ISSN 2522-6398 Vol. 4, Issue 11: 94-98, DOI: 10.25236/FER.2021.041117

Understanding the Effects of Automated Writing Corrective Feedback on L2 Writing Accuracy Across Proficiency Levels

Wang Shiyao

The University of New South Wales Randwick, New South Wales, Australia NSW2052

Abstract: *Exploring the effects of automated writing evaluation (AWE) is essential for maximizing the potential of using technology for language learning. Notwithstanding the large amount of research centered on AWE, the differential effects of automated written corrective feedback (AWCF) on writing accuracy across various proficiency levels remain to be empirically clarified. This current study fills the void in the existing literature by taking a quantitative approach in exploring the efficacy of Grammarly on L2 writing accuracy with various language proficiency.*

Keywords: *Automated Writing Corrective Feedback, Grammarly, L2 writing*

1. Introduction

Writing corrective feedback (WCF) is described as empowering L2 learners with the grammatical consciousness of their written production [2][20]. The field of second language writing has witnessed a considerable rise in the literature concerning the effects of WCF in recent years. Meanwhile, the increasing power wielded by computer-assisted language learning has been allowing educators and teachers to further explore the field of second language acquisition. One emerging trend is the introduction of automated writing evaluation (AWE) systems, that is developed on the premise of Artificial Intelligence (AI) which largely enhances the accuracy and flexibility of automatic corrective feedback [1] [18] [25].

2. Literature review

It is a general belief among L2 writing researchers that instructor CF can lead to an improvement of learner's explicit knowledge as well as written performance, while some still call into its lack of immediacy[12]. Besides, instructor WCF on L2 writing is vulnerable to fluctuation mediated by individual differences, such as attitude, motivation, and anxiety [7] [10] [13]. Put differently, teacher WCF is easily affected by affective variables, such as the instructors' favoritism towards students, students' fear of making mistakes under instructors' attention, stressful relationships between the teacher and the student.

On the other hand, automated writing evaluation (AWE) is increasingly embraced by classroom-based writing instruction and routinely offered as the solution to the problem typically associated with teacher CF [26][27]. It boasts an alleviation of instructional burden and allows for quick analysis on grammar with the benefit of more room for attention on content and discourse [5][15].

Research on the relationship between L2 writing and the introduction of AWCF has been widely conducted. For example, in a study conducted by Kim, the use of Google Docs was found to be positively associated with students' writing skills[14]. Another quasi-experimental study that concerned with the effect of automatic writing assessment (AWE) that approximates ACF, found a significant superiority for students who received AWE when examining the error rate development [21]. It is confirmed, to a large extent, that researchers agree on the positive association between AWE and writing accuracy. Moreover, Van Beuningen et al. [23], Frear [8], and Li et al. [18] referred ACF as an effective tool in improving the grammatical accuracy of EFL learners, with a shift of emphasis from errors at the lexical level to syntactic level.

The limitation of AWCF, however, lies in the absence of human pedagogy, the over-correction of

ISSN 2522-6398 Vol. 4, Issue 11: 94-98, DOI: 10.25236/FER.2021.041117

errors, and the inability to address individual differences [21]. According to Wang et al. [24], a significant improvement was found in EFL learners' writing accuracy and learner autonomy awareness level with the intervention of the AWE tool, Correct English. However, regarding the process in which students responded to the provision of AWCF, 'vague' was the word that frequently appeared, which indicates places where further explanation is needed. Similarly, a study conducted by Shintani et al. [22] discredited ACF, as they found EFL learners, especially those with a lower level of English proficiency, fail to correct their linguistic errors in later writing practices.

Although confirming the efficacy of teacher WCF has always been of great interest to many L2 researchers [6] [7], the differential effects of the AWE tools on students with different proficiency levels remain to be empirically clarified [19]. Therefore, based on the discussion above, this current work seeks to examine the following questions:

(1) What is the difference between EFL learners at different English proficiency in terms of the effects of AWCF application on their writing performance?

(2) To what extent can the variable (English proficiency) mediate the relationship between the effects of AWCF and writing performance?

3. Instrument and data collection

In this study, we employed Grammarly was as the automatic feedback provider, which is known as a powerful error-correcting tool for English writing [9]. To investigate the effects of AWCF on L2 writing development, accuracy was rated using the error-free clauses ratio (EFCR), which was measured by dividing the number of error-free clauses by the total number of clauses in text. Errors were analyzed comprehensively in this study, incorporating all errors in syntax, morphology, lexical choice, spelling and so on. The errors within a clause were regarded as one error, and the same type of errors occurs in different clauses were grouped as separate errors.

