English Teaching, Vol. 72, No. 3, Autumn 2017
DOI: 10.15858/engtea.72.3.201709.91

Effective Applications of Automated Writing Feedback in Process-based Writing Instruction

Won-Young Koh
(Ewha Womans University)

Koh, Won-Young. (2017). Effective applications of automated writing feedback in process-based writing instruction. *English Teaching, 72*(3), 91-118.

This study investigated the appropriate ways in offering immediate automated writing feedback within the framework of process-based writing pedagogy by comparing relative effects of two different automated writing evaluation (AWE) system application types on improving writing performance. The experiment took an initiative step in elucidating *at what point* in the process-based writing stages AWE feedback is best to be served. The research is conducted to confirm whether providing an instant language-related feedback whenever EFL students call for will either interfere the development of content indeed – as it has been expected by process-based writing approach, or bring improvement in students' writing. Two application types, namely non-continuous feedback (NCF) and continuous feedback (CF) group - are differentiated in terms of in which point of the writing stages students are enabled to get access to the AWE system. With the purpose of the study, a total of 20 students participated. The findings revealed that CF group did not receive language-related AWE feedback to the point of distracting the development of their content. Furthermore, CF group significantly outperformed NCF group in overall writing product, especially on the dimension of grammar and content. Students also expressed a positive attitude toward receiving instant language-related feedback via AWE system.

Key words: process-based writing approach, English writing feedback, automated writing evaluation system, immediate language-related feedback

1. INTRODUCTION

The significance of feedback has not been underestimated in both L1 and L2 composition classrooms ever since the advent of process-oriented approach in writing instruction (Barlett, 1982; Bitchener & Ferris, 2012; Witte & Arndt, 1991). According to process-based writing pedagogies, instructors' main role is to guide students through the writing process in generating

and refining ideas (Hyland, 2003). This is essentially based on the notion that writers' message is more pivotal than the grammar of the written piece. While a large number of researchers agreed the importance of content in delivering the intended meaning, it was not without criticisms about the issue of underestimating accuracy (Celce-Murcia, 1991a; Ferris, 2002). In written discourse, rather a high level of accuracy is required for effective communication to occur (Little, 1994) and errors are much less tolerated compared to oral conversation (Johnson & Roen, 1989). Although writing is not merely about producing grammatically correct sentences, accuracy cannot be sacrificed for the sake of fluency (Byrd & Reid, 1998).

The apparent importance of accuracy in L2 and foreign language (FL) writing promoted writing researchers to consider providing feedback on language-related issues to a reasonable degree throughout the writing process. As amply manifested in studies of second language writing, feedback on language features has been desperately called for EFL learners (Badger & White, 2000; Bitchener & Ferris, 2012). Students have been struggling to use appropriate surface-level morphological and syntactic features due to their short linguistic intuitions and resources (Bitchener & Ferris, 2012; Kim & Kim, 2005; Shin, 2008). Even for advanced EFL writers such as master's or PhD students who have readily prepared themselves with ideas to convey lacks English competence which at times interfere the flow of writing. Thus, content isn't their primary concern while they strive to use pertinent linguistic features in conveying their intended meaning clearly to readers (Frodesen, 2001). Difficulties in putting down appropriate language features to deliver the message brings attention to what language-related feedback can do to support EFL learners in writing meaningful sentences throughout the writing process. When it comes to feedback on grammar thus, the question is not *whether* to provide it but *how best* to provide it.

In responding to how best to provide feedback on grammar, a number of research addressed the issue of when and in what certain point writing teachers should attend to errors in the writing process (Bitchener & Ferris, 2012). Since the process approach, at its core, emphasize developing writers' content and arguments, form-focused written corrective feedback has been preferred to be reserved for the final stage of the process (Bitchener & Ferris, 2012). The premature feedback on accuracy is discouraged to support learners to fully engage in further revising and expanding of their ideas on writing (Bitchener & Ferris, 2012; Zamel, 1982, 1985). Although process-based writing pedagogy encouraged instructors to provide feedback primarily on content in preliminary drafts, the suggested *Content-Before-Form* feedback (Bitchener & Ferris, 2012) is not without doubts. Since students' writing is not entirely representative of the problems they have *while* they are writing - especially in using appropriate language features (Frankenberg-Garcia, 1999), feedback received *after* students had finished writing their drafts seems to have little influence on improving learners' writing (Leki, 1990; Peterson, 2003, 2010). Feedback as it is rarely provided within the process of making appropriate language choices (Frankenberg-Garcia, 1999), language-related feedback rather needs to be addressed while students are composing their drafts. Thus, research suggests offering EFL writers with

immediate support on language-related features whenever it is called for, regardless of the writing stages (Li, Link, & Hegelhemier, 2015; Hartshorn, Evans, Merrill, Sudeweeks, Strong-Krause, & Anderson, 2010). Forasmuch as EFL writers lack a sense of correctness or intuitions of the grammatical knowledge in conveying their ideas clearly, immediate feedback on those features can help students not be exhausted in search of the appropriate language. Immediate feedback by here means treatment of error that comes right after a student has responded to a task (Shute, 2008).

For reasons of practicality, however, it is almost infeasible for instructors to support students with immediate feedback throughout the entire writing process (Calfee & Miller, 2007; Ferris & Hedgcock, 2005; Maeng, 2010). This is where automated writing evaluation (AWE) system stands out as a breakthrough. AWE, established upon artificial intelligence integrated with natural language processing, not only evaluate and score written products but also provides immediate feedback through real time database (Shermis & Burstein, 2003; Warschauer & Ware, 2006). Initial AES systems were developed and used within the confines of assessment setting to grade students' essays (Chen & Cheng, 2008; Page, 2003). The confirmed high reliability of AWE system in testing situation (Attali, 2013; Attali & Burstein, 2006; Elliot, 2003; Shermis & Hamner, 2013; Ware & Warschauer, 2006) motivated researchers to seek into the educational value of the program (Chen & Cheng, 2008; Grimes & Warschauer, 2010; Li, Link, Ma, Yang & Hegelheimer, 2014).

Studies on AWE system thus further expanded its scope by conducting research that verifies the usefulness of applying AWE system in classroom setting. The research result supported the efficacy of the system by affirmatively providing students with appropriate language-related feedback along with more chances to revise and practice writing (Attali, 2004; Choi & Lee, 2010; Kellogg, Whiteford, & Quinlan, 2010; Li et al., 2015; Moon & Pae, 2011). Instructor can also concentrate on providing individualized comments on content/ organization rather than spending great deal amount of time responding to grammar errors (Ferris, 1999). Nevertheless, due to the technical defect, AWE system is computationally adept at providing feedback mostly on surface-level language-related features rather than on higher level concerns (Weigle, 2013). In short, AWE system has limitations in suggesting appropriate feedback on content, organization or style of the written product. Therefore, immediate AWE feedback is strongly encouraged to be applied as a complement to instructor feedback which is expected to have a potential in efficiently improving EFL learners' development ideas as well as accuracy (Chen & Cheng, 2008; Kellogg, et al., 2010; Li, et al., 2015; Shermis & Burstein, 2003; Warshauer & Ware, 2006).

Embracing automatic writing evaluation system needs more careful consideration to sufficiently capitalize on its immediate feedback. Although a handful of research verified the effectiveness of AWE feedback in comparison to the condition in which made no use of the system (Li et al., 2014; Li et al., 2015; Link et al., 2014), research stopped short of exploring *at*

what point in the process-based writing stages AWE feedback should be provided. Due to the limitation of AWE system in providing feedback mostly on surface level language-related issues, a mere implementation of AWE system – without consideration on in which writing stage it should be provided – has a risk of students paying too much attention to language features. If so, AWE system might as well not be applied since it runs counter to what process-based writing pedagogy suggest to not overly concentrate on language features. Therefore, more systematic research on pedagogically sound applications of automated writing feedback should be done to properly argue the effectiveness of embracing AWE system within the framework of process-based writing pedagogy (Ranalli, Link, & Chukharev-Hudilainen, 2017).

