INNOVARE JOURNAL OF EDUCATION

Vol 11, Issue 2, 2023, 14-19

ISSN: 2347-5528

Research Article

Potential Impact of Automated Writing Evaluation (AWE) on EFL Learners' Writing Accuracy: Learners' Attitudes in Focus

Mahshid Asfa and Ehsan Rezvani

English Department, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

Abstract

The present study aimed to examine the potential impacts of Automated Writing Evaluation (AWE) on Iranian English as Foreign Language (EFL) learners' writing accuracy. Another study objective was to identify the learners' attitudes toward using AWE in English classes. As such, 50 female EFL students at the intermediate level of language proficiency were chosen via convenience sampling to participate in the present study. They were then randomly divided into two groups of Control Group (CG) and the Experimental Group (EG), each consisting of 25 learners. Before the treatment, a researcher-made writing composition pre-test was administered to both groups. Next, the learners in the EG received the treatment in which an AWE system was employed to check the learner's written productions. Next, learners in the CG were exposed to the conventional pen-and-paper methods of English writing classes. Then, a writing composition post-test was administered. The performances of learners on pre and post-tests were scored based on accuracy in terms of syntactic errors. After administering the post-test, EG learners were asked to fill in a questionnaire on AWE.

Moreover, they were asked two questions in a semi-structured interview on their attitudes toward applying to the AWE program. The researcher recorded and later transcribed the learners' answers for further analysis. The independent sample *t*-test revealed that Iranian EFL learners' writing accuracy significantly improved after receiving AWE. Furthermore, the Chi-square test results, in conjunction with the interview results, showed that learners had positive attitudes toward AWE in writing classes. The study will have implications for the second language (L2) writing practices of EFL teachers and students.

Keywords: automated writing evaluation, l2 writing, syntactic accuracy, attitude

Introduction

Today, the computer has become a writing instrument and an interaction vehicle for many individuals. This movement has revolutionized the educational practice of writing instructors over the past decades (Chen & Cheng, 2008; Noroozi et al., 2021). Instructors have utilized different electronic writing media such as word processors, e-mail exchanges and bulletin boards in their training. Recent progress in Automated Writing Evaluation schedules (AWE) has attracted many instructors to execute this novel technology for scoring and measuring learners' writing (Li et al., 2015; Wilson et al., 2022).

For learners, electronic writing media do writing and rectifying activities much faster and simpler. As a result, learners are more inclined to reconsider their compositions (Moseley, 2006). For instructors – particularly those who think that writing is a recursive process and that learners must incessantly rewrite, reconsider and rectify their writing to amend their essays-involving learners in repetitive exercises may appear incumbent. However, instructors may become tired and lose their excitement after focusing on correcting papers and individually purveying special feedback if they instruct an important number of learners. Thanks to the utilization of AWE, instructors may be more inclined to give writing tasks more commonly to learners (Wang, 2015).

Several studies over the past two decades have determined the benefits of AWE programs: The immediate corrective feedback due to revisions (Phillips, 2007); the enhanced motivation of learners (Chou & Chung, 2013), the writing of longer texts with fewer errors (Grimes, 2008), useful alterations in learner conceptions, including the conception that writing is a recessive process rather than a linear process (Moseley, 2006). However, researches on the influence of AWE programs are inconclusive. Furthermore, several studies on AWE programs (e.g., Grimes, 2008) were run in English as a Second Language (ESL) rather than English as a foreign language (EFL) environments. For instance, while Vantage My Access® and ETS Criterion® are presently the most frequent AWE utilizations, research has rarely been conducted due to these two schedules in EFL classroom settings. Therefore, the present study aims to investigate AWE's effects on the writing accuracy of Iranian EFL learners to bridge this gap.

Review of the Related Literature

Acquiring L2 writing ability is one of the most problematic abilities for L2 students since it needs the proficiency of a diversity of linguistic, cognitive, and sociocultural abilities (Barkaoui, 2007). Some researchers, including Ferris and Roberts (2001) and Hyland and Hyland (2001), stated that, in addition to encouraging L2 students to write frequently, they should also be given helpful and suitable feedback and encouragement. As a result, as Meng (2013) stated, error correction has played a significant role in language instruction. Whatever the suggested type of feedback, some prior studies have shown that written corrective feedback (WCF) can improve second-language students' writing accuracy and efficiency (see, e.g., Bitchener, 2008; Bitchener & Knock, 2008).

