
ENHANCING WRITING SKILLS OF EFL LEARNERS THROUGH AUTOMATED FEEDBACK: AN EMPIRICAL INVESTIGATION

Pariyanto ; Novalita Fransisca Tungka
Universitas 17 Agustus 1945 Surabaya, Indonesia
Universitas Sintuwu Maroso Poso, Poso
pariyanto@untag-sby.ac.id ; novalita@unsimar.ac.id

Received: December 24, 2023; Accepted: January 20, 2024; Published: January 30, 2024

Abstract- *This study examines the effectiveness of an automated writing evaluation (AWE) tool in improving the writing skills of lower-proficiency EFL learners. The study involved 34 non-English major university students who received AWE feedback from the WhiteSmoke Writing Assistant during their writing practice sessions. The quality of their writing was evaluated using pre-and post-tests, and their perceptions of the feedback were gathered through a questionnaire. The results showed a significant improvement in the quality of students' writing after using AWE feedback. Additionally, the students perceived the AWE feedback as positively impacting the quality of their writing. These findings suggest that AWE feedback is a promising tool that can be incorporated into the writing class program to enhance the quality of students' writing. Writing teachers can use the tool to evaluate students' compositions and provide targeted feedback to help them improve their writing skills. Overall, the study highlights the potential of technology-based writing tools to facilitate language learning and improve EFL learners' writing proficiency.*

Keywords: *automated writing evaluation; writing skills; feedback; writing class program; technology-based writing tools*

Corresponding Author:

Pariyanto
Universitas 17 Agustus Surabaya
Jl. Semolowaru 45, Surabaya

INTRODUCTION

In recent years, second language writing has seen a growing interest in using technology-based tools to support writing instruction and assessment. One such tool is automated writing evaluation (AWE), which provides instant feedback on written texts' grammatical, lexical, and organizational features. AWE technology has been increasingly integrated into language teaching and learning contexts, and a growing number of studies

have investigated its efficacy in enhancing writing skills and providing more efficient and practical feedback to language learners (Bai & Hu, 2016; Hockly, 2019; Huang & Renandya, 2018; Koltovskaia, 2020; Ranalli et al., 2016; Tian & Zhou, 2020; Wilson et al., 2017; Zhang, 2020).

Many of the previous studies focused more on the validity and reliability of the AWE (e.g., Bai & Hu, 2016; Shermis & Burstein, 2013), and few studies investigated the effect of specific AWE feedback on the quality of EFL learners' writing (e.g., Hegelheimer, Dursun, & Li, 2016; Stevenson, 2016). Understanding the effect of particular AWE feedback on the quality of students' writing is essential since it may inform the writing teachers to consider incorporating AWE feedback into their teaching.

The research gap in previous studies is the need for more investigation into the effect of specific AWE feedback on the quality of EFL learners' writing. While there is substantial literature on the validity and reliability of AWE, there is a lack of research on the specific impact of particular AWE feedback on the quality of EFL learners' writing. This gap is significant because understanding the effect of particular AWE feedback on the quality of students' writing is crucial for informing writing teachers about the potential benefits of incorporating AWE feedback into their teaching. This gap presents an opportunity for future research to delve into the specific effects of different types of AWE feedback on the writing quality of EFL learners, thereby providing valuable insights for writing instruction and assessment. The present study is aimed at addressing such a research gap by exploring the effect of AWE feedback on the quality of EFL learners' writing and EFL learners' perception.

While AWE technology holds promise for various advantages, its efficacy concerning improving writing skills and enhancing feedback quality remains a subject of contention within the academic discourse among researchers and practitioners. Certain investigations have yielded affirmative outcomes, elucidating that AWE feedback facilitates learners in recognizing and rectifying errors, elevates the overall caliber of their writing, and fosters heightened motivation and engagement. Conversely, alternative scholarly inquiries have voiced apprehensions about the precision and dependability of AWE systems, the constraints inherent in their feedback mechanisms, and the potential

deleterious consequences associated with an undue reliance on technological interventions in writing instruction.

AWE systems, underpinned by computational linguistic algorithms, can analyze written texts and provide comprehensive feedback on various facets of writing, including grammar, vocabulary, syntax, and coherence. Feedback can encompass error correction, improvement suggestions, or scoring based on predefined criteria (Ducasse, 2023; Hinkel, 2004; Khezrlou, 2023; Kirsznner & Mandell, 2009; Lee, 2020). AWE technology is engineered to offer more objective and consistent feedback compared to traditional assessment methods such as teacher feedback or peer review, thereby reducing the time and effort required for grading and providing feedback (Chen & Cheng, 2008; Hegelheimer et al., Z, 2016).

