# Perspectives

# Perspectives on Using Automated Writing Evaluation Systems to Provide Written Corrective Feedback in the ESL Classroom

*Johanathan Woodworth & Khaled Barkaoui*

*While feedback is widely considered essential for second language (L2) writing development (Bitchener & Ferris, 2012), teachers may not always be able to provide their learners with immediate and frequent corrective feedback. Automated writing evaluation (AWE) systems can help respond to this challenge by providing L2 learners with written corrective feedback (WCF) instantaneously and as frequently as needed both inside and outside the ESL classroom. Proponents of the use of AWE systems argue that these systems can facilitate more writing practice, increase learner motivation and accuracy, and promote learner autonomy. Critics argue that AWE systems cannot give individualized feedback, are prone to errors, can diminish the role of the teacher, and warp students' notions of good writing. As a compromise, it is recommended to use feedback from AWE systems to complement, rather than replace, teacher WCF. In this perspectives paper, we discuss the main benefits and drawbacks of using AWE to provide WCF in the ESL classroom. We conclude by arguing that, when used judiciously and effectively to complement teacher feedback, WCF from AWE systems can support teachers' work and enhance learners' writing motivation and development in the ESL classroom.*

*Alors qu'un consensus existe pour dire que la rétroaction est essentielle pour le développement de la rédaction en langue seconde (L2) (Bitchener & Ferris, 2012), les enseignants ne sont peut-être pas toujours en mesure de fournir une rétroaction corrective immédiate et fréquente à leurs apprenants. Les systèmes d'évaluation automatique de rédaction peuvent aider à répondre à ce défi en fournissant aux apprenants de L2 une rétroaction corrective écrite instantanée et de façon aussi fréquente que nécessaire à la fois en classe et hors de la classe d'ALS. Les défenseurs de l'utilisation des systèmes disent que ces systèmes peuvent encourager la pratique de la rédaction, augmenter la motivation et la précision de l'apprenant et promouvoir l'autonomie de l'apprenant. Les critiques avancent que les systèmes d'évaluation automatique ne peuvent pas donner de rétroaction personnalisée, sont susceptibles de faire des erreurs, peuvent diminuer le rôle de l'enseignant et déformer la perception des étudiants quant à ce qui constitue une bonne rédaction. On recommande, comme compromis, d'utiliser la rétroaction*

*des systèmes d'évaluation automatique comme un complément, plutôt que comme un remplacement de la rétroaction corrective écrite de l'enseignant. Dans cet article donnant des perspectives, nous discutons des principaux avantages et inconvénients de l'utilisation de l'évaluation automatique de rédaction pour fournir de la rétroaction corrective écrite dans la classe d'anglais langue seconde. Nous arrivons à la conclusion que, utilisée de manière judicieuse et efficace pour compléter la rétroaction de l'enseignant, la rétroaction corrective écrite issue des systèmes d'évaluation automatique de rédaction peut soutenir le travail des enseignants et augmenter la motivation et le développement de la rédaction des apprenants dans la classe d'ALS.*

Practice and feedback play a central role in second language (L2) writing development (Barkaoui, 2007). Research suggests that writing frequently can raise learners' awareness about the conventions of L2 texts, help with the automatization of processes such as lexical retrieval, and compensate for the often short time of instruction (Bitchener & Storch, 2016; Lee, 2014). However, frequent writing practice by itself is not sufficient; learners need also to receive feedback on their writing to improve (Barkaoui, 2007; Bitchener & Storch, 2016; Ferris & Roberts, 2001). For example, research has shown that written corrective feedback (WCF) can facilitate learners' uptake and retention of correct forms for writing development (Ellis, 2009).

