ATLANTIS PRESS
Advances in Social Science, Education and Humanities Research, volume 378
6th International Conference on Education, Language, Art and Inter-cultural Communication (ICELAIC 2019)

The Effectiveness of Online Automatic Evaluation on College English Writing*

Meigen Yu
School of Foreign Languages
China West Normal University
Nanchong, China

***Abstract*—Online Automatic Writing Evaluation has been developed for more than 50 years and studied by many scholars and researchers. This paper discusses its effectiveness: the application of AWE can stimulate students' writing motivation and writing enthusiasm, promote them to improve the content, organization structure and the correction of mistakes in grammar and spellings, improve their writing ability, and free teachers from teaching work. On the other hand, it has some limitations: its feedback mainly focuses on shallow language forms such as words, grammar and sentence structure, and lacks feedback on the cohesion of text content, deep structure and deep semantics. For the improvement of students' writing, it is suggested to use the combination of multiple feedback or evaluation methods.**

Keywords—online automated evaluation; effectiveness; writing

I. INTRODUCTION

College English writing is an important index to measure students' comprehensive English level and application ability. Although writing is stressed and students write more, student's writing ability even hasn't improved as it is expected. In China, the current college English teaching situation is that teachers have a heavy teaching task. Most of the basic English teaching stages are taught in large classes. The traditional EFL writing adopts the product approach, not process approach. The feedback to students' writing is scores or short comments by teachers. As a result of some difficulties in reviewing students' compositions in detail, teachers can not give effective feedback to students' compositions. What's more, there are few overall evaluations of students' compositions, but only with one score or grade, lacking of sustained evaluation (Lu Yanhui, et al, 2010). Considering this situation, more and more universities and colleges employ various automatic writing evaluation systems to reduce teachers' stress in scoring students' compositions (Elena Cotos & Sarah Huffman, 2013).

The system of Automatic writing evaluation (AWE), also called automated essay scoring (AES), began its development in the United States in the early 1960. In North America, many subject courses are taught in large classes. The large scale of students makes essay writing review a very difficult problem. In this context, the United States, Canada and other countries have developed some computer-aided peer review systems, such as Peer Mark, Calibrated Peer Review (CPR), Peer Review from A to Z for Education (PRAZE) and Peerceptiv (Xu Jingfen, Zhu Xi, 2019). For more than forty years, AWS or AWE has been studied abroad, especially in America. In recent years, the research about AWE has gradually gained the attention in China (Liang Maocheng, Wen Qiufang, 2007; Lu Yanhui, etal, 2010; Liu Rongjun, et al, 2013; Zhong Caishun, 2013; Yang Xiaoqiong, Dai Yuncai,2019; YuYang, et al, 2019; Xu Jingfen, Zhu Xi, 2019; Li Guangfeng, 2019).

*Fund: Elite Project of China West Normal University: Research on Oral Grammatical Error Correction (Project Number, 17YC438).

II. AUTOMATED ESSAY SCORING SYSTEM

The history of Automated Essay Scoring (AES) or Automated Writing Evaluation (AWE) can be traced back to the 1960s in America with the development of Page Essay Grade (PEG), a program that used multiple regression analysis of measurable features of text, such as essay length and average sentence length, to build a scoring model based on a corpus of essays previously graded by hand (Page, 2003). With the widespread use of computers and the Internet, Artificial intelligence and natural language processing technologies have developed quickly, the 1990s witnessed the emergence of more AWE systems such as E-Rater, Intellimetric and Intelligent Essay Assessor (Warschauer & Ware, 2006). AWE systems as Intelligent Essay Assessor (IEA), IntelliMetricTM and Electronic Essay Rater (E-rater) developed to reduce teachers' workload by automating the scoring of student essays, analyzed the quality of texts by examining language at the surface level (Page, 2003). Modern AWE software, such as Criterion (Educational Testing Service), MY Access! (Vantage Learning), and Intelligent Essay Assessor (Pearson Knowledge Technologies), employ natural language processing techniques to enable more complex analyses of writing for performance-specific feedback. These products' scoring and feedback affordances are promoted as being capable of meeting the needs of L2 learners, writing teachers, and institutional administrators (Cotos & Huffman, 2013).

