners' written work. Despite this recognition, there has been a dearth of comprehensive research concerning the distinct influences of teacher feedback and Automated Essay Scoring (AES) feedback on the writing quality of learners. Addressing this gap, the present study employs a mixed-method approach, integrating both qualitative and quantitative methodologies. By meticulously examining AES feedback, teacher feedback, writing revision logs, and conducting interviews, this investigation identifies noteworthy differentials in the multidimensional aspects of writing quality attributed to these two distinct feedback modalities. Primarily, within the realm of syntactic complexity, lexical richness, fluency, and accuracy, the cohort exposed to teacher feedback demonstrated notably superior performance relative to their AES feedback counterparts. Secondarily, an assessment of revised text quality revealed compelling insights. The ultimate version of the text, stemming from the AES feedback group's iterative revisions, exhibited marked enhancements in terms of accuracy, total word count, and average word length. In contrast, the initial and final drafts of the teacher feedback group unveiled discernible disparities in vocabulary intricacy, accuracy, total word count, and average word length. Evidently, while the ultimate version did not witness a significant surge in average word length or total word count, it showcased heightened vocabulary sophistication and enhanced accuracy in relation to the initial draft. This study underscores the value of judiciously deploying these two categories of feedback within the landscape of writing instruction. The nuanced benefits of each feedback type can be strategically harnessed to suit distinct writing contexts, thereby augmenting the caliber of learners' written compositions.*

***Keywords***

*Teacher Feedback, AES Feedback, Text Quality, mixed-method approach*

## 1. Introduction

As an indispensable practice in writing instruction to sustain L2 learners' writing performance, feedback is widely utilized by L2 writing teachers to inform students of their writing problems and weaknesses so that students can improve their writing performance in both local and global aspects. Feedback is a crucial aspect of teaching English writing, and the quality of learners' writing depends on the writer's ability to write and the effectiveness of the reader's feedback (Zhou & Li, 2022). Writing feedback research is a hot topic of writing teaching research in the past two decades (Ferris & Roberts, 2001). Traditionally with teacher feedback as the main form, with the development of the times, in recent years with the help of research tools and automatic grading system for students' writing correction and feedback system is also gradually applied to the classroom, also has become another hot spot of research.

In the current literature, many researchers have studied teacher feedback and AES feedback on English teaching writing. Some researchers believe that teacher feedback has occupied a dominant position in foreign language teaching writing. In the current classroom environment where teachers' authority is valued, learners are more willing to accept teacher feedback to revise their compositions, believing that face-to-face communication with teachers will make it easier to understand the feedback content. Other researchers disagree, however, and their findings suggest that AES feedback not only saves time and effort compared to teacher feedback and other types of feedback, but also that the data show that learners who receive AES feedback consistently report higher levels of writing quality. However, some scholars have also questioned the usefulness of AES feedback in the context of real second language writing and whether students truly understand the questions raised by the feedback. However, some scholars have also questioned the usefulness of AES feedback in the context of real second language writing and whether students truly understand the questions raised by the feedback.

Despite the debate over Teacher feedback and AES feedback, however, these two feedback methods have been widely used in the real classroom of second language writing. Due to the current rapid development of computer network technology, AES feedback has also been applied to the second language classroom and learning. In previous studies, AES feedback has been less explored, with some studies only exploring the effectiveness of AES feedback, but fewer studies comparing teacher feedback to AES feedback. Therefore, it is necessary to explore the effectiveness of AES feedback and teacher feedback. Secondly, It is not difficult to find that previous studies mostly compare the nature and categories of teacher feedback and AES feedback, or study on the writing quality of a certain form of feedback, and rarely involve the comparison of different feedback methods on learners' writing performance. In addition, considering that most studies focus on teacher feedback and there are few studies on automatic online feedback based on computers and networks, we report a quasi-experimental study, which concentrated on and tracked what effects of teacher feedback and AES feedback had on Chinese EFL learners' writing performance. In addition, different from the most existing feedback studies that have used a single or limited measures to evaluate students' writing production, our study

employed multiple measures (i.e., accuracy, linguistic complexity, fluency, content and organization quality) due to the possible trade-off among them, which could advance the current knowledge in this area. Pedagogically, this study would enable feedback practitioners to reflect on their current feedback practices and optimize their feedback practices or initiate some pedagogical innovations to help students make progress in different dimensions of L2 writing, which is regarded as a sustainable goal in L2 writing pedagogy.

## 2. Literature Review

### *2.1 Teacher Feedback*

In the process of English writing teaching, feedback refers to the way that readers provide for learners to revise their compositions and improve their English writing ability (Hyland, 2006). Teacher feedback is one of the most important ways of English writing teaching feedback. Many experts and scholars believe that teacher feedback can promote learners to pay attention to the differences between interlanguage and target language, and also promote learners to reorganize language structure (Lee, 2014). In early stage of teacher feedback research, the idea that the teacher feedback can improve students' writing proficiency is not accepted in the beginning. Some scholars even strongly oppose correcting mistakes in students' compositions. They believe that teachers' feedback is vague, arbitrary and ineffective in writing teaching (Truscott, 2004). It is not only failing to enhance the accuracy of students' writing language, but also not conducive to promoting students' writing fluency and their quality of writing. Nevertheless, many scholars hold the opposite opinion that the teacher feedback is beneficial to help learners improve their writing proficiency. Chandler (2003) find that students who received direct feedback do well in revising the original text than those who received indirect feedback, when writing a new text, the latter have much higher accuracy in their language. In addition, In the past decade, many scholars exclude native English speakers and took Arabic, Chinese and Indonesian students as subjects to explore the effect of teacher correction in their second language writing (Ahemd, 2015; Sibo, 2016; Yustina, 2016). Most of the findings demonstrate positive attitudes on the effects of the teacher feedback. For example, written feedback can help students improve the effect of using prepositions. For Chinese students, besides content and grammar modification, they are more likely to get extended evaluation from teachers (Satoko & Kyoko, 2017). On the whole, these studies demonstrate the effectiveness of teacher feedback on learners' second language writing performance.

