EasyChair Preprint

№ 173

From Feedback to Revision–How Can Collaborative Peer Feedback Make a Difference?

Lu Huang

May 25, 2018

From Feedback to Revision—How Can Collaborative Peer Feedback Make a Difference?

Lu Huang
South China Normal University
Guangzhou, Guangdong
lhuang1@scut.edu.cn

Abstract: Automated writing evaluation(AWE) has been used more and more often in ESL writing classrooms as it can provide instant feedback in addition to an automated score. However, many studies reveal that AWE alone might not fulfill its potential in helping students to improve writing due to its limitations. This study aims to find out how effectively AWE can be used in ELT writing classroom when collaborative peer feedback is incorporated into it.

A total of 80 EFL undergraduate students in a university in south China participated in the study and were randomly assigned into either the experimental or control group. Participants in the experimental group conducted mobile-assisted collaborative peer feedback after getting their AWE feedback while students in the control group just used AWE system autonomously. A paired t-test indicated that collaborative peer feedback combined with AWE led to increased revision and positively impact students' revision types and quality of texts. The experimental group's revision significantly improved in terms of both holistic and analytic measurement. Further interview with students also reveals that students perceived comparatively more favorably when the collaborative feedback was conducted to facilitate students' revising process. This research suggests a implementation of collaborative peer feedback incorporating into AWE system which has pedagogical implications for EFL instruction.

Key Words: Collaborative peer feedback, AWE

I. INTRODUCTION

In Chinese universities, writing competence is regarded as one of the most important components of language learning and students need it in order to succeed in both academic and social settings.

Providing feedback on writing has been widely acknowledged to be able to improve students' writing abilities. Various forms of feedback has been employed in teaching practice: teacher written feedback, oral conferencing, peer feedback or combination of teacher and peer feedback.

With the development of technology, Automated Writing Evaluation (AWE) has been designed to provide feedback. It is becoming more and more sophisticated and able to evaluate students' writing at lexical, syntactic and discursive levels in addition to offering automated scores. Pigaiwang, a web-based AWE system has been increasingly used in Chinese universities because it "saves teachers time in assessing writing, enables more writing practice,

and supplements writing instruction" (Roscore et al, 2017).

However, both researchers and teachers are skeptical about the instructional and assessment values of AWE. A major concern is that AWE does best in identifying errors—particularly errors in grammar, mechanics and spelling and "by diverting students' attention from content development to formal attributes, AWE may lead to negative effects on students' writing behavior" (Cheville, 2004)

However, other researchers have exactly the opposite opinion and claimed that form-focused feedback was a necessary component to L2 writing instruction(Hyland 2000). Interestingly, teachers' understanding of what constitute useful feedback may in fact run counter to students' understanding. Many students have expressed a preference for error correction in their essays. (Ferros. 1999). Several studies revealed that students primarily expect to get feedback on form, though they have an interest in receiving feedback on content as well .

Anyway, a more pressing questions is not whether AWE should be used but how this new technology can be used to achieve more desirable learning outcomes while avoiding potential harms (Chen & Cheng 2008) the pedagogical effectiveness of using technology in various learning contexts. In addition to the research on how AWE can be redesigned to tackle its limitations, more classroom-based research on pedagogical practices with AWE is certainly needed.

Incorporating peer feedback has become a common practice in both L1 and L2 writing classrooms. Most ESL students appear to enjoy peer feedback and find it helpful. Peer feedback can be provided more frequently and more quickly than feedback by the instruction (falchikov and Goldfinch 2000) and can foster collaboration skills, communication skills and personal responsibility (liu and Careless 2006). More importantly, Miao, Badger & Zhen(2006) reported that peer feedback brought higher meaning changes in revisions.

The concerns with AWE system, together with the lack of evidence of the use of AWE feedback in ESL writing classrooms and its impact on students' writing, prompted the current study. Specifically, the present study aims to examine the use and the impact of the automated feedback through a naturalistic classroom-based approach and explore the pedagogical design of incorporating collaborative peer feedback with AWE in an ESL writing curriculum.

The motivation for incorporating collaborative peer feedback stemmed from the following assumptions: 1) Implementation of AWE requires pedagogical designs to be beneficial ; 2) Collaborative peer feedback can make up for the drawback of AWE 3) writing instruction should address all aspects of the writing process, including grammar, content, word choice, etc.