Automated Essay Scoring system uses computer technology to evaluate and score the compositions. Its use in foreign countries is earlier than that in China. Several factors suggest why automating scoring might be desirable: (i) practicality: essay grading is costly and time-consuming; (ii) consistency: essay grading is somewhat subjective in nature,

and consistency may sometimes suffer; and (iii) feedback: Providing feedback to a student is important, and automated scoring can provide ways of generating specific suggestions tailored to the needs of the author (Lonsdale & Strong-Krause, 2003). The application of AES system is based on the two goals: first, automatic scoring for large-scale examinations; second, writing teaching as a feedback tool. In China, "Pigai. org" software is used widely in universities and colleges. And a lot of scholars and researchers have done many studies about the effectiveness on the application of "Pigai. org." software. Even though AWE system has been widely employed abroad and in China, its effectiveness has been the focus of a heated debate.

III. Related Research on the Effectiveness of AWE

In 2004-2005, Warschauer & Grimes (2008) conducted a mixed-methods exploratory case study by using AWE software to learn about how AWE is used in classrooms and how that usage varies by school and social context. By interviewing, questionnaire and classroom observation of four middle school students, the study found that the application of AWE software engaged students in autonomous activity while freeing up teacher time. Instead of sitting idly at the end of a writing session, faster writers were engaged in revising and resubmitting for higher scores while slower writers continued to work on their first draft. Also, almost all the revisions addressed mechanics, not content or style. In our observations, virtually all the revisions we saw students making were of spelling, word choice or grammar, not content or organization. Attali (2004), Leah Rock (2007), Warschauer & Grimes 2008) found after using AWE software, learners showed improvement mainly at the level of grammar and mechanics.

The feedback of automatic scoring system can obviously stimulate students' interest in writing, enrich writing content and improve language quality (Cai Jigang，2011). Zhou Yishu (2013) studied 60 Sophomores who used multiple feedback methods. He found that the feedback of the network feedback group was relatively evenly distributed, especially in the text and the content of the article, which was significantly more than that of the peer feedback group. Because of its convenient operation mode and rich network resources, the network feedback played an important role in improving the structure of the article, enriching the content and modifying the words, but it also played an important role in the details of language expression (such as spelling, grammar, etc.). However, such aspects as spelling and grammar did not get enough attention. Jiang Yuhong (2005) thinks that online peer evaluation modification has great influence in promoting students to improve the content, layout, organization structure, style and the correction of grammatical errors. Yin Xiaojuan and Jia Yonghua (2015) conducted a study on 136 non-English majors who used Pigai network and found that students' writing improved significantly after using online automated evaluation system for one academic year. Yang Xiaoqiong and Dai Yuncai (2015) conducted a study on 135 non-English majors with different English proficiency who used the Pigai network to write. They found that the autonomous writing teaching mode based on the Pigai network can effectively improve the overall level of students' English writing, stimulate students' English writing learning motivation and improve their self-efficacy, and reduce students' writing anxiety. The study by Chen Feng et al (2010) showed that students like online timely feedback and comments so that they can make repeated modifications according to the online prompts until they are satisfied with their scores. This also shows that the online feedback effectively stimulates students' enthusiasm for writing, increases the frequency of composition modification, improves students' vocabulary and grammar level in writing, helps students improve their writing level, and effectively reduces teachers' teaching burden. On the other hand, due to the technical characteristics of the software itself, it can give students more help at the level of vocabulary and grammar, but it has limited help at the level of text structure, content logicality and coherence.

The previous studies proved that more research about AWE system is effective in evaluating students' writing, but it also has its limitations. Zhang Hua's (2019) study showed that automatic feedback is mainly helpful for the improvement of writing ability in terms of shallow language forms such as words, grammar and sentence structure, and lacks feedback on the cohesion of text content, deep structure and deep semantics, so it needs manual feedback to make up for the defects of automatic feedback. Teachers may realize AWE can help them reduce the heavy work burden but can not completely replace them. There are still some problems in autonomous writing modification based on online automatic feedback. Zhong Caishun(2015) found: first, learners' effective modification is very small. Most of the learners only pay attention to the spellings, wording and simple grammar of words, and can not go deep into the text and content; second, learners' motivation of self-directed writing modification is not strong, and they should be driven by utilitarian goals and lack of sustainability; third, learners cannot effectively use the modification tips provided by online system.

IV. Conclusion

As discussed above, AWE system is widely used and has been studied for many years. The application of AWE can stimulate students' writing motivation and writing enthusiasm, promote them to improve the content, organization structure and the correction of mistakes in grammar and spellings, improve their writing ability, and free teachers from teaching work. Also, it has some limitations. For example, AWE mainly focuses on shallow language forms such as words, grammar and sentence structure, and lacks feedback on the cohesion of text content, deep structure and deep semantics. For the improvement of students' writing, it is suggested to use the combination of multiple feedback or evaluation methods (such as teacher's feedback, peer feedback and AWE).

