

Impact of Research-Informed Integration of AWE Among Chinese EFL Undergraduates

Shu Huang¹ & Dan Chen¹

¹ School of Foreign Languages, Chengdu University of Information Technology, China

Correspondence: Shu Huang, School of Foreign Languages, Chengdu University of Information Technology, China.

Received: December 10, 2023 Accepted: February 5, 2024 Online Published: February 29, 2024

doi:10.5539/ijel.v14n2p26 URL: <https://doi.org/10.5539/ijel.v14n2p26>

Abstract

Recent years have seen growing attention to the use of automated writing evaluation (AWE) in the L2/EFL writing classrooms. While it is generally agreed that teachers should provide scaffolding to students when introducing AWE to the class, a paucity of research has an explicit focus on investigating methods for optimizing the integration of AWE feedback. To fill the gap, the present study proposed a research-informed integration of AWE based on the literature and explored empirically its impact on students' writing performance and their perceived usefulness of AWE feedback in the context of Chinese EFL undergraduates. Data of the study include writing scores and student responses to a questionnaire. The study found that the student participants made significant improvements in content, organization, and holistic score, but not in accuracy, and they seemed to perceive the automated feedback more positively after the intervention. The findings of the study can contribute to knowledge regarding the integration of AWE feedback and provide insights to teachers who are interested in utilizing AWE in L2 writing classrooms.

Keywords: automated writing evaluation (AWE), writing performance, L2 writing, efficacy of the automated feedback, writing classrooms

1. Introduction

Automated writing evaluation (AWE), also referred to as the computer-generated feedback, is defined as “the ability of computer technology to evaluate and score written prose” (Shermis & Burstein, 2003, p. 8). Theoretically, AWE systems can provide real-time feedback, motivate students to revise their written texts, improve the revision quality and hopefully the writing performance (Xu & Zhang, 2022); from a realistic perspective, AWE systems can reduce the workload of teachers in revising compositions, enabling them to spend more time on course preparation so as to improve the quality of teaching. Consequently, AWE systems have been increasingly applied in English writing classrooms across the world. Studies on the pedagogical use of AWE feedback, however, remain insufficient (Stevenson & Phakiti, 2014; Xu & Zhang, 2022). While it is generally agreed that AWE feedback can promote revision quality, its transfer effect on writing performance remains underexplored (Wilson & Roscoe, 2020; Xu & Zhang, 2021). Meanwhile, despite the consensus that student engagement with AWE feedback can be influenced by the implementation methods (Chen & Cheng, 2008; Li et al., 2015), and that teachers should provide guidance and training before the formal adoption of the AWE system in the writing instruction (Huang & Renandya, 2020; Koltovskaia, 2020; Xu & Zhang, 2022), few studies have involved coaching of the AWE feedback use beyond the training for the AWE software/system, and no study has formally examined the effect of the coached use of AWE feedback.

With this in mind, the present study proposed an integration of AWE based on the literature. Situated in the context of a group of Chinese undergraduate students, it explored its impact on students' writing performance and their perceived usefulness of the feedback. The findings of the study can contribute to knowledge regarding the integration of AWE feedback.

2. Literature Review

2.1 L2 Research on AWE Feedback

One thread of recent studies on AWE feedback has focused on validating the AWE score and the feedback. These studies often compare computer feedback with that given by human raters/teachers. The accuracy rate of AWE

feedback is usually not very high. The precision rate for Criterion, for instance, is merely 57.5% (Dikli & Bleyle, 2014). These studies indicate that AWE feedback cannot replace teacher feedback but rather as a supplement to facilitate L2 writing instruction (Bai & Hu, 2017; Cheung, 2015).