### *2.2 AES Feedback*

Based on Constructivism, Formative Assessment Theory and The Process Approach, the online Automatic English Writing Evaluation System evaluates and scores compositions by using computer technology and artificial intelligence (Wang, 2014). In recent years, the English writing teaching in China tends to use the online automatic correction system, which can help teachers save the time of correction, so that teachers can devote themselves to teaching more efficiently.

In the empirical research, many researchers have given positive comments on the Automatic Writing

Evaluation system. The results reveal that this system can cultivate students' independent writing ability and stimulate students' interest in writing. Zhou's study showed that AES feedback has a significant effect on the improvement of learners' syntactic ability in writing (Zhou, 2015). The addition of AES feedback and automatic scoring system feedback to the writing feedback process, as well as the matching of a massive information base as a source of writing information, can significantly enhance the learners' interest in writing, improve the quality of the writing text, and realize the effective combination and mutual complementation between various feedback methods (Zhou & Li, 2022). But some studies have also shown that online evaluation systems do not have a significant impact on learners' writing quality, and that multiple feedback methods combined with manual feedback are needed to significantly improve the quality of learners' written texts (Zhang, 2019).

*2.3 Sociocultural Theory*

Inspired by Marxist dialectics, Vygotsky establishes the Sociocultural theory, which includes the theories of Zone of Proximal Development (ZPD), Scaffolding, Internalization and Mediation (Vygotsky, 1978). In traditional English teaching classes, the writing procedure is always arranged as this order. Teachers publish writing assignments firstly, then students begin to write. After that, teachers revise and mark drafts. However, sometimes students couldn't modify their writings better after receiving feedback. The effect of corrective feedback doesn't receive teachers' timely concern. The Sociocultural theory could change the deadlock of traditional writing feedback which is flowing in one direction to realize a real development of students' writing proficiency. A multiple feedback mechanism which is structured on the basis of the Sociocultural theory could provide students a preferable way to interact with peers, intelligent English writing evaluation systems and teachers (Xi, 2020). Teachers' guidance and scaffolding can also function effectively. Due to its emphasis on a dynamic interaction in language learning, while writing teaching is a relatively static process, the application of the Sociocultural theory in English writing teaching, especially on writing feedback is not always the case. Some scholars make research from the perspective of EFL writing curriculum transformation and second language writing learning strategies to realize the interaction among learners' language, thinking and society and help them achieve a collaborative development of both language and cognition (Zhang & Sun, 2014; Chen, 2015). In recent years, investigations started to focus on the feedback. For example: the way of feedback on college English writing is reconstructed and feedback system on college English writing is built up. In this thesis, the teacher and automatic evaluation system provide students a series of different feedback according to their ZPD to facilitate their internalization of knowledge through scaffolding and increase their writing ability.

*2.4 Teacher Feedback or AES Feedback, which one is more Effective in L2 Writing?*

A number of scholars have also conducted controlled trials of teacher feedback versus systematic feedback, with varying conclusions (Li, 2019). Some studies have found that teacher feedback improves the quality of learners' writing, whereas systematic feedback has not done so. Zhou (2013) argued that compared to other feedback methods, teacher feedback can improve students' discourse and

generalization skills, while AES feedback can arouse interest in writing and improve the skill of using words.

All of the above studies have dealt with teacher feedback and AES feedback, but the main focus has been on the nature and type of teacher feedback and AES feedback. In addition, most of the studies have focused on teacher feedback, and research on the teacher feedback paid more attention to exploring different forms of feedback and the effectiveness of feedback, but less attention has been paid to web-mediated AES feedback (Stevenson, 2004), and most of the studies did not address the comparison of the two forms of feedback on learners' writing performance. Therefore, this study attempted to conduct a writing instruction experiment to explore the specific effects of teacher feedback and AES feedback on the quality of non-English major learners' writing through a combination of qualitative and quantitative research methods. The research mainly explores the following questions:

(1) What are the effects of the teacher's and AES feedback on the language quality of students' compositions?

(2) How well do students accept the teacher 's and AES feedback?

## 3. Methodology

### *3.1 Context and Participants*

In this study, the first-year undergraduate students at a major university in Northeast China will be the subjects of the study, which consists of two classes with a total of 60 students. The subjects have similar grades in English at enrollment, have studied English for basically similar lengths of time, and none of them have ever traveled to an English-speaking country. Two classes were selected for the study to receive teacher feedback and AES feedback. Teacher feedback is the online feedback given by the instructor of the participant's class in response to the classroom writing texts submitted online by the students; AES feedback refers to the scores on the essays and the revisions made to each sentence, as well as the suggestions, that are instantly generated by using the online correction software (Criticize.com) by calculating the discrepancies between the students' compositions and the standardized corpus. Each with 30 students as teacher feedback group and AES feedback group.

### *3.2 Data Collection*

The work of data collection mainly consists of the following two steps: one is to collect the data of the writing experiment; the other is to collect the data of the interview.

**Table 3.1 Procedures of Data Collection**

| WEEK | INTERVIEW QUESTIONS |
|---|---|
| Week1 | Writing the first draft |
| Week2 | According to the feedback, revise |

| | |
|---|---|
| Week3 | Rewriting |
| Week4 | According to the feedback, revise again |
| Week5 | Writing the final draft |

3.2.1 Writing Tests

An immediate writing in the usual classroom will be chosen as the research material for this study. The writing will be IELTS essay task2, because Argumentative writing is recognized as a reliable and popular instrument to evaluate L2 learners' writing proficiency in academic contexts. As such, Chinese tertiary EFL learners tend to be required to complete argumentative writing in various well-established English proficiency examinations, including IELTS, TOEFL, College English Test Band 4/6 (CET-4/6), and Tests for English Majors Band 4/8 (TEM-4/8).