Both theory and research highlight the important role of WCF in L2 development, in general, and in L2 writing development. Cognitive, sociocognitive, and sociocultural theories all agree about the positive role of WCF in L2 writing development (Lee, 2017). In sociocultural theory, for instance, WCF is an important tool for scaffolding L2 development, particularly if the feedback is immediate and contingent (Bitchener & Storch, 2016; Lavolette et al., 2015). Cognitive theories, such as the cognitive-information processing model, also highlight the important role of WCF for the development of implicit L2 knowledge (e.g., Cotos, 2011; Bitchener & Ferris, 2012; Dekeyser, 1998; Schmidt, 1995, 2001). From a cognitive-information processing model perspective, immediate feedback and metalinguistic feedback can facilitate noticing-with-awareness (Schmidt, 1995), a requisite step for noticing-with-understanding, which is, in turn, vital for learning (Schmidt, 2001). Explicit noticing can occur when WCF provides negative feedback and focuses the learner's attention on the forms and meanings in the input. As such, WCF can facilitate the conversion of declarative knowledge into automatized procedural knowledge (Bitchener & Ferris, 2012). WCF can lead learners to notice contrary evidence in their

　　　　　　　　　　　　JOHANATHAN WOODWORTH & KHALED BARKAOUI

production which can make them more cognitively engaged and more likely to modify their output (Cotos, 2011), which can lead to the conversion of explicit knowledge to implicit knowledge.

As will be explained below, in addition to negative feedback, AWE systems provide explicit metalinguistic explanations of language mistakes in students' writing. Research shows that salient and meaningful feedback is that which draws the learners' attention to the error and explains it (e.g., Jeon & Kaya, 2006). For example, an error code alone may lack salience or communicative value for some learners because these learners may not be able to resolve the code. The metalinguistic explanation can help these learners resolve the errors.

Additionally, several studies have found that L2 learners prefer and request frequent and immediate corrective feedback on multiple drafts (Lee, 2017). However, teachers struggle to do so due to a lack of time and/ or resources (Lee, 2014). Since writing practice and feedback are essential for ESL writing development, there has been a steady increase in the use of automated writing evaluation (AWE) systems to provide L2 learners with WCF on their writing (Stevenson & Phakiti, 2014). Research suggests that an AWE system can alleviate teacher workload (Tang & Rich, 2017), facilitate L2 writing development by supporting learners' reflective use of feedback (Yannakoudakis et al., 2018, p. 252), and promote learner autonomy (El Ebyary & Windeatt, 2010). However, AWE systems are not without their limitations; therefore, this paper argues that AWE systems should be integrated with circumspection to complement teacher feedback in order to mitigate their limitations.

## Overview of Automated Writing Evaluation (AWE)

An automated writing evaluation (AWE) system is a suite of tools within a single interactive program to help facilitate writing development. AWE systems typically combine automatic written corrective feedback (AWCF) to promote noticing language errors, with an automated essay scoring (AES) algorithm that evaluates writing quality, a management system to provide learners with multiple drafting opportunities, and a collection of writing resources such as a dictionary, thesaurus, *Writers' Handbook* and other resources for self-access (Chen & Cheng, 2008; Stevenson & Phakiti, 2014). Typically, AWE systems have a built-in set of topics organized by grade level and writing genres (e.g., descriptive, expository, narrative, persuasive), and has functionality for teachers to add their own topics, add external resources to create integrated writing tasks, create and manage writing portfolios, and other class management tasks (Ranalli et al., 2017).

AWE systems should not be confused with online grammar checkers such as *Grammarly* (www.grammarly.com) and *LanguageTool* (languagetool. org). These grammar checkers miss some critical features of AWE systems.

To elaborate, although some grammar checkers can provide instantaneous feedback and metalinguistic explanations of some grammatical mistakes, they cannot be moderated by the teacher, do not evaluate writing quality, and do not include any portfolio and class management tools. Their focus is on correcting grammatical mistakes rather than writing development. Therefore, they are not discussed in this paper (for research on grammar checkers, see Ghufron, 2019; Nova & Lukmana, 2018; O'Neill & Russell, 2019; Park & Yang, 2020).