4. Participants and Procedure

This study selected 24 participants at various writing proficiency levels from a language program that provides specialized training for IELTS tests. All of the participants were English language majors who come from various universities in Tianjin. The native language of the students participating in this study is Chinese, and no one has ever lived in English-speaking countries for a long time. The pre-test and post-test are performed in the same way —students are asked to write an argumentative essay according to a topic, and are we tested for the accuracy of their language by using EFCR. According to a pre-test focusing on accuracy we administered to them, they are divided into three groups: students whose scores fall into the range of 0.7-1 are considered as the high-proficiency group, whose scores of 0.4-0.7 are regarded as the intermediate-proficiency group, while students with a score of 0.4 or below are viewed as the low-proficiency group. Complete instructions regarding how to use Grammarly as a tool to assist in English writing were given throughout the whole 3months. To control the relevant variables, we do not allow teachers to give any corrective feedback on writing during the three months. Students can judge whether to accept the answers given by the tool or not.

5. Data analysis

Quantitative analyses were used to address the research questions. The quantitative data in this study was encoded in a spreadsheet, and then was analyzes with the use of SPSS software 25.0 to gain descriptive and inferential statistics. We conducted paired t-tests to determine the difference between the accuracy scores in the pretest and posttest among students from different English proficiency levels.

Results

Table 1. Low proficiency students' EFCR during the pretest and the posttest.

Low proficiency	x	SD	mean difference	t	P	Cohen's *d*
pre-test	0.46	0.13	0.02	0.509	0.626	0.180
post-test	0.44	0.14				

* p<0.05 ** p<0.01

ISSN 2522-6398 Vol. 4, Issue 11: 94-98, DOI: 10.25236/FER.2021.041117

Table 2. Intermediate proficiency students' EFCR during the pretest and the posttest

Intermediate	x	SD	mean difference	t	P	Cohen's *d*
pre-test	0.61	0.06	-0.23	-13.748	0.000**	4.861
post-test	0.84	0.05				

* p<0.05 ** p<0.01

Table 3. High proficiency students' EFCR during the pretest and the posttest.

High proficiency	x	SD	mean difference	*t*	*P*	Cohen's *d*
pre-test	0.80	0.08	-0.05	-1.871	0.104	0.661
post-test	0.85	0.05				

* *p*<0.05 ** *p*<0.01

The figure presented in Table 1 shows the results of a T-test comparing the mean scores of low proficiency students' writing accuracy measured by EFCR during the pre-test and post-test. There is a slight improvement in students' accuracy level as the mean difference -0.23 shows. The t-test results (t=0.509, p=.626; t= -1.017, p>0.05) indicate that the differences between pre-test and post-test in terms of EFCR is not significant. Thus, it is concluded that the effect of Grammarly on L2 writing accuracy is not statically significant.

As shown in Table 2, the mean difference (-0.23%) demonstrates that the accuracy of the post-test is lower than the pre-test. The t-test results (t = -13.748, P = 0.000) indicate that the difference in terms of the percentage of error-free clauses is significant for the intermediate level students. the T-test showed significant difference (P < 0.05). The specific comparative difference shows that the average value of the pre-test (0.61) will be significantly lower than that of the post-test (0.84). Hence, the findings posit the use of AWCF (Grammarly) has a significant effect on accuracy for intermediate students.

In Table 3, the mean differences (-.05) demonstrates that the accuracy of high proficiency students experienced a slight increase. In the meantime, the t-test results of accuracy (t=-1.87, p=.104) mean that the difference between the pre-test and post-test is not significant. Thus, we did not find a significant effect on accuracy for the high proficiency group.

6. Discussion and Conclusion

This study inquires into the relationship between L2 proficiency and the effects of AWCF which response to the call from Bitchener and Ferris [3] for more WCF research centered on L2 proficiency. To the best knowledge of the researcher, the investigation with regard to the use of AWE tools has been limited to a single course level [4] [16] [17]. Thus, this research fills the void in the existing literature by taking a comparative perspective on the effects of AWCF on students with various L2 proficiency levels.

Overall, the research yielded the following major findings: (i) the implementation of AWCF can positively improve learners' writing performance in terms of accuracy across three language proficiency levels: low, intermediate, and high (ii) The improvement of intermediates' written proficiency is the most pronounced with a significant difference. However, the accuracy of low and high proficiency learners did not show a significant change. The pedagogical benefits of adopting Grammarly are confirmed in this study, which lends further support to Guo et al. [11]'s finding concerning the positive role played by Grammarly in reducing L2 learners' written error. Besides, the result concerning low proficiency learners is in accordance with the argument of Lin and Griffith [19] that students with lower L2 proficiency are unable to use ACF to generate correct grammatical forms because they lack a comprehensive understanding of the target structure. As regards the non-significant difference for high-proficiency learners between the pre-test and post-test, we propose that they enjoy sophisticated grammar knowledge, which leads to limited room for improvement. Their mistakes are often appeared as a result of carelessness, in conjunction with some logical or syntactic errors that cannot be detected and corrected by the AI-powered machine. In a sense, it also points out the deficiency of the information processing ability of this software. However, caution should be exercised when generalizing results beyond the scope of the present study, given the limited sample we included and the fact that we can not rule out individual factors that may affect this study.

