p-ISSN: 2797-4537 e-ISSN: 2809-4980

Providing Online Feedback Using an AI Proof-reading Tool to Enhance Student Surface-Level Writing Skills

Ester Widawati Tedjo[1*]

[1]Office Administration Study Program, Vocational Faculty of Widya Mandala Catholic University, Surabaya

Abstract: This study explores the use of an AI proofreading tool in providing feedback on surface features of student writing. It examines the effect of ProofWriter online feedback compared to teacher feedback on students' writing improvement, and investigates which aspect of students' surface-level writing skills improves the most after using the proofreading tool for two-months. Participants were 20 fourth-semester secretarial students taking English Business Correspondence II course. The 'online group' students used ProofWriter online tool to get feedback on surface-level aspects of their writing, and the 'non-online group' students got feedback from the class teacher while practicing to write English business letters. A pre-test and post-test were administered to show the students' performance in writing the business letters. The letters were graded using the ESL Composition Profile rubric, and then means of the pre-test and post-test were calculated and T test was applied. A questionnaire about using ProofWriter online tool was distributed to the online group students. The results showed the online group students significantly increased their surface-level writing skills in a slightly higher level than those of the non-online group. They also improved the most in grammar for the surface features of writing.

Keywords: written corrective feedback; proofreading tools; surface-level errors

INTRODUCTION

Written corrective feedback (WCF) is an important factor in the process of second or foreign language learning. Although it may have powerful and major influences on learning and achievement, a wide range of factors, such as the type of feedback and the way it is provided, can contribute to how effective the feedback is (Hattie and Timperley, 2007; O'Flaherty, 2016). It might be for this reason why there is still no clear consensus among researchers on providing WCF effectively and regarding students' feedback preferences, though there has been a great deal of research conducted.

Studies on WCF have shown contradictory results. Some studies demonstrate the significance and positive impact of WCF. Ferris (2004) pointed out that there are errors that students can never address on their own. Based on the results of their studies, Ferris and Hedgcock (2005), Yang, Badger and Yu (2006), and Lee (2008) concluded that students view teacher feedback as both critical to improving their writing skills and more valuable than self or peer feedback. Evan, Hartshorn, and Strong-Krause (2011) suggested that students will benefit most from WCF if the writing tasks are timely, constant and manageable. In other words, it is a relatively short writing, perhaps a paragraph, limited to around ten minutes, and assigned three or four times a week. In that context, teachers may expect students to address all or nearly all errors in their writing. Irwin (2017) showed that

*Corresponding author: esttedjo@ukwms.ac.id
To cite this article: Tedjo, E.W. (2022). Providing Online Feedback Using an AI Proofreading Tool to Enhance Student Surface-Level Writing Skills. *Ebony --- Journal of English Language Teaching, Linguistics, and Literature, 2 (1)* 2022, pp. 37--45.

regarding students' expectation, students prefer teachers to provide direct WFC and to attend to all errors. Several other researchers also agreed that WCF can facilitate improved accuracy in L2 students' writing (Kang and Han, 2015; Hashemian and Farhang-ju, 2018; Lim and Renandya, 2020).On the other hand, Truscott (1996) indicated that WCF has little positive impact and in some cases it even has negative impact. Truscott, therefore, claimed that WCF should be abandoned. Similarly, O'Flaherty (2016) argued despite the fact that students feel the teacher needs to provide WCF, they do not actually make use of the feedback.

In addition to the impact of WCF, studies have also considered whether or not to address all types of errors in the feedback. Again, researchers disagree on being either focused or unfocused in providing WCF. Among those researchers who are in favor of WCF, Halimi (2008), Zhang et.al. (2021) and Halim et al. (2021), similar to Irwin (2017), suggested unfocused WCF. They indicated that students prefer feedback on all errors, including surface-level errors, in their writing. Halimi (2008) pointed out that both teachers and students agree that accuracy is very important, and showing where the error is and giving cues or editing symbols about how to correct it is the technique preferred. Different views, however, are given by some other researchers who suggest a more focused approach to WCF. For examples, Bitchener (2008), Ellis et.al. (2008) and Sheen, Wright and Moldawa (2009) recommended teachers focus on particular grammatical issues such as articles and past tense. Rafoth (2004) proposed addressing only the most important errors since too much information can be overwhelming, while Harris and Silva (1993) suggested teachers concentrate on global errors which include minor mistakes affecting meaning (for example, "the teacher is bored" instead of "the teacher is boring").