While there are currently many different AWE systems, the most widely used are *Criterion* by Educational Testing Service (ETS) (ETS, n.d.; see Attali, 2004, for a description of *Criterion*), *Project Essay Grade* (*PEG*) by Measurement Inc. (Measurement Incorporated, n.d.; see Wilson & Roscoe, 2019, for a description of PEG), *MyAccess!* by Vantage Labs (Vantage Learning, n.d.; see Chen & Cheng, 2008, for a description of MyAccess!), and *Pigai* (Pigai, 2017; see Huang & Renandya, 2020, for a description of Pigai). With all these systems, learners can write and submit as many drafts as they wish or are allowed. When learners submit their written responses, the system analyzes them and provides immediate feedback on specific language and other writing aspects. Different AWE systems provide feedback on different aspects of writing, but all supply information on language use. The analysis tools in *Criterion*, for example, detect errors not only in grammar (e.g., fragments, run-on sentences, subject-verb agreement), but also usage (e.g., wrong article, wrong form of a word, preposition error), and mechanics (missing comma, compound words, spelling). They also highlight discourse elements that may be desirable or problematic in an essay (e.g., introductory material, theses statement, main ideas), give feedback and suggestions on how to improve the text (e.g., reduce repetition of words, inappropriate words or phrases, passive voice), and provide a score on each component along with a holistic score (Burstein et al., 2004). Figure 1 provides an example of holistic and analytic (or trait) scores from *Criterion*.

Figure 1

Screenshot of Criterion® Score and Trait Levels
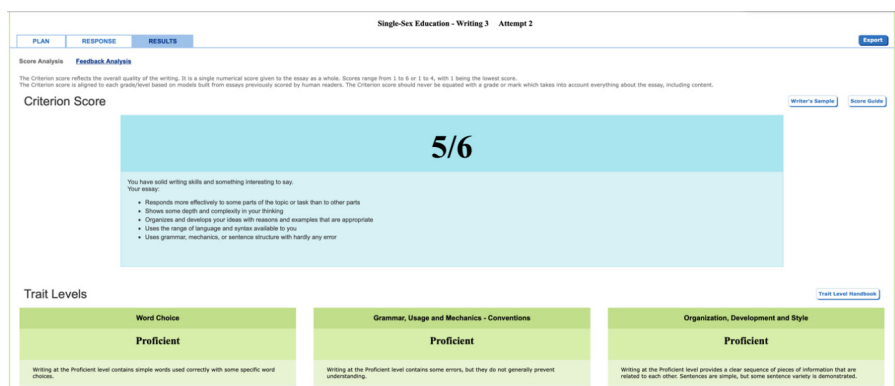


Different AWE systems have different views, but typically, they show the number of different categories of language errors in a submission and highlight them to draw learners' attention to the errors. The learner can get further information about the feedback by clicking on the highlighted error. For example, *Criterion* may highlight a sentence and indicate that the subject and verb of the sentence do not agree. If the student clicks on the highlighted error, *Criterion* provides an explanation such as the following: "Proofread the sentence to make sure the subject agrees with the verb; it appears that the verb *eat* should be in the infinitive form. Consider adding the word *to*." In addition, the feedback is linked to a detailed metalinguistic explanation of the mistake with examples. Using this information from the AWE system, the learner can revise their essay and resubmit it to the system for further feedback. Students can repeat this process as many times as they wish or are permitted. The system saves each draft, and learners and teachers can see detailed reports for each draft creating a portfolio of the learners' work.

## Benefits of Using AWE in the ESL Classroom

Research on the effects of WCF provided by AWE systems in the ESL classroom illustrates three main benefits. Such feedback can alleviate teacher workload, improve learners' L2 development, and promote learner autonomy and motivation to write. The first benefit is that when using AWE systems, teachers may not need to spend as much time correcting and giving feedback on surface errors. In classroom-based studies of AWE feedback researchers suggest that with the integration of AWE in the ESL classroom, teachers can shift the focus of their feedback to other aspects of writing such as content, argumentation, and writing processes and strategies (Wang et al., 2013; Warschauer & Grimes, 2008) or focus more time and energy on

classroom instruction and writing practice (Truscott, 1996). In a study of 10 teachers using AWE in their classes, Tang and Rich (2017) reported that all the participants felt that AWE reduced their marking workload and allowed them to provide more targeted writing instruction.