DOI: https://dx.doi.org/10.22159/ijoe.2023v11i2.47035. Journal homepage: https://innovareacademics.in/journals/index.php/ijoe.

Acknowledgment: We would like to thank the IJOE reviewers for their meticulous reading of our article and their valuable comments. **Authors' Contributions:** Both authors have equally contributed to this paper. **Conflict of Interest:** The authors declare that there is no conflict of interest of any sort. **Funding Source:** This research did not receive any kind of funds.

Correspondence concerning this article should be addressed to Ehsan Rezvani, English Department, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran. **Email:** e.rezvani@khuisf.ac.ir

Nevertheless, as the significance of learning second or foreign languages enhanced around the universe, the frequency of the L2 students also enhanced in classes; as an outcome, catering Written Corrective Feedback (WCF) has become a hard task for L2 instructors because catering the WCF is really time-consuming (Salteh & Sadeghi, 2012). In this regard, some researchers suggested that technology can solve this problem and feedback given by computers can decrease the instructor's workload (Chappelle, 2001). Nonetheless, feedback provided by a computer cannot replace the instructor's feedback, but it can be a helpful addition to it (Lavolette et al., 2015).

The primary challenge writing instructors face is ensuring that students transfer knowledge and skills from one writing assignment to the next, learn from their mistakes, and develop both writing fluency and accuracy. This is something that most writing teachers claim to be true, and Leki (1992) shares this belief. Unfortunately, despite teacher instructions and needed reconsiderations, students usually indicate little or no development in their writing. As Howrey and Tanner (2008) expressed, students often do not learn to take charge of their writing and usually ignore instructor feedback on final drafts.

On the other hand, the study has indicated that Computer-Assisted Language Learning (CALL) could efficiently assist L2 students in developing their information about the English Language and its relevant abilities (Blake, 2000; Noroozi et al., 2020). Another significant issue assuming the feedback given by a computer is what some methods of SLA, containing the communication approach, usage-based approaches, and sociocultural assumption, have offered and that is feedback on language requirements to be instant to be efficient.

Different second language learning notions can support the effectiveness of CALL. As Warschauer (2005) illustrated, Vygotsky's sociocultural notion can be utilized to CALL. According to Vygotsky (1981, as cited in Warschauer, 2005), the whole stream and construction of mental works will change through intermediation or the incorporation of instruments or meditational means. Warschauer (2005) noted computers as an instance of those meditational means. CALL can also cater to communicative learning setting and collaborative writing in the L2 learning, modalities that are consistent with the social learning facet of Vygoskyian's notion (Warschauer et al., 1996) as well as Ellis's (1999) view of interactionist SLA (cited in Soltanpour & Valizadeh, 2017). Given the significant role of CALL in educational settings, the present study aimed to investigate the effects of AWE on the writing accuracy of Iranian EFL learners.

Empirical Background

Different assessments have been utilized to measure more explicitly the usefulness of AWE in assisting learners in revising. Though they appeared to have purveyed some conceptions of the impacts of AWE, the findings were far from decisive.

For example, Attali (2004) utilized the frequency of submissions as an index of learners' utilization of an AWE system and detected that 71% of the compositions were given just once. The frequency of submissions purveyed a harsh assessment of the number of learners' utilization of AWE, but any conclusions about how learners utilized AWE or profited from utilizing it based just on the frequency of submissions would be wrong. To address the issues focusing on the process of utilizing AWE, an increasing number of researchers have trusted classroom perceptions and interviews with instructors and learners.

Chen and Cheng (2008) investigated the usefulness of MY Access! in measuring writing development in three EFL writing classes, mainly via oral questions with the instructors and the learners. By observing the process, the writers were capable of showing a sharp comparison between the three classes in terms of the instructor's utilization and necessity of utilizing the system and the learners' answers and probable links between the instructor's educational decisions and the learners' conception. Meantime, the process-oriented method as well permitted the scholars to indicate the complication in assessing the impacts of AWE. While the research indicated that some learners observed the automated feedback as ambiguous and formulaic, others held that it was useful for determining and rectifying grammatical and mechanical errors, especially for learners with a lower level of English competency (Li et al., 2015).

