The Impact of Using Automated Writing Feedback in ESL/EFL Classroom Contexts

Ameni Benali[1]

[1] York University, Toronto, Canada

Correspondence: Ameni Benali, York University, Toronto, Canada.

Abstract

It is undeniable that attempts to develop automated feedback systems that support and enhance language learning and assessment have increased in the last few years. The growing demand for using technology in the classroom and the promotions provided by automated- written-feedback program developers and designers, drive many educational institutions to acquire and use these tools for educational purposes (Chen & Cheng, 2008). It remains debatable, however, whether students' use of these tools leads to improvement in their essay quality or writing outcomes. In this paper I investigate the affordances and shortcomings of automated writing evaluation (AWE) on students' writing in ESL/EFL contexts. My discussion shows that AWE can improve the quality of writing and learning outcomes if it is integrated with and supported by human feedback. I provide recommendations for further research into improving AWE tools to give more effective and constructive feedback.

Keywords: writing corrective feedback, automated writing evaluation, human feedback, automated scores, second language writing

1. Introduction

Written corrective feedback is frequently used in language learning classrooms. It involves identifying students' errors (Liu et al., 2017), and it can be provided by peers, teachers or machines (Mehrabi-Yazdi, 2018). Feedback can be provided on different aspects of writing, including "accuracy, fluency, complexity, mechanics, cohesion, coherence, reader/writer interaction and content" (Knoch, 2011, as cited in Liu et al., 2017, p. 504).

There is evidence that feedback leads to successful revisions of drafts (Ferris & Roberts, 2001), and that it is associated with benefits in writing progress for EFL and ESL students (Stevenson & Phakiti, 2014). Research has also shown that written corrective feedback can lead to writing development in that it helps learners process, retain and uptake correct forms in writing (Ellis, 2008).

Research suggests different effective feedback types and strategies that can help learners edit their errors and improve the quality of their writing. It has been claimed that indirect (or implicit) feedback, which refers to identifying students' errors without providing any support or suggestions to correct them, can be more helpful to high proficiency students who do not need a lot of assistance because it helps them think critically about their mistakes, which can lead to improvement not only in their accuracy but also in their learning ability (Ferris & Roberts, 2001).

Direct (or explicit) feedback, on the other hand, refers to feedback that "not only calls attention to the error but also provides a specific solution to the problem" (Bitchener & Ferris, 2012, p. 148). It has been reported, however, that choosing between direct or indirect feedback depends on pedagogical considerations, including the level of the students, students' needs, and the nature of errors (Ranalli, 2018). Similarly, recent studies have shown that focused feedback that covers predetermined, specific error types (and are mostly based on learners' needs) is more valuable than unfocused feedback (Ferris et al., 2013) because it targets single error types chosen for learning goals, such as English articles (a, an, the) (Ranalli, 2018). Moreover, from a sociocultural perspective, effective corrective feedback needs to be dialogic and interactive, in which the reviewer can accommodate the students' needs and provide them with continuous support. This can guide the reviewer to the origin of the error, providing the students with the assistance they need. In fact, dialogic feedback is said to help students get further explanations about their errors, engage in interactive feedback and negotiate meaning with teachers and peers, which can promote second language learning (Merhawi, 2018; Woodworth & Barkaoui, 2020).

Research suggests that diagnostic feedback is valuable because it focuses more on specific detailed features rather than on general skills and abilities and thus can be more useful for students' improvement (Liu et al., 2017). Furthermore, there have been calls in the research to use formative feedback in addition to summative feedback. In formative assessments, students can be engaged in a process of drafting, revising and editing their essays, which can improve the quality of their writing and can inform instructional practices (Mellati & Khademi, 2018). Many contemporary researchers have argued that immediate feedback given to students during their ongoing courses can be more effective and constructive because it is used to increase writing production, as compared to delayed feedback that is provided at the end of courses and consequently becomes less relevant (Gibbs & Simpson, 2015).

However, despite the benefits that writing corrective feedback is thought to provide, for many reasons ESL/EFL learners may not always benefit from feedback provided by their teachers. First, providing formative feedback can be daunting and time-consuming for teachers because it requires frequent revising. As a result, students may end up receiving holistic feedback that focuses on "surface-level features" like grammar and spelling but not on content (Wang et al., 2020, p. 2). Second, it could be challenging for teachers to give instant frequent feedback inside and outside the classroom (Woodworth & Barkaoui, 2020). Research has shown that because of lack of training and writing assessment literacy in general, some teachers struggle to use feedback efficiently to improve their students' writing outcomes even if they hold positive beliefs about formative feedback (Crusan et al., 2016).