To highlight the effective AWE utilization practice that maximizes its pedagogical value (Link et al., 2014), studies should be conducted to examine in which stage of the writing process AWE is best to be applied. Whether it is better to be implemented throughout the overall writing stages or be applied in certain point of the writing stages remains inconclusive. Thus, the effects of two different AWE application types - differed in access period to AWE system, was compared and examined in response to *how much* language-related feedback is beneficial and *at what point* in the process of writing should language-related feedback be served (Bitchener & Ferris, 2012; Byrd & Reid, 1998). Therefore, the present study aims to advance written corrective feedback and second language writing research by identifying how to effectively implement AWE feedback throughout the composing stages to support EFL learners in developing their writing proficiency. To achieve the research goal, the study was conducted by comparing the effects of two different AWE application types on Korean EFL writing performance and perceptions on AWE as well as immediate feedback. Two application types are distinguishable in terms of in which point of the writing stages students are enabled to get access to the system.

Putting it all together, this study aims to identify effective AWE utilization practices within process-based writing stages. The current research is to confirm whether providing a simple instant language-related feedback whenever students call for will either interfere the development of content indeed - as it has been expected by current process-based writing research, or bring positive influence on the writing proficiency. The two treatment groups - namely continuous feedback (CF) and non-continuous feedback (NCF) group - differed in the restriction in getting access to AWE system, and they were developed based on previous theoretical and empirical research on written corrective feedback. As for the NCF group, they have access to AWE only once before they submit each draft to instructor. This is established upon the current paradigm of process-oriented writing pedagogy that promotes limited attention to be paid on language-related issues and encourages error correction to be deferred. On the other hand, the CF group utilizes AWE whenever they want to without restrictions in the frequency and access period. CF group is formed based upon the previous empirical research which revealed L2 and EFL learners in need of sufficient amount of grammar support from the early stages of writing (Hajian, Farahani, & Shiraizi, 2014).

In search for the most effective and appropriate ways to provide instant feedback in the process of EFL students' execution of writing, two central research questions were established. The first research question concerns the effects of different frequency applications of automated writing feedback in the writing process on improving Korean EFL writing performance. Going further, the second research question is posed to question students' opinion on AWE feedback and whether it has brought changes to students' perceptions on immediate feedback. The followings are the two research questions:

1) How do the different application types of automated writing feedback affect EFL learners' writing performance?
   (1) What is the continuous feedback groups' frequency pattern in receiving automated writing feedback during the process of writing?
   (2) What are the differences between the continuous feedback and non-continuous feedback group in the overall writing performance?
2) How do the different application types of automated writing feedback influence EFL learners' perceptions on the value of usefulness of immediate AWE feedback?
   (1) In what way do the different application types of automated writing feedback influence EFL learners' perceived effectiveness on AWE feedback?
   (2) In what way do the different application types of automated writing feedback influence EFL learners' perceived effectiveness on receiving immediate feedback?

2. REVIEW OF PREVIOUS STUDIES

2.1. Perspectives on Written Corrective Feedback on Accuracy

There have been various issues regarding written corrective feedback, and one of the biggest questions dealt among researchers was "whether or not error feedback helps students to improve in written accuracy over time" (Ferris, 2004, p. 56). In process-oriented writing pedagogy, error feedback was reserved for the final stage of the writing process based on the belief that early excessive attention to error corrections may short-circuit writers' cognitive ability in developing and expanding their content (Ashwell, 2000; Bitchener & Ferris, 2012; Zamel, 1982, 1985). Especially Truscott persistently have argued over the inefficacy of error feedback in improving students' writing ability (Truscott, 1996, 1999, 2004, 2007, 2009; Truscott & Hsu, 2008).

Despite the importance of pertaining well-developed content in writing, de-emphasizing written corrective feedback particularly on accuracy of form and structures has not been without criticisms since a reasonable degree of accuracy is required to convey writers' intended meaning (Celce-Murcia, 1991b; Ferris, 2002). That is, rather a high level of correctness and

accuracy is required for effective communication in written discourse to occur (Byrd & Reid, 1998; Little, 1994). Thus, studies moved onto examining the effects of written corrective feedback in improving accuracy of students' drafts overtime and it was amply manifested by research done in process-based composition class (Ashwell, 2000; Fathman & Whalley, 1990; Ferris, 2006; Ferris & Roberts, 2001).

Based on the studies that confirmed the positive effect of written corrective feedback on improving students' accuracy, researchers and instructors directed their attention to exploring how and when to provide written corrective feedback, rather than concerning about whether or not to provide it. In terms of how to provide feedback, researchers have consistently insisted to delay language-related feedback until the final stage of students' writing. Studies note that learners' excessive attention on accuracy is believed to short-circuit students' cognitive capacity in processing content-related issues (Bitchener & Ferris, 2012), which in result obstructs improvement in writing (Zamel, 1982, 1985). Hence, the *Content-Before-Form* feedback (Bitchener & Ferris, 2012) has been preferred.

Yet the *Content-Before-Form* pattern is disputable when concerning the inherent differences between L1 and L2 writers. That is, L2 students should be provided with pedagogical suggestions that are distinct from their L1 peers (Bitchener & Ferris, 2012; Ferris, 2002; Horowitz, 1986; Leki, 1990; Zhang, 1995). What is more, contradicting results on the perhaps overstated concerns over L2 writers' restricted ability in cognitively processing both types of feedback assure researchers of learners' capability to attend to both content and form feedback simultaneously (Ashwell, 2000; Bitchener & Ferris, 2012; Fathman & Whalley, 1990; Ferris, 1997). Thus, several studies argued focus on form to take place at varied stages of composing process to capitalize its support in developing their ideas, rather than reserving it for later stage (Badger & White, 2000; Frodesen & Hoten, 2003). This ongoing discussion on when and how to provide written corrective feedback on language features in the writing process thus call for further studies to get a conclusive evidence, which present study embarks upon.

In conferring about how best to integrate feedback on grammar into the overall writing process, empirical studies revealed students' strong belief in the value of written corrective feedback (Enginarlar, 1993; Ferris, 2006; Ferris & Roberts, 2001; Ferris, Brown, Liu & Stine, 2011; Leki, 1991). Learners appealed for feedback on accuracy since they have been subject to struggles with language features throughout the writing process. Studies conducted exclusively on EFL contexts especially denotes the struggles EFL writers face in appropriately using surface-level morphological and syntactic features, due to their lack of access to linguistic resources and intuition of grammatical rules (Bitchener & Ferris, 2012; Kim & Kim, 2005; Shin, 2008). Especially EFL writers who have readily manifested with ideas to convey, noted grammar rather than content to be their primary concern in putting down words appropriately and meaningfully in the intended way (Frodesen, 2001).

To provide written corrective feedback that can improve both accuracy and content of the

writing, research further argued learners in need of 'real-time' feedback on the questions that emerge at the moment they are struggling to put their ideas down on paper (Frankenberg-Garcia, 1999). Since considerable depth of cognitive engagement of mental formulations and revisions of texts occur mostly in the drafting stage, before shared with other readers or instructors (Leki, 1990; Peterson, 2010; Schoonen, Snellings, Stevenson & Van Gelderen, 2009), instant feedback has been called for. Thus, immediate written corrective feedback provided during the process of writing is expected to bring some noticeable effect in improving learners' overall writing quality. Unfortunately, only a dearth of study examined whether immediate feedback on language features provided in drafting stages enhance learner's accuracy of writing and whether real-time feedback throughout the writing process truly disrupts the development of content as it has been suggested by current process-writing pedagogy.