In the EFL setting, Wang et al. (2013) examined the influence of utilizing AWE on freshmen writing with a team of 57 learners from a university. They utilized a quasi-experimental pre-posttest study design and the findings indicated an important diversity between the treatment group and the control group in terms of writing accuracy, with the treatment group determining clear writing achievements in terms of writing accuracy and student autonomy cognizance. In arguing the pedagogical implications, they offered that instructors should be more actively engaged in training learners' structure and instructing learners' patterns of writing so that learners learn how to enhance their language accuracy and how to amend their writing content and organization.

Palermo and Wilson (2020) touched upon the AWE system MI Write and reported the results of a mixed-methods study that examined the integration and implementation of AWE with writing instruction at the middle-school level, investigating AWE integration within both a traditional process approach to writing instruction and with strategy instruction based on the Self-Regulated Strategy Development model. Both instructional contexts were examined in terms of fostering growth in students' first-draft writing quality across successive essays, as well as students' and teachers' experiences and perceptions of teaching and learning with AWE. Their findings revealed that learners exhibited growth in first-draft writing performance and at comparable rates. Moreover, qualitative analyses of interview data revealed that AWE's influence on instruction was similar across contexts.

In a very recent study, Wilson et al. (2022) examined a novel approach to formative writing assessment encompassing an AWE system called MI Write. More precisely, they discovered elementary teachers' perceptions and implementation of MI Write and changes in students' writing performance in three genres. Teachers reported that MI Write was usable and acceptable, useful, and desirable; however, teachers tended to implement MI Write in a limited manner. Moreover, multilevel repeated measures analyses showed that students tended not to increase their performance in all genres. Their findings echoed the significance of utilizing scalable formative assessments to evaluate and adjust core instruction.

Objective

To fill the existing gap in the literature on the application of AWE in the Iranian EFL context, this study attempted to examine the effects of AWE on the writing accuracy of Iranian EFL learners as well as their attitudes toward implementing AWE in L2 writing classes.

Research Questions

This study aimed to answer the following questions:

1. Does use Automated Writing Evaluation (AWE) have a significant effect on Iranian EFL learners' writing accuracy?
2. What are Iranian EFL learners' attitudes toward using AWE in English classes?

Methodology

Design of the Study

The study was an experiment with a two-group pre-test and post-test design. In fact, in the present study, a quasi-experimental design was conducted and the main variables were AWE and writing accuracy as the independent and dependent variables, respectively.

Participants

The research population was female students enrolled in English classes at a language institute in Isfahan, Iran. The sample consisted of 50 female participants chosen via convenience sampling from the population of 100 learners at the language institute. First, 50 learners at the intermediate level of language

proficiency were selected, and an Oxford Quick Placement Test (OQPT) (Allan, 1992) was administered to ensure their homogeneity. Then, the sample was randomly divided into control (CG) and experimental (EG) groups. The experimental group (25 learners) was taught by using the Automated Writing Evaluation (AWE) software and the control group (25 learners) was taught using the traditional pen and paper method. English lessons were given to students two sessions a week and most students rarely had a chance to practice more English. The first language of learners was Persian; their ages ranged from 19 to 23. The participants were informed of the study's objectives and signed an informed consent form. They were also ensured that the study results would not affect their end-of-the-term final evaluations.

Instruments and Materials

The following instruments were used in the study:

Placement test. OQPT was administered to select intermediate learners. Oxford University Press and Cambridge ESOL developed the OQPT, a flexible measure of English language proficiency. This test consists of 60 multiple-choice items on vocabulary (30 items) and grammar (30 items), and learners with scores ranging from 0 to 10 are considered beginners; the learners with scores of 11 to 17 are deemed breakthrough; learners with scores of 18 to 29 are considered elementary; Pre-intermediate students have 30 to 39 points; intermediate students have 40 to 47 points; advanced students have 48 to 54 points, and; proficient students have 55 to 60 points. Based on the band scores of OQPT, advanced learners in the three intact classes of the institute were chosen as the study's participants.