AWE tools or systems that have been under development and use since the 1960s (Chen & Cheng, 2008) are thought to help overcome these challenges and therefore alleviate teacher workload by giving the students more opportunities to receive immediate continuous feedback even outside the classroom (Woodworth & Barkaoui, 2020). AWE tools are promoted for their capability of providing diagnostic feedback that covers different aspects of writing, including semantics, syntax, grammar, word choice and content, in addition to enabling scoring options that help teachers grade and evaluate students' essays with a focus on linguistic elements. These tools allow learners to use multiple drafting and writing resources like thesauri (Chen & Cheng, 2008) and are equipped with features that allow teachers to create and guide writing tasks (Ranalli et al., 2016). According to Woodworth & Barkaoui (2020), the most widely used AWE tools are *Criterion*, *Project Essay Grade, My Access* and *Piagi,* all of which provide multiple drafting and instant feedback on different aspects of writing, and help learners and teachers keep portfolios of each student's work.

In this paper I discuss the effectiveness of AWE tools in improving the quality of writing, and I examine the role of classroom teachers in the presence of these tools. Accordingly, the research questions are as follows:

1) What are the benefits of using automated written feedback in the EFL/ESL context?

2) What are the shortcomings of using automated written feedback in the EFL/ESL context?

2. Discussion

2.1 Advantages of Using AWE in EFL/ESL Classroom Contexts

In general, research on the effectiveness of AWE identifies four benefits. Automated feedback can improve the quality of students' writing, enhance learning, promote learner autonomy and motivation to write, and ease the burden on teachers.

In addition to providing summative feedback, AWE systems can provide writers with formative feedback by giving them opportunities for multiple drafting, which can raise students' awareness of their errors and weaknesses hence improve their writing outcomes (Chen & Cheng, 2008). In fact, some of these systems function as writing supporting tools as they are supplied with different online writing resources (word banks and thesauri) and editing features (grammar and spelling) (Chen & Cheng, 2008). It has also been argued that these systems promote learner autonomy by providing student writers with opportunities to practice, write, evaluate and revise even a large number of essays in a self-regulated learning environment (Chen & Cheng, 2008). Furthermore, research shows that the immediacy of automated feedback can improve the quality of students' writing (Stevenson & Phakiti, 2014). Fast responses in addition to the types of comments and scores generated by these machines can increase students' competition and motivation to practice writing (Cheng, 2017). AWE programs are also thought to be helpful for teachers who can spend quality time teaching their students about different aspects of writing instead of spending most of the time rating students' essays and giving them feedback (Stevenson & Phakiti, 2014).

Some empirical studies have been conducted to examine the impact of using AWE feedback on EFL/ESL students' witting. Li et al. (2015) carried out a study to investigate the impact of using the AWE system on writing accuracy in two academic ESL writing courses during a 15-week semester in a midwestern university in the United States. Data were gathered qualitatively through interviews with instructors and ESL students and quantitatively through collecting the number of submissions for each paper recorded by the system and AWE error reports. Four

instructors and 70 ESL students participated in the study. The students were required to use the system for the revision and submission of the final draft, and they were encouraged to use the AWE outside the classroom. The teachers used *Criterion,* an AWE system that provides corrective feedback (in grammar, usage, mechanics and styles), highlights errors, and provides explanations and/or suggestions for corrections.

The data analysis showed that both the instructors and the students held positive views regarding the impact of AWE tools in improving the essay quality, but it also showed that the instructors and the students were not satisfied with the quality of feedback that was provided by the system in terms of content and organization. The findings also showed that *Criterion* led to increased revisions and that automated feedback helped the students improve their writing accuracy from first to final drafts.

It is difficult to draw conclusions about the impact of using automated corrective feedback on students' writing from this study. First, the AWE system might not have been the only reason for improving students' linguistic accuracy. In fact, the instructors used different strategies and instructional methods to implement this system in the classroom (like asking their students to achieve a minimum automated score before submitting their papers) that may have played an important role in improving students' linguistic accuracy and their ability to highlight and correct their errors. Research also suggests that students' writing improvement can result from the learning and teaching process and not necessarily from the implementation of these systems (Warschauer & Ware, 2006).

