A Study of Applying Automated Assessment in Teaching College English Writing Based on Juku Correction Network

Yusheng Wang (✉)
Nanjing College of Information Technology, Nanjing, China
`Wangyusheng@njcit.cn`

Abstract—With the continuous advancement of modern network technology, the drawbacks of the traditional English writing course teaching mode has become increasingly prominent, and the automated scoring system has gradually been used in the writing course. This paper proposes a college English writing teaching model based on Juku Correction Network, and we have conducted empirical research on the use of Juku Correction Network in college English writing teaching. The research results show that the teaching model based on Juku Correction Network can effectively improve the overall level of students' English writing, and stimulate students' English writing motivation.

Keywords—Juku Correction Network, English writing, writing ability

1 Introduction

In the traditional classroom-based writing teaching model, students lack individualized learning, passively waiting for teachers to assign the composition, correct it and wait for the score. However, due to the heavy teaching tasks of college English teachers and the time-consuming nature of the laborious task of correcting compositions, teaching of writing is also a major challenge for the teachers in China. The result is that the training of student writing is insufficient, and the teacher's feedback is not timely, leading to the lack of motivation to learn English among many Chinese vocational college students [1]. Therefore, it is necessary to find a novel way to overcome the stagnancy. The new way lies in integrating the advantages of modern network technology into the traditional teaching mode and creating a teaching mode for writing suitable for foreign language learners that is conducive to improving students' writing ability and stimulating their motivation.

The purpose of this study is to establish a college English writing teaching model based on Juku Correction Network to improve students' English writing ability. This study involves two major questions: Can this correction network effectively help improve students' English writing level? What is the students' evaluation of the correction network?

2 Research Background of Automated Evaluation System

In recent decades, as corpora and the internet have become powerful tools for assisting English teaching, the automated scoring system has gradually been used in writing course. Automated Essay Scoring (AES) is the use of computer technology to evaluate and score essays. [2] The development of the automated writing evaluation system began in the United States in the 1960s. Its representative is Page Essay Grade (PEG) developed by Professor Ellis Page of Duke University in the United States.

The working of automated scoring involves the following steps. First, a certain number of manually reviewed essays are collected as samples, the text is processed, and a small corpus is established; secondly, the essays in the corpus are analyzed, and classified according to different score segments. Shallow text features of different categories of composition, for example, lexical density (an important indicator of lexical richness), average sentence length (an important part of sentence complexity), composition length, number of relative pronouns, change in word length, collocation, etc. are extracted; then, the above text indicators are used as the evaluation criteria for the machine to automatically score and analyze other essays. The results of the analysis are objective, but the method also has its shortcomings. First, the shallow text features are some indirect variables, which do not truly reflect the quality of the composition; second, it is difficult to analyze the deep text features. There is no semantic analysis of the composition, no analysis of the structure and content. Hence, there is no provision to recognize digressions or plagiarism. The first generation PEG (Project Essay Grader) system and the BETSY (Bayesian Essay Test Scoring System) system developed by Rudner et al. in 2002, are examples of shallow text feature analysis [3].

In view of the limitations of this shallow text feature extraction technology, more and more automated assessing systems use Latent Semantic Analysis (LSA) technology to calculate the semantic similarity by statistical methods, counting the frequency of occurrence of words in the text. The semantic similarity between the measured text and the sample is obtained. Latent Semantic Analysis is an automated scoring mechanism for the Intelligent Essay Assessor (IEA), which was developed by Pearson Knowledge Analysis Technology in the late 1990s. The IEA pays attention to the analysis of the composition content, and uses the latent semantic analysis method in information retrieval to automatically score the composition. The English essays of American high school students were automatically evaluated by the IEA, and the scores were positively correlated with the scores of the reviewers significantly. Subsequently, J. Burstein, M. Chodorrow and C. Leacock et al. proposed the E-rater system around 2000. The system uses natural language processing technology to intelligently score essays based on corpus [4].