2.2. Prior Research on Applying AWE Feedback in Process-oriented Writing Approach

As the need for immediate written corrective feedback gradually stands out, studies went in search for finding appropriate ways to provide instant feedback. In response, computer-generated feedback emerged as a feasible way to provide real-time comments. Rapid development of computer technology enabled implementation of computer-generated feedback in evaluating written products and providing formative feedback, which is referred to as Automated Essay Scoring (AES) (Shermis & Burstein, 2003) or Automated Writing Evaluation (AWE) (Warschauer & Ware, 2006) system. AWE system has been approved to not only produce reliable and valid score (Chen & Cheng, 2008; Grimes & Warschauer, 2010; Li et al., 2014) but more importantly to provide relevant and appropriate written corrective feedback on accuracy (Attali, 2004; Choi & Lee, 2010; Kellogg et al., 2010; Li et al. 2015; Moon & Pae, 2010). AWE has gained its attention by being timely and constant in providing feedback on language-related features immediately after students' submission of writing drafts (Li et al., 2015).

Several attempts have been made to employ AWE feedback in L2/FLwriting instruction class and to reveal its effectiveness on improving learners' overall writing performance. In result, research revealed the positive influence of AWE on overall writing quality, particularly in improving students' accuracy of the essay. That is, AWE feedback has been found to have a potential in improving students' accuracy of writing by providing sufficient amount of feedback on language-related issues (Attali, 2004, Choi & Lee, 2010; Kellogg et al., 2010; Li et al., 2015; Moon & Pae, 2011). To further confirm this newly developed technology in terms of pedagogical perspective, studies observing how instructors and learners perceive the AWE system were progressed. Based on their responses, the program has been revealed to be beneficial in classroom settings, in terms of freeing up time and energy for instructors to highly engage in providing individualized feedback on content/organization and motivating learners to

further revise their writing (Chen & Cheng, 2008; Li et al., 2014; Li et al., 2015; Link et al., 2014; Moon & Pae, 2011; Warschauer & Grimes, 2008). Moreover, it has been supported by a number of writing instructors to be a manageable device (Grimes & Warschauer, 2010; Link et al., 2014). To compensate the limitation of AWE feedback in providing feedback somewhat insufficiently on content and organization, research strongly argued to capitalize merits of AWE system by employing it along with instructor's individualized feedback (Li et al., 2015).

Despite the value of AWE feedback, a dearth of study made good use of the system by applying it to deal with the unresolved question of *how* and *at what point* in the writing process should feedback on accuracy be provided. To date, Kellogg et al. (2010) is the only study that initiated to respond to how to provide feedback on language features by controlling the amounts of AWE feedback. The research aimed to investigate the influence of varied amounts of AES feedback on L1 college students' holistic scores, essay length, total errors, and error scores. In an attempt to differ the amount of feedback, participants were randomly assigned to three different groups comparable in quantities of *Criterion*® feedback they receive: no feedback, intermittent, and continuous feedback group. Participants wrote three essays and were required to complete two drafts for each essay grounded on process-based writing approach. While intermittent group received feedback after they completed first draft of the second practice essay, continuous group received AWE feedback in all the first draft of three assigned essays. The study revealed continuous feedback group producing significantly less errors in mechanics, usage, and grammar from first to second draft for each practice essay. Furthermore, students in continuous feedback group even had gains in 'style' by avoiding excessive word repetition (Burstein & Wolska, 2003). Thus, the research is particularly noteworthy in revealing positive effects of frequent immediate feedback on language features in improving not only accuracy of writing but also fluency in some degree. While Kellogg et al. (2010) shed some light on how to provide feedback, the unrefined research design lacks theoretical foundation on the reason behind the construction of the established experimental groups.

To advance research on how to appropriately provide instant language-related feedback within the framework of process-based writing pedagogy, the current study aims to examine a rather unexplored issue of when and how much to provide AWE feedback within the process-based writing instruction. The research is expected to reveal whether providing instant AWE feedback will indeed disrupt the development of the content, as it has been expected by process-based writing approach, or have somewhat positive influence in improving learners' writing. The research is conducted by investigating the effectiveness of two different AWE application types on improving EFL learning' writing performance and EFL students' perceptions on instant AWE feedback. To accomplish the aim of the study, two groups varied in the access period the AWE system.

3. RESEARCH METHOD

3.1. Participants

A total of 21 EFL students were recruited from several private universities in Republic of Korea. The participants consisted of 18 female and 3 male students of mixed majors, and their age ranged from 21 to 26. One student's data was excluded because of incompletion of writing assignments. Thus, the data from 20 students were analyzed for the study. With regard to participants' English learning background, none of them had an experience of living or studying abroad in English-dominant countries for more than a year. Also, all participants had at least one practice to write argumentative essay. Each continuous feedback (CF) and non-continuous feedback (NCF) group consisted of 10 students. Students were initially required to take pre-test within 30 minutes which is to complete argumentative writing in response to a question that asks of one's opinion about a topic. Students were randomly assigned to either NCF and CF group based on the pre-test score. Pre-test were rated based on five components - content, organization, grammar, vocabulary, and mechanics - with each on a scale of one to five, making a total of 25 points. To establish the comparability between the two groups, the homogeneity of the two treatment groups' writing performance was confirmed by conducting independent-samples *t*-test in which significance level set at .05 (Table 1).

TABLE 1

Homogeneity of Non-Continuous Feedback and Continuous Feedback Group

Rating dimension	Group	*M*	*SD*	*t*	*p*
Content	NCF	3.65	.67	.23	.82
	CF	3.65	.47		
Organization	NCF	3.75	.44	-.47	.65
	CF	3.85	.63		
Grammar	NCF	3.45	.46	-.23	.82
	CF	3.50	2.51		
Vocabulary	NCF	3.30	.53	.19	.85
	CF	3.25	.34		
Mechanics	NCF	3.50	.53	.23	.82
	CF	3.45	.54		
Total	NCF	17.70	.60	.00	1.00
	CF	17.70	1.97		

3.2. Instruments

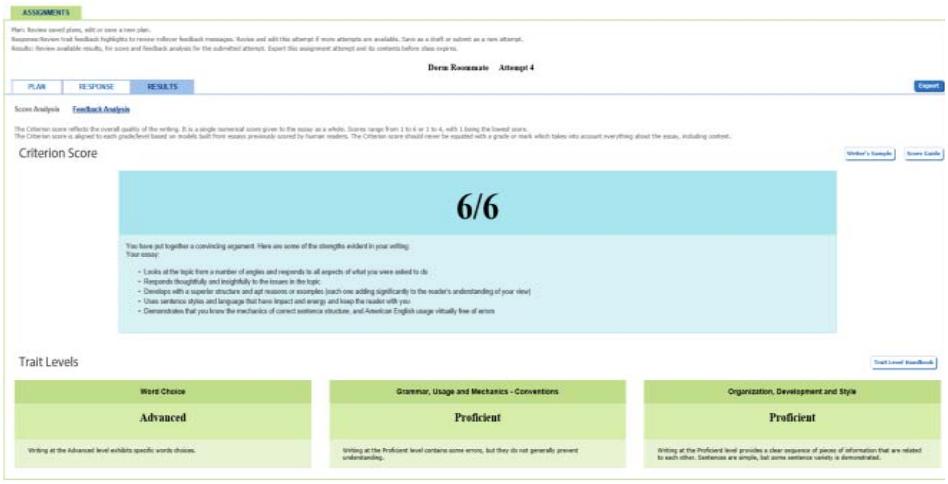
The materials for the study included automated writing evaluation system - *Criterion*®, pre- and post-test, and pre- and post-questionnaires. To provide participants with immediate AWE

feedback, *Criterion*® was utilized for the present study. The *Criterion*® Online Writing Evaluation Service, developed by Educational Testing Service (ETS), is a web-based instructor-led writing tool that helps students plan, write and revise their essays. It gives writers with score reports and instant diagnostic feedback based on their rating program - *e-rater*®, functioning based on natural language processing (Burstein, Chodorow, & Leahcock, 2003, 2004). This particular program has been chosen based on previous studies which revealed high validity and reliability of the system in providing appropriate language-related feedback (Attali, 2013; Attali & Burstein, 2006; Elliot, 2003; Shermis & Hamner, 2013; Ware & Warschauer, 2006).