With the continual advancement of technology, particularly in Natural Language Processing (NLP) and Latent Semantic Analysis (LSA), software that automatically analyzes students' writing has gained popularity. This software is referred to as Automated Essay Scoring (AES), Automated Essay Evaluation (AEE), or Automated Writing Evaluation (Hockly, 2019). Although researchers use different terms to denote this software, this study prefers the term AWE as it encompasses more text genres and provides formative evaluation, not merely scoring as suggested in AES. The software compares students' writing to a vast database of similar genres. The analyses conducted by the software include syntax, text complexity, vocabulary range, writing style, spelling, and punctuation. A score and improvement suggestions are also provided. AWE platforms such as Criterion, Write & Improve, MY Access, Grammarly and WhiteSmoke are commonly employed in English language teaching (Elliot et al., 2013 for details).

WhiteSmoke Writing Assistant is an online-based software tool that utilizes NLP technology to check grammar, style, spelling, and punctuation. It also provides a writing score based on several metrics, error explanations, and a thesaurus. This software offers two products, Desktop Premium and White Smoke Web, and three pricing packages: Web - 5/month, Premium - 6.66/month, and Business - \$11.50/month (see www.whitesmoke.com). Unlike other grammar checker software, such as Grammarly, this software does not provide a free version. However, at the time of this study, the

software was freely available for the Android version. Therefore, the WhiteSmoke Writing Assistant utilized in this study was the free Android version.

Student perceptions of automated feedback have been a topic of contention in the literature. Some studies (Parra & Calero, 2019; Fang, 2010; Li et al., 2015; Ma, 2013; Tsuda, 2014) have reported that students have favored opinions toward computer-generated feedback in their writing. However, contrasting findings on this topic are presented by other studies (e.g., Chou et al., 2016; Chen & Cheng, 2008; Grimes & Warschauer, 2010). Chen and Cheng (2008) found that scholars were not inclined toward using automated feedback in writing classes. Although research on students' perceptions of Automated Writing Evaluation (AWE) feedback suggests that students with lower language proficiency levels may find AWE feedback beneficial (e.g., Chen & Cheng 2008), there is a shortage of studies conducted among non-English majors and low proficiency English as a Foreign Language (EFL) learners.

This article aims to bridge this research gap by investigating the impact of AWE feedback on a cohort of lower-proficiency EFL students. It seeks to provide EFL students and educators with insights into the functionality of AWE feedback within their educational context and the perceptions of Indonesian EFL students towards such software. Understanding specific learning and teaching practices is crucial as it contributes to the efficacy and preparation of learning and teaching scenarios. Specifically, this study explores the extent to which AWE feedback enhances the quality of students' writing and the students' perceptions of the automated feedback in writing class.

METHODS

Research Design

The research design employed in this study was a mixed-method approach, which was a methodology that combined both quantitative and qualitative research elements to provide a more comprehensive understanding of the research problem (Berg-Schlosser, 2012; Creswell, 2014; Sreejesh & Mohapatra, 2014). This approach is particularly beneficial in educational research as it allows for a more nuanced

understanding of complex phenomena by capitalizing on quantitative and qualitative research strengths.

The quantitative component of the research primarily dealt with the participants' writing performance. This was measured through pretest and post-test scores, allowing for an objective evaluation of the participants' writing skills before and after the intervention. Using pretest and post-test scores is a common method in educational research for measuring improvement in a particular skill or knowledge area over time. In this case, it provides a quantifiable measure of the improvement in the participants' writing performance due to the intervention.

On the other hand, the qualitative component of the research focused on the participants' perception of AWE feedback used in the writing class. This involved gathering data on the participants' thoughts, feelings, and experiences with the AWE feedback. This was done through open- and closed-ended survey questions. The qualitative data provided a deeper understanding of the participants' experiences and perceptions. It could offer valuable insights into how the participants received and interpreted the AWE feedback and how it influenced their learning process.

The participants in this study were 34 university students enrolled in a 16-week English course at Adhi Tama Institute of Technology Surabaya. This course was required for and offered to first-year non-English-major students. It was a test-oriented English course before the students were introduced to the TOEFL preparation course in the following semester. The student's English proficiency was low (below 400 in Paper-Based TOEFL). The average age of the students was 20. Males, who were from various regions of Indonesia's East Java, predominated among the student participants. The students were chosen as the participants of this research as they were readily accessible and consented to participate.

