Artificial Intelligence Promotes the Evolution of English Writing Evaluation Model

Zheqian Su¹, Liu Miao¹ and Jiang Man²*
¹Changchun University of Chinese Medicine Foreign Language Department, Changchun, 130117
²Changchun University of Chinese Medicine Planning and Finance Division, Changchun, 130117
*18146034@qq.com

Abstract: On the background of “Internet +” era, the domestic higher education is showing the trend of artificial intelligence. The reliability and scientificity of computer intelligent evaluation are further carried out, and the mode of intelligent evaluation and data analysis in optimizing the precise teaching of English writing is explored, which can lay a foundation for the large-scale use of the technology. Based on data-driven theory, this study further analyzed the role of AI in promoting in-depth learning by comparing AI writing review model with manual review model.

Key words: Artificial Intelligence; English Writing; Evaluation Model

1. Introduction
The research on second language learning tools has a long history, from the initial phonograph English phonetics learning to radio news, from the promotion of cassette recorders in traditional English classrooms to situational writing teaching in bilingual laboratories[1], from the application of multimedia technology in English writing to AI-assisted reading in the context of mobile-interconnected era, the ever-changing information technology plays a revolutionary role in the process of second language acquisition, which also makes the mode of second language writing more and more intellectualized and humanized.

2. Research Background
2.1. Concepts of “Artificial Intelligence”
Artificial Intelligence (AI) is an interdisciplinary subject that simulates human intelligence. It mainly simulates and reproduces human thinking process[2]. The term “artificial intelligence” was first introduced at the Dartmouth Society in 1956. Four scholars, John McCarthy, Marvin Minsky, Allen Newell and Herbert Simon, then defined “Artificial Intelligence” as a discipline. Its main task is to establish the theory of intelligent information processing, and then design some computing systems that can show some approximate human intelligent behavior.

2.2. Application of Artificial Intelligence in English Writing Review
For more than half a century, Artificial Intelligence (AI) has been widely used in all parts of the world and in various fields, especially in second language intelligent writing and evaluation. The application of AI in the field of writing can be divided into four stages: one is to make courseware in high-level language; the other is to make courseware with writing system as work; the third is to apply multimedia network information technology in a large area; and the fourth is to combine expert system with AI program application [3].

The application of AI in the field of writing review has also experienced the following three stages: the first stage focuses on the development and application of the primary intelligent evaluation system. In the 1960s, Ellis Page and others in the United States developed the world's first automatic Writing Evaluation (AWE) system, Page Essay Grade (PEG), which only provided the score of the composition without any other form of feedback. The second stage focuses on the intermediate intelligent feedback. In the 1990s, on the basis of the AWE system, there appeared a feedback evaluation system featuring scanning, evaluation, diagnosis and scoring. However, due to the lack of computer corpus and other factors, there were many problems in the surface of feedback information [4]. The third stage is the advanced intelligent network evaluation system. At present, the Automated Essay Grading (AEG) is widely used in the world. The system plays an important role in the evaluation of TOEFL examinations as well as the manual evaluation. In addition, Writing Roadmap developed by McLaugher, Peer Mark system launched by Turnitin, Calibrated Peer Review system of the University of California, Peerceptiv system developed by the University of Pittsburgh and Peer Scholar system developed by the University of Toronto in Canada have been widely used in the process of English writing review[5]. This kind of review system is based on big data technology, which can comprehensively evaluate the author's linguistic skills, and at the same time examine the author's cognitive ability, and provide real-time feedback [6].

In China, the development and application of AI in the second language writing evaluation system started relatively late, until the end of 2013, China officially ushered in the era of 4G network. According to “the China Internet Education Trend Report 2013-2014”, the use rate of mobile learning among the major groups was more than 80%. By the end of 2014, the total number of educational APPs in China had exceeded 70,000. With the opening of mobile-interconnected 4G, the application of artificial intelligence in second language learning has been gradually promoted. Unlike the traditional manual marking, the real-time marking and feedback system for compositions embedded in mobile phones has been recognized by teachers'universities. Take “Correcting Network (Sentence Cool)” as an example, it is based on corpus and cloud computing and trains scoring engine. It automatically corrects students' compositions according to the scoring standards of CET-4 ,CET-6, National Matriculation Test and High school Entrance Examination, gives the scoring and gives the scoring in grammar and vocabulary. Feedback hints should be given where the expression is not standardized and suggestions for students'revision should be given.

