International Journal of Advanced Culture Technology Vol.11 No.4 42-47 (2023)
DOI https://doi.org/10.17703/IJACT.2023.11.4.42

IJACT 23-12-6

Effect of Using QuillBot on the Writing Quality of EFL College Students

Hye Kyung Kim

Lecturer, Institute of Liberal Education, Incheon National Univ., Korea
otter0219@naver.com

Abstract

The majority of research on Automated Writing Evaluation (AWE) programs has focused primarily on Grammarly, whereas QuillBot and its use in English as a Foreign Language (EFL) classrooms remains limitedly explored. This study examined the effectiveness of using QuillBot on the writing quality of college students. A total of 26 participants took pre- and post-writing tests, and four analytical tools were applied to assess their writing quality in terms of syntactic complexity, lexical diversity, lexical richness, and readability. Results of the syntactic complexity analysis across the four indices demonstrates that the syntactic complexity of EFL writing increased significantly, and substantial differences were observed in lexical richness and readability. These results suggest that QuillBot can compensate for the drawbacks of Grammarly and assist EFL writers in improving their overall writing quality.

Keywords: *Automated Writing Evaluation, AI Software, Corrective Feedback, Grammar Checker, Paraphraser, Syntactic Complexity, Lexical Diversity, Lexical Richness, Readability*

1. INTRODUCTION

The growing demand for efficient language feedback from second language (L2) teachers and learners has sparked interest in the educational application of ubiquitous artificial intelligence (AI) technology. Current learners, characterized as digital natives, tend to prefer instant corrective feedback [1]. For L2 writing teachers, providing comprehensive corrective feedback, which is usually complex and time consuming, is inevitable [2]; therefore, teachers have difficulty allocating time effectively and prioritizing feedback. Nonetheless, with the advent of AI technology, automated writing evaluation (AWE) tools have been used in EFL classrooms to alleviate teachers' workload and time constraints and facilitate students' writing.

In line with this trend, many recent studies have focused on Grammarly, a popular online grammar-checking tool that has demonstrated significant gains and pedagogical benefits [3]. However, in contrast to the scholarly focus on Grammarly, Ji noted that research on QuillBot, a potentially beneficial tool that can compensate for Grammarly's drawbacks, is lacking [4]. Specifically, researchers have only reviewed QuillBot tools, and there is scarce empirical evidence on the effects of employing QuillBot in EFL writing instruction [4,5]. In addition, it remains unclear whether QuillBot is appropriate for enhancing writing quality [4]. Thus, this study fills this literature gap by empirically investigating the effect of QuillBot on the writing quality of EFL college students in terms of syntactic complexity, lexical diversity, lexical richness, and readability.

Manuscript received: October 8, 2023 / revised: November 2, 2023 / accepted: November 5, 2023
Corresponding Author: otter0219@naver.com
Tel:*** - **** - ****
Lecturer, Institute of Liberal Education, Incheon National Univ., Korea

2. THEORETICAL BACKGROUND

As mentioned previously, Grammarly positively influences L2 writing development, such as reducing grammatical errors and improving writing confidence [1]. Recent research has also offered pedagogical considerations when implementing AWE tools and suggestions for future research. Moreover, given the nature of Grammarly's feedback, which focuses on grammar-oriented errors, Grammarly lacks support for lexical and syntactic revisions [4], implying the need for an alternative practical program. Therefore, QuillBot is an effective solution as it includes a paraphrasing tool as well as grammatical advice. Students will appreciate the immediacy of QuillBot's grammar checker if teachers instruct them on how to use the tool accurately [4]. Furthermore, QuillBot's paraphraser may allow learners to acquire and expand their vocabulary and sentence structures by exploring rephrased versions of texts. Paraphrasing skills are highly related to syntactic and lexical knowledge, which influences writing quality [4,6].