The second benefit is that feedback from AWE can have a positive effect on learner accuracy. Ranalli et al. (2017), for instance, found that learners using feedback from an AWE system were able to correct errors in their own writing 55% to 65% of the time. While some may find this correction rate insufficient, Ranalli et al., (2017) posited that the correction rate by learners may be acceptable if the goal of WCF is to support grammatical development, since "even inaccurate [WCF] can cause learners to notice linguistic forms, which potentially facilitates acquisition" (p. 22). Training students on how to evaluate and use WCF from AWE systems, for example, may lead them to pay more attention to and review grammatical rules related to the feedback they receive. Research also suggests that this can help students learn to apply these grammatical rules in their own writing independently and to evaluate and adopt only the feedback they deem useful. For example, Zhang (2017) observed that a learner noticed that the AWE system mistook the word "means" for a verb when the student intended to use it as a noun. The learner, recognizing the limitations and fallibility of the AWE system, consulted a dictionary and clarified the usage with her peers. Likewise, Liao (2016) reported that several low-proficiency learners in his study self-accessed the *Writer's Handbook* looking for metalinguistic explanations and other resources provided by the AWE system when they had difficulty revising their texts because they were unable to comprehend the error messages from the AWE system.

A third benefit is that the use of AWE systems can promote learner autonomy, which can help shift the focus in the ESL classroom from teaching to learning, increase learners' motivation to write more, and enhance learners' awareness of writing as a process that involves multiple drafts and revisions (El Ebyary & Windeatt, 2010). Revision is a central process in cognitive-information processing models of writing, such as the one proposed by Hayes (2012). In a study by Chen and Cheng (2008), one of the teacher participants reported that the use of *MyAccess!* facilitated more drafting and revising behaviour among learners. Similarly, El Ebyary and Windeatt (2010) found that AWE encouraged learners to reflect on their writing and plan before writing. They described how some learners exercised agency by adopting only the feedback they deemed useful and deciding if and how the feedback would be incorporated in revisions. Likewise, Zhang and Hyland (2018), in a case study of two L2 learners, found that by continually reminding learners of the subprocesses of planning, drafting, and revising and providing them with continuous feedback, AWE systems led both learners to adopt a process approach to writing and to revise their texts more frequently.

# Issues and Concerns with the Use of AWE

While AWE has gained traction, it has also been criticized for several limitations. The key limitations of AWE systems noted in the literature are that feedback from such systems tends to be more error-prone than teacher feedback, may not be consistent with teacher feedback, may devalue the social dimension of writing, and/or is not adapted to learners' L2 proficiency level, needs or goals. First, AWE systems can miss or miscode some errors, which can result in telling learners that they made a language error when they did not, giving wrong feedback on the correctly identified error, or not identifying some errors. For example, in a recent study that examined the accuracy of error identification and classification by *Criterion*, Lavolette et al. (2015) found that only 75% of the errors that *Criterion* identified were correct; 14% were correctly identified as an error but were misclassified; 11% of the errors identified were for structures that were already correct; and at least 46% of the errors in the corpus were altogether missed (p. 58). Shermis and Burstein (2003) cautioned that the low accuracy of current AWE systems may influence learners' willingness to use the resulting feedback. The second issue is that WCF from AWE systems may not be consistent with, or may even contradict, teacher assessment and feedback. In a study by Chen and Cheng (2008), learners noticed some discrepancies between teacher and AWE scores, which led them to question and distrust AWE scores and feedback.

A third limitation of AWE systems is that they do not attend to the social and communicative dimensions of writing. Herrington and Moran (2001) warned that writing to a machine or an algorithm, as learners do when they write to an AWE system, creates an unnatural writing environment. A position statement of the 2004 Conference on College Composition and Communication, which was reiterated 10 years later in 2014, stated that writing to machines devalues human communication and reduces the validity of the assessment. Lastly, from a sociocultural theory perspective, for WCF to be effective, it needs to be adapted to the learner's zone of proximal development (ZPD), which is defined as the distance between what a learner can accomplish alone and what that learner can achieve with the support of more capable experts, peers, and/or cultural artefacts (Vygotsky, 1987). However, because AWE systems are not adaptive, they supply the same feedback to all learners regardless of their needs, goals, L2 proficiency level, or readiness to engage with the feedback. Likewise, AWE systems currently provide only unfocused comprehensive feedback, which assumes that all errors deserve equal attention. However, some language errors are more serious because they can obstruct communication. In addition, as Ferris (2012) and Bitchener and Knoch (2009) have suggested, it is advisable to provide L2 students with selective feedback that focuses on treatable grammar errors that occur in a rule-governed way, and that can be fixed when learners review and follow the rules governing the errors, such as sentence fragments, subject-

verb agreements, and run-on sentences. Unfortunately, current AWE systems cannot adapt their feedback to students' needs, L2 proficiency level, or goals.