3. Review of Domestic and Foreign Research

3.1. Research Trends at Home and Abroad

By searching the keywords of “Artificial Intelligence” and “Writing” on CNKI, we found that 285 articles were collected in the whole network, especially the most concentrated ones published after 2013, showing an upward trend as follows: 2013 (3 articles), 2014 (7 articles), 2015 (22 articles), 2016 (34 articles), 2017 (71 articles), 2018 (102 articles). A deep search of the subject words of “Artificial Intelligence” and “English Writing” reveals that a total of 12 articles are included in CNKI.

Through searching the Scopus English academic literature database for topics, abstracts and keywords “AI” and “writing”, 50 English academic documents were collected from 2000 to December 2018. The overall research shows an upward trend, including 2013 (1 article), 2014 (6 articles), 2015 (2 articles), 2016 (3 articles), 2017 (7 articles), 2018 (12 articles). Through visual subject analysis, it is found that “Artificial Intelligence News Writing” or “Robot News Writing” occupies the mainstream, from which we can see that the news media have begun to lay out “Artificial Intelligence Robot”
instead of “Artificial” in the industry for news reporting.

From the perspective of current research at home and abroad, the research on the application of artificial intelligence (AI) in second language writing mainly focuses on the following three directions: the relationship between AI and the psychological dynamics of second language writers, the relationship between AI and the achievement of second language writing, and the relationship between AI and the teaching of second language writing.

3.2. A Study on the Relationship between AI and the Psychological Dynamics of Second Language Authors

Some domestic scholars have done some research on teaching practice by taking the “Correcting Network” as an example. It is found that AI system can automatically improve information updating and pushing according to users’ usage characteristics[7]. At the same time, based on large data theory and corpus construction, this kind of intelligent marking system can complete multi-dimensional quality feedback of vocabulary, grammar and content for L2 writers’ individualized writing, thereby eliminating L2 writers’ quality feedback. The author’s anxiety and conflict in the process of writing (Tang Jinlan et al., 2012; Wang Ying et al., 2012).

Hua Lulu, Chen Lin and Sun Mengmeng(2017) believe that AI system can provide individualized and situational learning resources for L2 idiom writers, avoid the uniformity of feedback information in the process of reviewing, and play an effective role in stimulating the author's learning motivation in the process of L2 writing [8]. Other domestic scholars believe that AI assessment system can stimulate students’ potential interest in English writing teaching and promote second language writers to shift from shallow learning to deep learning [9]. At the same time, AI system based on cloud computing and speech recognition technology can also monitor the changes of the ability and mood of second language writers, and provide teachers with first-hand learning analysis data by generating relevant data and graphs [10].

Some foreign scholars have deconstructed AI review system from the perspective of cognitive psychology. AI relies on big data to make the traditional static review process become a dynamic information intensive feedback process. Through monitoring the author’s writing ability and emotional changes, the diachronic data tracking is completed [11]. Schunn et al. (2016) pointed out that the current North American Writing Peer Assessment Computer Aided System (CAIS) can provide real-time feedback on the writing motivation and quality of L2 writers, and complete diachronic information interaction and follow-up survey by combining background database.

Some scholars have found that the current AI review system is not perfect. They believe that some online English composition review websites only provide vocabulary and grammar improvement suggestions in information feedback, but do not provide relevant examples. To some extent, this will cause secondary anxiety for second language writers (Shi Xiaoling, 2012; Jiang Yan, Ma Wulin, 2013; Zhang Li, 2013).

3.3. The Relationship between AI and Second Language Writing Achievement

At present, AI rating system is different from traditional manual scoring and computer closed scoring. It evaluates and analyses students’ second language writing in an open form. The effective application of AI rating system is positively correlated with writing achievement and writing motivation [12]. Huang Jing & He Huaqing (2018) found that AI intelligent feedback had a positive effect on students’English writing quality, especially in vocabulary [13].

3.4. The Relationship between AI and Second Language Writing Teaching

Scholars at home and abroad generally believe that traditional second language writing is affected by the mother tongue environment to a certain extent, which is often neglected by manual review. AI evaluation system can evaluate writing ability according to the social, linguistic and cultural environment in which the author lives [14-15], which makes the teaching of second language writing more humanized and provides a more effective communication bridge between teachers and...
students[16]. Dikli & Bleyle (2014) conducted a study of AI reviews and found that L2 writers have a positive attitude towards the feedback provided by AI system. According to Siddhartha Ghosh (2019), AI assessment systems such as AEG can correct the target language deviation based on the cognitive characteristics of second language learners, avoid the influence of mother tongue on second language writing and provide more authentic corpus usage.