Additionally, when using Grammarly, EFL teachers take an appropriate approach and method to maximize technology use [3]. O'Neill and Russell employed techniques to encourage students to actively consider their choices and explain their reasoning, instead of passively accepting the program's suggestions [1]. This study uses journals as a tool to avoid undue reliance on AI and to critically apply the technology. The students are expected to examine the written piece critically by reflecting on feedback from the teacher and QuillBot, and then make revisions based on informed decisions, ultimately having more responsibility for the outcome of their writing.

Finally, O'Neill and Russell recommended that Grammarly be effectively used in conjunction with teacher feedback [1]. Fahmi and Cahyono argued that Grammarly cannot merely replace instructor feedback, as the students still require assistance from teachers to enrich their writing content [3]. The researchers indicated that Grammarly can be superior in terms of grammatical support, whereas teachers can contribute to content and organization. However, the consolidation of feedback types in AWE programs remains understudied. Therefore, in this study, the QuillBot tools were applied in tandem with teacher feedback.

3. METHODS

3.1. Participants and the Learning Context

Participants were 26 Korean EFL college learners enrolled in a required English course. The course was designed to develop key academic writing skills through the writing process. Prior to the study, the students were instructed in writing fundamentals, such as paragraph construction and dictionary consultation. The class comprised 20 males and 6 females of various ages and majors. None of the students had lived in English-speaking nations or had taken any college-level writing courses.

3.2. QuillBot

QuillBot (https://quillbot.com/) provides AI-generated paraphrasing and grammar check tools. The Grammar Checker provides indirect and direct feedback on grammar, typographical errors, improper word usage, and punctuation. Once the text is entered in the box, the errors are automatically underlined in red. The Grammar Checker provides direct feedback when the cursor is hovered over the underline. The paraphraser differentiates QuillBot from Grammarly. As illustrated in Figure 1, once an original text is pasted in the left-hand box and the "Paraphrase" button is clicked, the altered text appears in the right-hand box, with different colors indicating which words or sentence structures have been modified. Users can explore alternative results by hitting the "Rephrase" button several times. QuillBot also has a selection of paraphrasing modes, such as Standard and Fluency, which allow users to compare outputs in terms of text fluency and readability. Moving the Synonyms slider to the left results in a paraphrase with fewer changes, whereas moving it to the right generates a creative output with additional modifications.

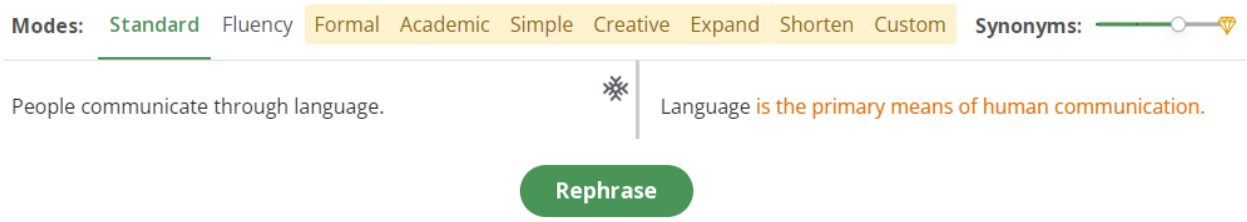

Figure 1. QuillBot-generated paraphrase

3.3. Procedure

In this study, a process-oriented writing model was adopted from Tribble [7]. After studying a writing model under their teacher's guidance, the students spent approximately three weeks writing and revising texts on one topic; four writing topics were covered throughout the semester. Figure 2 illustrates the entire writing process followed in this study. In the pre-writing stage, the students generated and organized ideas and constructed detailed outlines, followed by the second stage, which comprised writing the first draft. The students revised and edited their essays in two ways during the post-writing stage. Using a worksheet containing the criteria, the students self-reviewed the first draft while taking corrective and reflective notes on the errors or improvements they made to write the second draft accordingly.