In this study, we are interested in how students engage with both peer and AWE feedback in the process of revision and we set out to answer the following questions:

1 How does collaborative peer feedback impact students' revision after the use of AWE?

2 How students' perception of the effectiveness of AWE were affected by the use of collaborative peer feedback?

II. LITEREATURE REVIEW

Automated writing evaluation (AWE) has been under development since the 1960s and gained more and more attention since the mid-1990s. With the advance of artificial intelligence, in particular natural language processing and intelligent language tutoring systems, AWE systems, such as Criterion by Educational Testing Service and MY Access! by Vantage

Learning are able to conduct sophisticated analyses including lexical complexity, syntactic variety, discourse structures, grammatical usage, word choice, and content development. They provide immediate scores as well as diagnostic feedback. In addition, major AWE systems are now web-based and equipped with a variety of online writing resources and editing features which make them writing assistance tool for students.

With the increasing use of AWE in writing classes, there have been large amount of research concerning the validity of AWE as an assessment tools for both formative and summative assessment purposes. For the purpose of this study, we focus mainly on AWE's role as formative assessment tool. For ESL writing, formative feedbacks support process-writing approaches that emphasize the need for multiple drafting through a scaffold of prompts, explanations, and suggestions. (Hyland 2003) By providing scores and diagnostic feedback on various rhetorical and formal aspects of writing for every essay draft submitted to their scoring systems, students can then use the assessment results and diagnostic advice to revise their writing as many times as they need.

Contrary to expectations of teachers and students, automated feedback is reported to be of limited help during students' revising process. Students complain that automated feedback provides formulaic comments and generic suggestions for all the submitted revisions. Many studies indicate that AWE is useful only for the revision of formal aspects of writing but not of content development (Cheville, 2004; Grimes & Warschauer, 2006; Yang, 2004;). Students generally spend as little time as possible on automated feedback and their revision of writing focus mainly on spelling, punctuation, and grammatical errors. Attali (2004) reported that over two thirds of the students in his study did not produce a second draft after receiving AWE feedback.

In spite of the above problem, classroom-based AWE study found two main benefits of using AWE: increased motivation to practice writing for students and easier classroom management for teachers. （ Grimes and Warschauer 2006 ） For students, getting immediate response was a strong motivator to practice writing. AWE programs are at least valuable as a supplement to writing instruction (Shermis & Burstein, 2003b; Ware, 2005; Warschauer & Ware, 2006). Therefore, classroom-based naturalistic study is needed to investigate how AWE can be implemented in writing instruction. The present study aims to fulfill this purpose by exploring the use of collaborative peer feedback incorporating into AWE. .

Socio-cognitive view of writing regard it as a communicative, meaning-making act. Writing requires not only linguistic ability for formal accuracy but, more importantly, meaning negotiation with readers for genuine communicative purposes. Computers, no matter how intelligent, are not real "readers" and no genuine, meaningful communication is likely to happen. Peer feedback, on the contrary, has potential for exchanging reactions, questions, and responses from authentic readers (Leki, 1990, Mittan, 1989). Furthermore, researchers have claimed that participating in peer feedback may help build critical skills which in turn may assist in analyzing and revising one's own writing (Leki, 1990b)

Despite the positive findings of peer feedback, both researchers and teachers are cautious about incorporating a peer feedback approach. For one, students may misunderstand the purpose of peer feedback and may even feel uncomfortable about the process. (Zhang 1995). Students may feel inadequate in providing feedback as they are still in the process of learning English. Therefore, the present study use collaborative peer feedback instead of individual written peer feedback.

The purpose of this study is to explore the factors affecting students' intention to revise their draft after getting the feedback and how AWE can be used to achieve more desirable learning outcomes while avoiding potential harms.

. III. RESEARCH DESIGN AND DATA COLLECTIONS

The participants in this study were two classes of 80 Chinese undergraduate students majoring in architecture in a major university in South China. Students were studying College English as an integrated course of listening, reading, speaking and writing.

