OTESSA
PROCEEDINGS

Comparison of English Teacher Feedback and Automated Writing Feedback on the Quality of English Language Learners' Essay Revision

Feifei Han (iD)
Griffith Institute for Educational Research
Griffith University

Zehua Wang
School of Foreign Languages
Shaanxi Xueqian Normal University

Correspondence:
Feifei Han
Office of Pro-Vice-Chancellor (Arts, Education and Law), Griffith Institute for Educational Research, Griffith University

Email: feifei.han [at] griffith.edu.au

Abstract
This study compared the effects of teacher feedback (TF) and online automated feedback (AF) on the quality of revision of English writing. It also examined the strengths and weaknesses of the two types of feedback perceived by English language learners (ELLs) as a foreign language (FL). Sixty-eight Chinese students from two English classes participated the study. The two classes received TF and online AF (Scoring Network) respectively upon completion of their draft essays. While the two classes did not differ on the English writing proficiency, the class receiving TF obtained significantly higher scores on essay revision, indicating the better effect of TF. Students' responses showed that, overall, TF was more positively commented upon because the encouraging words motivated students to revise. In contrast, the students receiving online AF criticized the Scoring Network for their difficulty to comprehend the feedback they were provided. The results suggest that English teachers may consider using TF as a major source of feedback in English writing for ELLs in China.

Keywords: teacher feedback, automated writing feedback, English language learners, Chinese university students, essay writing, quality of revision

Introduction

Hyland and Hyland (2006) maintain that feedback plays a crucial role in students' writing processes due to its potential to improve the quality of writing through revision. With the advancement of information and communication technology, there is a boom of automated writing feedback (AF) software. AF is able to provide feedback on students' writing in various aspects, including mechanics, spelling, vocabulary, collocation, and grammar. Hence, the application of AF in writing classes may reduce teachers' workload in terms of providing feedback (Warschauer & Ware, 2006). In this sense, AF seems to be especially suitable as a pedagogical tool for college English teachers in China, because they often have to teach multiple large classes with more than 100 students in total. It is unrealistic for teachers to regularly provide detailed feedback on students' English writing. While the online AF platform enjoys popularity amongst English teachers to fully or partially replace teacher feedback (TF) in the writing classes, limited research has compared the effects of online AF and TF on the quality of revision in FL writing classes (Stevenson & Phakiti, 2014). Hence, the first aim of the current study was to conduct a quasi-experiment to compare the effects of online AF versus TF on the quality of revision in English writing amongst English language learners (ELLs) in China. Moreover, from a pedagogical point of view, it is important to know how students perceive the strengths and weaknesses of online AF and TF in their revision processes so that students' opinions can be considered by teachers to make wise decisions as to whether and how to use online AF in College English teaching in China. Investigation on students' perceptions of strengths and weaknesses of AF and TF formed an additional research aim of our study.

The present study sought to answer two research questions:

- To what extent did the quality of revision differ between students receiving TF and receiving online AF?
- What were the strengths and drawbacks of TF and online AF perceived by Chinese ELLs?

Methods

Research Design

The study adopted a mixed-methods design. A quasi-experiment was used to answer the first research question, whereas a qualitative method, which collected the data through an open-ended questionnaire, was used to answer the second research question.

The Participants

A total of 68 Chinese first year students, who majored in English Education in Early Childhood, participated in the study.

Instruments

The Writing Tasks

The first writing task ("my first day in the university") examined if the students in the two classes had the same level of English writing proficiency. The second writing task ("my ideal job") asked the students to complete an initial draft first and revise their texts according to either TF or AF

they received. Both writing tasks required students to produce an English text that was at least 150 words long.

The Online AF Platform

The online AF platform used in this study was called "Scoring Network" (http://www.pigai.org/). It is a cloud-based online service for evaluating English writing by Chinese ELLs.

The Open-Ended Questionnaire

The open-ended questionnaire asked students to list up to three aspects of strengths and weaknesses of either TF or AF depending on which type of comments they received.