Based on the rationale behind the compensatory intervention suggested in previous studies [3,4,8], the students received a combination of human and AI feedback for the final revision. While the teacher offered feedback on content and organization, the students received QuillBot feedback on grammar, vocabulary, syntax, and mechanics. The borders of the bottom-left cell in Figure 2 are in bold, indicating the feedback mode.

<table>
<tr><th>Stage</th><th>Teaching-Learning Activities</th><th>Learning Tools</th><th>Learning Outcomes</th></tr>
<tr><td>Pre-writing</td><td colspan="3">Creating an outline</td></tr>
<tr><td colspan="4">↓</td></tr>
<tr><td>Writing</td><td colspan="2">Writing a paragraph</td><td>Draft 1</td></tr>
<tr><td colspan="4">↓</td></tr>
<tr><td rowspan="4">Post-writing: revising and editing</td><td colspan="2">Self-reviewing
(Content, organization, language, and mechanics)</td><td>Checklist and journal</td><td>Draft 2</td></tr>
<tr><td colspan="2">Blended feedback</td><td rowspan="3">Journal</td><td rowspan="3">Final draft</td></tr>
<tr><td>Instructor</td><td>QuillBot</td></tr>
<tr><td>Content and organization</td><td>Language and mechanics</td></tr>
</table>

Figure 2. Process-oriented writing consolidating QuillBot with the instructor feedback

While improving their writing, the students maintained reflective journals to avoid passively accepting instructor and machine feedback and to fully articulate their final decisions. The students were instructed to pay close attention to the comments and errors flagged by the teacher and QuillBot and to decide whether to accept the feedback. If they accepted the feedback, they were required to report why, what, and how they revised it. Meanwhile, if they disagreed with the feedback, they were required to provide a justification.

3.4 Data Collection and Analysis

The students took 30-min identical pre- and post-writing tests at the beginning and end of the term to assess whether their writing quality had improved. They were asked to write a paragraph about a person who had made a difference in their lives in the following order: (1) pre-writing (5 min), (2) composing (20 min),

and revising (5 min).

Using corpus-based text analyzers, the collected timed writing was analyzed in terms of syntactic complexity, lexical diversity, lexical richness, and readability. L2SCA was selected to calculate syntactic complexity given that the analyzer has already been widely used by researchers, such as Lu [9] and Valizadeh and Soltanpour [10], for its reliability. Guided by Valizadeh and Soltanpour [10], 5 of the 14 indices were used in the present study, considering students' writing proficiency. The five syntactic complexity measures are presented in Table 1.

Table 1. Five syntactic complexity indices

Measure	Label	Definition
mean length of clause	MLC	number of words divided by number of clauses
mean length of sentence	MLS	number of words divided by number of sentences
mean length of T-unit	MLT	number of words divided by number of T-units
clause per sentence	C/S	number of clauses divided by number of sentences
verb phrases per T-unit	VP/T	number of verb phrases divided by number of T-units

Rather than measuring lexical diversity using a traditional index, such as type/token ratios (TTRs), this study calculated the number of types using WordSmith 4.0, as suggested by Vermeer [11]. In addition, VocabProfilers_Web VP Classic, a validated tool developed by Laufer and Nation [12], was used to gauge lexical richness, which accounts for the percentage of words the learners used at different vocabulary frequencies by dividing students' words into four vocabulary frequency levels: the top 1,000 word families (K1), the second 1,000 (K2), the Academic Word List (AWL), and words excluded from the previous lists (Off-list).

Finally, Coh-Metrix L2 readability, developed by Crossley, Greenfield, and McNamara [13] and assessed as a comprehensive predictive tool [4], was employed to measure the readability of students' written productions. Data were statistically analyzed using SPSS.