## Integration of Teacher and AWCF in the ESL Classroom

Critics are concerned that the promotion of AWE may lead to the automation of writing instruction (Warschauer & Grimes, 2008; Weigle, 2013), which can reduce teachers' autonomy, independence, and control over their work (Iskander et al., 2010). Due to the criticisms mentioned above, a growing number of scholars (e.g., Chen & Cheng, 2008; Li et al., 2015; Warschauer & Grimes, 2008; Zhang & Hyland, 2018) are advocating a hybrid approach that combines feedback from the teacher and AWE. In this approach, teachers can facilitate the adaptation of feedback from the AWE system to learners' needs and development level, reinforce the social and communicative dimensions of writing, and present the benefits and limitations of AWE to learners in a balanced manner. Additionally, the integration of the two feedback systems may help address some of the limitations of teacher feedback.

Firstly, to be effective and to facilitate L2 development, WCF needs to be appropriate to learners' needs and L2 proficiency level. Previous research on the effectiveness of WCF has examined the distinction between direct and indirect feedback and their effectiveness in raising learners' metacognitive awareness. However, AWE systems provide only two forms of indirect feedback: generic and specific. In generic feedback, the same message appears when a category of error is detected without offering any specific prescription for remedies. For instance, *Criterion* provides the same message when it detects a fragment: "This sentence may be a fragment. Proofread it to be sure that it has at least one independent clause with a complete subject and predicate." Specific feedback, conversely, incorporates some components of the text to give a specific recommendation. For instance, when *Criterion* detects a "confused" word error, it incorporates the original text in its feedback: "You have used 'a' in this sentence. You may need to use 'an' instead" (Ranalli, 2018, p. 3). Although both types of feedback may focus the learner's attention on the error and provide metalinguistic explanations, without further scaffolding, learners may find such feedback frustrating, especially when they cannot understand or resolve the issue (Li et al., 2015; Wang et al., 2012; Warschauer & Ware, 2006). Even though the learners may make revisions, feedback that is noticed but not understood may yield revisions that are not internalized (Storch, 2010). The opportunities afforded by feedback from the AWE system may go unheeded if they are beyond an individual learner's developmental level. Therefore, teachers are advised to mediate AWE feedback by giving specific, individualized, and concrete advice on how to improve writing and scaffolding classroom instruction to fill in gaps in learners' linguistic and metalinguistic knowledge.

Secondly, teachers can complement AWE by considering the social and communicative dimensions of writing. AWE systems evaluate learner essays against proxies of writing traits without considering the potential impact of the text on human audiences in real-world contexts (Cotos, 2010). As a result, Wang and Jiang (2015) found that the learners in their study felt that the AWE system did not meet their desire to have meaningful communications and that they were discouraged because their "expressions were not 'understood' by the machine" (p. 97). The usage of AWE as the sole feedback mechanism can frustrate learners since there is no dialogic process between the writer and the audience. In a study that examined learner perceptions of AWE's usefulness in the classroom, Huang (2014) found that learners preferred the meaningful negotiations facilitated by teacher feedback through classroom interactions, discussions, clarifications, and meaning making, to AWE feedback. Feedback by teachers involves corrective interaction to negotiate meaning within a social context, which makes the input meaningful for learners (Brown, 2000).

Thirdly, teachers can manage learner expectations by making decisions about how and when to use AWE by being fully and critically aware of its limitations and affordances. Zhang (2017) found that some learners expressed doubts about some of the error codes they received from an AWE system, while others were confident about false positives. However, inaccurate WCF from AWE systems may still be useful for learners when they are aware of these limitations and taught how to evaluate and use such feedback in conjunction with teachers' mediation of AWE feedback. WCF from AWE systems can encourage learners to be more cognizant of the writing and revision processes if learners receive training on how to evaluate and use such feedback (Lavolette et al., 2015). Grimes and Warschauer (2010) posited that erroneous feedback is more frequently "maleducative when it is presented as authoritative and when no human expertise is available to override dubious scores and feedback" (p. 31). In cases of inaccurate feedback and error codes from AWE systems, the teacher can help remove self-doubts and help learners move forward with appropriate strategies to double-check errors and the feedback they receive.