Many studies have shown that these automated scoring systems have many advantages, such as reliability and objectivity. Nevertheless, they are mainly aimed at English native speakers. The study also shows that there is a significant difference between scoring compositions by English as native speakers and EFL learners. Based on IEA, Chinese scholars developed LEES (LSA-based EFL Essay Scorer), and conducted experiments in Chinese college students' fourth and sixth band English essays. They found that the LEES scores positively correlated with the scores of the fourth

and sixth graders slightly, and failed to achieve the significant results of the IEA. The main reason is that in English-speaking essays, no matter whether their thoughts are smooth flowing, most sentences do not have serious grammatical errors. However, the compositions of Chinese English learners, especially elementary learners, are full of various grammatical errors, not to mention all those misspellings and wrong use of collocations [5]. In the evaluation of foreign language writing, the use of language, especially in terms of syntax, takes a relatively large proportion. This makes the automated evaluation system developed mainly for English native speakers not suitable for scoring compositions by learners of English as a foreign language.

Therefore, it is extremely important to develop a set of automated systems suitable for Chinese students' composition scoring. Although Chinese research on automated writing evaluation system has started late, some scholars have begun to make active and beneficial explorations in developing an automated English writing scoring system suitable for Chinese students [6]. Liang Maocheng's "Automated Grading System for Large-scale Exam English Composition" was developed in 2005, and the "Ice Fruit English Intelligent Composition Review System" was jointly developed by Zhejiang University in 2009, and Juku Correction Network (pigai.org) was developed by Beijing Ci Network Technology Co., Ltd. in 2010. The software developed by Professor Liang Maocheng is mainly for the scoring in various English tests, mainly based on the statistical analysis of shallow features, and the scoring model has a certain distance from the practical system. In fact, the former two belong to a stand-alone version of the scoring software, lacking the support of online dictionary and corpus.

At present, the relatively stable and mature system of automated scoring is Juku Correction Network. The research and development of this system not only pays attention to the objectivity and reliability of automated composition scoring, but also considers the development of supplementary functions, integrating scores, feedback reports and writing assistant functions in one system, and becomes a handy assistant for both teachers and students. The Internet corpus-based Juku Correction Network is more suitable for Chinese learners who use English as a foreign language. Through corpus comparison, it can identify most Chinglish (Chinese English) expressions. The most common errors Chinese learners make is the use of Chinese English.

3 The Main Features of Juku Correction Network

Juku Correction Network is an online automated correction service system based on corpus and cloud computing, by calculating the distance between the student composition and the standard corpus to generate student essay scores, comments and content analysis. As a result, the students modify the essay according to the prompts of the correction network, and submit it multiple times until it reaches the satisfaction level, thus guiding and assisting students to improve their writing ability.

Juku Correction Network evaluates students' compositions comprehensively and intelligently through six dimensions: content relevance, text structure, sentence, collocation, vocabulary and fluency. It regards student compositions as a learner corpus (Learner Corpus). The scores of each essay consist of 192 sub-dimensions. By com-

paring the students' essays with the standard corpus, the measured distances are converted into scores, comments and feedback on the students' compositions through mapping. The process of correcting the entire composition of a student is similar to a teacher correcting the student's composition from multiple dimensions. This gives an overall feedback and conducts a "sentence-by-sentence comment" to point out the problem of each sentence. The system is based on Web2.0. It supports teachers to correct student essays through the network, and teachers can make further corrections based on the system prompts. Juku Correction Network scoring engine is an automated English scoring system developed by integrated natural language processing technology, corpus analysis technology and educational measurement technology [7]. The system can automatically extract the data indicators reflecting the English writing level in the composition. At present, there are 15 different categories of corpora in the correction network, and the corpus used is updated daily.

Through the "perfect composition" function, in accordance with the "sentence analysis" correction instructions, students modify the essay, submit it to the system again and again for immediate analysis. In the whole process of writing and correcting, the correction network plays the role of a teacher of writing. The whole writing process actually conforms to the process writing method advocated by Flower et al, i.e. a cycle of outlining, brainstorming, writing independently, and checking revisions [8]. Writing is a communicative activity. We should attach importance to students' initiative in the process of writing. In writing teaching, we should pay attention to good revision to help students to express their thoughts successfully through polishing and modification [9].