Some domestic scholars have compared the differences between traditional English composition marking and AI marking: traditional manual marking occupies a large amount of teachers’ personal time, and can not make personalized learning plans according to individual problems in the process of marking; AI marking system makes teachers more effective in the teaching process, and teachers can target teaching according to the data summarized by the intelligent system. Learning intervention to achieve the goal of accurate teaching [17]. Through empirical research, Wu Yong and Zhang Wenxia (2016) found that L2 learners are more inclined to adopt teacher feedback. The feedback provided by AI is only for reference. Teacher feedback combined with AI feedback can help stimulate learners’ autonomous learning consciousness.

From the point of view of domestic and foreign research, the research on the application of AI evaluation system in writing only focuses on the two methods of questionnaire survey and interview, few studies involve text analysis (Huang Jing, Zhang Wenxia, 2014; Wu Yong, Zhang Wenxia, 2016).

4. An Empirical Study of AI Evaluation System in English Writing

Second language writing is a complex psychological cognitive process, which is influenced by many factors. For second language idiom writers, writing anxiety is a negative emotional experience in the process of second language writing [18]. It has a significant negative correlation with writing ability and writing achievement[19]. Nowadays, based on big data and cloud computing, AI Intelligent Writing Review System can control the second language writers’ learning style, cognitive tendency, learning strategies, motivation types and humanistic accomplishment. It not only reduces the writing anxiety of second language learners to a certain extent, but also stimulates the writers' writing interest and motivation [20]. In order to verify the effectiveness of AI assessment system in English writing, this study, based on data-driven theory, adopts the method of stratified cluster random sampling to select 2017 undergraduates (N=200) from a Medical University in China. In the process of College English writing, the experimental group adopted the marking system of AI, while the control group adopted the traditional form of teacher’s manual marking. The study lasted for one semester and 10 writing exercises were conducted. There was no difference between the experimental group and the control group in terms of writing topics, writing time and writing requirements. The full score of the composition was 15 points. The evaluation process followed the unified scoring criteria of CET-4. In this study, FLCAS and the self-designed College English Writing Competence Scale were used to measure the two scales. The Cronbach coefficient of the former is 0.87 and the Cronbach coefficient of the latter is 0.79, which shows that the two scales have good internal consistency. The scoring criteria were Likert 5-point scale, and the data were analyzed by SPSS 22.0.

Table 1: Independent sample t-test of English writing achievement, writing anxiety and writing ability between the experimental group and the control group in the post-test

|                          | Number | Mean value | Standard Deviation | t       | Sig.  |
|--------------------------|--------|------------|--------------------|---------|-------|
| **English Writing Achievements** | 100    | 10.56      | 3.82               | 5.49**  | 0.031 |
| Experimental Group       |        |            |                    |         |       |
| Control Group            |        | 8.94       | 2.77               |         |       |
| **English Writing Anxiety** | 100    | 3.06       | 0.21               | 7.17*** | 0.047 |
| Experimental Group       |        |            |                    |         |       |
| Control Group            |        | 4.12       | 0.36               |         |       |
| **English Writing Ability** | 100    | 4.22       | 0.31               | 4.96**  | 0.019 |
| Experimental Group       |        |            |                    |         |       |
| Control Group            |        | 3.54       | 0.46               |         |       |

*P<0.05; **P<0.01; ***P<0.001
Before the beginning of the experiment, two groups of subjects were tested on their writing performance, writing ability and writing anxiety. It was found that there was no significant difference between the two groups (p>0.05). After a semester’s teaching experiment, independent sample t-test showed as in table 1 that the English writing ability and writing achievement of the experimental group were significantly higher than those of the control group (p<0.01), while the writing anxiety of the experimental group was significantly lower than that of the control group (p<0.001).

Through paired sample t-test (table 2), the writing performance of the experimental group in the post-test was significantly higher than that in the pre-test (p< 0.01), and the writing ability of the experimental group was also significantly higher (p<0.01). Growth (p<0.05), the writing anxiety of the experimental group in the post-test was significantly lower than that of the pre-test (p<0.01).