Finally, the combination of teacher and AWE feedback can benefit teachers and learners alike because teacher WCF is not without its own problems. Lee (2004) noted that teachers sometimes rely on error codes and direct/indirect feedback with little provision for oral or written metalinguistic feedback. While research has shown conflicting results concerning the effectiveness of direct and indirect feedback (Bitchener & Knoch, 2009; Shintani & Ellis, 2015), metalinguistic feedback, like that provided by AWE, has been shown to help learners develop an awareness of grammatical and linguistic rules (Shintani & Ellis, 2013) because metalinguistic explanations increase the explicitness of WCF (Lee, 2017) and are easier to understand than metalinguistic clues such as error codes alone. In addition, research on teacher WCF suggests that it can be inconsistent, unsystematic, unclear, inaccurate, overly critical, and/

or delayed (Abel et al., 2018; Lee, 2008; Truscott, 1996). In contrast, AWE is very consistent and not affected by factors that can influence or bias human judgment and feedback, such as being tired, halo effects (where human judgment of one aspect of writing is affected by their judgment of other aspects), stereotyping (where one's impressions about a particular group influence their judgment of individuals in that group), and other sources of inconsistency and bias (Wang & Jiang, 2015). Therefore, a combination of teacher WCF and AWE may be better able to meet learners' needs and level of knowledge, revisit and build on prior learning, and be more consistent.

## Concluding Remarks

The effective integration of AWE requires teachers and learners to be cognizant of its advantages and limitations and the various factors that can affect the use of feedback from AWE in the ESL classroom (Liu & Kunnan, 2016). When used judiciously and effectively, AWE can support teachers' work, promote learner autonomy, and enhance learners' motivation and L2 writing development. However, the successful use of AWE calls for a shift in perspectives about technology among teachers, learners, and administrators, and greater administrative support (Warschauer & Grimes, 2008). Because AWE systems are still in their early stages of development, it is essential for teachers to raise learners' awareness of the functional value of AWE and to scaffold and motivate its use. This can be done by clarifying the benefits and limitations of AWE and scaffold and mediate feedback from AWE systems through teacher-learner interactions adapted to individual learners' needs and proficiency levels so that even underachieving learners can comprehend and use AWE feedback. The potential for AWE's use in (and outside) the ESL classroom may be profound as technology-enhanced assessment and feedback can contribute positively to transforming traditional teaching, learning, and assessment.

However, this is unlikely to happen without a deep and critical understanding of AWE and its implications for how we conceive of and facilitate learning, teaching, and assessment. As Hyland (2003) has asserted, "everything we do in the classroom, the methods, and materials we adopt, the teaching styles we assume, the tasks we design, are guided by both practical and theoretical knowledge, and our decisions can be more effective if that knowledge is explicit" (p. 1). The integration of AWE can only be a powerful catalyst for teaching and learning if supported by theory and empirical evidence. Because the effectiveness of all feedback practices is contingent, future research should provide more empirical evidence on how feedback from AWE systems affects learners' writing process so that teachers can better understand and improve the integration of AWE in their specific contexts to support the writing development of their learners. Teachers can, for example, conduct action research studies to document and examine the

use and effectiveness of feedback from AWE systems for learners in their own teaching contexts. For instance, teachers can examine the efficacy of indirect and specific WCF from AWE systems on the L2 writing development of learners with different proficiency levels in order to reveal more nuanced information about different ways to implement AWE for a targeted group of learners. Findings from such action research can inform instruction and shed more light on the effects of AWE feedback in specific contexts.

## Acknowledgement

## The Authors

Johanathan Woodworth is a doctoral candidate at the Faculty of Education, York University, Canada. Within the broader scope of second-language (L2) acquisition, his research interests include technology as a mediating tool, gamified learning, and corrective feedback for L2 development.