When students submit their writing, the total criterion score along with three trait levels (word choice; conventions; organization, development and style) are initially reported. The criterion score reflects the overall writing quality given on a six-point scale. The three-trait level indicates whether your writing skills for each trait are either in the Developing, Proficient, or Advanced level (Figure 1).

FIGURE 1

Examples of *Criterion*® in Providing Criterion Score and Trait Levels

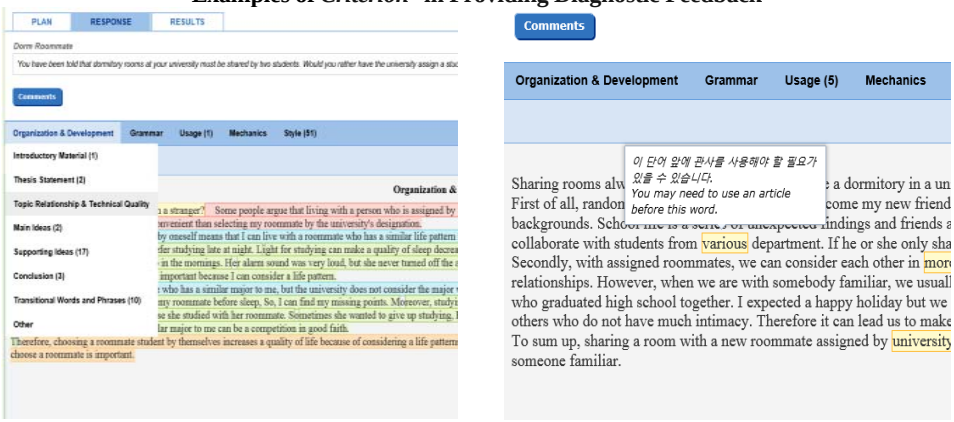


Additionally, when students move back to the drafting page, a detailed diagnostic feedback is given across five different categories: grammar, usage, mechanics, style, and organization and development. Each dimension consists of detailed scoring components, covering in total of forty-five different subcategories. When students tab on the indicated subcategories, the system highlights errors in question within the draft. While feedback on surface-level language-related issues are delivered in somewhat direct manner ("*You may need to remove this article.*"), other areas are indicated with brief and rather indirect suggestion or explanation for correction (Zhang, 2016), which in result relatively lacks detailed instructions on what to and what not to do ("*This*

verb may be incorrect. Proofread the sentence to make sure you have used the correct for of the verb.") (Figure 2).

FIGURE 2

Examples of *Criterion*® in Providing Diagnostic Feedback



The materials for the study included pre-and post-argumentative writing test to examine the relative effect of different applications of automated writing feedback on improving EFL learners' writing performance. Writing topics for both pre- and post-test were selected from topics equipped in the *Criterion*® service for accurate grading and solid feedback. The level of difficulty of the selected writing topics was controlled based on the informed standard level embedded in the program (Kellogg et al., 2010). Argumentative essay was selected since it is recognized as one of the necessary and fundamental text types that college-level students should practice, regardless of the students' background (Bridgeman & Carlson, 1984). Participants were given 30 minutes to complete each test (see Table 2).

TABLE 2

Pre- and Post-Test Writing Topics

Test	Topic
Pre-test	Do you agree or disagree with the following statement? People always learn from their mistakes. Use specific reasons and details to support your answer.
Post-test	"Children waste far too much time playing games when they could be involved in more constructive activities" – E.Gorkin. Do you agree or disagree with the opinion stated above? Support your position with reasons and examples from your own experiences, observations or reading.

To analyze students' perceptions on the value of immediate AWE feedback on language-related issues, pre- and post- questionnaire survey was provided to both NCF and CF group. Questionnaire items were adapted and modified from Chen and Cheng (2008) and Li et al.

(2015) studies and are composed of six-point Likert scale questions and open-ended questions. As for six-point Likert-scale items, one referred to “strongly disagree”, two for “very disagree”, three for “somewhat disagree”, four for “somewhat agree”, five for “very agree” and six referred to “strongly agree”.

3.3. Data Collection Procedure

The data collection procedure began with students’ completion of pre-test essay and pre-treatment questionnaire, along with teacher’s introduction on the entire treatment and instruction on AWE system. To help students to become familiar in using AWE, detailed instructions on how to utilize the system were given. Then students attended five sixty-minute sessions of process-based writing instructions for four weeks on how to write an argumentative essay. During the experiment procedure, students were required to write two different writing tasks. Both tasks had the identical text type – argumentative essay – as the pre- and post-test to exclude any influence of genre. Students were required to engage in process writing by completing three drafts for each task (first, second, and final draft). Each draft was submitted to the instructor via e-mail and instructor sent back drafts to students with individualized comments on it, primarily focusing on ‘content’ and ‘organization’. Topics were all selected from *Criterion*® to control the level of difficulty level with that of the pre- and post-test essay. Unlike the pre- and post-test, there was no time limit in completing the task (Table 3).

TABLE 3

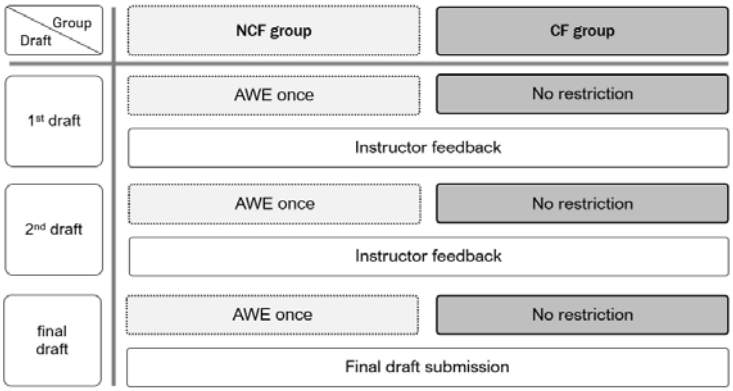
Writing Task Topics

Task	Topic
Task 1	It has been said, “ Not everything that is learned is contained in books.” Compare and contrast knowledge gained from experience with knowledge gained from books. In your opinion, which source is more important? Why?
Task 2	You have been told that dormitory rooms at your university must be shared by two students. Would you rather have the university assign a student to share a room with you, or would you rather choose your own roommate? Use specific reasons and details to explain your answer.

While both CF and NCF group received teacher feedback after they submit each draft to the instructor, they differed in frequency of receiving immediate feedback from AWE and access period to the system. As for CF group, they had access to AWE system without restrictions and received feedback whenever they needed. On the other hand, the NCF group had access to AWE only once before they submit each draft to the instructor. As for the last session, students completed post-test essay along with post-treatment questionnaire. The summary of the whole procedure is delineated in Figure 3.

FIGURE 3

Feedback Provision Procedure



3.4. Data Analysis

The participants' essays are measured based on analytic scoring rubric, adopted and elaborated referring to Cohen (1994) and Hedgcock and Lefkowitz (1992). The rubric contains five components: content, organization, grammar, vocabulary, and mechanics. Each component consists of one to five points, which makes maximum of 25 scores in total. Based on the obtained analytical scores, SPSS version 21 independent-samples *t*-test was performed for between-group analysis to examine the relative effects of two different applications of AWE on improving overall writing performance.

Two raters participated for the current study in measuring students' pre- and post-test essays. As for the two raters, bilingual researcher who holds a master's degree in English education and who is currently enrolled in doctoral program of English education participated. To reach agreement in scoring approaches, the raters met offline and first reviewed the analytic scoring rubric. After confirming their understanding of the rubric, they rated first three essays and discussed their reasons for assigning the scores. Both raters made some changes in scores until they reached consensus. The scores obtained from two raters were analyzed to confirm inter-rater reliability by utilizing Pearson correlation, SPSS version 21. Based on the Pearson's correlation coefficients, the levels of reliability were found to be significantly high (Table 4).