Another primary focus of AWE research is examining AWE's effectiveness in writing classrooms (Barrot, 2023; Liao, 2016; Wilson & Roscoe, 2020; Xu & Zhang, 2022). Overall, studies have agreed that using AWE can motivate learners to revise (Grimes & Warschauer, 2010), leading to higher revision quality (Li et al., 2015). However, only a small number of studies have investigated whether the effect on writing quality could be transferred to newly submitted essays (Barrot, 2023; Wang et al., 2013; Wilson & Roscoe, 2020; Xu & Zhang, 2022), the findings of which have so far remained controversial. Barrot (2023), for instance, found that Grammarly had positively affected students' writing accuracy. On the other hand, Xu and Zhang (2022) found that the higher-proficiency group seemed to make more errors in the final test after one semester of intervention. Wilson and Roscoe (2020) found no difference in writing performance between the pre-test and the post-test. The researchers ascribed the disappointing finding to insufficient practice opportunities and suggested implementing AWE within the curriculum. These studies indicated the necessity for further investigation of the transfer effect of AWE use.

L2 Studies on students' attitudes towards AWE feedback have generally yielded positive results (e.g., Chen & Cheng, 2008; Li et al., 2015; Ranalli, 2018; Xu & Zhang, 2022), although mixed findings have emerged. Some researchers report that higher-proficiency students might not appreciate the value of the AWE feedback as their lower-proficiency classmates (Chen & Cheng, 2008). However, the higher-proficiency learners in Zhang (2020) were found to hold the most positive attitude towards *Pigai* feedback. Zhai and Ma (2021) examined the influencing factors of student acceptance of AWE feedback using a Technology Acceptance Model (TAM). Based on SEM analysis of responses from 448 Chinese students, the study found that students' usage intention of AWE feedback could be influenced by a comprehensive set of factors, including the perceived trust of the AWE reliability, facilitating conditions, AWE self-efficacy, and system features.

2.2 Pedagogical Suggestions over the Integration of AWE Feedback in L2 Writing Classrooms

Despite the wide adoption of AWE systems in writing classrooms (Zhang & Zhang, 2018), very little research has been conducted to examine how best to facilitate students' use of AWE feedback. Chen and Cheng (2008) explored how three teachers' different implementations of *MY Access!* influenced their students' perceived usefulness of the feedback. The study suggests that teacher facilitation is necessary to avoid causing frustration to the students. In a more recent attempt, Zhang and Hyland (2022) examined the effect of an integrated approach involving AWE feedback, peer feedback, and student feedback on student engagement among 33 EFL students over one semester. The researchers found that the approach effectively promoted students' behavioral, affective, and cognitive engagement, but the study was of qualitative design and did not examine the impact of the approach on writing performance.

Although no study has formally outlined a training plan for the AWE feedback use, several pedagogical suggestions have been proposed in the literature (e.g., Bai & Hu, 2017; Jiang & Yu, 2020; Koltovskaia, 2020; Xu & Zhang, 2022; Zhang & Hyland, 2018; Zhang & Hyland, 2020). Depending on their different orientations, the suggestions can be grouped into the student-oriented and the task-oriented categories. The student-oriented suggestions can be further divided into the affective ones (i.e., involving beliefs and perceptions) and the cognitive ones (i.e., related directly to strategies in addressing the feedback). A summary of the training suggestions can be seen in Table 1. So far, few attempts have been made to incorporate the suggestions. It should be noted that some suggestions are found controversial. For example, it is proposed that a threshold score for submission should be set (Chen & Cheng, 2008; Xu & Zhang, 2022), but some researchers found that this operation might have a washback effect (Jiang & Yu, 2020).