The two groups were asked to complete a composition talking about the opinion of the phenomenon of the use of phones by secondary school students. This study chooses this topic for the following two reasons. First of all, in EFL classrooms in Chinese universities, teachers' teaching practice is always based on test-oriented, so students tend to write arguments based on the topic is written on. More importantly, the two groups of students are faced with going abroad in the second and third years of college, because IELTS writing is a large-scale standardized test in mainland China, its writing questions have high reliability and validity. Specifically, they draw on general education and students' daily life experiences and are therefore considered familiar and fair to each student, which guarantees that the difficulty of such topics is largely consistent. All participants were required to complete each writing task within 40 minutes, and the writing length was expected to be no less than 200 words, in accordance with the requirements of IELTS. During the test period, they are not allowed access to any external resources, such as dictionaries and textbooks.

3.2.2 Revise and Feedback

The experiment lasted for five weeks, during which all students were given the same writing assignment. After students submit their first draft to teachers and an online feedback system, they receive revision suggestions and can voluntarily choose whether to revise their writing. The assignment, submission, feedback, and revision of the two groups' essays were all conducted online. For the AES Feedback group, once students submit their essays to the grading system, the system instantly gives them a score and suggestions for changes to the entire text. For the teacher feedback group, teachers used the correction function of Critique.com to correct the essays submitted by students online, including word spelling, grammar usage, sentence errors, and chapter coherence, and to write comments. After that, the two groups of students made revisions based on their respective critiques and completed the revisions with feedback and submitted their final drafts. It should be noted that all drafts of this experiment should be submitted. The researchers graded the draft with reference to the IELTS essay marking criteria.

3.2.3 Semi-Structured Interview

Two semi-structured interviews will be conducted for this study. First interview Before the formal experiment, this study will select two students from each class and replace them with S1, S2, S3 and S4 respectively. 4 students (2 students in each class) with high and low scores were selected as interview subjects with reference to their admission scores. After submitting the final draft in the fifth week, the researchers selected four students (numbered S5, S6, S7, and S8) from the two classes with high and low scores for a second semi-structured interview. The interview content was approved by the students and transcribed after the end. The schedule and content of the first interview are as follows. The first interview mainly analyzes students' expectation level of writing ability from a personal perspective. The topic of the writing and the feedback preferences they expect to receive are also interviewed. The outline of the first interview is shown in Table 3.1:

**Table 3.2 The Outline of the First Semi-Structured Interview**

| PURPOSE | INTERVIEW QUESTIONS |
|---|---|
| Expected writing level | What level of English writing do you want to achieve? |
| Feedback preference | Do you want to receive feedback on your composition?<br>What kind of feedback do you want to receive? |

In the second interview, four students were selected based on their scores in the final draft. Survey their emotional attitude towards the feedback they receive and their evaluation of the strengths and weaknesses of the feedback. In addition, they were asked to rate whether their personal writing skills had improved. The second semi-structured interview is shown in the Table below.

**Table 3.3 The Outline of the Second Semi-Structured Interview**

| INVESTIGATION DIRECTION | INTERVIEW QUESTIONS |
|---|---|
| Attitude towards feedback | After receiving feedback, what do you think of the teacher or AES feedback? |
| Evaluation about feedback (Advantages and disadvantages) | Do you accept scores and comments of the evaluation? |
| Improvement of writing proficiency | Do you think the quality of your writing has improved since the revision? |

In the first interview, through the interview of students' expectation level and preference for feedback, we find out which of these two feedback methods is more suitable for students, and help students understand and sort out the content of feedback. This process promotes the internalization of students' external knowledge of English writing through scaffolding and consolidate students' writing foundation. In the second interview, the researchers collected the subjects' acceptance and opinions of the two feedback methods after receiving feedback.

*3.3 Data Analysis*

At the end of the experimental phase, all experimental data were collected, including transcripts of writing assignments and semi-structured interviews. Wolfe et al. (1998) divided language indicators into four categories: lexical complexity, syntactic complexity, accuracy and fluency, and Kuiken et al. (2008) finely divided the measures of complexity, fluency and accuracy according to different indicators. This study will use L2SCA as well as L2LCA to digitize the text of the collected compositions and counted the total number of words (W), the total number of T units (T), the total number of clauses (C), the writing time (M), the total number of error-free T units (EFT) and the number of vocabulary types in each piece of writing and enters the data into Microsoft Word. Syntactic complexity (MLT, C/T, MLC) and lexical complexity (MSTTR-50) were calculated (Révész & Kourtali, 2016; Ong & Zhang, 2010). In addition, accuracy is measured by calculating the error-free T-unit rate (EFT/T) and Error/100. Fluency is measured by calculating the M/S, M/T, M/W. Secondly, the number of errors in the essays was counted by the researcher, and the number of errors was firstly done by a combination of correction software Grammarly and manual correction, and the final error number of each composition was based on the agreement. Finally, these data were entered into SPSS 26.0. And analyzed them using an independent-sample T-test and Paired-Samples T-Test for descriptive statistical analysis of learners' writing accuracy, fluency, and complexity. Traditional research on second language writing generally measures output in terms of manually assessed composition scores (Abrams & Byrd, 2016). Drawing on related studies, this study intends to examine the impact of feedback approaches on learners' language quality using holistic indicators, which are often measured in terms of fluency, syntactic complexity, lexical complexity, and accuracy (Johnson, 2017; Kuiken & Vedder, 2007; 2008; Zhang & Jiang, 2020).

The complexity, accuracy and fluency measures employed in this research are shown in Table 3.4. Cho (2015) proved that these measures are valid by a number of researchers

**Table 3.4 Measurement Indicators Status Table**

| MEASURES | COMPUTATION |
|---|---|
| **Syntactic Complexity** | |
| MLT | Total number of words divided by T-units |
| C/T | Number of clauses divided by number of T-units |
| MLC | Number of words divided by number of clauses |
| **Lexical Complexity** | |
| MSTTR | Average TTR of all successive segments of 50 words |
| **Accuracy** | |
| EFT/T | Total number of error-free T-units divided by total number of T-units |
| Errors/100 | Total number of errors divided by total number of words multiplied by 100 |
| **Fluency** | |
| W/S | Number of words divided by number of sentences |
| W/T | Number of words divided by total number of T-units |
| W/M | Number of words divided by writing time |

## 4. Result and Discussion

### *4.1 Comparative Text Analysis on Students' Writing Performance after Receiving Feedback from AES and the Teacher*

In this chapter, descriptive statistical analysis is conducted by comparing the influence of the two types of feedback on learners' writing performance and the quality of learners' writing before and after receiving feedback.