Second, students' dissatisfaction with the quality of feedback on content and organization could have been influenced by their teachers' practices and perspectives; the study reported that some teachers who did not think that the AWE system could help improve content and organization of essays provided their students with feedback on these aspects and asked them to use AWE only to get feedback in grammar and mechanics.

Finally, the study did not use a control group to measure and compare any potential improvement in the students' writing outcomes. Consequently, further studies that use controlled variables (teachers providing feedback) or control groups are required to draw accurate conclusions about the effectiveness of AWE systems in improving students' writing, not only in form but also in content.

Wang et al. (2020) conducted a study to investigate changes in the quality of students' use of text evidence from the first draft to the revised draft based on feedback messages they received from AWE systems. The study also examined students' perceptions and implementation of the feedback messages they received. Seven teachers from 7 different schools and 143 (3rd to 8th grade) students participated in the study, which took place in Louisiana (in the United States). The assessment required students to "show understanding of text(s) by writing a multi-paragraph response" (Wang et al., 2020, p. 5) that used evidence from the text. Also, among other dimensions, students' literary analysis response was scored for use of "clear reasoning supported by relevant text-based evidence in the development of the topic" (Wang et al., 2020, p. 5). Data were collected qualitatively by analyzing students' essays across first and second revised drafts. Learners also completed a survey about their perceptions on the effectiveness of *eRevise* upon submitting their revised draft. *eRevise* was the AWE system used in this study and was designed to rate responses, provide feedback to students on Response-to-Text-Assessment, and assess students' ability to reason about texts in their writing. The students were asked to respond to a prompt and support their arguments with ideas. To assess students' ability to reason about texts in their writing, *eRevise* provided three levels of feedback: completeness (by directing students to use more evidence and provide more details for each piece of evidence), explanation (by guiding them to explain their evidence), and connection (by directing students to connect their evidence to the overall argument).

The data analysis revealed that the students held positive views toward using *eRevise.* The findings also showed that most of the students made changes to their essays in that their use of text evidence improved from first to revised drafts and that this progress was in line with the feedback they received from *eRevise*. However, only 18% of the students showed substantive progress in essay quality.

This study confirmed the claims that AWE systems can provide writing assistance for students and that automated multiple drafting can improve the quality, organization, and content of the essay. However, it seems premature to draw any firm conclusions about the benefits of using automated corrective feedback to improve writing outcomes. First, this study showed improvement in the quality of one single essay, and further studies are needed to show that automated corrective feedback can be transferred to other writing situations and contexts and thus can lead to better writing outcomes on a long-term basis.

The studies discussed above showed that student writers hold positive views toward using AWE systems. They also showed that AWE can improve the accuracy of students' essays as well as the content and organization of students' writing. However, because of limitations in the two studies, it is difficult to evaluate the extent to which the two examined AWE systems contributed to the improvement of essay quality and/or content. Consequently,

more rigorous studies that (1) investigate the effectiveness of AWE tools for both content and form, (2) use a control group to compare and measure the impact and change in students' essays, and (3) examine writing outcomes on a long-term basis are required to support claims about the effectiveness of these tools.

2.2 AWE Tools: Challenges and Shortcomings

Despite the benefits brought by AWE, there are some claims in the literature about the limitations and challenges of implementing automated feedback in writing classes. First, the quality of AWE is called into question because these tools lack "human inferencing skills and background knowledge" (Stevenson & Phakiti, 2014, p. 52), and thus they cannot read, understand and evaluate meaning (Chen & Cheng, 2008). Furthermore, the fact that students write to machines instead of a real audience is said to strip writing from its meaningful communicative function because these texts will be treated as codes "devoid of any sociocultural contexts" (Chen & Cheng, 2008).

Moreover, feedback generated by machines may not be focused (Ranalli, 2018). In fact, it has been argued that feedback generated by automated systems like *Criterion* is unfocused in that it targets multiple types of errors as opposed to focused feedback that targets single types of errors for learning purposes. Research suggests that unfocused comprehensive feedback may "cognitively overload low proficiency learners" (Ranalli, 2018, p. 657) who, even if they can notice the errors identified by the AWE system, may not be able to understand them (Merhabi-Yazdi, 2018).