Table 1. Summary of the training suggestions of AWE feedback

Category		Suggestions	Supporting Literature
Student-oriented	Affective	Establish positive perceptions that AWE can serve a helpful writing tool in the writing and revision process	Zhang & Hyland, 2018; Zhang, 2020
		Be informed about merits and demerits of AWE feedback	Koltovskaia, 2020; Zhang, 2020
	Cognitive	Self-regulated use of external resources (e.g., dictionary, teacher, peers)	Bai & Hu, 2017; Jiang & Yu, 2020; Koltovskaia, 2020;
		Learn to question and analyze the feedback critically	Koltovskaia, 2020
Task-oriented		Acquire other revision strategies and cognitive strategies to respond to AWE feedback	Zhang & Hyland, 2018; Zhang, 2020
		Used with teacher feedback and/or peer feedback	Xu & Zhang, 2022; Zhang, 2020
		Mandatory multiple-revision policy to optimize student agency	Li et al., 2015; Liao, 2016; Xu & Zhang, 2022
		Assign reflections on the use of automated feedback	Koltovskaia, 2020

The present study, therefore, proposes a classroom-based investigation to address the gaps above. The study would explore the effect of one research-informed implementation of AWE feedback on students' writing performance and their perceived usefulness of the feedback. Specifically, the research questions are:

RQ1. How did the research-informed integration of AWE feedback affect students' writing performance?

RQ2. How did the research-informed integration of AWE feedback affect students' perceived usefulness of AWE feedback?

3. Method

3.1 The AWE System *Pigai*

The AWE system applied in the present study is the largest locally designed web-based AWE platform *Pigai* (<http://www.pigai.org/>). It is a commercial online AWE platform developed explicitly for Chinese EFL learners by Beijing Ciwang Technology Co., Ltd., and is widely adopted among Chinese university students (Zhang & Hyland, 2018). According to its web page, it has been used by more than 1000 schools or universities and has assessed about 1 billion compositions by November 2023. In addition to a holistic score and general comments on the student writing's global issues (i.e., content and organization), the system gives written feedback on grammar, collocation, lexical use, and mechanics sentence by sentence. The system also features in its ability to highlight collocation mistakes in the writing of Chinese students caused by negative transfer of their L1, examples being "learn ...knowledge". Another popular feature of *Pigai* regards the multiple synonyms it provides to facilitate students' vocabulary learning (Yao, 2021). Compared with the more well-known AWE tool *Criterion*, it has a high precision rate for mechanics errors (98.07%) (Bai & Hu, 2017), which significantly surpasses the 50% reported for *Criterion* (Dikli & Bleyle, 2014), and the precision rate regarding grammar errors was 58.71%, only 5% lower than the 63% reported for *Criterion*. While the overall precision rate of *Pigai* was not very high (45.77%), Bai and Hu (2017) found that *Pigai* was reliable in detecting conjunction errors, misuse of articles, verb forms, and subject-verb agreement issues and that it was a valuable tool to supplement the teaching in the Chinese EFL writing classroom.

3.2 Participants and Context

The study was conducted in the context of a 16-week mandatory course called *College English*, targeting non-English-major students in a southwestern university in the Chinese mainland. While the course is not a specific writing course, the teaching of writing is considered an essential component in the curriculum, and one class of 45 minutes is allocated for teaching writing. According to the course outline, students would write 4 compositions in the semester. The study adopted convenience sampling to recruit one intact class of first-year students (N = 52) taught by the first author. The students had used *Pigai* the previous semester but had not received any coaching on its use.

3.3 Instruments

To answer the research questions in the present study, two types of instruments were developed: 1) one writing prompt for the pre-test and one for the post-test; 2) a questionnaire survey on learners' perceived usefulness of AWE feedback.

The two writing prompts were obtained from College English Test-Band 4 (CET-4), a standardized test administered by the Higher Education Department of the Chinese Ministry of Education. For each prompt,

students would be asked to write an essay on the topic in 30 minutes. To ensure the equivalence in difficulty, both prompts demanded the writing of an argumentative essay, the genre often tested in CET-4.

The questionnaire (Appendix A) was developed to probe learners' perceptions of *Pigai*. It consists of six items on a five-point Likert scale based on Xu and Zhang (2022), with 1 indicating Strong Disagreement and 5 Strong Agreement. The Cronbach's alpha score was .837, demonstrating good reliability. Therefore, a composite variable was produced by summing up the six items for further comparison.