ISSN 2522-6398 Vol. 4, Issue 11: 94-98, DOI: 10.25236/FER.2021.041117

References

[1] Bai, L., & Hu, G. (2017). In the face of fallible AWE feedback: How to students respond? Educational Psychology, 37(1), 67–81. doi:10.1080/01443410.2016.1223275
[2] Bitchener, J., & Knoch, U. (2010). The contribution of written corrective feedback to language development: A ten month investigation. Applied Linguistics, 31(2), 193–214. doi:10.1093/applin/amp016
[3] Bitchener, J., & Ferris, D.R. (2012). Written corrective feedback in second language acquisition and writing. New York, NY: Routledge.
[4] Chen, C.-F.E., & Cheng, W.-Y.E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. Language Learning & Technology, 12(2), 94–112.
[5] Dikli, S., & Bleyle, S. (2014). Automated essay scoring feedback for second language writers:How does it compare to instructor feedback?Assessing Writing, 22, 1–17. doi:10.1016/j.asw.2014.03.006
[6] Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). Journal of Second Language Writing, 8, 1–11.
[7] Ferris, D. R., Liu, H., Sinha, A., & Senna, M. (2012). Written corrective feedback for individual L2 writers. Journal of Second Language Writing, 22(3), 307–329. https://doi.org/10.1016/j.jslw.2012.09.009
[8] Frear, D. (2012). The effect of written corrective feedback and revision on intermediate Chinese learners' acquisition of English (Unpublished doctoral dissertation). The University of Auckland.
[9] Grammarly. (2021). Grammarly: Free online writing assistant. Retrieved February 15, 2021, from https://www.grammarly.com/
[10] Goldstein, L. (2006). Feedback and revision in second language writing: Contextual, teacher, and student variables. In K. Hyland & F. Hyland (Eds.), Feedback in second language writing: Contextual and issues (pp. 185–205). Cambridge University Press.
[11] Guo, Q., Feng, R., & Hua, Y. (2021). How effectively can EFL students use automated written corrective feedback (AWCF) in research writing? Computer Assisted Language Learning, 1–20
[12] Hartshorn, K. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy? TESOL Quarterly, 44, 84-109. https://doi.org/10.5054/tq.2010.213781
[13] Havranek, G., & Cesnik, H. (2001). Factors affecting the success of corrective feedback. EUROSLA Yearbook, 1, 99–122
[14] Kim, S. (2010). Revising the revision process with Google Docs. In S. Kasten (Ed.), TESOL classroom practice series (pp. 171–177). Alexandria, VA: TESOL Publications.
[15] Lai, Y.H. (2010). Which do students prefer to evaluate their essays: Peers or computer program. British Journal of Educational Technology, 41(3), 432–454. doi:10.1111/j.1467-85352009.00959.x.
[16] Lavolette, E., Polio, C., & Kahng, J. (2015). The accuracy of computer-assisted feedback and students' responses to it. Language Learning & Technology, 19(2), 50–68.
[17] Lee, J., & Hegelheimer, V. (2012). A hybrid use of Criterion® and teacher feedback in process writing. Paper presented at the EUROCALL Conference, Gothenburg, Sweden.
[18] Li, J., Link, S., & Hegelheimer,V. (2015). Rethinkingthe role of automated writing evaluation (AWE) feedback in ESL writing instruction. Journal of Second Language Writing, 27, 1–18.
[19] Lin, S. M., & Griffith, P. (2014). Impacts of online technology use in second language writing: A review of the literature. Reading Improvement, 51(3), 303–312.
[20] Loewen, S. (2012). The role of feedback. In S. Gass, & A. Mackey (Eds.), The Routledgehandbook of second language acquisition (pp. 24–41). New York: Routledge.
[21] Ranalli, J. (2018). Automated written corrective feedback: How well can students make use of it? Computer Assisted Language Learning, 31(7), 653–674. doi:10.1080/09588221.2018.1428994
[22] Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. Language Learning, 64(1), 103–131.
[23] Van Beuningen, C. G., De Jong, N., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. Language Learning, 62(1), 1–41.
[24] Wang, Y. J., Shang, H. F., & Briody, P. (2013). Exploring the impact of using automated writing evaluation in EFL university students' writing. Computer Assisted Language Learning, 26(3), 234–257.
[25] Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. Language Teaching Research, 10(2), 157–180. doi:10.1191/1362168806lr190oa
[26] Weigle, S.C., & Malone, E. (2016). Assessment of English for academic purposes. In K. Hyland & P. Shaw (Eds.), The Routledge handbook of English for academic purposes (pp. 608– 620).New York, NY: Routledge.

ISSN 2522-6398 Vol. 4, Issue 11: 94-98, DOI: 10.25236/FER.2021.041117

[27] Yu, B. (2015). Incorporation of automated writing evaluation software in language education: A case of evening university students' self-regulated learning in Taiwan. International Journal of Information and Education Technology, 5(11), 808–813. doi:10.7763/IJIET.2015. V5.616