One of the concerns attributed to feedback provided by AWE systems is the issue of explicitness or the "extent to which feedback is direct or indirect" (Ranalli, 2018, p. 654). It has been reported that AWE systems' use of direct or indirect feedback is not based on "pedagogical considerations" like human corrective feedback, but it is based on the "technological capacities" of these systems (Ranalli, 2018, p. 655). Another issue is the "inaccuracy of error flagging" (Ranalli, 2018, p. 655). In fact, research has shown that accuracy rates of some types of errors provided by some AWE systems like *Criterion* can fall below 50%. Research shows that students tend to respond more frequently to accurate rather than inaccurate error flagging (Ranalli, 2018). The other factor influencing the extent to which students use and learn from automated feedback is that most AWE tools do not take students' individual differences (like language ability, first language, educational background and prior knowledge of grammar) into consideration (Ranalli, 2018).

Another limitation of AWE is that some of these systems cannot provide dialogic feedback. In other words, unlike human feedback, students cannot get further explanations of their errors from these tools and may not be able to find the origin of their error, which can hinder learning and improvement (Merhabi-Yazdi, 2018). Moreover, automated corrective feedback in content and subjective areas that is essential in improving students' writing skills (like reasoning, thinking and argumentation) may not be rigorously evaluated because these aspects are "still in early phases of development" (Wang et al., 2020). Furthermore, since most AWE systems focus on surface linguistic features, they are said to lead to negative washback effects on students' writing goals. Instead of attempting to improve their writing skills in both form and content, students may shift their attention to acquiring only those aspects of writing that are evaluated by these systems, such as grammar correctness, and therefore may neglect other essential areas, such as content and meaning (Stevenson & Phakiti, 2014).

Finally, there is evidence that a lack of understanding and implementation of AWE systems by teachers and/or students can lead to limited progress in the students' writing outcomes. Students may not benefit from the revising features provided by AWE systems either because they are not motivated to use them or because they lack sufficient knowledge and instruction on how to use them (Wang et al., 2020).

Some researchers have conducted empirical studies to investigate the shortcomings of using AWE tools in EFL/ESL classrooms. Chen and Cheng (2008) examined the effectiveness of AWE as a pedagogical tool in three college EFL writing classes in Taiwan. *My Access!* was the AWE tool used throughout a semester to improve students' writing and to ease the burden on teachers. The program provided feedback on formative and summative assessment and offered writing support to students. All the teachers adopted a similar "process-writing approach, including model essay reading activities followed by language exercises and pre-writing, drafting and revising activities" (p. 98). Three instructors and 68 Taiwanese English major EFL students at upper-intermediate levels participated in the study. Data were collected through end-of-course questionnaires, group interviews with 16 students, individual interviews with two instructors, and students' essays with feedback provided by the tool.

Data analysis showed that the students did not hold positive views about the program. The findings also revealed that teachers' instructions influenced students' perceptions of the effectiveness of the AWE system in improving their writing, and that student writers appreciated integrating human and automated feedback. The results showed that *My Access!* failed to provide useful feedback in some aspects, like developing ideas and coherence, whereas

human feedback was perceived as being meaningful and specific and as taking the social and communicative context into consideration.

The Chen and Cheng (2008) study confirmed many claims in literature about the shortcomings of AWE systems in improving only some aspects of writing. It also confirmed the importance of integrating automated and human feedback to achieve better writing outcomes. However, it should be noted that the findings of the study were based only on students' perceptions of the effectiveness of the AWE system. Consequently, if students' first and revised essays were compared, results could have been more accurate.

Ranalli (2018) investigated the influence of feedback accuracy and explicitness and other factors on students' ability to use *Criterion* to correct their errors. The study took place in two ESL writing courses in a midwestern university for 16 weeks. Eighty-two students from the high-intermediate level and advanced-low level participated in the study. To examine the influence of these factors on students' ability to use automated feedback, an error correction task was designed to observe students' response type and explicitness and accuracy of the feedback. Data analysis showed the importance of feedback explicitness in students' success in correcting errors. In fact, learners reported that responding to specific recommendations and feedback generated by *Criterion* was more helpful and easier than responding to generic feedback mainly because some of the participants lacked knowledge of grammatical terminology. The results also revealed that the absence of an accuracy determination led to relatively more appropriate and effective corrections.