Writing test (pre-test and post-test). The researcher-made pre and post-tests were designed to measure the student's achievement in writing accuracy. All learners in two groups had 50 minutes to write compositions in each pre- and post-test phase. The writing compositions were descriptive, and each article was about 150 words. The topic was: Describe the most embarrassing moment of your life and the total marks in the pre and post-tests were 20. After collecting the writing compositions, the researcher examined and counted the syntactic errors manually.

Questionnaire. The third instrument was a researcher-made questionnaire constructed based on previous studies. It was administered to the participants to obtain information about their attitudes toward AWE. The questionnaire was comprised of 25 items in total. Each item included a statement about students' attitudes towards using the AWE on the Likert Scale (strongly disagree, disagree, neutral, agree, and strongly agree). In the questionnaire, the learners were asked to choose one choice for each item. Before being officially used in the research, the questionnaire was piloted with twenty students with similar backgrounds and English proficiency levels to test the instrument's reliability. In fact, the questionnaire was pilot-tested and Cronbach's alpha was used to measure the reliability of the items used in the Likert scale. The reliability index was found as .72.

Interview. The interview is one of the most significant data-collecting instruments in the qualitative study. Interviews can allow researchers to examine phenomena that are not directly observable, such as students' self-reported experiences or attitudes (Gass & Mackey, 2000). Thanks to the interview, the researcher can achieve data from particular participants through face-to-face meetings, which can help the learners by clarifying the questions or clearly defining the learners' replies (Thuy, 2012).

In the present study, the interview was designed to explore the students' attitudes toward AWE software and explain a more comprehensive understanding of the results; therefore, it was given only to the students in the experimental group after the treatment. The students in the experimental group were required to answer the following interview questions:

1. Do you think using the Automated Writing Evaluation (AWE) software to improve writing accuracy in English classes is effective or ineffective?
2. Why do you think using the Automated Writing Evaluation (AWE) software to improve writing accuracy in English classes is effective or ineffective?

Procedures

The present study's objective was to investigate AWE's effects on the writing accuracy of Iranian EFL learners at the intermediate level of language proficiency. This being so, 50 female EFL students participated in the present study. They were intermediate learners based on their scores on a proficiency test. The learners were divided into two groups of EG and CG randomly. Each group consisted of 25 learners. Before the treatment, all learners were asked to sit for a writing composition test that fulfilled the objectives of the pre-test. All learners in two groups had 50 minutes to write compositions in a descriptive genre. After one week, the treatment started and continued over a three-month academic term. For the learners in the EG, an AWE system was employed and served as both a formative and summative assessment instrument. The program allows for multiple revisions and editing when applied for formative learning. Learners can revise their compositions multiple times based on the analytic assessment results and diagnostic feedback given to each essay draft submitted to the program. When run for summative assessment, the system is configured to provide a single submission with an overall assessment result (Chen & Cheng, 2008). Learners in the CG were exposed to the conventional methods of English classes. Essentially, students had to write about the textbook's topics and deliver them to the teacher for future rating. After the treatment, learners in two groups took part in the post-test. Accuracy in terms of syntactic mistakes was used to grade and compare student performances on pre-and post-tests. After administering the post-test, learners in EG were asked to fill in the questionnaire. They had 25 minutes to answer the questions on their attitudes toward AWE. In order to examine the quantitative information obtained from tests and questionnaires, all the quantitative data went under statistical computation by SPSS software. The researcher explained the purposes and the importance of the interview to all students in the experimental group. The students were asked two questions on their attitudes towards the application of AWE software. The length of the interview was different from seven to ten minutes. The researcher recorded and later transcribed the learners' answers for qualitative data analysis.

Results

The first research question was: Does the use of Automated Writing Evaluation (AWE) have a significant effect on the writing accuracy of Iranian EFL learners?

In order to answer the first research question, at first, it was necessary to compare the mean scores of the control and experimental groups' performances on pre-tests. Table 1 indicates the mean scores of two groups of learners on pre-tests.