4. RESULTS AND DISCUSSION

4.1 Effectiveness of Using QuillBot on EFL Writers' Syntactic Complexity

As shown in Table 2, the post-test mean scores showed slight differences across all five measures. Paired *t*-test analysis revealed significant improvements in MLS ($t = -3.761$, $p < .05$), MLT ($t = -2.421$, $p < .05$), CS ($t = -2.545$, $p < .05$), and VPT ($t = -2.326$, $p < .05$). Overall, these results indicate that the blended feedback using QuillBot provided in this study positively affected EFL writers' compositions of syntactically more complex structures. This result is partly supported by McCarthy *et al*. [6], who demonstrated that paraphrasing ability is associated with syntactic knowledge. The output of the QuillBot paraphraser may have increased students' awareness in terms of different sentence types and patterns, allowing them to manipulate and construct complex structures in their writing.

Table 2. Paired t-tests of syntactic complexity per measure (*N* = 26)

Measure	Test	M	SD	*t*	*p*
MLS	Pre	11.18	2.15	−3.761	.001
	Post	12.75	2.35		
MLT	Pre	10.26	2.24	−2.421	.023
	Post	11.31	2.41		
MLC	Pre	7.18	1.04	−.644	.525
	Post	7.35	1.06		
CS	Pre	1.57	.27	−2.545	.017
	Post	1.74	.21		

VPT	Pre	1.76	.43	−2.326	.028
	Post	1.98	.45		

4.2 Effectiveness of Using QuillBot on EFL Writers' Lexical Diversity and Richness

To check whether there was any change in the lexical diversity level from using QuillBot, the number of types was counted, which showed a large increase from 694 words in the pre-test to 761 words in the post-test. This result indicates that QuillBot, when used in conjunction with teacher feedback, can enrich L2 writers' lexical diversity.

As seen in Table 3, lexical richness was computed, and *t*-tests were conducted to examine statistical differences. No significant mean differences for the first 1,000 most frequent words (K1), second 1,000 most frequent words (K2), and the first two combined frequency bands (K1+K2) were detected. However, the academic word list (AWL) band demonstrated a statistically significant improvement ($t = -3.289$, $p < .01$). These results suggest that blended feedback using QuillBot can promote students' productive use of academic words. The students appear to have engaged in a process of continually evaluating their word choices over time by critically assessing and benchmarking QuillBot's vocabulary recommendations and applying these lexical items in their writing practice.

Table 3. Paired *t*-tests of lexical richness per measure (*N* = 26)

Measure	Test	*M*(%)	*SD*	*t*	*p*
K1	Pre	90.22	4.82	.951	.351
	Post	89.18	3.64		
K2	Pre	3.76	2.28	−.481	.635
	Post	4.03	2.43		
K1+2	Pre	93.98	4.55	.975	.339
	Post	93.21	3.08		
AWL	Pre	1.75	1.64	−3.289	.003
	Post	2.97	1.57		
Off-list	Pre	4.44	3.84	.612	.546
	Post	3.97	2.73		

4.3 Effectiveness of Using QuillBot on the Readability of EFL Writing

The paired *t*-test result for the readability of students' writing is presented in Table 4. With an increase from 27.89 to 30.98, the students' readability scores demonstrated a significant improvement ($t = -2.234$, $p < .05$). The Coh-Metrix L2 readability tool incorporates variables that better reflect the psycholinguistic and cognitive processes of L2 texts, one of which is closely related to the text comprehension process: coherence and meaning construction [4]. Instructor assessment focused on a coherent structure, which included a logical flow of concepts and clear explanations with adequate information. Teacher feedback can result in students' sufficient use of transition signals, which can improve coherence. Moreover, QuillBot can provide feedback on inconsistent use of nouns and pronouns, increasing the likelihood of achieving coherence. Thus, the combination of QuillBot and teacher feedback may enhance the logical and structural coherence of students' writing quality, resulting in a higher readability score.