Another important issue in examining WCF is how the feedback is given. In this regard, WCF can be categorized into three broad types: direct WCF, indirect WCF and metalinguistic WCF (Ellis, 2009). Direct corrective feedback refers to a practice of supplying the correct form above or near the error, crossing out unnecessary words(s) and inserting missing word(s). In indirect corrective feedback teachers indicate the occurrence of an error by underlining or circling the error, or by highlighting it in some way. Metalinguistic corrective feedback requires teachers to provide an error code or a brief metalinguistic comment or clue to indicate the nature of the error, or to assign numbers to errors and then give brief grammatical descriptions at the bottom of the paper. In short, the three types of WCF differ in their degree of explicitness, and metalinguistic WCF has similar features as indirect WCF in that it withholds the correct form and promotes self-correction.

The researchers who advocate direct WFC believe that it is more constructive to students because of these three reasons: (1) students are less likely to be confused; (2) the feedback gives students valuable resources to predict and prevent similar errors in the future, and (3) students will find it more interesting and immediate. Liu (2008) stated that direct WCF is better than indirect WCF for lower proficiency students. Simply highlighting the errors may be confusing for the students since they are unable to self-correct the errors. Van Beuningen, De Jong, and Kuiken (2012) claimed that direct WCF and indirect WCF are suitable for different types of errors. Direct WCF works more effectively with grammatical errors whereas indirect WCF is better with non-

grammatical errors. On the other hand, the advocates of indirect WCF have revealed that this approach is better since it requires students to engage in self-guided learning and thus promotes greater cognitive experience and enhances long-term acquisition (Ferris, 2004; Bitchener and Knoch, 2008; Hashemian and Farhang-ju, 2018).

The debate on the effectiveness of WCF has been continuing up to the present. In practice when it comes to student writing, teachers often question themselves as to whether or not they have to address all types or errors. Should teachers attend to only deep-feature or global errors? Or, should they also address surface-level errors? According to Chomsky's theory (Chomsky, 1965), the surface structure is the actually produced structure (that is the sentence as it is written), while the deep structure is the abstract structure that expresses the semantic contents of the sentence and that allows the readers to know what the sentence means. The surface structure deals with syntactic, lexical and graphophonic systems whereas the deep structure is concerned with semantic, schematic and pragmatic systems. For most writing teachers, providing WCF is a large investment of time and energy, and what's more when it should cover surface-level errors (errors in grammar, usage, spelling, punctuation and mechanics) that might be quite many in a piece of student writing. Moreover, some teachers and researchers consider surface-level errors to be trivial as long as they do not affect the meaning. In their view, teachers and students should not be so obsessed with surface-level errors or problems of sentence accuracy that they neglect more global issues such as idea development. That might be true to some extent, yet for particular professions like in banking and business, problems with accuracy at sentence level can create a negative impact on how people are perceived (Beason, 2001).

Nowadays the advancement of technology has also made great changes in feedback practices. Computer-mediated feedback, mobile-mediated feedback, computer tools and artificial intelligence (AI) tools can facilitate students in getting WCF and improving their writing. There have been a few empirical studies examining the provision of WCF using technology. Bitchener, East and Cartner (2010) showed that direct and indirect WCF can be effectively provided on-line. Similarly, Elola and Oskoz, (2016) indicated that computer-mediated feedback offers effective approaches to engaging students in improving their writing. It is, therefore, interesting to find out how technology-aided feedback can assist teachers to deal with surface-level errors in student writing.

The current study aims at investigating the effectiveness of using an AI proofreading tool, ProofWriter, to provide feedback on surface-level aspects of student writing, and examining the effect of ProofWriter online tool feedback compared to teacher feedback on students' writing improvement. Further, the study intends to reveal which aspect of students' surface-level writing skills improves the most after using the proofreading tool for a two-month period.