Table 1

The Mean Scores of Control and Experimental Groups' Performances on Pre-tests

Factor	Group	*M*	*SD*	*SEM*
Pre-test	Control	17.4	1.8	.36
	Experimental	17.5	2.1	.42

Note. $N = 25$

Table 1, the mean scores of the control and experimental groups' performances on pre-tests were 17.4 and 17.5, respectively.

Moreover, an independent sample *t*-test was conducted to compare the mean scores of learners' performances on pre-tests. Table 2 indicates the results of the independent sample *t*-test due to the learners' performances on pre-tests.

Table 2
The Results of Independent Sample t-Test Due to the Control and Experimental Groups' Performances on Pre-tests

		Levine's test for equality of variances		t-test for equality of means						
		F	Sig.	*t*	*df*	Sig. (2-tailed)	*MD*	*SED*	95% Confidence interval of the difference	
									Lower	Upper
Pre-test	Equal variances assumed	.121	.73	-.217	48	.82	-.12	.55	-1.23	.99
	Equal variances not assumed			-.217	46.89	.82	-.12	.55	-1.23	.99

Note. $p < .05$

Table 2, the difference between the mean scores was not significant as the value of sig was bigger than .5 (sig = .82). After that, it was required to compare the mean scores of two groups of learners' performances on post-tests. Table 3 displays the mean scores of two groups of learners on post-tests.

Table 3
The Mean Scores of Control and Experimental Groups' Performances on Post-tests

	Group	*M*	*SD*	*SEM*
Pre-test	Control	17.5	1.7	.35
	Experimental	18.6	1.4	.28

Note. $N = 25$

Table 3 indicates the mean scores of the control and experimental groups' performances on post-tests were 17.5 and 18.6, respectively.

Finally, an independent sample *t*-test was conducted again to compare the mean scores of two groups of learners' performances on post-tests. Table 4 displays the results of the independent sample *t*-test due to the learners' performances on post-tests.

Table 4
The Results of Independent Sample t-Test Due to the Control and Experimental Groups' Performances on Post-tests

		Levine's test for equality of variances		t-test for equality of means						
		F	Sig.	*t*	*df*	Sig. (2-tailed)	*MD*	*SED*	95% Confidence interval of the difference	
									Lower	Upper
Post-test	Equal variances assumed	3.433	.07	-2.28	48	.02	-1.0	.45	-1.95	-.12
	Equal variances not assumed			-2.28	46	.02	-1.0	.45	-1.95	-.12

Note. $p < .05$

Table 4, the difference between the mean scores of the control and experimental groups' performances on post-tests was significant as the value of sig was smaller than .5 (sig = .02). This result indicates that using Automated Writing Evaluation had a significant effect on writing accuracy of Iranian EFL learners. Thus, the first research hypothesis, which stated that "Using Automated Writing Evaluation (AWE) does not affect on writing accuracy of Iranian EFL learners," was rejected.

The second research question was: What are Iranian EFL learners' attitudes toward AWE in English classes? All learners in the experimental group were asked to fill in the questionnaire. Learners had 25 minutes allocated time to answer the questions on their attitudes toward AWE. The chi-Square test was used to answer the second research question. Table 5 indicates the results of the Chi-Square test.

Table 5
The Results of the Chi-Square Test on EFL Learners' Attitudes toward AWE

Value	*df*	Asymp. Sig. (2-sided)
53.217[a]	96	.95

Note. a. 75 cells (60.0%) have an expected count of less than 5. The minimum expected count is .40.

Table 5 indicates the results of the Chi-square test were not statistically significant (sig = .95). This means that EFL learners in the experimental group had positive attitudes toward AWE.

Moreover, as it was mentioned before, learners in the experimental group were interviewed to know their attitudes toward AWE. The learners' attitudes, both positive and negative opinions toward AWE, were as follow:

All participants were very positive towards AWE. They believed that computers are substantial in all aspects of life and expressed that the first time they experienced a writing task through AWE, they became very agitated. They said that their interest in writing English enhanced, and they had never experienced writing composition so much fun. Moreover, they said that machine scoring, such as AWE software, should be integrated with the instructor's comments for a writing class. Some students expressed that the AWE software enhanced their motivation in language learning. They said that the program assisted them in comforting the process of learning through corrective feedback and this motivated them to better their language writing competency.