3.4 Procedure

The main study took place over 14 weeks. In the first week, the teacher administered the pre-writing test and the questionnaire to the students. Then, the teacher trained the students on the *Pigai* use (See Table 2).

Table 2. AWE training for the Experimental Group in the first week

Stage	Operations	Purpose
1	Teacher holding a whole-class discussion on the merits and demerits of <i>Pigai</i>	To inform students of the merits and demerits of AWE feedback; Establish positive perceptions that AWE can serve a helpful writing tool in the writing and revision process;
2	Teacher modeling how to use AWE feedback to revise a sample essay, and explaining the technical terms	To guide students to question and analyze the feedback critically; to instruct on other revision strategies and cognitive strategies to respond to AWE feedback
3	Teacher holding a whole-class discussion on the role of AWE in the revision	Establish positive perceptions that AWE be a helpful writing tool in the writing and revision process

In the next 12 weeks, the students completed 4 assignments following the research-informed procedure (See Table 3). In the last week, the students completed the post-writing test and the questionnaire.

Table 3. Procedure for each of the 4 writing tasks during the 12 weeks

Step	Experimental Group	Highlights
1	Student submitting the writing via <i>Pigai</i>	
2	Teacher holding whole-class discussion and giving oral feedback on sample essays (content and organization); explaining criteria of good essays	Used with teacher feedback
3	Teacher modelling how to use the AWE feedback to revise using a sample essay of the assignment	Prepare students for the ability to question and analyze the feedback critically; provide other revision strategies and cognitive strategies to respond to AWE feedback
4	Students revising independently and submitting again on <i>Pigai</i> for teacher grading	Ensure self-regulated use of external resources (e.g., dictionary, teacher, peers)
5	Students completing the reflective report reflecting on the use of AWE feedback	Assign reflections on the use of automated feedback

3.5 Data Collection and Data Analysis

Data of the study included students' scores in the writing test before and after the intervention, and their responses to the questionnaire survey.

To answer the first research question regarding the effect of the intervention on students' writing performance, SPSS 26.0 was used to compare students' scores in content, organization, accuracy, and overall. Before conducting the planned statistical analyses, the data were tested for normality of distribution. The first author and the second author individually scored 20% of the essays on content (0–5), organization (0–5), and overall (0–15), and discussed the disagreements. Then, another 20% of the essays were graded, and the inter-rater agreement reached over 90% for all the categories. After that, the first author graded the rest.

Writing accuracy was measured by error-free T-units ratio (EFT/T), namely the number of error-free T-units divided by the number of T-units (Xu & Zhang, 2022), which is a commonly used measure of linguistic accuracy (Polio, 1997). The number of T-units was identified by the web-based L2 Syntactic Complexity Analyzer (Ai & Lu, 2013; Lu, 2010, 2011; Lu & Ai, 2015). The errors were manually flagged without identifying the categories (Xu & Zhang, 2022). The first author and the second author independently circled errors in 10 compositions and found that the agreement rate was 95%. Based on the discussion, the first author flagged the errors in the rest of

the compositions.

As to the second research question, i.e., the effect of the intervention on students' perceived usefulness of the feedback, student responses to the questionnaire were examined using SPSS 26.0.

4. Results and Discussion

The present study explored whether a research-informed integration of AWE feedback could impact students' writing performance and perceived usefulness of the feedback and generated positive findings in general regarding the two research questions.

4.1 Effectiveness of the Intervention on Writing Performance

The Kolmogorov-Smirnov normality tests showed that the p values for writing accuracy (post-test) were greater than 0.05, which suggests normality, but the p values for the rest were below 0.05, indicating that these data were not normally distributed. Therefore, four Wilcoxon signed-rank tests were conducted to compare the scores in the pre-test and the post-test.