4.1.1 Effects of Different Feedbacks on Writing Complexity

This study will use the final drafts of the two groups in an independent sample t-test analysis to explore the effect of different feedback methods on the complexity of learners' writing outputs. The teacher feedback group will be represented by 1 group and the AES feedback group will be represented by group 2. Since complexity consists of two main aspects, lexical and syntactic complexity, this study will examine these two aspects. In this study, MLT, MLC, and C/T were used as measures of syntactic complexity, and MSTTR was used to measure lexical complexity, with larger values of this value representing higher complexity of writing. The results of the analysis are shown in the table below.

**Table 4.1 Independent Sample T-Test Results of Writing Output Complexity Between Different Groups**

| Measure Type | Measure | Group 1 | | Group 2 | | t | df | p |
|---|---|---|---|---|---|---|---|---|
| | | Mean | SD | Mean | SD | | | |
| Syntactic complexity | MLT | 11.29 | 2.49 | 12.70 | 2.63 | -1.27 | 27.00 | 0.152 |
| | MLC | 7.06 | 0.24 | 7.23 | 0.20 | -0.54 | 27.00 | 0.597 |
| | C/T | 1.24 | 0.16 | 0.35 | 0.23 | 10.80 | 27.00 | 0.000*** |
| Lexical complexity | MSTTR | 0.64 | 0.02 | 0.49 | 0.03 | -17.73 | 38.00 | 0.020*** |

*$p$<0.05.

*Notes.* MLT=dividing the total number of words by the total number of T-units. MLC=dividing the total number of words by the number of clauses. C/T=dividing the total number of clauses by the number of T-units. MSTTR= average type-token ratio of all successive segments of 50 words.

As can be seen from the above figure, it is first analyzed in terms of the case of syntactic complexity. In the aspect of MLT, the mean value of group 1 was 11.29 (SD=2.49), the mean value of group 2 was 12.70 (SD=2.63). From the two different feedback ways, there was no significant difference on MLT ($p$ = 0.152 > 0.05). The mean value of group 1 on MLC was 7.06 (SD = 0.24), and the mean value of group 2 was 7.23 (SD = 0.20), It can be seen that there was no significant difference on MLC ($p$ = 0.597 > 0.05). In C/T, the mean value of group 1 was 1.24 (SD = 0.16), while the mean value of the group 2 was 0.35 (SD = 0.23). From the two different writing tasks, there was a significant difference on C/T ($p$ = 0.000 > 0.05). And the higher C/T ratio represents the higher writing complexity, so it can be shown that the complexity of writing outputs is higher in group 1 than in group 2.

In terms of lexical complexity, MSTTR was chosen as the measure. For lexical complexity, the mean value of MSTTR with group 1 was 0.64 (SD = 0.02), while the mean value in task with group three was 0.49 (SD = 0.03). there was no significance difference in statistics between the two writing compositions ($p$ = 0.020 < 0.05). And since the mean value of group 1 is higher than that of group 2, it can be inferred that the lexical complexity used by group 1 is slightly higher than that of group 2 at the time of writing.

4.1.2 Effects of Different Feedbacks on Writing Fluency

In order to investigate the effects of the two types of feedback on writing fluency, this study used three measures of W/M, W/T and W/S in order to compare the effects of the two different types of feedback on the learners' writing fluency, where higher ratios of these three means that the learners are more fluent in writing. The following table shows the performance of the two groups of subjects in terms of W/M, W/S and W/T.

Table 4.2 Independent Sample T-Test Results of Writing Output Fluency Between Different Groups

| Measure | Group 1 | | Group 2 | | t | df | p |
|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | | | |
| W/T | 12.27 | 2.49 | 13.40 | 3.00 | -1.19 | 39.00 | 0.213 |
| W/M | 6.71 | 1.33 | 5.80 | 1.12 | 2.09 | 39.00 | 0.043* |
| W/S | 13.83 | 3.27 | 13.90 | 3.50 | -0.91 | 39.00 | 0.339 |

*P*<0.05.

*Notes.* W/T = dividing total number of words by number of T-units. W/M = dividing total number of words by writing time. W/S = dividing total number of words by number of sentences.

As far as the first measure of fluency is concerned, the mean value obtained by group 1 was 12.27 (SD= 2.49), and the mean score obtained by group 2 was 13.40, (SD=3.00), and according to the results of the independent t-test, it was shown that with regard to the W/T measure, the fluency of the two groups of subjects was not significantly affected by the different types of feedback (t=-1.19, p=0.213>0.05). Secondly, for the second measure of fluency, W/M, the mean value by group 1 was 6.71 (SD= 1.33), and the mean value obtained by group 2 was 5.80 (SD=1.12), and the data showed that there was a significant difference in scores on W/M between group 1 and group 2 (t=2.09, p=0.043<0.05). By the last measure of fluency, W/S, the mean value of group 1 is 13.83 with SD is 3.27 and the mean value of group 2 is 13.9 with SD is 3.50. Since t=0.339>0.05, the data shows that in terms of W/S, group 1 does not show a significant difference in fluency with group 2. It can be concluded that the ratio of W/M is higher in group 1 compared to group 2 and so the writing is more fluent.

*4.2 Effects of Different Feedbacks on Writing Accuracy*

In order to avoid the one-sidedness of a single measure, the two most commonly used accuracy measures were used in this study: the ratio of error-free t-units to the total number of t-units (EFT/T), with a larger ratio representing higher accuracy; and the number of errors per 100 words (Errors/100), with a larger ratio representing lower accuracy. In this paper, the results of independent t-test were obtained with the help of SPSS 26.0 software. The following table gives the performance of writing output accuracy between different groups.