Khaled Barkaoui is an associate professor at the Faculty of Education, York University, Canada. His current research and teaching focus on second-language (L2) assessment, L2 writing, L2 program evaluation, longitudinal and mixed methods research, and English for Academic Purposes (EAP).

## References

Abel, S., Kitto, K., Knight, S., & Shum, S. B. (2018, January). Designing personalised, automated feedback to develop students' research writing skills. In *ASCILITE 2018-Conference Proceedings-35th International Conference of Innovation, Practice and Research in the use of Educational Technologies in Tertiary Education: Open Oceans: Learning Without Borders.*

Attali, Y. (2004). Exploring the feedback and revision features of Criterion. *Journal of Second Language Writing, 14*, 191–205.

Barkaoui, K. (2007). Teaching writing to second language learners: Insights from theory and research. *TESL reporter, 40*(1), 35–48.

Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing.* Routledge.

Bitchener, J., & Knoch, U. (2009). The relative effectiveness of different types of direct written corrective feedback. *System, 37*, 322–329.

Bitchener, J., & Storch, N. (2016). *Written Corrective Feedback for L2 Development.* Multilingual Matters Limited.

Brown, H. D. (2000). *Principles of Language Learning and Teaching* (Vol. 4). New York: Longman.

Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The Criterion online writing service. *AI Magazine, 25*(3), 27–36.

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.

Conference on College Composition and Communication (CCCC). (2004). *A position statement on teaching, learning, and assessing writing in digital environments*. Retrieved 2020-08-04 from https://cccc.ncte.org/cccc/resources/positions/digitalenvironments

Conference on College Composition and Communication (CCCC). (2014). *A position statement on*

*writing assessment*. Retrieved 2020-08-04 from https://cccc.ncte.org/cccc/resources/positions/writingassessment

Cotos, E. (2010). Automated writing evaluation for non-native speaker English academic writing: The case of IADE and its formative feedback.

Cotos, E. (2011). Potential of automated writing evaluation feedback. *Calico Journal, 28*(2), 420–459.

Dekeyser, R. M. (1998). Beyond focus on form: Cognitive perspectives on learning and practising second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on Form in Classroom Second Language Acquisition* (pp. 42–63). Cambridge University Press.

Educational Testing Service (ETS). (n.d.). *Criterion® online writing evaluation service.* http://www.ets.org/Criterion

El Ebyary, K., & Windeatt, S. (2010). The impact of computer-based feedback on students' written work. *International Journal of English Studies, 10*(2), 121–142.

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal, 63*, 97–107.

Ferris, D. R. (2012). Written corrective feedback in second language acquisition and writing studies. *Written Corrective Feedback in Second Language Acquisition and Writing, 45*, 1–220.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be. *Journal of second language writing, 10*(3), 161–184.

Ghufron, M. A. (2019, April). Exploring an automated feedback program 'grammarly' and teacher corrective feedback in EFL writing assessment: Modern vs. traditional assessment. In *Proceedings of the 3rd English Language and Literature International Conference.*

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).

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*, 369–388.

Herrington, A., & Moran, C. (2001). What happens when machines read our students' writing. *College English, 63*(4), 480–499.

Huang, S. J. (2014). Automated versus Human Scoring: A Case Study in an EFL Context. *Electronic Journal of Foreign Language Teaching, 11*, 149–164.

Huang, S., & Renandya, W. A. (2020). Exploring the integration of automated feedback among lower-proficiency EFL learners. *Innovation in Language Learning and Teaching, 14*, 15–26.

Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of second language writing,* 12(1), 17–29.

Iskander, M., Kapila, V., & Karim, M. A. (2010). *Technological developments in education and automation.* Springer Science & Business Media.

Jeon, E. H., & Kaya, T. (2006). Effects of L2 instruction on interlanguage pragmatic development. *Synthesizing research on language learning and teaching,* 165–211.

Lavolette, E., Polio, C., & Kahng, J. (2015). The accuracy of computer-assisted feedback and students' responses to it. *Language, Learning & Technology, 19*(2), 50–68.

Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing, 13*(4), 285–312.