TABLE 4

Inter-rater Reliability of the Pre- and Post-treatment Essay Scores

Test	*r*	*p*
Pre-treatment	.96	.00
Post-treatment	.90	.00

As for perceived effectiveness of AWE feedback, post-questionnaire six-point Likert scale items were coded. The coded data were statistically analyzed to investigate the differences in perceived effectiveness of AWE feedback between the CF and NCF group. As for open-ended questions, students' responses were categorized and counted for its frequency along with further analyses done to examine the differences between the two group.

4. RESULTS AND DISCUSSION

4.1. Effects of Different Application Types of Automated Writing Feedback on Writing Performance

4.1.1. The continuous feedback group's frequency pattern of utilizing AWE feedback

To gain detailed information on how participants make use of automated writing evaluation system when they were allowed to get access to it whenever they are in need, continuous feedback group's AWE utilization pattern was examined. The pattern was investigated in terms of the number of submissions to AWE system (Zhang, 2016) at each writing stage. Number of submissions is noteworthy due to its portrayal of how much and what kinds of efforts students put into revisions (Attali, 2004; Li et al., 2015). To that end, the CF group frequency in accessing to the AWE system was counted for each writing phase in completion of their first and second writing task (Table 5).

TABLE 5

Continuous Feedback Group AWE Utilization Pattern

Task	Draft	N	*M*	*SD*	Min.	Max.
First	First	10	2.8	.79	2	4
	Second	10	1.9	.57	1	3
	Third	10	1.1	.32	1	2
Second	First	10	2.7	1.06	2	5
	Second	10	1.9	.88	1	4
	Third	10	1.2	.42	1	2

Unlike the concern over students' overt attention paid to language-related issues when they were allowed to have access to the AWE system whenever they want to, the descriptive statics of the maximum value revealed that EFL students did not received AWE feedback of more than 5 times over a single drafting stage. More interestingly, the descriptive statistics of the CF group AWE utilization pattern in both first and second task revealed students' preference to receive immediate feedback most frequently while they write their first draft (M = 2.8 and 2.7

respectively). Furthermore, the frequency in use of AWE gradually decreased as they move on to the final draft. The result indicates that EFL learners are in favor of obtaining detailed instant feedback in preliminary drafts, which somewhat goes against what was suggested in process-oriented writing pedagogy to encourage feedback on language-related issues to be deferred until the last period (Sommers, 1982; Zamel, 1982, 1985). The empirical evidence in students' preference to receive feedback on earlier drafts calls in need of careful pedagogical considerations on how to embrace feedback on language-related issues in the process of writing, as in terms of when and how to provide it.

4.1.2. Relative effects of varied application types of AWE feedback on overall writing quality

To closely investigate the relative effects of different application types of automated writing feedback on improving students' writing performance, between-group analysis was performed. To that end, analytic scores of the NCF and the CF group's post-test essay were compared and statistically examined by conducting multiple independent *t*-tests at the significance level of .05 (Table 5).

The dimension of content, grammar and the total score were the three area that was revealed to have significant differences between the NCF and CF group. As for the *content* dimension, the CF and NCF group each obtained mean scores of 4.05 and 4.55 and the mean difference was found to be statistically significant (t = -2.76, p = .01). In the dimension of *grammar*, the mean score of the CF group (M = 4.20) was significantly higher than that of the NCF group (M = 3.85) in which p-values were lower than the chosen significance level of .05 (t = -3.13, p = .01). Going in tandem, the total score of the CF group (M = 21.25) was revealed to significantly outperform the NCF group (M = 19.20) (t = -2.19, p = .04). The outperformance of CF group in the dimension of not only grammar but also content and the total score indicates that receiving immediate language-related feedback in any stage of the writing process rarely downgrades the quality of writing. Contrary to the concern posed by process-based writing pedagogy (Bitchener & Ferris, 2012; Zamel, 1982, 1985), the result indicates that students in CF group did not pay too much attention to language-related issues to the degree of interfering the development of content. To provide EFL students with instant language-related feedback whenever they call for in any stage of the writing process thus seems beneficial to some degree in improving EFL learners' writing quality.

Unlike most other previous studies which did not revealed of significant positive influence of AWE on increasing content level (Kellogg et al., 2010; Li et al., 2015; Moon & Pae, 2011), the result demonstrated statistically significant improvement in content area. In appreciation of AWE feedback capable of providing real-time feedback on language-related features, EFL writers seems to have benefited by using their time and freed up cognitive capacity to concentrate largely on elaborating their ideas, rather than struggling to correct surface-level

language features (Ashwell, 2000; Fathman & Whalley, 1990; Ferris, 1997). Thus, in contrast to second language writing researchers' concern over providing premature feedback on accuracy, EFL students were capable of simultaneously sharing their cognitive processing capacity in both meaning- and form-related issues, which seems to have yet interfere learners' expansion of ideas (Zamel, 1982, 1985). Furthermore, the appropirate grammar - with the help of AWE feedback - seems to have made the essay more readable by removing errors that can disrupt in conveying the meaning clearly and appropriately to intended readers (Hinkel, 2013, Manly & Calk, 1997; Nunan, 2005; Weaver, 2008).

TABLE 6

Effects of Different Application Types of Automated Writing Feedback on Overall Writing Quality

Rating Dimension	Group	*N*	*M*	*SD*	*t*	*p*
Content	NCF	10	4.05	.39	-2.76	.01
	CF	10	4.55	.44		
Organization	NCF	10	4.05	.28	-0.85	.41
	CF	10	4.15	.24		
Grammar	NCF	10	3.85	.42	-3.13	.01
	CF	10	4.20	.26		
Vocabulary	NCF	10	3.80	.42	-1.56	.14
	CF	10	4.05	.28		
Mechanics	NCF	10	4.10	.39	-1.57	.14
	CF	10	4.40	.46		
Total	NCF	10	19.20	2.61	-2.19	.04
	CF	10	21.25	1.42		

Note. NCF = Non-continuous Feedback, CF = Continuous Feedback

Unlike the aforementioned results, the analyzed result revealed statistically insignificant differences between CF (M = 4.15) and NCF (M = 4.05) group in *organization* dimension (t = -0.85, p = .41). The dimension of *vocabulary* also portrayed no significant differences between the two group, despite CF group receiving higher score in descriptive statistics (M = 4.05) than NCF group (M = 3.80) (t = -1.56, p = .14). Insignificant differences between the NCF and CF group in the dimension of *organization* and *vocabulary* seem to have retrieved due to *Criterion*® being insufficient in providing feedback regarding 'vocabulary' and 'organization/development'. This leads EFL writers to solely rely on teachers' feedback, in which there are no differences in frequency of receiving comments from instructor between the two groups.

Although AWE system is successful in providing feedback on language-related issues, no significant differences were revealed in the dimension of *mechanics* between the CF (M = 4.40) and the NCF group (M = 4.40) (t = -1.57, p = .14). When examined in which area students mostly received AWE feedback regarding mechanics, they predominantly had errors in

'spelling' and 'missing comma'. It seems that the 'spell-check' function equipped in *Criterion*® was capable enough to support both NCF and CF group in revising those simple mechanical errors. Since both NCF and CF group had access to the function throughout their drafting stages, no significant differences between the two group can be attributable to students' flexible use of the 'spell-check' function (Kellogg et al., 2010). Although NCF group still had more chances to use the function due to no restrictions in accessing to AWE program, minor errors in mechanics seems to in less need of a frequent feedback to be corrected.

4.2. Learner Perceptions on Automated Writing Feedback

4.2.1. Perceived effectiveness of automated writing feedback on writing performance

Closed- and open-ended questions in post-questionnaire were analyzed to explore EFL learners' overall perception on automated writing feedback and to examine differences between the NCF and the CF group in the perceived effectiveness of automated writing feedback.