Students were given assignments of writing through Pigaiwang, an AWE system run by Juku Company. The system is currently used by many universities in China. Similar to other AWE system like Criterion and My Access, Pigaiwang can provide students with holistic and analytic scores along with corrective feedback that covers various aspects of writing. It highlights errors and provide brief suggestions for correction. The system allows for multiple revision and editing. Students can revise their essays multiple times based on the analytic assessment results and diagnostic feedback given to each essay draft submitted to the program. Fig.1 is a screenshot of feedback from Pigaiwang, illustrates how errors are highlighted and explanations and suggestions for corrections are provided in the system. Each student can access their own accounts and detailed explanations for grammar and language use can also be found in their accounts.

One class of students served as control group and the other class served as experimental group. For the control group, students worked with the AWE autonomously, submitting their essay drafts and revising their writing as many times as they need based on the automated scores and feedback they receive for each draft. For the experimental group, however, after submission and getting the feedback, students are asked to post their essay drafts in Shimo Document, a tool of collaborative writing, discuss and review each draft. The students then had to revise their essays based on the AWE's feedback along with their peers' comments before submitting second draft.


Far away from the ancient times when man first took shape, a bond between man and environment had come into being.（到现在应该已经形成完了）
As time goes by, man's dependence on the environment has become increasingly apparent. That man was ruled by the natural disasters in the ancient times gradually develop into the situation that the natural resources are assumed by man. Though the main role（？） on earth has transformed,man must be alive along with what environment applies.（读不通，中式英语？）
Nevertheless, so rapidly does man develop that the environment can't control it（does 有必要用吗） within its capability.When man is aware of the fact that they are doing harm to the nature, the nature has been deeply injured.
Why do some countries set up laws to protect animals?Because some animals are dying out due to man's destroying.Panda is our treasure.According to surveys, there are less than 1600 pandas on earth,in which there are 333 living in China.（用法有点奇怪）This is owing to the habitats lessened.
Not only animals,but also the（删） climate gets problems.As it is,people produce millions of tons CO2（会写 carbon dioxides 更好）and other harmful gases,which do harm to the air and result in the temperature increasing.What is worse,because of the higher degree,the sea level rises,flooding a lot of cities by the seaside.
So many problems have occurred that We need to take measures.Development is necessary,but during the time, we are supposed to consider how to decrease the damage to the environment.We can produce habitats for animals and take green traffic tools such as bicycles to reduce fuel-burning.Government can also make strategies to protect the environment.More ways to protect the earth should be dug out.

总体思路很清晰，中心突出，段落过渡自然
运用举例论证，使得文章更有说服力
短语及从句运用丰富
后面部分用了很多冠词 the，有些 the 似乎不是必要使用的应删去，否则显得重复累赘了。
最后一段简单句比较多，应穿插一两句复杂句。
有些词出现了很多次如 due to，可以用同义词进行替换。

Fig. 1. Sample of feedback from Pigaiwang

Research on effect of AWE feedback on student writing either examine the written products or the writing processes such as editing time, rates of revisions, etc. To serve the purpose of the study, we focus not only on students' written products but also students' writing and revising process.

Both qualitative and quantitative data were collected in the study. The data included the students' responses to the questionnaire made by the researcher, interviews with the students and the students' writing samples along with the scores and feedback generated by Pigaiwang and the peers.

Qualitative data, semi-structured interviews with students, were collected in order to examine students' use of Pigaiwang and their perceived usefulness and problems. Quantitative data, including the number of submissions for each paper recorded by the system and AWE error reports, were collected to analyze the influence of AWE feedback on the students' writing practice and the change of accuracy on the paper drafts.

The questionnaire was conducted online from Sojump, a website specializing in making questionnaires, and was distributed to students who could respond either by computer or from their mobile devices. The questionnaire is mainly about the students' perceived effectiveness of Pigaiwang for writing improvement, with a primary focus on the adequacy and helpfulness of its automated scores, comments and advice. The questionnaire contained both multiple-choice questions using a Likert scale and open-ended questions. In total, 75 out of 80 students responded to the questionnaire.

To explain the questionnaire results and clarify the learning process with the AWE program, two focus group interviews with the students from each class were conducted. The students participating in interviews were all volunteers: ten from experimental group and ten from control group, Each interview lasted approximately 30 minutes and was conducted in Mandarin Chinese. The interviewees were asked to talk about how Pigaiwang was used in their writing classes, how they felt about the value of the program, and what factors affected their perceived effectiveness of using the program. All the interviews were audio-taped and then transcribed for Nvivo 9.0 to code. In addition, the students' writing samples along with their AWE scores and feedback in their online accounts were use.