Data Collection

The data of this study were collected from two different instruments. The quantitative data was collected through pre-tests and post-tests (normally called quasi-experiments, Muijs, 2004; Phakiti, 2014). The study lasted for two sessions. In the first session, the students were asked to write a paragraph based on the writing task prompt

(see Appendix 1 for details) for 30 minutes. The prompt was chosen as more and more students were using smartphones for their assistance, and it was relevant to their context. As the student participants were low-proficiency EFL learners, only one paragraph (approximately 150 words) was expected for students to write based on the given task prompt. After completing the paragraph, students submitted it on Google Classroom, the online platform they used for activities after the class. This draft was a pre-test to inform students' initial writing performance.

In the following session, the students were asked to download WhiteSmoke Writing Assistant on their smartphones. The instructor conducted a 15-minute demonstration of using WhiteSmoke Writing Assistant to familiarize students with the features. After that, the students were asked to upload or copy the text written in the first session and make necessary revisions on their smartphones based on the feedback provided by the WhiteSmoke Writing Assistant. After 30 minutes of revision time, the students were asked to resubmit their revised draft (used as a post-test) to Google Classroom and to complete the questionnaire on how they perceive the automated feedback given by the WhiteSmoke Writing Assistant.

Table 1. The Experimental Procedure of the Study

Session	Activities
1	Students wrote the first draft responding to the writing task prompt (Pre-test)
2	a. Students revised the first draft based on the automated feedback provided by WhiteSmoke Writing Assistant (Post-test) b. Students completed the questionnaire.

In addition, the qualitative data was collected through a questionnaire. The questionnaire was designed to explore students' perceptions of using WhiteSmoke Writing Assistant for their writing. Exploring students' perceptions is one of the characteristics of what Creswell (2014) defines as a Case Study. It contained 14 questions, 13 closed-ended and 1 open-ended. A combination of closed-ended and open-ended questionnaire items accommodated more responses that might not be found in a single type of question item. The questionnaire was made using Google Forms and distributed to students via Google Classroom.

Data Analysis

To answer research question no. 1, students' pre-test and post-test writing were scored using Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey's (1981) ESL Composition Profile rubric. It contains five categories: content, organization, vocabulary, language use, and mechanics. The researcher and his colleague teacher each scored three of the students' written texts at random to ensure that the scoring was accurate. The Cronbach's alpha coefficient obtained was 0.98, a high-reliability coefficient. A paired-sample t-test was conducted to determine whether there was a difference in the pre-test and post-test scores, as suggested by Muijs (2004) and Phakiti (2014).

To answer research question no. 2, a questionnaire containing 5-point Likert Scale questions (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree), dichotomous questions (yes and no), and an open-ended question (to comment; see Appendix 1) was analyzed. While the closed-ended questions were coded using the Likert Scale, the open-ended questions were analyzed using the 'grounded theory' method (see Creswell, 2014 for the grounded theory method in detail). The researcher and his colleague's teacher coded five randomly chosen student comments to establish reliability in coding. 97% of the inter-rater agreement was achieved for comment classification.

FINDINGS AND DISCUSSION

Finding

The quality of students' writing after using AWE feedback

The students' writing scores on the pretest before using AWE feedback and the posttest after using it on students' writing tasks can be seen in Table 2.

Table 2. Students' Writing Scores on the Pretest and the Posttest

	Pretest		Posttest	
	M	SD	M	SD
Content	23.52	1.54	25.02	1.62
Organization	13.97	1.26	15.14	1.87
Vocabulary	14.32	1.40	16.05	1.34
Language Use	14.35	2.15	18.85	2.25
Mechanic	3.52	.50	4.61	.49
Total Score	69.70	5.45	79.70	6.24

As can be seen from Table 2, students' writing scores in each category, content, organization, vocabulary, language use, and mechanics on the pretest and the posttest experienced changes. A Paired-samples t-test analysis was conducted to see if the changes were significant. Table 3 shows a Paired sample t-test analysis of the student's writing scores on the pretest and the post-test.

Table 3. Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	M	SD	SEM	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair Pretest-Posttest	-10.00	2.76	.47	-10.96	-10.96	-9.03	33	.00

As displayed in Table 3, there was a significant difference in the student's writing scores for the pretest (M = 69.70, SD = 5.45) and the posttest (M = 79.70, SD = 6.24), $t(33) = -9.03, p < .05$. It indicates that the use of WhiteSmoke Writing Assistant improved the quality of students' writing.