1) Responses to closed items

In examining EFL learners' perceptions on automated writing feedback in improving their overall writing performance, item 1 through 5 asked whether AWE system was beneficial in improving each analytical dimension of writing, namely content, organization, grammar, vocabulary, and mechanics. Item 6 was on the participants' perceived effectiveness on *Criterion*® score and trait levels while item 7 inquired of students' willingness to utilize AWE in future. The relative differences between the CF and NCF group responses in using AWE system were investigated by comparing and conducting independent-samples *t*-test in which significance level was set on .05. (Table 7).

Out of 5 different items that asked students' perceived effectiveness of automated writing feedback in improving each analytic rating dimension, item 3 (grammar) and item 5 (mechanics) were the two dimensions that students expressed in favor of using automated writing evaluation system, in which the mean score was higher than point 4 (*somewhat agree*). As for *grammar* (item 3) particularly, the difference in the mean score of 4.65 and 5.10 for each NCF and CF group was found to be statistically significant based on the independent-samples *t*-test result (t = -2.31, p = .03). This goes in tandem with what was obtained in Li et al. (2015) and Moon and Pae (2011), which demonstrated students' positive reactions to the use of AWE corrective feedback specifically in resolving language-related issues. CF group showing significantly more favorable attitude in receiving AWE feedback regarding grammar further indicates learners' preference in receiving fair amount of feedback on grammar as well as the merits of AWE system in providing appropriate level of feedback on grammar.

TABLE 7

Perceived Effectiveness of Automated Writing Feedback

	Questionnaire Items	NCF(N=10)		CF(N=10)		*t*	*p*
		M	*SD*	*M*	*SD*		
1.	AWE is beneficial in composing topic-related content.	2.50	.85	2.70	.68	-0.58	.57
2.	AWE is beneficial in composing logically well-organized writing.	3.30	.48	4.00	.82	-2.33	.11
3.	AWE is beneficial in composing grammatically-correct sentences.	4.65	.47	5.10	.39	-2.31	.03
4.	AWE is beneficial in using appropriate vocabulary.	3.40	.70	3.60	.84	-0.58	.57
5.	AWE is beneficial in writing accurate spelling and punctuation.	4.40	.52	4.60	.52	-0.87	.40
6.	*Criterion*® score and trait levels were beneficial in revising drafts.	3.40	1.08	3.75	.72	-0.86	.40
7.	I would like to use AWE in the future.	4.60	.52	5.00	.82	-1.31	.21

Note. NCF = Non-continuous Feedback, CF = Continuous Feedback

For item 6 which asked whether *Criterion*® score and reported trait levels are beneficial in revising drafts, the NCF and CF group each responded in the mean score of 3.40 and 3.75, with no statistical mean difference retrieved from independent-samples *t*-test ($p = .40 > .05$). Students expressed slightly unfavorable in those scores mainly due to overly generous grading, as it will be discussed in the following section. Overall, the retrieved result from item 7 indicates EFL learners' in both group having positive attitude toward the future use of *Criterion*®.

2) Responses to open-ended items

For further examination on EFL learners' perceived effectiveness of AWE feedback, responses from open-ended questions were categorized and counted for frequency. Open-ended items required students to write down advantages and disadvantages of AWE feedback. To investigate whether there are differences in learners' perceptions depending on how frequently and freely learners had access to AWE system, frequency for each categorized response were compared in terms of different AWE application type group. Results of open-ended questions are organized in Table 8.

EFL learners' comments on the advantages of AWE feedback were categorized into five different ideas. Students were in favor of AWE mostly in terms of being able to identify their frequently-made mistakes or weak points, with the response frequency of 7 for both CF and NCF group Learners in both group, with the response frequency of 3, also considered AWE feedback positively in allowing them to weigh up their English writing ability. The CF group students were in larger preference to AWE feedback in terms of

being provided with the feedback at any time they want (Moon & Pae, 2011), with the response frequency of 8 compared to that of 6 for NCF group. The CF group was further indicated positive attitude toward AWE feedback which helped them to increase their motivation, confidence, and a sense of accomplishment while revising drafts (Li et al., 2015), with the response frequency of 5 compared to that of 3 for NCF group. What is more, one student in CF group stated that submitting drafts through AWE before handing it in to the instructor made one to release the stress of showing the drafts to the others.

TABLE 8

Comments on Advantages and Disadvantages of Automated Writing Feedback

	Comments	Frequency	
		NCF	CF
Advantage	Beneficial in identifying one's frequently-made mistakes or weak points	7	7
	Provision of feedback at any time one wants	6	8
	Increase in motivation, confidence, and a sense of accomplishment in revising drafts	3	5
	Beneficial in weighing up my English writing ability.	3	3
	Lightened burden of showing my writing to others by being assessed to a machine	-	1
Disadvantage	Lacks detailed explanations on how to revise marked errors	7	9
	Overly generous in marking	3	1
	Arbitrary scoring criteria or standard	2	1

Note. NCF = Non-continuous Feedback, CF = Continuous Feedback

Unlike the aforementioned positive attitude toward the system, students made some negative remarks on AWE feedback on three different ideas. Both groups of students were mostly dissatisfied with AWE in terms of lacking detailed instructions on how to revise marked errors in which NCF and CF group responded with the frequency of 7 and 9 respectively. This was also shown in Moon and Pae (2011) and Zhang (2016) where students expressed dislike and frustration in AWE of lacking explanation on how to correct errors. The reasons behind the CF group expressing greater dissatisfaction on this issue may seem to stem from more frequent access to AWE system which increases the proportion of receiving more feedback than the NCF group. Furthermore, both groups made additional comments on the disadvantages of AWE of being overly generous in grading and providing arbitrary scoring criteria.

4.2.2. Influence of AWE on the perceived effectiveness of immediate feedback

Students further reported their perceptions on whether they prefer to receive frequent AWE

feedback and whether they are in favor of having no restrictions in accessing to the system, in which the analysis organized in Table 9. The result revealed that EFL learners preferred to receive immediate feedback via AWE system as much as they want (*frequency*) and whenever they are in need throughout the entire drafting stages (*access period*).

TABLE 9

Perceived Effectiveness of Frequency and Access Period of Automated Writing Feedback

	Group	Questionnaire Items	*N*	*M*	*SD*
Frequency	NCF	I would like to receive AWE feedback without any restrictions on frequency.	10	5.00	.82
	CF	No restriction in frequency of receiving AWE feedback was beneficial.	10	4.80	.63
Access Period	NCF	I would like to receive AWE feedback at any period.	10	4.40	.84
	CF	Being able to access to AWE and receive feedback at any period was beneficial.	10	4.80	.79

Note. NCF = Non-continuous Feedback, CF = Continuous Feedback

Although advocates of process-based writing approach were heavily reluctant to provide too much written corrective feedback on accuracy out of concern for distracting students' reformulation and development of content (Sommers, 1982; Truscott, 1996, 2007; Zamel, 1982, 1985), EFL learners expressed in need of sufficient feedback at any time they are writing. Thus, straining not to provide feedback on accuracy from earlier writing stages should pedagogically be reconsidered in EFL context.

5. CONCLUSION

The present study was conducted to examine the effects of different application types of automated writing feedback on Korean EFL writing, utilizing *Criterion*® which is capable of providing students with instant feedback while they are writing drafts (Attali, 2013; Attali & Burstein, 2006; Elliot, 2003; Shermis & Hamner, 2013; Ware & Warschauer, 2006). Despite the verified effectiveness and perceived usefulness of AWE system in instructional setting (Li et al., 2014; Li et al. 2015; Link et al., 2014; Moon & Pae, 2011; Warschauer & Grimes, 2008), only a handful of research explored further into examining *at what point* in the process-based writing stages AWE feedback should appropriately be provided (Kellogg et al., 2010). Since AWE system is computationally adept at providing instant feedback mostly on sentence-level correctness rather than on higher level concerns (Weigle, 2013), a mere implementation of AWE can incite learners to overly concern about their language-related issues (Zamel, 1982, 1985). For AWE feedback to effectively function within the framework of process-based

writing pedagogy (Ranalli et al., 2017), more systematic research is to be conducted. In other words, careful considerations on *at what point* in the writing stage AWE feedback be applied needs to be examined, so as not to risk EFL students to pay too much attention to surface-level language features throughout their overall process of writing. Therefore, the current study has implication in identifying pedagogically sound applications of AWE practice that can properly be embraced in process-based writing instruction (Link et al., 2014).