Some learners expressed that their confidence increased after writing the compositions with the help of the AWE program. They sensed less stress because they knew that they had committed fewer errors. Moreover, their self-esteem enhanced significantly in utilizing online writing and automatic corrective feedback programs. They expressed that they became more inclined to try writing English online. Two learners referred to independent learning, which happened to them in English classes. They expressed that they experienced autonomous learning when the AWE program was presented to the classroom. They said that the computer freed the learners to some extent from the instructor's job of scoring compositions. Moreover, learners believed that they could modify more after utilizing AWE software.

Almost all participants referred to the immediate corrective feedback which was provided by the computer. They expressed that they were more comfortable when the AWE software assisted them in detecting their errors. They also expressed that when they observed the immediate corrective feedback from the system, they were very pleasant and sensed that they wrote a perfect composition. Some learners expressed that the use of the AWE program could enhance the development of writing competency in terms of both accuracy and fluency. Moreover, they reported that "We sense that the other language skills such as speaking and reading skills have been amended via employing AWE program in English classes." "Some learners believed that "English learners have lots of grammatical errors and therefore they lose score." They expressed that the error analysis of grammar by AWE software was helpful for them as it developed their knowledge of grammar greatly. In spite of the above-mentioned positive attitudes, two learners expressed that utilizing the AWE program was not adequate for them and they required their instructor's assistance, as well.

Discussion

In this study, the effects of using AWE on the writing accuracy of Iranian EFL learners were investigated. The results revealed that using AWE had positive effects on the writing accuracy of EFL learners in Iranian contexts. In contrast to the previous studies, the finding of the present study does not agree with Braine's (1997) research that the conventional environment promotes more progress in writing than utilizing computers.

In spite of this disagreement, the findings agree with the study of Bayraktar (2001). The researcher ran a meta-analysis of the effectiveness of computer-assisted instruction (CAI) on learner success in secondary and college science education compared to the conventional method of instruction. Findings indicated a positive impact for CAI utilization when applied in tutorial patterns, with individual computer utilization, and when utilized as a complement to conventional instruction.

The findings of other studies (Cheng & Cheng, 2008; Grimes & Warschauer, 2010; Klobucar et al., 2013; Palermo & Wilson, 2020) also suggest that automated evaluation system plays a pivotal role in L2 writing classrooms, although teachers are not completely convinced of the accuracy of the computer-generated feedback. This argument is consistent with the research results reported by Elliot and Klobucar (2013) on automated essay evaluation and the teaching of writing. "Writing is a complex socio-cognitive construct… automated essay evaluation does not read the essay but is trained to behave as a human rater would" (p. 20).

The second research question focused on the Iranian EFL learners' attitudes toward the AWE program in English classes. The results of the questionnaire indicated that Iranian EFL learners had positive attitudes toward the AWE program in English classes. Moreover, the findings obtained from the interview indicated that learners perceived AWE as a useful tool in language learning, especially in writing skills, as they experienced the program as positive and their motivation, self-esteem and independent learning were enhanced. Learners also expressed that they were more relaxed when AWE catered corrective feedback for their written productions. Learners also expressed that the use of the AWE program could increase the development of language skills such as reading and writing and the development of grammar sub-skill. In spite of the mentioned positive opinions, some learners expressed that using the AWE was not adequate for them and they required their instructor's assistance, as well. In line with our findings, Wilson et al. (2022) offered the application of computer software to serve the objective of the independent enhancement of writing abilities, especially for EFL authors. According to Williams (2005), if the application of a computer program is attentively modeled, it can suggest to learners both help and independency in the writing process.

However, the findings of the second research question are not consistent with the study of Yu and Yeh (2003). They showed that most learners expressed that the corrective feedback from My Access was helpful just for the first review and that the following equivalent and frequent corrective feedback was ineffective. Our findings are also supported by the study of Chou and Chung (2013). They examined non-English majors' conceptions of the application of My Access. Many learners perceived that the diagnostic corrective feedback assisted them in considering their individual writing difficulties. The immediate scoring process also stimulated the learners to rectify their mistakes. Chou and Chung concluded that utilizing AWE is useful for EFL learners at a lower level of English proficiency.