**Table 4.3 Independent Sample T-Test Results of Writing Output Accuracy Between Different Groups**

| Measure Type | Measure | Group 1 | | Group 2 | | t | df | p |
|---|---|---|---|---|---|---|---|---|
| | | Mean | SD | Mean | SD | | | |
| Accuracy | EFT/T | 0.31 | 0.13 | 0.25 | 0.08 | 2.67 | 39.00 | 0.011** |
| | Errors/100 | 6.70 | 2.52 | 10.01 | 3.23 | -3.42 | 39.00 | 0.001*** |

$p < 0.05$

*Notes.* EFT/T = dividing total number of error-free T-units; Errors/100 = total number of errors divided by total number of words multiplied by 100.

From the above table, it can be seen that there are some differences between the different feedback methods on the learners' writing performance in terms of accuracy. First of all, as far as the first measure of accuracy, EFT/T, is concerned, learners had a mean value of 0.31 (with SD=0.13) in the writing experiment with teacher feedback and a mean value of 0.25 (SD = 0.08) in the AES feedback group, so that the mean value of the teacher feedback group was higher than that of the AES feedback group, and thus it can be inferred that the teacher feedback group had a better accuracy than the AES feedback group. In addition, the table above also shows that there is a significant difference in the aspect of EFT/T between the two groups (t=2.67, p=0.011<0.05), which suggests that in terms of the ratio of this value, the difference between the teacher feedback group and the AES feedback group in terms of accuracy is extremely significant. Secondly, for the second measure of accuracy, Errors/100, the subjects' value in the teacher feedback group had a mean of 6.70 (SD=2.52), while the other group's scores had a mean of 10.01 with SD is 3.23. As stated in the introduction, a higher ratio of Errors/100 represents a lower level of accuracy, which suggests that group 1 was more accurate than group 2. Also, there is a significant difference between the scores of the two groups in this measure (p=0.001<0.05). In conclusion, there is a significant difference between the two groups in terms of accuracy, and the group that has been corrected by the teacher is significantly more accurate than the group with AES feedback.

*4.3 Comparative Text Analysis on Students' Writing Performance before and after Receiving Feedback from AES and the Teacher*

In this section, the text before and after receiving feedback is compared between the two groups of subjects, so as to find out the influence of feedback on learners' writing quality.

4.3.1 Comparison of Text Quality Between the First and Final Drafts of AES Feedback Group

This study used a paired t-test to compare the text quality of the first and final drafts of the AES feedback and teacher feedback groups, respectively, to examine whether there were significant differences before and after revisions. If there were differences, on which quality elements were the differences. In this subsection of the study, language quality involves three aspects of language complexity, fluency, and accuracy, and in order to show the feedback differences more clearly, three items such as total word count, average word length, and verb word length are added to measure in this subsection. The specific table is shown below.

**Table 4.4 Results of the First and Final Drafts of the AES Feedback Group**

| Measure | First Draft | | Final Draft | | F | p |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | | |
| W/M | 3.90 | 0.83 | 4.38 | 0.82 | 3.041 | 0.046 |
| W/T | 14.42 | 3.00 | 13.29 | 2.49 | 1.036 | 0.362 |
| W/S | 15.93 | 3.62 | 14.85 | 3.39 | 0.441 | 0.646 |
| MLT | 14.42 | 3.00 | 13.29 | 2.49 | 1.036 | 0.362 |
| MLC | 9.53 | 1.84 | 9.08 | 1.25 | 0.341 | 0.713 |
| C/T | 1.53 | 0.26 | 1.46 | 0.18 | 0.931 | 0.400 |
| MSTTR | 0.77 | 0.03 | 0.75 | 0.03 | 1.598 | 0.212 |
| EFT/T | 0.20 | 0.08 | 0.28 | 0.10 | 5.040 | 0.010 |
| Errors/100 | 7.29 | 2.36 | 7.21 | 2.34 | 12.802 | 0.000 |
| Total words | 208.7 | 58.73 | 233.29 | 52.48 | -24.100 | 0.006 |
| Average word length | 4.48 | 0.30 | 4.62 | 0.28 | -0.030 | 0.007 |
| Verb length | 4.33 | 0.25 | 3.76 | 0.31 | -0.027 | 0.063 |

$p<0.01$

*Note*. W/M = total number of words divided by writing time. W/T = the number of words divided by number of T-units. W/S = the total number of words divided by number of sentences.

MLT = dividing the number of words by the number of T-units. MLC = dividing the number of T-units. C/T = dividing the number of clauses by the number of T-units.

MSTTR = average TTR of all successive segments of 50 words.

EFT/T = total number of error-free T-units divided by T- units. Errors/100 = total number of errors divided by total number of words multiplied by 100.

First of all, from the data in the above figure, it can be seen that there is a significant difference between the first and final drafts of the AES feedback group in terms of W/M ($p=0.046<0.05$). And the mean value of the first draft is significantly lower than that of the final draft, this data indicates that the total number of words in the revised article is more than that of the article before the revision, and W/M, as an important measure of fluency, also indicates that the output of the final draft is more fluent than that of the first draft. There was a significant difference between the two groups of compositions in terms of EFT/T ($p=0.010<0.05$). A larger value of this ratio represents a higher level of linguistic accuracy; thus, it can be inferred that the accuracy of the first draft is lower than that of the final draft. This suggests that revision based on AES feedback is an improvement in composition accuracy. Secondly, in the Errors/100 measure, there is also a significant difference between the first draft and the final draft in the system feedback group ($p=0.000<0.05$), the larger the value means the lower the

accuracy, as can be seen from the above figure, the mean value of the final draft is lower than the first draft, which indicates that the quality of the text after systematic modification is more accurate than that of the text written for the first time. In addition, the total number of first drafts and final drafts was also significantly different ($p=0.006<0.05$). The total number of first drafts was significantly lower than the total number of final drafts, which indicates that the length of the final drafts in the system feedback group was significantly increased, and the length of the text is one of the most important indicators to test the fluency of the composition, which suggests that the modifications based on the system feedback led to an increase in the fluency of the composition. There was also a significant difference in the learners' mean word length between the first and final drafts ($p=0.007<0.05$), and it was found during the interviews that the learners tended to use more complex vocabulary and long words instead of the common words when revising, believing that it would make the essay look more complex and easier to score high marks, thus leading to a significant increase in the use of long words in the final drafts, and an increase in mean word length, and word length is also an indicator of an indicator of lexical complexity, so this indicates a significant increase in lexical complexity in the final draft.