Data Collection Procedure

The quasi-experiment was conducted in the three English lessons. In the first English lesson, participants in the two classes completed and submitted their essays entitled "my first day in the university" to Scoring Network, which calculated scores as students' English writing proficiency. The results of the two-sample t-test showed that there was no significant difference of the English writing proficiency, $t(2, 66) = -0.12$, $p = .42$, Cohen's $D = 0.03$. In the second English lesson, both groups drafted a response to "my ideal job." The students assigned with the AF option wrote their essays into Scoring Network directly. The students assigned to TF handed in their essays to their English teacher. The English teacher spent approximately one week to complete giving the feedback. In the third English class, the students were instructed to revise their essays using either TF or AF. They were given 30 minutes to complete their revision. After the English teacher collected the students' revised essays, they were given 10 minutes to complete the open-ended questionnaire for their perceptions of the strengths and/or drawbacks of the feedback they had received. All the revisions were scored by Scoring Network to represent the quality of the revision.

Data Analysis

To answer the first research question, we used the two-sample t-tests to compare the revision scores of the two classes. To answer the second research question, we conducted the thematic analyses of students' responses to the open-ended questionnaire.

Results and Discussion

Revision Quality of Students Receiving TF and AF

The two-sample t-test showed that the quality of revision of students who received TF ($M = 77.79$, $SD = 3.46$) was higher than that of students who received AF ($M = 75.31$, $SD = 5.78$), $t(2, 66) = 2.15$, $p < .05$, Cohen's $D = 0.52$). Because the students in the two classes did not differ in terms of their English writing proficiency as shown by their essay writing scores in the pre-test, the significantly better essay scores of the revised texts of the students receiving TF seems to suggest that TF was more effective in helping students revise their essays.

Perceptions of TF and AF

The participants had mixed feelings towards both TF and AF. The most frequently mentioned strength of TF was that TF had balanced comments on both the positive and negative aspects of the writing. The second frequent strength of TF was about the encouraging words used by

the English teacher. The third frequently mentioned strength towards TF was clarity and easiness of the language by the English teacher.

When looking at the positive comments on AF, we found that the participants predominantly focused on AF's ability of providing corrective feedback. The second most frequently mentioned strength of AF was that it provided suggestions for synonyms and detailed explanations of the differences between the synonyms. The next most frequently mentioned strength of AF was that it ranked students according to their writing performance in relation to others in the class. This allowed the students to know their writing ability. While it is undeniable this function is useful and attractive, it does not seem to be closely related to the revision processes.

In terms of the shortcomings, only a few students mentioned some issues in TF. Four students believed that the English teacher sometimes did not give detailed explanations as the teacher simply underlined the sentences or highlighted the words. Four students commented that there were too many comments made by the English teacher. In contrast, as many as 21 students pointed out that AF was not always straightforward and comprehensible. They mentioned that the comments by AF were often quite ambiguous, which created barriers in the revising process. The second most frequently mentioned weakness of AF was that AF emphasized too much on the mechanic problems, such as punctuations and capitalization problems. The qualitative responses from the students seem to reflect that in general students hold more positive perceptions towards TF than towards AF, which might affect students' revision processes.

Authors' Contributions

F.H. and Z.W. conceived of and designed the research together. Z.W performed the experiments while F.H. analyzed the data and interpreted the results of the experiments. F.H. and Z.W. drafted, edited, and revised the manuscript, and approved the final version of the manuscript.

Open Researcher and Contributor Identifier (ORCID)

Feifei Han https://orcid.org/0000-0001-8464-0854
Zehua Wang

Ethics Statement

Before the data collection, ethical application was submitted to the ethics committee of School of Foreign Languages, Shaanxi Xueqian Normal University. The committee evaluated the study and deemed that the research would be part of classroom teaching. Hence, a written consent procedure was not required. However, all participants orally agreed to take part in the research on a voluntary basis. "All participants orally agreed to take part in the research on a voluntary basis" was documented by the ethics committee.

Conflict of Interest

The authors do not declare any conflict of interest.

Data Availability Statement

The data presented in this study are available on request from the corresponding author. The data are not publicly available due to potential ethical issues.

References

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching, 39*(2), 83–101. https://www.cambridge.org/core/journals/language-teaching/article/abs/feedback-on-second-language-students-writing/2A58FAA099F7628B459AE677F7098493

Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing*, *19*, 51–65. https://doi.org/10.1016/j.asw.2013.11.007

Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language Teaching Research, 10*(2), 157–180. https://doi.org/10.1191/1362168806lr190oa