Conclusion

The focus of the present study was to identify the effect of Automated Writing Evaluation (AWE) on the writing accuracy of Iranian EFL learners. It was also aimed at identifying the learners' attitudes toward using AWE in English classes. The findings indicated that AWE had significant effects on the writing accuracy of Iranian EFL learners and they had positive attitudes toward the program. Thus, it can be claimed that using AWE within process writing instruction and strategy instruction contexts improved their writing accuracy at comparable rates. Nonetheless, it is possible that there were differences in the ways that students in the two conditions subsequently revised their essays, as well as the way that students in each condition enacted their revision process.

It is, thus, vital for teachers to recognize that using AWE may result in their need to expend more effort to help students work with and around the limitations of AWE. This, in turn, will require effective professional development and support that goes beyond the technical knowledge required to use AWE. Moreover, EFL learners are encouraged to benefit from the findings of the present study to improve their L2 speaking by effective utilization of AWE. In other words, it can be claimed that if learners are familiar with the potentials of OCSs and their effects on their speaking and autonomous learning, then there are better chances of L2 development. Specifically, EFL learners who are reluctant or resistant to OCSs may fail to understand its benefits for their oral communication skills. Additional research should examine other aspects of writing, such as writing fluency and accuracy.

References

Attali, Y. (2004). Exploring the feedback and revision features of the Criterion service. http://www.ets.org/Media/Research/pdf/erater_NCME_2004_Attali_B.pdf

Barkaoui, K. (2007). Teaching writing to second language learners: Insights from theory and research. *TESL Reporter*, *40*(1), 35–48. http://ojsdev.byuh.edu/index.php/Issue1/article/view/1112

Bayraktar, S. (2001). A Meta-analysis of the effectiveness of computer-assisted instruction in science education. *Journal of*

Research on Technology in Education, 34(2), 173–188. https://doi.org/10.1080/15391523.2001.10782344

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, *17*(2), 102–118. https://doi.org/10.1016/j.jslw.2007.11.004

Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, *12*(3), 409–431. https://doi.org/10.1177/1362168808089924

Blake, R. (2000). computer mediated communication: A window on L2 Spanish interlanguage. *Language Learning and Technology*, *41*(1), 120–136. https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/973bed3f-8657-45ef-8717-bbd758105b45/content

Braine, G. (1997). Beyond word processing: Networked computers in ESL writing classes. *Computers and Composition*, *14*(1), 45–58. https://doi.org/10.1016/S8755-4615(97)90037-2

Chappelle, C. A. (2001). Innovative language learning: Achieving the vision. *ReCALL*, *13*(1), 3–14. https://doi.org/10.1017/S0958344001000210

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 and Technology*, *12*(3), 94–112. https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/77d263c9-dfcf-4382-b179-7002b43cd5ef/content

Chou, H. N., & Chung, C. M. (2013). EFL college students' writing performance and perceptions of the use of an automated writing evaluation system. *English Education*, *2*(1), 93–126.

Elliot, N., & Klobucar, A. (2013). Automated essay evaluation and the teaching of writing. In M. D. Shermis & J. Burstein (Eds.), *Handbook of automated essay scoring* (pp. 16–35). Routledge.

Ellis, R. (1999). *Learning a second language through interaction*. John Benjamins Publishing Company.

Ferris, D. R., & Roberts, B. (2001). Error feedback in L2 writing classes? *Journal of Second Language Writing*, *10*(3), 161–184. https://doi.org/10.1016/S1060-3743(01)00039-X

Grimes, D. (2008). *Assessing automated assessment: Essay evaluation software in the classroom*.

Grimes, D., & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *Journal of Technology, Learning, and Assessment*, *8*(6), 14–30.

Howrey, J. D., & Tanner, P. (2009). Writing portfolios: Empowering students, teachers, and the curriculum. in A. M. stoke (Ed.), *JALT2008 Conference Proceedings*. tokyo: JAlt.