METHOD

The participants were 20 fourth-semester students enrolled at a secretarial college in Surabaya, Indonesia. They were taking English Business Correspondence II course, and had passed Structure I, Structure II, Sentence Structure I and Sentence Structure II subjects. Thus, they had had overall knowledge of English grammar and sentence structure though

their English proficiency level varied as shown by their achievements in those subjects. All of them, however, were highly motivated and had positive attitude towards learning English Correspondence as an English writing course.

To collect the data for the study, the students were given a pre-test in which they were asked to write two English business letters, a direst request and a positive acknowledgment of request. These two types of letters are the topics of English Business Correspondence II course for the first half of the course. The pre-test was conducted in the first week (the first meeting) of the course.

The 20 participants were then divided into two groups. One group (the 'online group') would use ProofWriter online tool to get feedback on surface-level problems of their writing, and the other group (the 'non-online group') would get feedback from the class teacher while they were practicing to write the two types of letters during the first half semester of the course which lasted in two months effectively. There were totally eight letters, four direct requests and four acknowledgments, assigned to the students during this period. Coupons for ProofWriter online tool were given freely to the ten students of the online group. Each student was given 20 coupons for proofreading their letters. It means that the students had at least two ProofWriter coupons to proofread and get feedback on each business letter they made. For each letter assigned to the students to write, they had to revise it twice. Immediately after they submit the first version of a letter, they would get the first feedback. For the surface-level aspects of writing, the participants got the feedback either online or from the teacher as assigned to their group, while for the deep features all the students from both the online group and the non-online group got the feedback from the classroom teacher. The students had to revise the letter based on the feedback given and submit the second version of the letter to get another feedback if there was still some problem with the letter.

The post-test was administered a week after the students had taken the mid-term examination with these two types of business letters as the material. In the post-test, the students had to write also two letters, a direct request and a positive acknowledgment of request.

To get further information about how the participants using ProofWriter felt about the online tool, a questionnaire was distributed to the ten students belonging to the 'online group'. The questionnaire consisted of two types of questions. The first is nine open-ended questions; the subjects were supposed to give their response by writing their answers. The second type of questions also consists of nine items. Here the subjects were asked to give their opinion by circling one of the numbers in the scale which best represented their opinion.

The data collected from the pre-test and post-test showed the 20 students' writing performance, which took the form of English business letters. These letters were checked by three raters who were the English Correspondence class teachers, and they are graded using the ESL Composition Profile (Jacobs et.al., 1981) as the scoring rubric. The profile consists of five descriptors or main traits which are used to assess each letter focusing on the details of each writing aspect and content. The five traits are content, organization, vocabulary, language use, and mechanics. The profile also provides a proficiency scale (which scores

from 1 for excellent up to 4 for very poor) into which each trait fits. The raters assessed and gave scores to each trait based on the criteria for each score in the scale. After the letters were scored, the inter-rater reliability was computed using the Intraclass Correlation Coefficient (ICC) as a measure (McGraw and Wong, 1996).

Next, means of the pre-test and post-test were calculated for each group (the 'online group' and the 'non-online group'). T test was also applied to the scores to find out whether there is a significant difference between the pre-test and the post-test results.

In addition, the other set of data gathered by the questionnaire which were distributed to the subjects belonging to the 'online group' were tallied and analyzed to find out the students opinions on using ProofWriter online tool to get feedback on their writing.

RESULTS

The Intraclass Correlation Coefficient (ICC) for individual ratings of the students' writing in the pre-test and post-test were 0.708 and 0.749 respectively. This showed a sufficiently acceptable internal consistency of scoring using the ESL Composition Profile among the three raters.

The pre-test and post-test of the online group result in average score of 32 for the pre-test and 42.6 for the post-test. The pre-test data showed large differences among the ten subjects on three descriptors/traits, namely mechanics, language use and organization. These differences disappeared at the post-test due to the subjects' improvement in writing English grammatical sentences and a strong improvement in their ability to express ideas for the particular texts (direct request and acknowledgment of request).

Table 1. Means of Pre- and Post-Tests of the Online Group

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pretest Online	32.0000	10	3.01846	.95452
	Posttest Online	42.6000	10	3.37310	1.06667

Across all ten subjects of the online group, post-test increased 10.6 points on average after two months of writing and revision practices with ProofWriter online tool feedback. At 0.644, the correlation between the pre-test and post-test measurements of the online group is statistically significant (see table 2 and table 3). This is due to the fact that the increase was consistent across all the ten online group subjects. The score increase was within the range of 9 to 13.