Lee, I. (2008). Student reactions to teacher feedback in two Hong Kong secondary classrooms. *Journal of Second Language Writing, 17*(3), 144–164.

Lee, I. (2014). Revisiting teacher feedback in EFL writing from sociocultural perspectives. *TESOL Quarterly, 48*, 201–213.

Lee, I. (2017). *Classroom Writing Assessment and Feedback in L2 School Contexts.* Springer.

Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing, 27*, 1–18.

Liao, H. C. (2016). Enhancing the grammatical accuracy of EFL writing by using an AWE-assisted process approach. *System, 62*, 77–92.

Liu, S., & Kunnan, A. J. (2016). Investigating the Application of Automated Writing Evaluation to Chinese Undergraduate English Majors: A Case Study of" WriteToLearn". *calico journal, 33*(1), 71–91.

Measurement Incorporated (n.d.). *Automated essay scoring—Project Essay Grade (PEG®)*. https://www.measurementinc.com/products-services/automated-essay-scoring

Nova, M., & Lukmana, I. (2018, July). The detected and undetected errors in automated writing evaluation program's result. In *English Language and Literature International Conference (ELLiC) Proceedings* (Vol. 2, pp. 120–126).

O'Neill, R., & Russell, A. (2019). Grammarly: Help or hindrance? Academic learning advisors' perceptions of an online grammar checker. *Journal of Academic Language and Learning, 13*(1), A88–A107.

Park, J. H., & Yang, I. Y. (2020). Utilizing an AI-Based Grammar Checker in an EFL Writing Classroom. 응용언어학, *36*(1), 97–120.

Pigai. (2017). *Pigai: Write more, write better.* http://en.pigai.org/

Ranalli, J. (2018). Automated written corrective feedback: how well can students make use of it. *Computer Assisted Language Learning, 31*(7), 653–674.

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.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. *Attention and awareness in foreign language learning, 9*, 1–63.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge University Press.

Shermis, M. D., & Burstein, J. (2003). *Automated essay scoring: A cross-disciplinary perspective.* Mahwah.

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22*(3), 286–306.

Shintani, N., & Ellis, R. (2015). Does language analytical ability mediate the effect of written feedback on grammatical accuracy in second language writing? *System, 49*, 110–119.

Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing, 19*, 51–65.

Storch, N. (2010). Critical feedback on written corrective feedback research. *International Journal of English Studies, 10*, 29–46.

Tang, J., & Rich, C. S. (2017). Automated writing evaluation in an EFL setting: Lessons from China. *JALT CALL Journal, 13*(2), 117–146.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language learning, 46*(2), 327–369.

Vantage Learning. (n.d.). *My Access!—Virtual writing classroom.* http://www.vantagelearning.com/products/my-access-school-edition/

Vygotsky, L. S. (1987). *The collected works of LS Vygotsky: Vol. 1, Problems of general psychology (RW Rieber & AS Carton, Eds., N. Minick, trans.).* New York: Plenum Press.

Wang, H., Choi, I., Schmidgall, J., & Bachman, L. F. (2012). Review of Pearson Test of English Academic: Building an assessment use argument. *Language Testing, 29*, 603–619.

Wang, T., & Jiang, L. (2015). Studies on written corrective feedback: Theoretical perspectives, empirical evidence, and future directions. *English Language Teaching, 8*(1), 110–120.

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.

Warschauer, M., & Grimes, D. (2008). Automated writing assessment in the classroom. *Pedagogies: An International Journal, 3*(1), 22–36.

Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language teaching research, 10*(2), 157–180.

Weigle, S. C. (2013). English language learners and automated scoring of essays: Critical considerations. *Assessing Writing, 18*, 85–99.

Wilson, J., & Roscoe, R. D. (2019). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research, 58*(1), 87–125.

Yannakoudakis, H., Andersen, Ø. E., Geranpayeh, A., Briscoe, T., & Nicholls, D. (2018). Developing an automated writing placement system for ESL learners. *Applied Measurement in Education, 31*, 251–267.

Zhang, Z. (2017). Student engagement with computer-generated feedback: a case study. *Elt Journal, 71*(3), 317–328.

Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing, 36*, 90–102.