One implication of the study is showing the importance of explicitness in automated feedback mainly because most of the AWE corpus is based on generic feedback, which explains why in many studies that used AWE systems like *Criterion*, most of the feedback was not used by the participants (Ranalli, 2018). These findings can encourage AWE program designers to develop tools that provide either both explicit and generic comprehensive feedback or feedback that is specially designed to meet learners' needs and backgrounds. However, because of some limitations in the study, it is difficult to claim that the findings were accurate representations of the influence of the discussed factors on students' interaction with AWE systems. First, the participants did not correct their own errors; rather, they corrected their peers' errors, and this may have influenced the findings because there are claims in literature that finding someone else's errors is easier than finding one's own errors (Ferris, 2012). Second, the participants were awarded gift cards to use *Criterion*, which might have influenced their engagement with the tool. Consequently, further research that investigates students' engagement with these AWE systems is needed.

3. Conclusion

The discussion above shows that the effectiveness of AWE tools depends on two main factors: a hybrid approach that combines human and AWE feedback (Chen & Cheng, 2008), and an efficient and judicious use of AWE systems. First, as AWE tools generate the same errors and explanations for all learners regardless of their language abilities, L1 and background, teachers can redress this limitation by helping their learners understand errors that are beyond their comprehension, thus allowing them to benefit from the feedback (Woodworth & Barkaoui, 2020). Second, AWE's lack of sociocultural, communicative and dialogic aspects can be complemented by human feedback. Learners can continuously negotiate meaning back and forth with their teachers in the process of revising and writing new drafts, enabling them to "achieve the goal of writing for effective communication in terms of form and meaning" (Chen & Cheng, 2008, p. 108).

In addition, my discussion reveals that the effectiveness of AWE tools also depends on how well teachers are trained to use these tools. Teachers need to know when, why and how to use AWE systems based on their learners' goals and needs. A learner who is writing for a real audience may need more human feedback, but another learner who struggles with fossilized grammatical errors may need more practice using automated feedback. Finally, AWE program designers today offer various systems with different options, and even if it is not the teacher who selects the appropriate AWE tool for the learners, it is the teacher's role to guide the learners to make the best use of these tools to achieve better learning outcomes based on their specific needs.

4. Implications, Limitations and Recommendations

4.1 Implications

This paper has some implications for teachers, researchers, AWE program developers and policymakers, and writing pedagogy. It showed that automated written feedback can improve learners' writing quality across multiple revisions. It also showed the important role that teachers' guidance and pedagogical instructions play in increasing students' benefits from using these machines.

My review highlighted the role of students' engagement with these tools in improving the quality of their essays. Thus, a valuable future line of inquiry needs to examine the factors that drive students to either engage or disregard automated feedback.

In discussing advantages and concerns related to using these tools, I do not aim at deciding whether these tools should be used. Instead, I seek to use the findings to raise awareness of limitations and affordances of using these tools in EFL/ESL classrooms and to encourage researchers to further investigate these concerns and invite AWE program designers to consider them when designing AWE systems (Cheng & Cheng, 2008). Finally, introducing technology in learning in recent years raises debates over replacing the role of teachers with technology devices. My findings have revealed the importance of teacher intervention in helping students make best use of AWE tools, and they strengthen the claims that AWE systems should be used to support writing instruction instead of replacing writing teachers (Chen & Cheng, 2008).

4.2 Limitations

Despite the implications of this paper, it is difficult to draw any firm conclusions about the effectiveness of AWE in improving the quality of drafts upon multiple revisions or in leading to better writing or learning outcomes. First, methodological issues in the studies I have discussed make it difficult to assert that the findings are accurate representations of the usefulness of these machines in improving essay quality. In fact, the studies used different AWE systems (that vary in their affordances and limitations), different samples (with different age groups, proficiency level and educational background), and different variables (including the presence or absence of factors such as teacher intervention and control group). Consequently, I suggest that future research needs to focus on the most widely used tools for educational purposes and to include different variables and representative samples to support these conclusions.

Second, I made two assumptions in my discussion of the effectiveness of these tools. My first assumption was that corrective feedback is effective in improving essay quality. However, some researchers have questioned the effectiveness of corrective feedback mainly on a long-term basis (Hyland & Hyland, 2006, cited in Stevenson & Phakiti, 2014). Others have gone further, claiming that corrective written feedback has negative impacts on essay quality and learning in general (Truscott, 1996, cited in Stevenson & Phakiti, 2014). Therefore, future studies that investigate the effectiveness of corrective written feedback on a long-term basis are needed to support this assumption.