Concurrently, we collected quantitative data from the AWE system that include the number of submissions for each paper and the number of errors identified and classified by Pigaiwang in the students' first and final drafts. The Pigaiwang reports on the first and the final drafts were used to compare changes of accuracy in writing.

We analyze the quantitative data using SPSS 20.0. Descriptive statistics were reported , and paired sample t-tests were performed to investigate the differences between experiment group and the control group. To answer the question about the change of accuracy, we calculated the average number of errors identified by Pigaiwang on each draft. We also compared the participants' first drafts, second revised drafts with collaborative peer feedback were scored by Pigaiwang

IV. RESULTS AND DISCUSSIONS

By examining the emerging themes of students' interviews, we identified the major themes highlighting students' perceptions in terms of their expectations of the AWE system, their satisfaction of using it, positive and negative view of Pigaiwang and confidence in using the feedback and the improvement after using it.

In the control group, 2 out of 10 students expressed a high level of satisfaction and 5 students were not satisfied. All 5 students who expressed positive responses commented on the helpfulness of the AWE feedback. In the experimental group, 4 out of 10 students expressed a high level of satisfaction and 2 students were not satisfied. All 8 students who expressed positive responses commented on the helpfulness of the AWE feedback. Specifically, students in both groups frequently commented on how AWE feedback can be helpful for them to resolve grammatical and spelling issues in writing and can motivate them to revise.

However, students' perception on the usefulness of the feedback on organization were different in two groups. 8 out 10 students in control group felt AWE feedback on organization not helpful and need more clarification and guidance. However, only 3 out of 10 students in experimental group think that they need further guidance in terms of organization.

A possible reason why students feel that they need more guidance is that Pigaiwang only offers reminders for revision by highlighting the sentences that should be the thesis statement or the topic sentences. For students who are stilling struggling in using more sophisticated

organization in their writings, such feedback might not be clear and sometimes even confusing. And this knowledge about organization of the writing might be supplemented by discussion from the peer feedback.

TABLE I. OVERALL PERCEIVED EFFECTIVENESS OF USING PIGAIWANG

Variable	Control group	Experimental group
Very helpful	2(20%)	4（25%）
Moderately helpful	2(20%)	3（30%）
Slightly helpful	1(10%)	1（10%）
Not helpful	5(50%)	2（20%）

To examine the writing improvement after collaborative peer feedback, first drafts and the second drafts in the experimental group were measured holistically and analytically. As can be seen in Table 2, the mean scores of the second drafts (3.85) were higher than the first draft (3.29) when analyzed holistically. A further paired t-test shows significant difference between the two drafts. The findings reveal that collaborative feedback played an important role in improvement of second drafts. In terms of analytic scores, the mean scores of the second drafts were also higher than the first drafts in all traits. Significant differences were also found in all aspects. The results indicated that the experimental group with collaborative peer feedback improves not only in form but also in content.

TABLE II. IMPROVEMENT OF EXPERIMENTAL GROUP

	First drafts(SD)	Second Drafts(SD)	MD	t	P
Holistic	.60	.55	0.58	-9.00	.000
Content	.61	.68	0.76	-9.42	.000
Organization	.70	.63	0.66	-7.85	.000
Voice	.52	.62	0.72	-7.88	.000
Sentence Fluency	.58	.59	0.61	-6.34	.000
Word choice	.62	.55	0.38	-4.30	.000
Conventions	.65	.53	0.49	-5.45	.000

The mean differences between the two drafts in terms of content, i.e. content and organization were higher than form, i.e. sentence fluency, word choice and convention. These results confirmed the findings that peer feedback can bring higher meaning changes in revision (Miao, Badger &Zhen 2006). For the experimental group, participants' first and second drafts were compared in order to find out whether the revision after collaborative peer feedback in the second drafts improve significantly in terms of content and organization.

Example one illustrates the revision of content in the second drafts.

The first draft

Provided that our ancestors came to the society nowadays, they will surely find a strange phenomenon: people smile or cry in front of a shining brick, plug colorful strings into their ears, and cross roads with their heads down...

The second draft

If our ancestors could travel to our world, they would find some really strange phenomenon—people smiling or crying in front of a shinning brick, plugging colorful strings into their ears and crossing road without looking up.