As can be seen in Table 4, the four Wilcoxon signed-rank tests indicated that the post-test content scores were significantly higher than those in the pre-test, $z = -1.966$, $p = .049$, with a small effect size ($r = .19$); the organization scores were significantly higher after the intervention, $z = -3.775$, $p = .000$, with a medium effect size ($r = .37$); and significant increase was also noted in the holistic score, $z = -2.992$, $p = .003$, with a small effect size, $r = .29$. However, no significant difference was observed in the accuracy scores before ($Md = .8905$, $n = 52$) and after ($Md = .8910$, $n = 52$) the intervention. The findings showed that the intervention had enhanced the writing scores in content, organization, and overall, but not in accuracy.

Table 4. Wilcoxon signed-rank test results comparing students' writing scores before and after the intervention

	Pre		Post		Z	p
	Md	N	Md	N		
Content	3	52	4	52	-1.966	.049*
Organization	3	52	4	52	-3.775	.000**
Accuracy	.8905	52	.8910	52	-0.697	.486
Holistic score	7.5	52	8.38	52	-2.992	.003**

Note. * $p < 0.05$, ** $p < 0.01$.

The findings have extended previous research by examining the impact of the coached use of AWE on writing performance in terms of both low-level and high-level skills, indicating that careful integration of AWE feedback can enhance the efficacy of AWE use (Zhang & Zhang, 2018). Unlike Wilson and Roscoe (2020), the study found that the intervention significantly impacted the holistic score of students' writing. One possible explanation could be that the coached use of AWE feedback had equipped students with various external resources (Koltovskaia, 2020; Jiang & Yu, 2020), empowering them to employ the feedback (Zhang & Hyland, 2018; Zhang, 2020) and enhancing their perceived usefulness of the feedback, which enabled the students to make better use of the writing and revision opportunities in the writing classrooms. This explanation could also apply to the significant improvement in content and organization, suggesting that decoding sample essays could effectively reinforce writing knowledge.

The findings, however, failed to support the observation that AWE use had a significant transfer effect on students' writing accuracy, as noted in the previous literature (i.e., Barrot, 2023; Liao, 2016; Wang et al., 2013). Several factors may account for this. Above all, it might be related to the unsatisfactory overall precision rate of *Pigai* (45.77%) (Bai & Hu, 2017), which could hamper students' writing accuracy development. While Wang et al. (2013) and Barrot (2023) did not clearly report the precision rate of the AWE tool in their studies, *Criterion* employed in Liao (2016) was reported to have a precision rate of 57.5%, about 12% higher than that of *Pigai* (Dikli & Bleyl, 2014). In addition, a longer span of intervention may be necessary for the progress in accuracy to be observable because writing accuracy can pose a great obstacle for EFL learners (Hartshorn et al., 2010; Polio & Williams, 2009). Third, the use of distinct measures of accuracy may lead to different results (Polio, 1997). Instead of using error-counts (Wang et al., 2013; Liao, 2016), or the percentage of error-free clauses (Tian et al., 2022), the present study employed EFT ratio (Error-free T-units/T-units) to assess accuracy as in Xu and Zhang (2022). As the EFT ratio considers only the number of errors but not the severity of those errors on the overall quality of the writing (Barrot, 2023), it might not fully capture students' development in writing accuracy, which might explain why the higher-proficiency learners in Xu and Zhang (2022) were found to decrease in

accuracy in the post-test.

4.2 Influence on the Perceived Usefulness of AWE Feedback

To answer the second research question, namely how the intervention affected students' perceived usefulness of the AWE feedback, we grouped the composite variable, which is produced by summing up the number for the 6 items in the questionnaire, into 5 categories (i.e., below 10, 10–14, 15–19, 20–24, 25–30), and calculated their percentages via SPSS 26.0.

As can be seen in Table 5, the percentage of students who fell in the last two categories (“below 10” and “between 10 and 14”) was reduced to 0 after the intervention, and there was an increase of 5.8% in the category of “20–24”. Overall, the finding suggested that the intervention had enhanced students' perceived usefulness of the feedback, particularly those who held a highly unfavorable attitude towards the feedback.