4.3.2 Comparison of Text Quality Between the First and Final Drafts of Teacher Feedback Group

In addition, this study also compared the text quality of the initial and final drafts of the teacher feedback groups, and the survey data showed the following.

**Table 4.5 Results of the First and Final Drafts of the Teacher Feedback Group**

| Measure | First Draft | | Final Draft | | F | p |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | | |
| W/M | 4.24 | 0.71 | 4.54 | 0.96 | 3.039 | 0.056 |
| W/T | 12.39 | 2.37 | 14.70 | 2.63 | 1.307 | 0.364 |
| W/S | 13.58 | 3.28 | 15.66 | 3.15 | 0.367 | 0.548 |
| MLT | 12.38 | 2.43 | 14.70 | 2.63 | 1.026 | 0.433 |
| MLC | 9.10 | 1.24 | 9.35 | 1.40 | 0.432 | 0.013 |
| C/T | 1.64 | 0.27 | 1.95 | 0.40 | 0.945 | 0.248 |
| MSTTR | 0.57 | 0.02 | 0.67 | 0.03 | 1.602 | 0.022 |
| EFT/T | 0.29 | 0.13 | 0.27 | 0.10 | 5.042 | 0.019 |
| Errors/100 | 7.12 | 2.43 | 7.21 | 2.34 | 12.708 | 0.110 |
| Total words | 233.4 | 57.13 | 232.5 | 54.99 | 2.900 | 0.003 |
| Average word length | 5.53 | 0.32 | 4.52 | 0.30 | -0.031 | 0.009 |
| Verb length | 3.62 | 0.26 | 4.01 | 0.29 | -0.028 | 0.100 |

$p<0.01$.

*Note.* W/M = total number of words divided by writing time. W/T = the number of words divided by

number of T-units. W/S = the total number of words divided by number of sentences.
MLT = dividing the number of words by the number of T-units. MLC = dividing the number of T-units.
C/T = dividing the number of clauses by the number of T-units.
MSTTR = average TTR of all successive segments of 50 words.
EFT/T = total number of error-free T-units divided by T- units. Errors/100 = total number of errors divided by total number of words multiplied by 100.

According to the above figure, there is a significant difference between the first draft and the final draft of the teacher feedback group in terms of MLC ($p=0.013<0.05$). From the above figure, it can be seen that the mean value of the first draft is 9.10 (SD= 1.24) and the mean value of the final draft is 9.35 (SD=1.40), and the score of the first draft is smaller than that of the final draft, which proves that the level of syntactic complexity of the final draft is slightly higher than that of the first draft. Based on the interviews, it can be inferred that students will tend to use more complex clauses to improve the quality of their writing texts after being corrected by the teacher. Secondly, there is also a significant difference in lexical complexity ($p=0.022<0.05$), with a mean value of 0.57 (SD=0.02) for the first draft and 0.67 (SD=0.03) for the final draft, which proves that the revised final draft is a little bit more complex than the first draft. Thirdly, there was also a significant difference between learners' first and final drafts in terms of EFT/T ($p=0.019<0.05$), with larger values representing higher text accuracy, but the data showed that the mean value of the unaltered first drafts was lower than the mean value of the final drafts, which may be due to the fact that learners who had received the feedback utilized more complex clauses and words, but the learners did not have a good grasp of complex clauses and word This may be due to the fact that after receiving feedback, learners used more complex clauses and words, but learners did not master the use of complex clauses and words well, which contributed to the lower quality of the text. In terms of the total number of words, the mean of the first draft was 233.4 (SD=57.13) and the mean value of the final draft was 232.5 (SD=54.99), which means that the total number of words in the first draft was more than the number of words in the final draft and there was a significant difference between the two ($p=0.003<0.05$). The final draft length in the teacher feedback group did not increase but decreased, which may be due to the fact that the revisions in this group were concentrated on words and below, and when repeated revisions were unsuccessful, students tended to use avoidance strategies to remove unsuccessful revisions of expressions in order to improve their scores. Finally, average word length is also an important indicator of lexical complexity, suggesting that teacher feedback-induced revisions can also be helpful in improving lexical complexity.

## 5. Discussion

### *5.1 The Comparative Effects of AES and Teacher Feedback on Written Texts*

Firstly, this study compared the text quality of the final draft after two groups of subjects received different types of feedback. According to data analysis, Group 1 (receiving Teacher feedback) is

superior to Group 2 (receiving AES feedback) in terms of lexical complexity, syntactic complexity, fluency and accuracy. The reason may be that in the first interview, it was mentioned which form of feedback learners prefer. Among them, three students chose teacher feedback as their first choice. Secondly, in the second interview, four subjects showed a positive attitude towards the teacher's feedback and expected the teacher's feedback content. However, two participants did not agree with the content of AES feedback, and thought that AES feedback was not as good as teachers in terms of the overall arrangement and logic of the article. Secondly, in terms of revised content, although AES feedback can help students improve their writing skills in terms of grammar and paragraph arrangement, AES feedback is inferior to teacher feedback in terms of more comprehensive correction. The reason may be that the accumulation and use of vocabulary and content writing is a relatively long-term process, which mainly depends on the autonomy of students, and the teacher's feedback on vocabulary, grammar, and the overall organization of the text can improve the quality of writing. For example, AES feedback in the process of revising the text can give learners more suggestions on the expression of synonyms, but relatively simple, while teacher feedback can give more appropriate synonyms according to the context.