REFERENCES

[1] Attali, Y. (2004). Exploring the feedback and revision features of Criterion. In Proceedings of the National Council on Measurement in Education (NCME), San Diego, CA.

[2] Cai Jigang. (2011). A contrastive study of online peer feedback and online teacher feedback on Chinese college students' English Writing [J]. Foreign Language World. No. 2 : 65-72.

[3] Chen Feng, Xiang Peng & Yang Caiyu. (2010). An Analysis on the Metacognitive Strategy Training Effects for Students at Different Mitivation Levels — A Study Based on the Authentic Context[J]. Journal of East China University of Technology (Social Science).No.4 : 360 －364.

[4] Elena Cotos & Sarah R. Huffman, 2013. Learner Fit in Scaling Up Automated Writing Evaluation. International Journal of Computer-Assisted Language Learning and Teaching, 3(3), 77-98, July-September 2013.

[5] JiangYuhong (2005). The role of online peer review in learners writing ability development [J]. Foreign Language Teaching and Research.No.(3) : 226－230.

[6] Leah Rock, J. (2007). The impact of short-term use of CriterionSM on writing skills

[7] in ninth grade (Research report). Princeton, NJ: Educational Testing Service.

[8] Li Guangfeng. (2019). The impact of the integrated feedback on students' writing revision based on the AWE[J]. Foreign Language Education. No. 4:72-76.

[9] Liang Maocheng, Wen Qiufang. (2007). A Critical Review and Implications of Some Automated Essay Scoring Systems[J]. Technology Enhanced Foreign Language Education. No. 117:18-24

[10] Liu Rongjun, Wang Na & Zou Yanxun. (2013). An Empirical Study on Application of Network Technology in Teaching of College English Writing[J]. Modern Educational Technology. No.8: 81-86.

[11] Lonsdale, D. & Strong-Krause, D. Automated Rating of ESL Essays. http://wing.comp.nus.edu.sg/~antho/W/W03/W03-0209. pdf

[12] Lu Yan-hui, Tan Fumin & Peng Shun. (2010). Intelligent Writing Grading System in

[13] College English Writing: An Empirical Study[J]. Modern Educational Technology .No.6: 56-58.

[14] Page. E.B. 2003. Project Essay Grade: PEG. In Mark D. Shermis and Jill C. Burstein, editors, Automated Essay Scoring: A Cross-Disciplinary Perspective. Lawrence Erlbaum, Mahwah, NJ.

[15] Shermis, M. D. & J. Burstein. Automated Essay Scoring: A Cross-disciplinary Perspective [M]. Mahwah, NJ: Lawrence Erlbaum Associates, 2003.

[16] Warschauer, M. & D. Grimes. 2008. Automated writing assessment in the classroom [J]. Peda-gogies: An International Journal 3: 22-36.

[17] Warschauer, M. & Ware, P. 2006. Automated essay scoring, defining classroom research agenda. Language Teaching Research, 10(2), 1-24.

[18] Xu Jin-fen, Zhu Qian. (2019). A Review of the Research on Peerceptiv: A

[19] Computer-Assisted Peer Review[J]. Technology Echanced Foreign Language Education. No.186:10-16.

[20] Yang Xiaoqiong, Dai Yunca. (2015). An Empirical Study on College English Autonomous Writing Teaching Model Based on www.pigai.org [J]. Technology Enhanced Foreign Language Education. No. 162: 17-23.

[21] Yin Xiaojuan, Jiao Yonghua. (2015). A Study on the Reliability of Online Autonomous Scoring System in China — Exemplified by Juku Correcting Network [J]. Journal Of Minjiang University. No. 6：72-78.

[22] Yu yang, Yu Tao & Wang Xinghu. (2019). A study on the English Writing Teaching of Pigai network based on big data[J]. Journal of Northeast Agricultural University(Social Science Edition). No.2:82-88.

[23] Zhang Hua. (2019). The Effect of Man-Machine Multiple Feedback on Improving the Quality of English Writing Texts[J]. Technology Enhanced Foreign Language Education. No.186:34-39.

[24] Zhong Caishun. (2015). Research on automatic writing modification based on online automated feedback[J]. Journal of PLA University of Foreign Languages. No. 4: 82-88.

[25] Zhou Yishu. (2013). A contrastive research on feedback in teaching college English writing [J]. Foreign Language World. No.3: 87-96.