Table 5. The percentages of students in different score ranges before and after the intervention

Composite Variable	Pre-intervention Percentage (%)	Post-intervention Percentage (%)
Below 10	0	0
10–14	5.7	0
15–19	32.6	32.7
20–24	48.1	53.9
25–30	13.4	13.4

The finding indicated that teaching practice could influence students' perceptions of automated feedback (e.g., Chen & Cheng, 2008) and confirmed Zhai and Ma (2021) that facilitating conditions play an essential role in shaping students' perceptions of AWE feedback. It provided empirical support to the pedagogical suggestions such as establishing positive perceptions of AWE (Zhang & Hyland, 2018; Zhang, 2020) and raising students' awareness of the merits and demerits of AWE feedback (Koltovskaia, 2020; Zhang, 2020). More importantly, the finding revealed that the coaching of AWE use should be necessary for some students, who might otherwise dismiss the feedback as worthless. These students were likely to need help understanding the automated feedback or have a limited repertoire of revision strategies (Zhang & Hyland, 2018), which prevented them from benefiting from the feedback.

5. Conclusion

This study was a tentative attempt to explore ways of integrating AWE tools in L2 writing classrooms. By proposing an intervention based on the pedagogical suggestions in this regard and then examining its effectiveness on students' writing performance and perceived usefulness of the feedback in the context of Chinese EFL undergraduates, the study provided one possible recipe that may enhance the efficacy of AWE use in L2 writing classrooms. One limitation of the study is the absence of a control group, which prevents us from dismissing the possibility that other factors may have played a role in the notable improvements observed among the student participants.

However, this study has been one of the first attempts to link research and teaching practices concerning the application of AWE in L2 writing classrooms. It responds to the call to facilitate students in AWE use (e.g., Bai & Hu, 2017; Koltovskaia, 2020; Xu & Zhang, 2022) and provides support to the pedagogical suggestions proposed. While the study employed just one AWE tool, it can provide insights for teachers who want to utilize different AWE tools in other contexts (e.g., *Criterion*, *Grammarly*, *MY Access!*, *WriteToLearn*). Future research can include a control group to exclude factors that might confound the findings, and studies could be undertaken to explore the effects of the intervention facilitated by other AWE tools, especially those with a higher precision rate, and to examine whether the students' language accuracy can be significantly improved under that circumstance.

Acknowledgments

We would like to thank the editor and the peer reviewers for their valuable suggestions. We would also like to thank Chengdu University of Information Technology for the research grant.

Authors' contributions

Huang Shu was responsible for study design, data collection and the writing of the manuscript. Chen Dan was responsible for revising the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the research grant provided by Chengdu University of Information Technology [grant number JYJG2022100].

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

- Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. *Automatic Treatment and Analysis of Learner Corpus Data*, 249–264. <https://doi.org/10.1075/scl.59.15ai>
- Bai, L., & Hu, G. (2017). In the face of fallible AWE feedback: how do students respond? *Educational Psychology*, 37(1), 67–81. <https://doi.org/10.1080/01443410.2016.1223275>
- Barrot, J. S. (2023). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, 36(4), 584–607. <https://doi.org/10.1080/09588221.2021.1936071>
- Chen, C.-F. E., & Cheng, W.-Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, 12(2), 94–112.
- Cheung, Y. L. (2015). Feedback from automated essay evaluation systems: A review of selected research. *TESL Reporter*, 48, 17.
- Dikli, S., & Bleyle, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback? *Assessing Writing*, 22, 1–17. <https://doi.org/10.1016/j.asw.2014.03.006>
- Grimes, D., & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *The Journal of Technology, Learning and Assessment*, 8(6).
- Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *Tesol Quarterly*, 44(1), 84–109. <https://doi.org/10.5054/tq.2010.213781>
- Huang, S., & Renandya, W. A. (2020). Exploring the integration of automated feedback among lower-proficiency EFL learners. *Innovation in Language Learning and Teaching*, 14(1), 15–26. <https://doi.org/10.1080/17501229.2018.1471083>
- Jiang, L., & Yu, S. (2020). Appropriating automated feedback in L2 writing: experiences of Chinese EFL student writers. *Computer Assisted Language Learning*, 1–25. <https://doi.org/10.1080/09588221.2020.1799824>
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, 44, 100450. <https://doi.org/10.1016/j.asw.2020.100450>

- Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing, 27*, 1–18.
- Liao, H. C. (2016). Enhancing the grammatical accuracy of EFL writing by using an AWE-assisted process approach. *System, 62*, 77–92. <https://doi.org/10.1016/j.system.2016.02.007>
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*(1), 36–62. <https://doi.org/10.5054/tq.2011.240859>
- Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing, 29*, 16–27. <https://doi.org/10.1016/j.jslw.2015.06.003>
- Polio, C. G. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning, 47*(1), 101–143. <https://doi.org/10.1111/0023-8333.31997003>
- Polio, C., & Williams, J. (2009). Teaching and testing writing. *The Handbook of Language Teaching*, 486–517. <https://doi.org/10.1002/9781444315783.ch26>
- Ranalli, J. (2018). Automated written corrective feedback: How well can students make use of it? *Computer Assisted Language Learning, 31*(7), 653–674. <https://doi.org/10.1080/09588221.2018.1428994>
- Shermis, M. D., & Burstein, J. C. (Eds.). (2003). *Automated essay scoring: A cross-disciplinary perspective*. Routledge. <https://doi.org/10.4324/9781410606860>
- Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing, 19*, 51–65. <https://doi.org/10.1016/j.asw.2013.11.007>
- Wang, Y., Shang, H., & Briody, P. (2013). Exploring the impact of using automated writing evaluation in English as a foreign language university students' writing. *Computer Assisted Language Learning, 26*(3), 234–257. <https://doi.org/10.1080/09588221.2012.655300>
- Wilson, J., & Roscoe, R. D. (2020). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research, 58*(1), 87–125. <https://doi.org/10.1177/0735633119830764>
- Xu, J., & Zhang, S. (2022). Understanding AWE feedback and English writing of learners with different proficiency levels in an EFL classroom: A sociocultural perspective. *The Asia-Pacific Education Researcher, 31*(4), 357–367. <https://doi.org/10.1007/s40299-021-00577-7>
- Yao, D. (2021). Automated writing evaluation for ESL Learners: A case study of *Pigai* system. *Journal of Asia TEFL, 18*(3), 949. <https://doi.org/10.18823/asiatefl.2021.18.3.14.949>
- Zhai, N., & Ma, X. (2022). Automated writing evaluation (AWE) feedback: A systematic investigation of college students' acceptance. *Computer Assisted Language Learning, 35*(9), 2817–2842. <https://doi.org/10.1080/09588221.2021.1897019>
- Zhang, Z. V. (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing, 43*, 100439. <https://doi.org/10.1016/j.asw.2019.100439>
- Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing, 36*, 90–102. <https://doi.org/10.1016/j.asw.2018.02.004>
- Zhang, Z. V., & Hyland, K. (2022). Fostering student engagement with feedback: An integrated approach. *Assessing Writing, 51*, 100586. <https://doi.org/10.1016/j.asw.2021.100586>

Appendix A

The Heading to Appendix A

Questionnaire on Students' Perceived Usefulness of *Pigai* feedback

To what extent do you agree with the following items? (1 = Strong Disagreement; 2 = Disagreement; 3 = Neutral; 4 = Agreement; 5 = Strong Agreement)

1. Using the feedback function helped me understand my writing performance.
2. Using the feedback function helped me develop my writing performance.
3. It was easy for me to understand the feedback.

4. The feedback identified the problems in my writing.
5. *Pigai* provided comments that helped improve my grammar.
6. Using the feedback helped me improve my ability in vocabulary use.

Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).