*5.2 The Effects of AES and Teacher Feedback on Written Texts before and after Revision*

The second conclusion of this study was to compare the first and final drafts of the AES feedback group and the teacher feedback group by using the paired sample T-test. The results of the study found that the accuracy, total word length, and average word length of the text written after AES feedback were better than the first draft, possibly because AES feedback marked the misuse of words in the original text for revision, and helped students improve the expression of rich words and the use of complex words. Second, after teacher feedback, the vocabulary complexity and accuracy of the final draft improved, but the total number of words and the average word length decreased instead of rising. This may be due to the fact that teachers will pay attention to the use of complex vocabulary by students and give reasonable suggestions for revision. In addition, teachers will also pay attention to the whole paragraphs and content cohesion of the full text, and delete lengthy and repetitive expressions. As a result, the total number of words and the average word length of the revised text will decrease.

*5.3 Students' Perceptions to Feedback from AES and the Teacher*

Interviews and data show that learners' approval of AES feedback is not as high as that of teacher feedback. By comparing the final drafts of the two groups of subjects, it was found that the fluency, accuracy and complexity of the teacher feedback group were higher than that of the AES feedback group. Moreover, it was found through interviews that learners' acceptance of AES feedback was lower than that of the teacher feedback. This may be because in the current teacher-led classroom environment, students are more inclined to communicate with the teacher face to face after receiving feedback. And compared with AES feedback, teacher feedback gives clearer opinions on revision, and AES feedback is more inclined to modify the sentence. However, learners also hold a positive attitude towards AES feedback, believing that AES is of great help to learners' writing accuracy.

## 6. Conclusion

Teacher feedback is one of the most common forms of feedback in second language classrooms today, and AES feedback is the preferred form of feedback in second language feedback because of its time-saving and labor-saving benefits. Current research on feedback is uncommon regarding the comparison of these two types of feedback. In addition, there is less research on AES feedback on writing performance, and neither takes into account the different writing abilities of second language learners. Since L2 learners have different writing abilities, it is also necessary to explore the two types of feedback on the writing performance of L2 learners with different learning abilities. By comparing the online revision results of the teacher feedback group and the AES feedback group, this study found that the effects of the two feedback methods on text quality were reflected in different dimensions. The main findings include the following. First, the effects of AES feedback and teacher feedback on learners' writing quality were compared, and it was found that the syntactic complexity, lexical complexity, fluency, and accuracy of the teacher feedback group were higher than those of the AES feedback group. Second, in terms of text quality after revision, there were significant differences between the first and final drafts in the AES feedback group in terms of EFT/T, Errors/100, total word count, and average word length, with the revised final drafts showing a significant increase in each of these elements; whereas there were significant differences in the teacher feedback group in terms of MSTTR, EFT/T, total word count, and average word length, which were reflected in the fact that the average word length and the total number of words did not increase, but the final drafts had higher means than the first drafts in terms of lexical complexity as well as accuracy.

The above findings suggest that how to accurately use teacher feedback and AES feedback after writing is a topic worthy of further exploration and practice. First of all, as a result of the analysis, it is clear that the post-feedback writing texts were superior to the pre-feedback writing texts in both groups, both in the AES feedback group and in the teacher feedback group. For the current second language writing of college students who are not English majors, the model of full systematic approval and teacher approval can be used, and the feature of instant feedback of the AES system can be utilized to give feedback on each student's writing. Secondly, data analysis showed that the teacher feedback group performed better overall in writing than the AES feedback group. Therefore, teachers can select representative essays for accurate correction and explanation to improve the efficiency of correcting essays. Finally, the resources provided by the scoring system can be fully utilized to promote writing teaching, such as using the acrostic paragraph detection of essays provided by the scoring system to prevent plagiarism, as well as analyzing the commonality of each essay to save teachers' time for correcting and improve efficiency.

This study also has some shortcomings. It only examined the effects of two types of feedback on the revision of an essay, and in the future, the sample of subjects could be increased, and the number of subjects could be expanded for a larger systematic study.

**Reference**

Abrams, Z. I., & Byrd, D. R. (2016). The Effects of Pre-task Planning on L2 Writing: Mind Mapping and Chronological Sequencing in a 1st-year German Class. *System, 63*(1), 1-12. https://doi.org/10.1016/j.system.2016.08.011

Ahmed, A. S., & Al, A. (2015). The Effect of Written Corrective Feedback on Omani Students' Accuracy in the Use of English Prepositions. *Advances in Language and Literary Studies*, *6*(1), 61-71. https://doi.org/10.7575/aiac.alls.v.6n.1p.61

Bai, Y., & Wang, J. J. (2018). A Study on the Influence of Feedback Mode on the Process and Quality of Writing Revision. *JOURNAL OF NORTHEAST NORMAL UNIVERSITY (Philosophy and Social Sciences)*, (02), 134-139.

Carroll, D. W. (2008). *Psychology of Language. Beijing: Foreign Language Teaching and Research Press*.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, *12*(3), 267-329. https://doi.org/10.1016/S1060-3743(03)00038-9

Chen, A. M. (2015). A study of Second Language Writing Learning Strategies from the Perspective of Sociocultural Theory. *Journal of Liaoning Normal University (Social Sciences Edition)*, *38*(06), 807-812.

Cho, H. (2015). Effects of Task Complexity on English Argumentative Writing. *English Teaching*, *70*(2), 107-131. https://doi.org/10.15858/engtea.70.2.201506.107

Coffin, C., & J. Donohue. (2014). A language as social semiotic-based approach to teaching and learning in high education. *Language Learning*, *38*(5), 768-771. https://doi.org/10.1093/applin/amw054

Diki, S., & Bleyle, S. (2004). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback. *Assessing Writing*, (22). https://doi.org/10.1016/j.asw.2014.03.006

Ferris, D, R., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be. *Journal of Second Language Writing*, (10). https://doi.org/10.1016/S1060-3743(01)00039-X

Hu, X. W. (2015). The Influence of Online Essay Self-revision on College Students' English Writing Results. *Technology Enhanced Foreign Language Education*, (05).