Table 2. Correlations of Pre- and Post-Tests of the Online Group

		N	Correlation	Sig.
Pair 1	Pretest Online & Posttest Online	10	.644	.045

Table 3. T-Test of Pre-and Post-Tests of the Online Group

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	Pretest Online - Posttest Online	-10.60000	2.71621	.85894	-12.54306	-8.65694	-12.341	9	.000

Mean scores of the non-online group (the group with teacher feedback) are showed in table 4 below. The data indicated that the subjects increased in their post-test over the course of the study, on average about 7.3 points.

Table 4. Means of Pre- and Post-Tests of the Non-Online Group

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pretest Teacher	32.6000	10	3.16930	1.00222
	Posttest Teacher	39.9000	10	2.99815	.94810

As seen in table 5, the Pearson correlation between the pre-test and post-test measurements indicated a moderate degree of correlation (0.638), which is similar to that of the online group. The increase of scores (7.3) was also consistent across the ten subjects of the non-online group, but it was lower than the increase of scores of the online group.

Table 5. Correlations of Pre- and Post-Tests of the Non-Online Group

		N	Correlation	Sig.
Pair 1	Pretest Teacher & Posttest Teacher	10	.638	.047

Table 6. T-Test of Pre- and Post-Tests of the Non-Online Group

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	Pretest Teacher - Posttest Teacher	-7.30000	2.62679	.83066	-9.17909	-5.42091	-8.788	9	.000

DISCUSSION

Based on the results it could be concluded that the online group students significantly increased their writing skills with reference to the surface-level aspects. This study also found out that the students of the online group improved in their writing skills in a slightly higher level than those of the non-online group.

The participants' responses to the questionnaire showed that all of them had never used any online proofreading tool to get feedback on their writing before. The students felt they increased the most in terms of grammar in their writing. This is in line with the increased scores on the language use showed in their post-test. In addition, the students also thought that they improved in organization (idea arrangement) and vocabulary.

Though the online group students found it simple and easy to use ProofWriter, and they liked that fact that the feedback was given promptly. Some of them (60%) said that they frequently did not understand the feedback given. Most of the students (80%) thought that the feedback given was too general. The feedback just said that there was an inappropriate or incorrect use of, for example, articles or prepositions. When the students did not understand the real problem which had been indicated by the Proofwriter feedback, they would just guess, change the sentence (wrote another new sentence and proof read it again), or discuss with peers or the teacher. It could be for that reason that though the students found ProofWriter helpful for them in giving feedback to revise their writing, they preferred teacher feedback which could guide them in details to see the specific problem in their writing.

CONCLUSION

The present study has indicated that an online proofreading tool like ProofWriter, which was applied in this study, could help students increase their writing skills in term of surface-level aspects. Both the students' post-test scores and responses to the questionnaire showed that the students improved in grammar the most. The students also found it simple, easy and helpful using the proofreading tool. However, they still expected to get teacher feedback on their writing because in their opinion teacher feedback provided them with more detailed explanation to let them see by themselves the real problem with their sentences in writing English business letters. What the students like best about ProofWriter is that, unlike teacher feedback, the AI proofreading tool is fast in providing them with feedback. It needs to be noted that the small size of participants in the study should nevertheless be considered a limitation.

This study made use of one kind of proofreading tool, i.e. ProofWriter, to provide feedback on surface-level features of student writing. There are several other AI proofreading tools which can be used to provide feedback on written texts. It is suggested that future research in the similar field is carried out using other proofreading tools and it compares their effectiveness in helping students developing their writing skills. Further research may also investigate the impact of WCF provided by a proofreading tool on each separate surface feature of student writing.

REFERENCES

Beason, L. (2001). Ethos and error: How business people react to errors. *College Composition and Communication*, 53 (1), pp. 53-64.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing* 17, pp. 102–118.