Table 2: T-test of paired samples of English writing achievement, writing anxiety and writing ability between the experimental group and the control group in the pre-and post-test

| Rank | Mean value | Standard Deviation | t   | Sig. |
|------|------------|--------------------|-----|------|
| **Experimental Group** |            |                    |     |      |
| English Writing Achievements |            |                    |     |      |
| pretesting | 8.24 | 3.09 | 5.77** | 0.000 |
| post-test | 10.56 | 3.82 |          |      |
| **Control Group** |            |                    |     |      |
| pretesting | 8.13 | 2.61 | 1.76 | 0.036 |
| post-test | 8.94 | 2.77 |          |      |
| **Experimental Group** |            |                    |     |      |
| English Anxiety |            |                    |     |      |
| pretesting | 3.45 | 0.44 | 6.31** | 0.026 |
| post-test | 3.06 | 0.21 |          |      |
| **Control Group** |            |                    |     |      |
| pretesting | 3.42 | 0.41 | 1.91 | 0.269 |
| post-test | 4.12 | 0.36 |          |      |
| **Experimental Group** |            |                    |     |      |
| English Ability |            |                    |     |      |
| pretesting | 3.24 | 0.70 | 4.37** | 0.017 |
| post-test | 4.22 | 0.31 |          |      |
| **Control Group** |            |                    |     |      |
| pretesting | 2.97 | 2.61 | 1.83 | 0.176 |
| post-test | 3.54 | 0.46 |          |      |

*P<0.05;**P<0.01;***P<0.001

Pearson correlation analysis shows in table 3 that the effectiveness of the use of corrective web is positively correlated with English writing achievement and writing ability (p<0.01), negatively correlated with writing anxiety (p<0.01), and positively correlated with writing achievement (p<0.05), writing ability (p< 0.05), and has negative correlation with writing anxiety (p<0.01); the sub-dimension of feedback adoption was significantly positively correlated with English writing achievement and writing ability (p<0.01); and the sub-dimension of convenience was only positively correlated with English writing achievement (p< 0.05).

Table 3: Pearson correlation analysis of the effectiveness of the use of marking network with English writing achievement, writing anxiety and writing ability

|                           | Effectiveness of Net Usage | Duration | Feedback to Adopt | Convenience of Use |
|---------------------------|----------------------------|----------|-------------------|-------------------|
| **English Writing Achievements** | 0.937**               | 0.204*   | 0.547**           | 0.202*            |
| **English Writing Anxiety**       | -0.238**             | -0.229** | -0.166            | -0.089            |
| **English Writing Ability**       | 0.890**              | 0.216*   | 0.460**           | 0.126             |

*P<0.05;**P<0.01;***P<0.001
Therefore, it can be seen from the analysis that the effective use of the marking network promotes English writing achievement and writing ability, and also inhibits writing anxiety. The longer the marking network is used, the greater the improvement of students’ English writing achievement and writing ability, and the lower their writing anxiety to a certain extent. The more they are adopted, the better their writing performance and writing ability will be. The convenience of using the marking network can also promote the improvement of writing performance to a certain extent.

5. Epilogue
With the further popularization of artificial intelligence in the field of education, College English writing teaching has also ushered in the “two-way liberation” of teachers and students: on the one hand, teachers release more time through AI evaluation system, which can better observe students’ learning process and characteristics, formulate more detailed analysis of learning conditions and feedback of opinions; on the other hand, students can continuously improve through AI evaluation system. Self-writing ability, through machine review and feedback, finds the joy and confidence of writing in the process of repeated revision. Introducing artificial intelligence into college English writing teaching can not only deeply integrate teaching and learning, create a new mode of second language writing, but also stimulate the internal motivation of second language learners, fundamentally improve the quality of English teaching and effectively release the power of information technology.

ACKNOWLEDGMENTS
Our thanks to the projects supported by Research Project of the National Administration of Traditional Chinese Medicine Higher Education “Implementing University Study and Research”: Research on the Theory and Practice of the Teaching Reform of “New Model of Apprenticeship System” in Colleges and Universities of Chinese Medicine (No. GZ-10); Scientific Research Project of Jilin Vocational and Technical Education Association in 2018: Study on the “New Model of Apprenticeship System” in Chinese Medicine Vocational Colleges in Jilin Province (No. 2018XHZ024); Project of Social Science Foundation of Jilin Province in 2019: Investigation on Cultural Identity of Traditional Chinese Medicine in Jilin Province and Study on Promotion Path (No.2019B89); 2018 Changchun University of Chinese Medicine curriculum thinking education and teaching reform project.