Hyland, k., & Hyland, F. (2006). Feedback on second language students' writing. *Language teaching*, *39*(2), 83-101. https://doi.org/10.1017/S0261444806003399

Johnson, M. D. (2017). Cognitive Task Complexity and L2 Written Syntactic Complexity, Accuracy, Lexical Complexity, and Fluency: A Research Synthesis and Meta-Analysis. *Journal of Second Language Writing, 37*, 13-38. https://doi.org/10.1016/j.jslw.2017.06.001

Krashen, S. D. (1981). *Second Language Acquisition and Second language Learning*. Cambridge: Cambridge University Press.

Kuiken, F., & Vedder, I. (2007). Task Complexity and Measures of Linguistic Performance in L2 Writing. *International Review of Applied Linguistics*, *45*(2), 261-284. https://doi.org/10.1515/iral.2007.012

Kuiken, F., & Vedder, I. (2008). Cognitive Task Complexity and Written Output in Italian and French as a Foreign Language. *Journal of Second Language Writing*, *17*, 48-60. https://doi.org/10.1016/j.jslw.2007.08.003

Lee, I. (2014). Feedback in writing: Issues and challenges. *Assessing Writing*, (19), 1-5. https://doi.org/10.1016/j.asw.2013.11.009

Li, F. G. (2019). A Study on the Influence of Multiple Feedback on English Composition Revision Based on Automatic Evaluation System. *Foreign Language Education*, (04), 72-76.

Li, Z., Link. S., Ma, H., Yang, H., & Hegelheimer, V. (2014). The role of automated writing evaluation holistic scores in the ESL classroom. *System*, (44). https://doi.org/10.1016/j.system.2014.02.007

Liang, M. C. (2012). *Development of Automatic Scoring System for Large-scale English Composition Examination*. Beijing: Higher Education Press.

Liu, B., Wang, Y. K., & Zhang, J. L. (2017). The effect of Task types on Task Preparation and Output in Online English Writing. *Modern Foreign Languages*, (1).

Ong, J., & Zhang, L. J. (2010). Effects of Task Complexity on the Fluency and Lexical Complexity in EFL students' argumentative writing. *Journal of Second Language Writing*, *19*(4), 218-233. https://doi.org/10.1016/j.jslw.2010.10.003

Rahimi, M., & Zhang, L. J. (2018). Effects of Task Complexity and Planning Conditions on L2 Argumentative Writing Production. *Discourse Processes*, *55*, 726-742. https://doi.org/10.1080/0163853X.2017.1336042

Révész, A., Kourtali, N. E. et al. (2016). Effects of Task Complexity on L2 Writing Behaviors and Linguistic Complexity. *Language Learning*, *67*(1), 208-241. https://doi.org/10.1111/lang.12205

Satoko, HABAKARI., & Kyoko, O. I. (2017). Reexamining the Role of Corrective Feedback on Students' Writing in the EFL Context. *KATE Journal*, (27), 15-27.

Sibo, C., Hossein, N., & Qian, L. (2016). EFL learners' perceptions and preferences of written corrective feedback: a case study of university students from Chinese mainland. *Asian-Pacific Journal of Second and Foreign Language Education*, *1*(1), 1-17. https://doi.org/10.1186/s40862-016-0010-y

Stevenson, M., & Phakiti, A. (2004). The impact of e-feedback on the rivisions of L2 writers in an academic writing course. *Computers and Compositions*, (21). https://doi.org/10.1016/j.compcom.2004.02.003

Stevenson, M., Schoomen, R., & de Glopper, K. (n.d.). Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing*, (15). https://doi.org/10.1016/j.jslw.2006.06.002

Truscott, J. (2004). Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of Second Language Writing*, *13*(4), 337-344. https://doi.org/10.1016/j.jslw.2004.05.002

Vygosky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge: Harvard University Press.

Wang, J. J. (2006). General Attitude, Types of Feedback and Types of Error Correction: A Probe into the Written Feedback of College English Teachers. *Foreign Language Teaching Abroad*, (03), 24-30.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.Y. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. Honolulu, HI: University of Hawaii, Second Language Teaching and Curriculum Center.

Xi, J. (2020). Sociocultural Theory and Second Language Acquisition: Theory, Method and Practice. *Foreign Language World*, (02), 90-96.

Yin, R. K. (2011). *Qualitative Research from Start to Finish*. New York: The Guilford Press.

Yustina, P. K. (2016). THE EFFECT OF WRITTEN CORRECTIVE FEEDBACK ON HIGHER EDUCATION STUDENTS' WRITING ACCURACY. *Journal Pendidikan Bahasa dan Sastra*, *16*(2), 121-131. https://doi.org/10.17509/bs_jpbsp.v16i2.4476

Zhang, H. (2019). A Study on the Effect of Human-computer Multiple Feedback on Improving the Quality of English Writing Texts. *Technology Enhanced Foreign Language Education*, (02), 34-39.

Zhang, L., & Sheng, Y. (2015). Case Study of Feedback from Automatic Essay Grading System. *Technology Enhanced Foreign Language Education*, (05).

Zhang, L., & Sun, Y. Z. (2014). Reform Practice of English Major Writing Curriculum based on Sociocultural Theory. *Foreign Language World*, (05), 2-10

Zhou, K.L., & Li, M. R. (2022). A study on EFL Writing Effects of English Majors under Three Feedback Models. *Journal of Honghe University*, (04), 152-154.

Zhou, L. (2015). The Influence of Online English Writing Platform on College Students' Writing Syntactic Ability. *Technology Enhanced Foreign Language Education*, (9).

Zhou, Y. S. (2013). A Contrastive Study of Feedback in College English Writing. *Foreign Language World*, (03), 87-96.