To that end, the effect of two different AWE application types on writing performance was investigated. Application types were differentiated in terms of at what point in the writing stages learners were enabled to have access to the AWE system. Thus, the participants in the current study was classified into either non-continuous feedback (NCF) – where students had access to the system only once right before they submit their drafts which is in accordance to process-based writing pedagogy, or continuous feedback (CF) group – where students had no restrictions in accessing the system throughout the overall writing stages.

The research result indicated that although students were allowed to receive instant language-related feedback via AWE system whenever they need it (CF group), students did not receive feedback at a worrying level and there was no sign of students being interfered by frequent language-related feedback in developing their content. In fact, CF group significantly outperformed NCF group in overall writing product (Kellogg et al., 2010). Furthermore, students mostly had positive perceptions on AWE as well as instant feedback in improving their writing.

As for the relative effects of two different application types on overall writing performance, the CF group statistically outperformed the NCF group in the dimension of *content* and *grammar*. The result reveals that with the support of AWE system - which mainly provides feedback on language features - participants in CF group seems to have less concern over language-related issues. This enabled them to further capitalize the advantages of AWE feedback by using more of their cognitive capacity in developing their content. In addition, since AWE system primarily and favorably provides appropriate language-related feedback, receiving comparatively large amount of AWE feedback seemed to have positively influenced the CF group to outperform NCF group in the dimension of grammar (Attali, 2004, Choi & Lee, 2010; Kellogg et al., 2010; Li et al., 2015; Moon & Pae, 2011). This in turn implies that frequent feedback on accuracy from preliminary drafts, which has been discouraged in current process-oriented writing pedagogy, unlikely interferes with the development of ideas or disrupts improvement of writing performance at a worrying level. Since teacher feedback, at its core, primarily focus on organization and content of the writing, receiving continuous support of immediate feedback on grammar has a great potential in advancing EFL students' overall writing quality.

With regards to the perceived effectiveness of automated writing feedback on writing performance, CF group expressed slightly more favorable attitude than the NCF group on

grammar and *mechanics* dimension. What is more, the current study indicated EFL learners' in favor of being offered with sufficient amount of feedback earlier from preliminary drafts. Thus, the provision of feedback on language-related issues is likely to alleviate EFL learners' affective concerns over using appropriate language features to convey their ideas meaningfully.

From a pedagogical perspective, the result of the present study has several useful implications on Korean EFL process-based writing classroom context. First of all, implementing AWE system along with instructor feedback within the framework of process-based writing pedagogy is expected to significantly support EFL learners in improving their writing performance. With the help of AWE system in immediately correcting language-related issues, learners can possibly spend additional time and energy in developing content. This is especially noteworthy in that, in contrast to concerns suggested by SLA research, instant language-related seems unlikely to disrupt learners' expansion and reformulation of ideas but rather encourages improvement in overall writing performance. Furthermore, although students had no restrictions in receiving instant sentential-level feedback, students did not utilize the system at a worrying level. A relatively scant number of submission to the system suggests instructors to integrate AWE feedback in complementing instructor feedback without being overly concerned about students paying too much attention to language-related issues throughout the composing process

To generalize what was retrieved from the present study, the research should necessarily be expanded by carrying out a larger number of participants varied in proficiency level. Furthermore, a follow-up study should necessarily be conducted in regular writing courses to improve ecological validity as well as to provide long-term effects of AWE feedback (Li et al., 2014). Systematic research on EFL students' detailed engagement in utilizing AWE feedback and reasons behind their use or non-use of the provided feedback on subsequent drafts is expected to further establish a more concrete understanding of the effectiveness of instant AWE feedback in improving EFL learners' writing performance (Zhang, 2016). Nonetheless, the result is contributing by providing empirical evidence that immediate language-related feedback in the process of composing seems unlikely to interfere EFL writers' development of content as well as the improvement of overall writing performance.

REFERENCES

Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing, 9*(3), 227-257.

Attali, Y. (2004). Exploring the feedback and revision features of Criterion. Paper presented at the *National Council on Measurement in Education (NCME)* held between April 12 to 16,

2004, in San Diego, CA.

Attali, Y. (2013). Validity and reliability of automated essay scoring. In M. D. Shermis, & J. Burstein (Eds.), *Handbook of automated essay evaluation: Current applications and new directions* (pp. 181-198). New York: Routledge.

Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater® v.2. *Journal of Technology, Learning, and Assessment*, *4*(3), 1-30.

Badger, R., & White, G. (2000). Product, process and genre: Approaches to writing in EAP. *ELT Journal*, *54*(2), 153-160.

Barlett, E. J. (1982). Learning to revise: Some component process. In M. Nystrand (Ed.), *What writers know: The language, process, and the structure of written discourse* (pp. 345-363). New York: Academic Press.

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

Bridgeman, B., & Carlson, S. B. (1984). Survey of academic writing tasks. *Written Communication*, *1*, 247-280.

Burstein, J. C., Chodorow, M., & Leachock, C. (2003). Criterion online essay evaluation: An application for automated evaluation of student essays. In *Proceedings of the 15th Annual Conference on Innovative Applications of Artificial Intelligence*. Acapulco: Mexico.

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

Burstein, J. C., & Wolska, M. (2003). Toward evaluation of writing style: Finding overly repetitive word use in student essays. In *Proceedings of the 10th Conference on European chapter of the Association for Computational Linguistics-Volume 1* (pp. 35-42). Association for Computational Linguistics.

Byrd, P., & Reid, J. M. (1998). *Grammar in the composition classroom: Essays on teaching ESL for college-bound students*. Boston: Heinle and Heinle.

Calfee, R. C., & Miller, R. G. (2007). Best practices in writing assessment. In S. Graham, C. A. MacArthur & J. Fitzgerald (Eds.), *Best Practices in Writing Instruction* (pp. 265-286). New York: Guilford Press.

Celce-Murcia, M. (1991a). Discourse analysis and grammar instruction. *Annual Review of Applied Linguistics*, *11*, 459-480.

Celce-Murcia, M. (1991b). Grammar pedagogy in second and foreign language teaching. *TESOL Quarterly*, *25*(3), 459-480.

Chen, C., & Cheng, W. (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.

Choi, J., & Lee, Y. (2010). The use of feedback in the ESL writing class integrating automated essay scoring (AES). In *Proceedings of Society for Information Technology and Teacher*

Education International Conference, (pp. 3008-3012). Association for the Advancement of Computing in Education (AACE), Chesapeake, VA.

Cohen, A. D. (1994). *Assessing Language Ability in the Classroom* (2nd ed.). Boston, MA: Heinle and Heinle.

Corbett, A. T., & Anderson, J. R. (2001). Locus of feedback control in computer-based tutoring: Impact of learning rate, achievement and attitudes. In Proceedings of *ACM CHI2001 Conference on Human Factors in Computing Systems* (pp. 245-252). New York: Association for Computing Machinery Press.

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.

Elliot, S. (2003). Intellimetric: From here to validity. In M. D. Shermis & J. Burstein (Eds.), *Automated essay scoring: A cross-disciplinary perspective* (pp. 71-86). Mahwah, NJ: Lawerence Erlbaum Associates, Inc.

Enginarlar, H. (1993). Student response to teacher feedback in EFL writing. *System*, *21*, 193-204.

Eskey, D. E. (1983). Meanwhile, back in the real world…: Accuracy and fluency in second language teaching. *TESOL Quarterly*, *17*, 315-323.

Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178-190). Cambridge: Cambridge University Press.

Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly*, *31*(2), 315-339.

Ferris, D. R. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing, 8*(1), 1-10.

Ferris, D. R. (2002). *Treatment of error in second language student writing*. Ann Arbor, MI: University of Michigan Press.