Students' opinions of the AWE feedback in a writing class

Table 4 shows students' perceptions of the feedback provided by the WhiteSmoke Writing Assistant. The average response of 4.0 shows that students primarily found the software useful for raising the caliber of their writing.

Table 4. Students' Perception of WhiteSmoke Writing Assistant

	1	2	3	4	5	Average Response
WhiteSmoke Writing Assistant is useful in improving the quality of my writing	0 (0%)	0 (0%)	2 (5.9%)	14 (41.2%)	18 (52.9%)	4.4
WhiteSmoke Writing Assistant is easy to use	0 (0%)	0 (0%)	6 (17.6%)	15 (44.1%)	13 (38.2%)	4.2
WhiteSmoke Writing Assistant provides detailed feedback	0 (0%)	1 (2.9%)	6 (17.6%)	16 (47.1%)	11 (32.4%)	4.0
WhiteSmoke Writing Assistant provides suggestions that help improve my writing	0 (0%)	0 (0%)	4 (11.8%)	10 (29.4%)	20 (58.8%)	4.4
WhiteSmoke Writing Assistant provides a succinct explanation of errors I committed	0 (0%)	0 (0%)	4 (11.8%)	12 (35.3%)	18 (52.9%)	4.4

	1	2	3	4	5	Average Response
WhiteSmoke Writing Assistant is useful in improving the quality of my writing	0 (0%)	0 (0%)	2 (5.9%)	14 (41.2%)	18 (52.9%)	4.4
WhiteSmoke Writing Assistant is easy to use	0 (0%)	0 (0%)	6 (17.6%)	15 (44.1%)	13 (38.2%)	4.2
WhiteSmoke Writing Assistant provides detailed feedback	0 (0%)	1 (2.9%)	6 (17.6%)	16 (47.1%)	11 (32.4%)	4.0
WhiteSmoke Writing Assistant provides suggestions that help improve my writing	0 (0%)	0 (0%)	4 (11.8%)	10 (29.4%)	20 (58.8%)	4.4
WhiteSmoke Writing Assistant helps me understand English grammar rules	0 (0%)	0 (0%)	2 (5.9%)	16 (47.1%)	16 (47.1%)	4.4
Feedback provided by WhiteSmoke Writing Assistant is always helpful	5 (14.7%)	3 (8.8%)	7 (20.6%)	12 (35.3%)	7 (20.6%)	3.3
I agree with the suggestions provided by WhiteSmoke Writing Assistant	4 (11.8%)	3 (8.8%)	9 (26.5%)	5 (14.7%)	13 (38.2%)	3.5
I understand the explanation provided by WhiteSmoke Writing Assistant	1 (2.9%)	4 (11.8%)	9 (26.5%)	5 (14.7%)	15 (44.1%)	3.8
I do not have a technical problem with the use of the WhiteSmoke Writing Assistant application	3 (8.8%)	5 (14.7%)	6 (17.6%)	7 (20.6%)	13 (38.2%)	3.6
Average						4.0

Note: 1= strongly disagree, 2= disagree, 3= neutral, 4= agree, 5= strongly agree

Table 5. Students' Perception in General of WhiteSmoke Writing Assistant

	Yes	No
In general, WhiteSmoke Writing Assistant has a positive impact on the quality of my writing	33 (97.1%)	1 (2.9%)
WhiteSmoke Writing Assistant makes me more confident in writing English text	33 (97.1%)	1 (2.9%)
WhiteSmoke Writing Assistant increases the grade of my writing	33 (97.1%)	1 (2.9%)

Discussion

Students' responses on dichotomous questions, as displayed in Table 5, confirm students' perception discussed above, where students mostly view WhiteSmoke Writing Assistant as positively impacting the quality of their writing.

The analysis reveals a marked improvement in the quality of students' writing following the utilization of Automated Writing Evaluation (AWE) feedback provided by the WhiteSmoke writing assistant. This software enables students to significantly reduce errors in their written text, aligning with the findings of Liao's (2015) study on automated writing evaluation to diminish grammar errors in writing. Liao (2015: 12) posits that this positive outcome may be attributed to the 'repeated practice of linguistic elements facilitated by using the AWE system in a process-writing pedagogy.'

Furthermore, these findings corroborate the study by Chen and Cui (2022) and Wang, Shang, and Briody (2013) on the impact of AWE on English as a Foreign Language (EFL) university students' writing. They reported that exposure to AWE feedback improved students' writing, encompassing sentence structures and tenses and appropriate word usage. It was also observed that students predominantly exhibited a positive attitude towards the AWE application in terms of enhancing writing accuracy.