Klobucar, A., Elliot, N., Deess, P., Rudniy, O., & Joshi, K. (2013). Automated scoring in context: Rapid assessment for placed students. *Assessing Writing*, *18*(1), 62–84. https://doi.org/10.1016/j.asw.2012.10.001

Lavolette, E., Polio, C., & Kahng, J. (2015). The accuracy of computer-assisted feedback and students' responses to it. *Language Learning and Technology*, *19*(2), 50–68. https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/53b08da7-fa29-4ad8-8594-9ea68bf9ef3d/content

Leki, I. (1992). *Understanding ESL writers: A guide for teachers*. Boynton/Cook Publishers.

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

Meng, Y. (2013). *Written corrective feedback: A review of studies since Truscott* (Working Paper). *TESOL & Applied Linguistics. https://doi.org/10.7916/salt.v13i2.1339*

Moseley, M. H. (2006). *Creating recursive writers in middle school: The effect of a writing program on student revision practices*. Capella University.

Noroozi, A., Rezvani, E., & Ameri, A. (2021). Students' perceptions of the incorporation of flipped learning into L2 grammar lessons. *Teaching English with Technology*, *21*(1), 112–130. https://files.eric.ed.gov/fulltext/EJ1283678.pdf

Noroozi, A., Rezvani, E., & Ameri-golestan, A. (2020). The effect of flipped classrooms on L2 learners' development and retention of grammatical knowledge. *Turkish Online Journal of Distance Education*, *21*(4), 14–30. http://doi.org/10.17718/tojde.803344

Palermo, C., & Wilson, J. (2020). Implementing automated writing evaluation in different instructional contexts: A mixed-methods study. *Journal of Writing Research*, 12(vol. 12 issue 1), 63–108. https://doi.org/10.17239/jowr-2020.12.01.04

Phillips, S. M. (2007). *Automated essay scoring: A literature review*. TASA Institute, Society for the Advancement of Excellent in Education.

Salteh, M. A., & Sadeghi, K. (2012). Teachers' corrective feedback in L2 writing revisited: Concerns against and suggestions for its employment. *World Applied Sciences Journal*, *17*(3), 375–383. https://www.idosi.org/wasj/wasj17(3)12/16.pdf

Soltanpour, F., & Valizadeh, M. (2017). The effect of the collaboration of reflective notes with CALL on EFL learners' writing accuracy. *European Journal of English Language Teaching*, *2*(2), 22–52. https://oapub.org/edu/index.php/ejel/article/view/644

Thuy, N. N. (2012). *The effects of semantic mapping on vocabulary memorizing.*

Vygotsky, L. S. (1981). The genesis of higher mental function. In J. V. Wertsch (Ed.), *The concept of activity in Soviet psychology*. Sharp.

Wang, P. (2015). Effects of an automated writing evaluation program: Student experiences and perceptions. *Electronic Journal of Foreign Language Teaching*, *12*(1), 79–100. https://e-flt.nus.edu.sg/wp-content/uploads/2020/09/v12n12015/wang.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

Warschauer, M. (2005). Sociocultural perspectives on CALL. In J. L. Egbert & G. M. Petrie (Eds.), *CALL research perspectives*. Lawrence Erlbaum Associates, Inc.

Warschauer, M., Turbee, L., & Roberts, B. (1996). Computer learning networks and student empowerment. *System, 24*(1), 1–14. https://doi.org/10.1016/0346-251X(95)00049-P

Williams, J. (2005). *Teaching writing in second and foreign language classrooms*. McGraw-Hill.

Wilson, J., Myers, M. C., & Potter, A. (2022). Investigating the promise of automated writing evaluation for supporting formative writing assessment at scale. *Assessment in Education: Principles, Policy and Practice*, *29*(2), 183–199. https://doi.org/10.1080/0969594X.2022.2025762

Yu, Y. T., & Yeh, Y. L. (2003). Computerized feedback and bilingual concordance for EFL college students' writing. *Taipei, Taiwan: Crane*.

Received: 09 December 2022
Revised: 10 January 2023
Accepted: 21 January 2023