Bitchener, J., East, M. and Cartner, H. (2010). The effectiveness of providing second language (L2) writers with on-line written corrective feedback. Ako Aotearoa. Retrieved from https://ako.ac.nz/knowledge-centre/the-effectiveness-of-providing-second-language-l2-writers-with-on-line-written-corrective-feedback/

Bitchener, J. and Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, 12 (3), 409-431.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*, Cambridge: M.I.T. Press. Retrieved from http://www.chomsky.info/onchomsky/19720629.htm

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63 (2), pp. 97–107.

Ellis, R., Sheen, Y., Murakami, M. and Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System* 36, pp. 353–371.

Elola, I. and Oskoz, A. (2016). Supporting second language writing using multimodal feedback. *Foreign Language Annals*, 49 (1), pp. 58-74.

Evan, N. W., Hartshorn, J. K. and Strong-Krause, D. (2011). The efficacy of dynamic written corrective feedback for university-matriculated ESL learners. *System* 39 (2), 229-239.

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, pp. 49-62, doi:10.1016/j.jslw.2004.04.005.

Ferris, D. R., and Hedgcock, J. S. (2005). Teaching ESL Composition: Purpose, process, and practice. Mahwah, NJ: Lawrence Erlbaum Associates.

Halim, T., Wahid, R. and Halim, S. (2021). EFL students' attitudes toward corrective feedback: a study conducted at undergraduate level. *Saudi Journal of Language Studies* Vol. 1 No. 1, pp. 40-49, doi: 10.1108/SJLS-03-2021-0004.

Halimi, S.S. (2008). Indonesian teachers' and students' preferences for error correction. *Wacana*, Vol. 10 No. 1, pp. 50-71.

Harris, M. and Silva, T. (1993). Tutoring ESL students: issues and options. *College Compositions and Communication* 44 (4), pp. 525-537.

Hashemian, M. and Farhang-ju, M. (2018). Effects of metalinguistic feedback on grammatical accuracy of Iranian field (in)dependent L2 learners' writing ability. *Journal of Research in Applied Linguistics* Vol. 9 No. 2, pp. 141–161.

Hattie, J. and Timperley, H. (2007). The Power of Feedback. *Review of Educational Research* Vol. 77 No. 1, pp. 81-112, doi: 10.3102/003465430298487.

Irwin, B. (2017). Written Corrective Feedback: Student Preferences and Teacher Feedback Practices. *IAFOR Journal of Language Learning* Vol. 3 No.2, pp. 35-58.

Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., and Hughey, J. B. (1981). *Testing ESL Composition: A practical approach*. Rowley, MA: Newbury House.

Kang, E., and Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: a meta-analysis. *Modern Language Journal* 99, pp. 1–18, doi: 10.1111/modl.12189.

Lee, I. (2008). Student reactions to teacher feedback in two Hong Kong secondary classrooms. *Journal of Second Language Writing* 17, 144–164, doi: 10.1016/j.jslw.2007.12.001.

Lim, S. C., and Renandya, W. A. (2020). Efficacy of written corrective feedback in writing instruction: a meta-analysis. *The Electronic Journal for English as a Second Language* 24, pp. 1–26.

Liu, Y. (2008). The effects of error feedback in second language writing. *Journal of Second Language Acquisition and Teaching* 15, p. 65-79.

McGraw, K. O. and Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1, pp. 30–46, doi:10.1037/1082-989X.1.1.30.

O'Flaherty, D. (2016). Japanese High School Students' Attitudes Towards and Usage of Corrective Feedback on Their Written Work. *The Language Teacher* Vol. 40 No. 6, pp. 3-8.

Rafoth, B. (2004). Responding online. In Bruce, S. and Rafoth, B. (Eds.) *ESL Writers: A Guide for Writing Center Tutors* (pp. 1-17). Portsmouth, NH: Boynton.

Sheen, Y., Wright, D. and Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System* 37, pp. 556–569.

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

Van Beuningen, C., De Jong, N. H., and Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning* 62 (1), 1-41.

Yang, M., Badger, R., and Yu, Z. (2006) A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing* 15, pp. 179–200, doi: 10.1016/j.jslw.2006.09.004.

Zhang, T., Chen, X., Hu, J. and Ketwan, P. (2021). EFL students' preferences for written corrective feedback: Do Error Types, Language Proficiency, and Foreign Language Enjoyment Matter? *Front. Psychol.* 12:660564, doi: 10.3389/fpsyg.2021.660564.