The AWE feedback, an explicit type of corrective feedback, allows students to readily identify the errors they committed in their writing and contemplate revisions based on the provided improvement suggestions. As Ellis (2006) suggested, low-intermediate speakers of English as a second language benefited more from explicit than implicit corrective feedback on acquiring grammatical features. Theoretically, explicit learning yields explicit memories (Ellis, 2005). Explicit learning engenders conscious awareness, which, according to Ellis (2005, p.317), is pivotal in 'the initial consolidation of a unitary representation, and in order to 'bind features to form newly integrated objects,' attention is required

In terms of students' perceptions of the WhiteSmoke writing assistant, an overwhelming majority expressed positive sentiments. Students confidently acknowledged that the software provides constructive suggestions that contribute to the enhancement of their writing. Numerous students echoed the sentiment that it is a valuable tool for writing, with comments such as, "This application helps correct errors in my writing" and "This tool helps explain errors and suggestions for improvement."

Most students resonated with the synonym-related comment, "This application is great," which accounts for the high percentage of agreement (97.1%) with each statement. This revealed that the English as a Foreign Language (EFL) students

participating in this study exhibited significant improvements in the quality of their written texts after utilizing the automated feedback provided by WhiteSmoke. It is worth noting that such an improvement might not have been realized if the students harbored negative attitudes towards WhiteSmoke. However, one student found the software's feedback challenging to comprehend, commenting, "I am not sure about a few terms in the feedback." This student may have a limited vocabulary, hindering their understanding of the feedback provided by WhiteSmoke.

It is crucial to acknowledge the limitations of this study. Firstly, the research participants comprised a specific group of Indonesian EFL learners, which may limit the applicability of the findings to other learning contexts. Secondly, this study examined the free mobile version of the Automated Writing Evaluation (AWE), specifically the WhiteSmoke writing assistant. A premium AWE version may yield different effects on EFL students' writing. Lastly, students only wrote two drafts, including the revision. The effect on students' writing may vary with multiple drafts.

CONCLUSIONS AND SUGGESTIONS

To conclude, the present study indicated that AWE feedback by WhiteSmoke Writing Assistant effectively improves the quality of lower proficiency EFL students' writing and is perceived positively by the students. Positive perception by students regarding the AWE feedback by WhiteSmoke Writing Assistant is likely to contribute to its effectiveness in improving the quality of students' writing. Also, repeated practice of linguistic elements by students in the AWE feedback may trigger them to minimize the errors so that such improvement of students' text can take place.

Future studies with a more varied context of students, AWE applications, and multiple drafts may be required to validate the findings of the present study and the previous research. A different focus of study regarding AWE feedback may also be needed to describe a more comprehensive phenomenon of automatic feedback in the writing domain. It is significant as automation has become one of the central features of 21st-century education. This part should contain an answer or explanation to the problem of research.

REFERENCES

- 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>.
- Berg-Schlosser, D. (2012). *Mixed methods in comparative politics: Principles and applications*. Springer.
- Chen, C. F. E., & Cheng, W. Y. E. C. (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. <http://dx.doi.org/10125/44145>.
- Chen, M., & Cui, Y. (2022). The effects of AWE and peer feedback on cohesion and coherence in continuation writing. *Journal of Second Language Writing*, 57, 100915. <https://doi.org/10.1016/j.jslw.2022.100915>.
- Chou, H. N. C., Moslehpour, M., & Yang, C. Y. (2016). My access and writing error corrections of EFL college pre-intermediate students. *International Journal of Education*, 8(1), 144-161.
- Creswell, J. W. (2014). *Research designs. Qualitative, quantitative, and mixed methods approaches*. London: Sage.
- Ducasse, A. M. (2022). The Use of Technology for Redesigning L2 Language Assessments: Tasks, Rubrics, and Feedback in Emergency Remote Teaching Contexts. In *Technology-Assisted Language Assessment in Diverse Contexts* (pp. 212-227). Routledge. <https://doi.org/10.4324/9781003221463-17>.
- Elliot, N., Gere, A. R., Gibson, G., Toth, C., Whithaus, C., & Presswood, A. (2013). Uses and Limitations of Automated Writing Evaluation Software. *WPA-CompPile Research Bibliographies*. <http://comppile.org/wpa/bibliographies/Bib23/AutoWritingEvaluation.pdf>
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, 27(2), 305-352. <https://doi.org/10.1017/S027226310505014X>.
- Ersanlı, C. Y., & Yeşilel, D. B. A. (2023). The Use of Automated Writing Evaluation Tools to Foster ESL Writing Instruction. In *New Directions in Technology for Writing Instruction* (pp. 193-209). Cham: Springer International Publishing. https://doi.org/https://doi.org/10.1007/978-3-031-13540-8_10.
- Fang, Y. (2010). Perceptions of the computer-assisted writing program among EFL college learners. *Journal of Educational Technology & Society*, 13(3), 246-256.