Table 4. Paired *t*-test of readability (*N* = 26)

Measure	Test	*M*	*SD*	*t*	*p*
Coh-Metrix L2 Readability	Pre	27.89	6.73	−2.234	.035
	Post	30.98	6.02		

5. CONCLUSION

This study aimed to determine whether using QuillBot in the feedback process positively influences the writing quality of EFL college learners, as measured across four dimensions. The findings demonstrate that the use of QuillBot in a college EFL classroom significantly improved students' syntactic complexity measures. Regarding lexical diversity, students' writing was found to show greater lexical diversity after intervention with QuillBot. As with lexical richness, significant improvement was observed in academic vocabulary. Finally, another paired comparison analysis unveiled that QuillBot can enhance EFL writing readability. Overall, these results suggest that QuillBot, in place of Grammarly, can be applied as a compensatory tool in feedback sessions in an EFL setting, which consequently helps students enhance their writing quality. This study holds a limitation in that it used a one-group pre- and post-test design; thus, the study results are tentative. Future research should employ a more explanatory design to elicit in-depth responses concerning student perceptions of dual-feedback activities.

ACKNOWLEDGEMENT

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5B5A17043518)

REFERENCES

[1] R. O'Neill and A. M. T. Russell, "Stop! Grammar Time: University Students' Perceptions of the Automated Feedback Program Grammarly," *Australasian Journal of Educational Technology*, Vol. 35, No. 1, pp. 42–56, Mar. 2022.

[2] L. Junqueira and C. Payant, "I Just Want to Do It Right, but It's So Hard: A Novice Teacher's Written Feedback Beliefs and Practices," *Journal of Second Language Writing*, Vol. 27, pp. 19-36, Mar. 2015.

[3] M. A. Fahmi, and B. Y. Cahyono, "EFL Students' Perception on the Use of Grammarly and Teacher Feedback," *Journal of English Educators Society*, Vol. 6, No. 1, pp. 18-25, Feb. 2021.

[4] N-Y. Ji, "Effectiveness of AI Editing Software for Corrective Feedback in L2 Writing," *English Language Assessment*, Vol. 15, No. 2, pp. 31-52, Dec. 2020.

[5] J. Park, "Implications of AI-Based Grammar Checker in EFL Learning and Testing: Korean High School Students' Writing," *English Language Assessment*, Vol. 14, pp. 11-39, Dec. 2019.

[6] P. McCarthy, R. Guess, and D. McNamara, "The Components of Paraphrase Evaluations," *Behavior Research Methods*, Vol. 41, pp. 682-690, Feb. 2009.

[7] C. Tribble, *Writing*, Oxford University Press, pp. 38-39, 1996.

[8] D. Lazic, A. Thompson, T. Pritchard, and S. Tsuji, "Student Preferences: Using Grammarly to Help EFL Writers with Paraphrasing, Summarizing, and Synthesizing," *CALL for Widening Participation: Short Papers from EUROCALL 2020*, pp. 183-189, Dec. 2020.

[9] X. Lu, "Automatic Analysis of Syntactic Complexity in Second Language Writing," *International Journal of Corpus Linguistics*, Vol. 15, No. 4, pp. 474-496, Oct. 2010.

[10] M. Valizadeh and F. Soltanpour, "Focused Direct Corrective Feedback: Effects on the Elementary English Learners' Written Syntactic Complexity," *Eurasian Journal of Applied Linguistics*, Vol. 7, No. 1, pp. 132-150, Mar. 2021.

[11] A. Vermeer, "Coming to Grips with Lexical Richness in Spontaneous Speech Data," *Language Testing*, Vol. 17, No. 1, pp. 65-83, Jan. 2000.

[12] B. Laufer and P. Nation, "Vocabulary Size and Use: Lexical Richness in L2 Written Productions," *Applied Linguistics*, Vol. 16, No. 3, pp. 307-322, Sep. 1995.

[13] S. Crossley, J. Greenfield, and D. McNamara, "Assessing Text Readability Using Cognitively Based Indices," *TESOL Quarterly*, Vol. 42, No. 3, pp. 475-493, Sep. 2008.