Ferris, D. R. (2004). The "grammar correction" debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime...?) *Journal of Second Language Writing*, *13*, 49-62.

Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81-104). Cambridge: Cambridge University Press.

Ferris, D. R., Brown, J., Liu, H., & Stine, M. E. A. (2011). Responding to L2 writers in college writing classes: What teachers say and what they do. *TESOL Quarterly*, *45*, 207-234.

Ferris, D. R., & Hedgcock, J. S. (2005). *Teaching ESL composition: Purpose, process, and practice* (2nd ed.). Mahwah, NJ: Erlbaum.

Ferris, D. R., & Roberts, B. J. (2001). Error feedback in L2 writing classes: How explicit does it

need to be? *Journal of Second Language Writing, 10*, 161-184.

Frankenberg-Garcia, A. (1999). Providing student writers with pre-text feedback, *ELT Journal, 53*, 100-106.

Frodesen, J. (2001). Grammar in writing. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3rd ed.) (pp. 233-248). Boston, MA: Heinle & Heinle.

Frodesen, J., & Holten, C. (2003). Grammar and the ESL writing class. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 141-161). New York: Cambridge University Press.

Gardiner, J. (2012). Student attitudes to EAP grammar instruction. *Cambridge ESOL: Research Notes, 48*, 11-17.

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), 4-44.

Hajian, L., Farahani, A. A. K., & Shiraizi M. A. (2014). A study of students' and teachers' preferences and attitudes towards correction of classroom written errors in Iranian EFL context. *International Journal of Applied Linguistics & English Literature, 3*, 287-297.

Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudeweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly, 44*, 84-109.

Hedgcock, J., & Lefkowitz, N. (1992). Collaborative oral/aural revision in foreign language writing instruction. *Journal of Second Language Writing, 3*(1), 255-276.

Hinkel, E. (2011). What research on second language writing tells us and what it doesn't. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning*, Vol. 2 (pp. 523-538). New York: Routledge.

Horowitz, D. (1986). Process not product: Less than meets the eye. *TESOL Quarterly, 20*, 141-144.

James, C. (1998). *Errors in language learning and use: Exploring error analysis*. London: Longman.

Johnson, D. M., & Roen, D. H. (Eds). (1989). *Richness in writing: Empowering ESL students*. New York: Longman.

Kellogg, R. T., Whiteford, A. P., & Quinlan, T. (2010). Does automated feedback help students learn to write? *Journal of Educational Computing Research, 42*(2), 173-196.

Kim, Y., & Kim, J. (2005). Teaching Korean university writing class: Balancing the process and the genre approach. *Asian EFL Journal, 7*(2), 1-15.

Knoblauch, C. H., & Brannon, L. (1981). Teacher commentary on student draft: The state of the art. *Freshman English News, 10*, 1-4.

Leki, I. (1990). Coaching from the margins: Issues in written response. In B. Kroll (Ed.), *Second Language Writing* (pp. 57-68). Cambridge: Cambridge University Press.

Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals*, *24*(3), 203-218.

Li, Z., Link, S., Ma, H., Yang, H., & Hegelheimer, V. (2014). The role of automated writing evaluation holistic scores in the ESL classroom. *System*, *44*, 66-78.

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.

Link, S., Dursun, A., Karakaya, K., & Hegelheimer, V. (2014). Towards best practices for implementing automated writing evaluation. *CALICO Journal*, *31*(3), 323-344.

Little, D. (1994). Words and their properties: Arguments for a lexical approach to pedagogical grammar. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 99-122). New York: Cambridge University Press.

Maeng, U. (2010). The effect and teachers' perception of using an automated essay scoring system in L2 writing. *English Language and Linguistics*, *16*, 247-275.

Manly, J., & Calk, L. (1997). Grammar instruction for writing skills: Do students perceive grammar as useful? *Foreign Language Annals*, *30*(1), 77-83.

Moon, Y. I., & Pae, J. K. (2011). Short-term effects of automated writing feedback and users' evaluation of Criterion®, *Korean Journal of Applied Linguistics*, *27*(4), 125-150.

Mora-Flores, E. (2009). *Writing Instruction For English Learners: A focus on Genre*. California: Corwin Press.

Nunan, S. L. (2005). Forgiving ourselves and forging ahead: Teaching grammar in a new millennium. *English Journal*, *94*(4), 70-75.

Page, E. (2003). Project essay grade: PEG. In M. D. Shermis & J. Burstein (Eds.), *Automated essay scoring: A cross-disciplinary perspect*ive (pp. 43-54). Mahwah, NJ: Lawrence Erlbaum Associates.

Peterson, S. (2003). Peer response and students' revisions of their narrative writing. *L1-Educational Studies in Language and Literature*, *3*(3), 239-272.

Peterson, S. (2010). Improving student writing: Using feedback as a teaching tool. What Works? *Journal of Research into Practice*, *29*(2), 1-4.

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.

Schoonen. R., Snellings, P., Stevenson, M., & Van Gelderen, A. (2009). Towards a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. In R. M. Manchon (Ed.), *Writing in Foreign Language Contexts: Learning, Teaching, and Research* (pp. 77-101). Multilingual Matters.

Shermis, M. D., & Burstein, J. (2003). Introduction. In M. D. Shermis, & J. Burstein (Eds.), *Automated essay scoring: A cross-disciplinary perspective* (pp. xiii-xvi). Mahwah, NJ: Lawrence Erlbaum.

Shermis, M. D., & Hamner, B. (2013). Contrasting state-of-the-art automated scoring of essays. In M. D. Shermis, & J. Burstein (Eds.), *Handbook of automated essay evaluation: Current applications and new directions* (pp. 313-346). New York: Routledge.

Shin, S.-K. (2008). 'Fire your proofreader!' Grammar correction in the writing classroom. *ELT Journal*, *62*(4), 358-365.

Shute, V. (2008). Focus on formative feedback. *Review of Educational Research*, *78*, 153-189.

Sommers, N. (1982). Responding to student writing. *College Composition and Communication*, *33*, 148-156.

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

Truscott, J. (1999). The case for "the case for grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, *8*, 111-122.

Truscott, J. (2004). Evidence and conjecture: A response to Chandler. *Journal of Second Language Writing*, *13*, 337-343.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, *16*(4), 255-272.

Truscott, J. (2009). Arguments and appearances: A response to Chandler. *Journal of Second Language Writing*, *18*, 59-60.

Truscott, J., & Hsu, A. Y-P. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, *17*, 292-305.

Ware, P. D., & Warschauer, M. (2006). Electronic feedback and second language writing. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 105-122). New York: Cambridge University Press.

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

Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language Teaching Research*, *10*(2), 1-24.

Weaver, C. (2008). *Grammar to enrich & enhance writing*. Portsmouth, NH: Heinemann.

Weigle, S. C. (2013). English as a second language writing and automated writing evaluation. In M. D. Shermis & J. Burstein (Eds.), *Handbook of automated essay evaluations: Current applications and new directions* (pp. 36-54). New York, NY: Routledge.

Witte, R., & Arndt, V. (1991). *Process writing*. Harlow, UK: Longman.

Zamel, V. (1982). Writing: The process of discovering meaning. *TESOL Quarterly*, *16*, 195-209.

Zamel, V. (1985). Responding to student writing. *TESOL Quarterly*, *19*, 79-102.

Zhang, S. (1995). Reexamining the affective advantage of peer feedback in the ESL writing class. *Journal of Second Language Writing*, *4*, 209-222.

Zhang, Z. (2016). Student engagement with computer-generated feedback: a case study. *ELT Journal*, *3*(1), 317-328.

Applicable levels: Tertiary

Wonyoung Koh
Department of English Education
Ewha Womans University
52, Ewhayeodae-gil, Seodaemun-gu,
Seoul, 03760, Korea
Phone: 02-3277-2647
Email: odori1991@ewhain.net

Received on June 1, 2017
Reviewed on July 15, 2017
Revised version received on August 15, 2017