- 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).
- Hegelheimer, V., Dursun, A., & Li, Z. (2016). Automated writing evaluation in language teaching: Theory, development, and application. *Calico Journal*, 33(1), i-v.
- Hockly, N. (2019). Automated writing evaluation. *ELT Journal*, 73(1): 82-88. <https://doi.org/10.1093/elt/ccy044>.
- 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>.
- Jacobs, H. L. (1981). *Testing ESL Composition: A Practical Approach. English Composition Program*. Newbury House Publishers, Inc., Rowley, MA 01969.
- Kirsznner, L. G., & Mandell, S. R. (2011). *Writing first with readings: Practice in context*. Macmillan.
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) wprovided by Grammarly: A multiple case study. *Assessing Writing*, 44, 100450. <https://doi.org/10.1016/j.asw.2020.100450>.
- Khezrlou, S. (2022). Repeated academic writing with synchronous and asynchronous teacher electronic feedback: How are macro and micro aspects affected?. In *Technology in Second Language Writing* (pp. 63-79). Routledge. <https://doi.org/10.4324/9781003279358-5>.
- Lee, I. (2020). Utility of focused/comprehensive written corrective feedback research for authentic L2 writing classrooms. *Journal of Second Language Writing*, 49, 100734.
- Liao, H. C. (2016). Using automated writing evaluation to reduce grammar errors in writing. *Elt Journal*, 70(3), 308-319. <https://doi.org/10.1093/elt/ccv058>.
- 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. <https://doi.org/10.1016/j.jslw.2014.10.004>.
- Ma, K. (2013). Improving EFL Graduate Students' Proficiency in Writing through an Online Automated Essay Assessing System. *English Language Teaching*, 6(7), 158-167. <http://dx.doi.org/10.5539/elt.v6n7p158>.
- Muijs, D. (2010). *Doing quantitative research in education with SPSS*. London: Sage.

- Parra G, L., & Calero S, X. (2019). Automated writing evaluation tools in the improvement of the writing skill. *International Journal of Instruction*, 12(2), 209-226. <https://doi.org/10.29333/iji.2019.12214a>.
- Phakiti, A. (2014). *Experimental Research Methods in Language Learning*. London: Bloomsbury.
- Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, 37(1), 8-25.
- Renandya, W. A. (2020). *Instructed SLA: A Practical Guide For Teachers*. TEFLIN Publication Division: Soul.
- Stevenson, M. (2013). Shermis, MD, & Burstein, J. (2013). *Handbook of Automated Essay Evaluation: Current applications and new directions (Eds)*. Routledge: New York.
- Sreejesh, S., & Mohapatra, S. (2014). *Mixed method research design: an application in consumer-brand relationships (CBR)*. New York: Springer.
- Stevenson, M., & Phakiti, A. (2019). Automated feedback and second language writing. *Feedback in second language writing: Contexts and issues*. Cambridge University Press: Cambridge. <https://doi.org/10.1017/9781108635547.009>.
- Stevenson, M. (2016). A critical interpretative synthesis: The integration of automated writing evaluation into classroom writing instruction. *Computers and Composition*, 42, 1-16. <https://doi.org/10.1016/j.compcom.2016.05.001>.
- Tian, L., & Zhou, Y. (2020). Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context. *System*, 91(4), 1-14. <https://doi.org/10.1016/j.system.2020.102247>.
- Tsuda, N. (2014). Implementing Criterion (automated writing evaluation) in Japanese college EFL classes. *Language and Culture: The Journal of the Institute for Language and Culture*, 18, 25-45. <https://konan-u.repo.nii.ac.jp/record/566/files/K00490.pdf>.
- Wang, Y. J., Shang, H. F., & 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., & Ahmed, Y. (2017). Automated formative writing assessment using a levels of language framework. *Assessing Writing*, 34(4), 16-36. <https://doi.org/10.1016/j.asw.2017.08.002>.

Zhang, Z. V. (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing*, 43(1), 1 - 14. <https://doi.org/10.1016/j.asw.2019.100439>.