My second assumption was that students' needs and goals are aligned with teachers' pedagogical and instructional goals, whereas research has shown that teachers and students may have different perceptions of learning and writing goals (Chen & Cheng, 2008). Thus, AWE future studies need to take students' needs into consideration when investigating the usefulness of these tools.

4.3 Recommendations for Further Research and AWE Program Design

As I discussed earlier, machines are used widely for learning and teaching purposes, and AWE research about the benefits of using these tools is no longer focused on whether we should use them in the classrooms. Instead, future studies should aim at improving AWE program design to overcome the shortcomings of current systems and to ensure that these tools provide students with effective feedback that can help them improve their essay quality and writing outcomes even on a long-term basis.

There have been calls from researchers to activate the dialogue box that is available in some AWE systems like *Criterion* to overcome the lack of dialogic features of automated feedback. This feature can allow for negotiation and interaction between students and teachers throughout multiple revisions. These negotiations in the dialogue box can also help researchers trace the progress of negotiations and evaluate the quality of collaboration between teachers and students and its impact on writing progress (Mehrabi-Yadi, 2018).

Some researchers strongly recommend integrating insights from applied linguistics and computer technology to design AWE systems that are informed by language assessment and learning theories, pedagogies and practices. It has been argued that if these two fields continue to work separately, benefits for students and learning will remain limited (Xi, 2010). Other researchers have suggested integrating current AWE systems with other kinds of support like games and tutorials (e.g., Writing Pal), classroom instructional practices, and teacher intervention to overcome the shortcomings and limitations of the current systems. Until AWE systems develop, teachers can rely on these resources to fill this gap (Wang et al., 2020).

References

Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge. https://doi.org/10.4324/9780203832400

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.

Cheng, G. (2017). The impact of online automated feedback on students' reflective journal writing in an EFL course. *The Internet and Higher Education, 34*, 18-27. https://doi.org/10.1016/j.iheduc.2017.04.002

Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing, 28*, 43-56. https://doi.org/10.1016/j.asw.2016.03.001

Ellis, R. (2008). A typology of written corrective feedback types. *ELT Journal, 63*(2), 97-107. https://doi.org/10.1093/elt/ccn023

Ferris, D. R. (2012). Written corrective feedback in second language acquisition and writing studies. *Language Teaching, 45*(4), 446-459. https://doi.org/10.1017/s0261444812000250

Ferris, D. R., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing, 22*(3), 307-329. https://doi.org/10.1016/j.jslw.2012.09.009

Ferris, D., & 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

Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning and teaching in higher education. *Learning and Teaching in Higher Education, 1*(1), 3-31.

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

Liu, M., Li, Y., Xu, W., & Liu, L. (2017). Automated Essay Feedback Generation and Its Impact on Revision. *IEEE Transactions on Learning Technologies, 10*(4), 502-513. https://doi.org/10.1109/TLT.2016.2612659

Mehrabi-Yazdi, O. (2018). Short communication on the missing dialogic aspect of an automated writing evaluation system in written feedback research. *Journal of Second Language Writing, 41*, 92-97. https://doi.org/10.1016/j.jslw.2018.05.004

Mellati, M., & Khademi, M. (2018). Exploring Teachers' Assessment Literacy: Impact on Learners' Writing Achievements and Implications for Teacher Development. *Australian Journal of Teacher Education, 43*(6), 1-18. https://doi.org/10.14221/ajte.2018v43n6.1

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

Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2016). 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. https://doi.org/10.1080/01443410.2015.1136407

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, E. L., Matsumura, L. C., Correnti, R., Litman, D., Zhang, H., Howe, E., Magooda, A., & Quintana, R. (2020). eRevis(ing): Students' revision of text evidence use in an automated writing evaluation system. *Assessing Writing, 44*, 100449. https://doi.org/10.1016/j.asw.2020.100449

Warschauer, M., & Ware, P. (2006). Automated writing evaluation: defining the classroom research agenda. *Language Teaching Research, 10*(2), 157-180. https://doi.org/10.1191/1362168806lr190oa

Woodworth, J., & Barkaoui, K. (2020). Perspectives on Using Automated Writing Evaluation Systems to Provide Written Corrective Feedback in the ESL Classroom. *TESL Canada Journal, 37*(2), 234-247. https://doi.org/10.18806/tesl.v37i2.1340

Xi, X. (2010). Automated scoring and feedback systems: Where are we and where are we heading? *Language Testing, 27*(3), 291-300. https://doi.org/10.1177/0265532210364643

Copyrights