Example 2 illustrates how organization was revised after peer feedback

The first draft

After smart phone was born, the efficiency of our life had a rapid increase. We can see and chat with each other through a small screen without distant and dull transportation. A multitude of citizens go shopping without wallets in their pockets, for they can pay by scanning codes on the smart phones. Music, videos and games are no longer things that can only be enjoyed on computers or televisions.

The second draft

Smart phones make our life more efficient than ever before. We can chat online instead of traveling long distance. We can go shopping without bringing money. We can entertain ourselves wherever we are.

V. CONCLUSIONS

AWE, like any existing technology used in teaching, requires good pedagogical ideas and strategies for its implementation in writing classes. In writing instruction, assessing student writing and providing quality feedback are essential, yet they are often complex in nature and challenging to teachers. When AWE is used, this job does not necessarily become easier for teachers but may become more complicated, requiring more technological competence in working with AWE and more careful pedagogical designs to integrate AWE into writing instruction. ((Chen&Cheng 2008)

Clearly, collaborative peer feedback can help provide feedback on content development and organization while AWE system can provide feedback on grammar and mechanics. Integrating collaborative peer feedback in AWE in writing instruction can have strong impact on students' perception of how the AWE system can help them during revision. Quite a few students stated that they are happy to use AWE feedback to correct basic grammar errors, but they need further help on larger issues such as organization and rhetorical strategies

To include students in the grading process, students can actually observe and learn from each other when writing. They actively engaged in the writing process, which enables them to observe and learn from peers in writing as well as support peers in making text revisions.

Overall, collaborative peer feedback proves effective in ESL instruction. Combined with other forms of feedback, especially AWE, it will hopefully be able to provide timely and manageable feedback that will help students to improve their quality of writing as well as facilitate language learning as a whole. Writing teachers need to be fully aware of the limitations of AWE technology and make pedagogical choices to maximize effectiveness of AWE and to minimize its undesirable results.

ACKNOWLEDGMENS

I would like to thank South China University of Technology for funding the project under identification number TIN2010-21695-C02-01. This research work is made within the university.

REFERENCES

[1] Attali, Y. (2004). Exploring the feedback and revision features of Criterion. Paper presented at the National Council on Measurement in Education in San Diego, CA. April, 2004.

[2] Cheville, J. (2004). Automated scoring technologies and the rising influence of error. English Jour 93(4), 47-52.

[3] Chi-Fen Emily Chen and Wei-Yuan Eugene Cheng (2008) Beyond the design of automated writing evaluation: pedagogical practices and perceived learning effectiveness in EFL writing classes, Language Learning & Technology, June 2008

[4] Grimes, D., & Warschauer, M. (2006). Automated essay scoring in the classroom. Paper presented at the American Educational Association in San Francisco, California. April 7-11, 2006.

[5] Rod D. Roscoe, Joshua Wilson, Adam C. Johnson, Christopher R. Mayra 2017. Presentation, expectations, and experience: Sources of student perceptions of automated writing evaluation, Computers in Human Behavior (Jan. 2017)

[6] Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. Language Teaching, 39(2), 83-101..

[7] Leki, I. 1991. The preferences of ESL students for error correction in college-level writing classes, Foreign Language Annals. 24. 203-218

[8] Shermis, M. D., & Burstein, J. (2003b). Introduction. In M. D. Shermis & J. Burstein (Eds.), Automated Essay Scoring: A Cross-disciplinary perspective (pp. xiii-xvi). Mahwah, NJ: Lawrence Erlbaum.

[9] Truscott, J 1996 The case against grammar correction in L2 writing classes, Language Learning. 46 327-369

[10] P Ware, P. (2005). Automated writing evaluation as a pedagogical tool for writ Pandian, G. Chakravarthy, P. Kell, & S. Kaur (Eds.), Strategies and practices for improving learning literacy (pp. 174-184). Selangor, Malaysia:

[11] Yang, N. D. (2004). Using MyAccess in EFL writing. The proceedings of 2004 International Confere and Workshop on TEFL & Applied Linguistics (pp. 550-564). Taipei, Taiwan: Ming Chuan University.

[12] Yu-Fen Yang. 2011. A reciprocal peer review system to support college students' writing, British Journal of Educational Technology, 08 June 2011.