References
[1] Hu Jiasheng, Jin Yan. Theoretical and Practical Research on the Integration of Educational Technology and Foreign Language Curriculum [J]. Audiovisual Education in China, 2015 (4): 114-120.
[2] Luckin R, Holmes W. Intelligence Unleashed: An argument for AI in Education [EB/OL].https://www.pearson.com/news/blogs/CompanyBlog/2016/03/intelligence-unleashed-an-argument-for-ai-in-education.html, 2017-07-20.
[3] He Gaoda, Luo Zhongmin. Application of Artificial Intelligence in Foreign Language Teaching [J]. Foreign Language Audiovisual Teaching, 2008 (5): 74-80.
[4] Stevenson, M., Phakiti, A. The Effects of Computer-generated Feedback on the Quality of Writing [J]. Assessing Writing, 2014 (19): 51-65.
[5] A study on the mutual evaluation system of Christian D. Schunn. Peerceptiv to assist students in English writing [J]. Foreign Language Audiovisual Teaching, 2016 (8): 38-44.
[6] Siddhartha Ghosh.Online automated essay grading system as a web-based learning (WBL) tool in engineering education [M].Web-based Engineering Education: Critical Design and Effective Tools, 2010:53-62.
[7] Yang Xiaqiong, Dai Yunqian. Practical Research on the Teaching Model of College English Self-directed Writing Based on Correction Network [J]. Foreign Language Audiovisual Teaching, 2015 (2): 17-23.
[8] Hua Lulu, Chen Lin, Sun Mengmeng. Research on Artificial Intelligence Promoting English
[9] Zhang Qi. A Review of the Research on E-Learning Environmental Deep Learning in China [J]. Modern Distance Education, 2014 (5): 60-67.

[10] Liu Qingtang, Mao Gang, Yang Lin, Cheng Yun. Development and Prospect of Intelligent Teaching Technology [J], China Audiovisual Education, 2016 (6): 8-15.

[11] Wardle, E. & Roozen, K. Addressing the Complexity of Writing Development: Toward an Ecological Model of Assessment [J]. Assessing Writing, 2012 (2): 106-119.

[12] Yan Zhiming, Tang Xia Xia, Qin Xuan, Zhang Fei, Duan Yuanmei. The connotation, key technologies and application trend of educational artificial intelligence (EAI) -- Analysis of reports of "Preparing for the Future of Artificial Intelligence" and "National Strategic Plan for Artificial Intelligence Research and Development" [J]. Journal of Distance Education, 2017 (1): 26-35.

[13] Huang Jing, He Huaqing. Research on the influence of human-computer feedback on students' writing behavior [J]. Foreign Language Audiovisual Teaching, 2018 (12): 19-24.

[14] Behizadeh, N. & Engelhard Jr. G. Historical View of the Influences of Measurement and Writing Theories on the Practice of Writing Assessment in the United States [J]. Assessing Writing, 2011 (3): 189-211.

[15] Slomp, D. H. Challenges in Assessing the Development of Writing Ability: Theories, Constructs and Methods [J]. Assessing Writing, 2012 (2): 81-91.

[16] Deane, P. On the Relation between Automated Essay Scoring and Modern Views of the Writing Construct [J]. Assessing Writing, 2013 (1): 7-24.

[17] Gan Ronghui, He Gaoda. Learning Analysis and Foreign Language Teaching Research Prospects in the Big Data Era [J]. Foreign Language Audiovisual Teaching, 2016 (3): 40-45.

[18] Wu Yuhong, Gu Satellite. An Empirical Study on Cooperative Learning to Reduce English Writing Anxiety of Non-English Majors [J]. Foreign Language and Foreign Language Teaching, 2011(6):51-55.

[19] Guo Jidong. Effects of Writing Self-concept and Self-efficacy on College Students’ English Writing Achievements: Mediating Role of Writing Anxiety [J]. Journal of Foreign Languages, 2018 (2): 69-74.

[20] Wang Zhe, Zhang Yue. Big Data Approach and Innovative Research in Foreign Language Teaching: A Case Study of Millions of English Writing with the Same Topic in 2015 [J]. Foreign Language Audiovisual Teaching, 2015 (5): 3-8.