3.1 Sentence by sentence comments

Once the student essay is submitted, Juku Correction Network will immediately give the result of the correction (see Figure 1), and the result of the sentence review (Figure 2). The content of the "sentence by sentence" review includes "sentence error" tips, "spelling error" tips, and "suspected Chinglish", "low frequency warnings", "learning tips", "high score phrases", "collocation statistics" and "grammar check", etc. Students improve their essays according to the system tips until they are satisfied, and then accept "teacher reviews" or "peer reviews."

Fig. 1. Scoring page of Juku Correcting Network

Fig. 2. Correcting page on Juku Correcting Network

3.2 Teacher online comment

The correction network is based on the Web2.0 platform and supports the comment function. The content of the teacher's comment can be instantly displayed on the network for students to check. Teachers can click on the "reviews" at the end of each sentence to comment on this sentence, and they can make a better judgment on the student's composition. If the system does not recognize this erroneous sentence, the teacher can help the system to point out the problem of the sentence. Next time the similar errors will be identified by the system automatically. This is a very important part of the correction network and saves teachers a lot of repetitive work.

3.3 Traces of students modifications and progress

In addition to providing traces of each student's modification of the composition according to the prompts, the system also provides traces of the student's overall performance after each modification (as shown in Figure 3). The Figure 3 depicts the progress of the students' composition. The student's score was improved from the beginning of the 73 points to the final 85 points after 8 times modifications by the students. The teacher is able to follow the progress of the student in the writing process through the diagram.
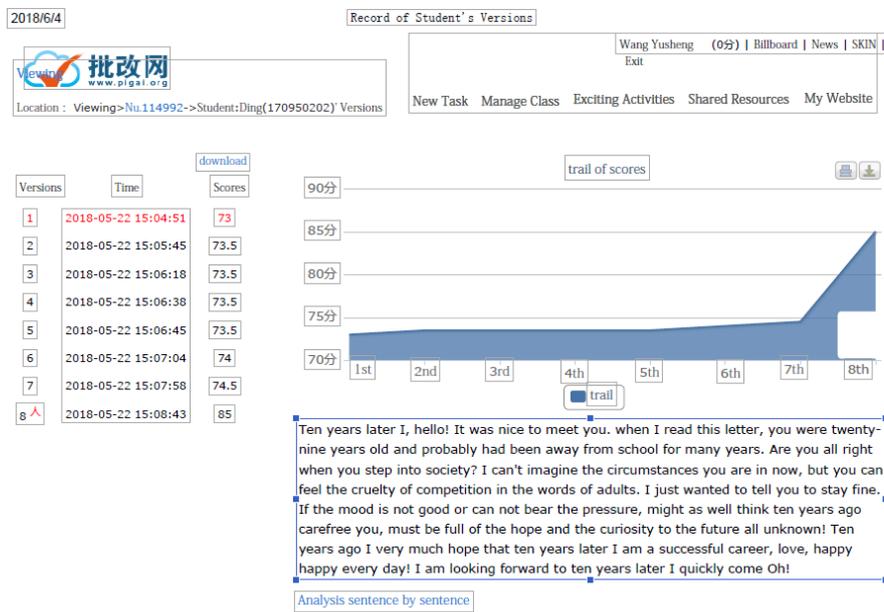

Fig. 3. Record of students' versions

3.4 Multi-dimensional analysis to fully diagnose students

The correction network selects several dimensions from the 192 assessment dimensions for judging student composition, such as "vocabulary richness", "spelling correct rate", "average word length", "average sentence length", "clause density", and list out measured values of the students in various dimensions to provide a data reference for the analysis of the composition. Through the data in each dimension, the weakness of the composition can be easily seen, and the overall situation of learning can be monitored. For composition number 1097014, the overall statistics of the task are as follows: A total of 46 essays were received, totaling 5,980 words, 382 sentences. The longest sentence was 52 words; the shortest sentence was 2 words; the average score was 76.0 points. A total of 145 errors were found in the students compositions. Among them, spelling mistakes, subject-predicate mismatches, etc. were the most common mistakes students made. In addition, the following dimensional analyses

were conducted: average word length, sentence length distribution, average sentence length, and total number of clauses. Teachers can have a much clear picture of the strength and weakness of the students instantly.

As illustrated in Figure 4, students generally use short sentences, and the number of students with 6-10 words is as high as 24%. The total number of students with sentences of fewer than 15 words is 55%.
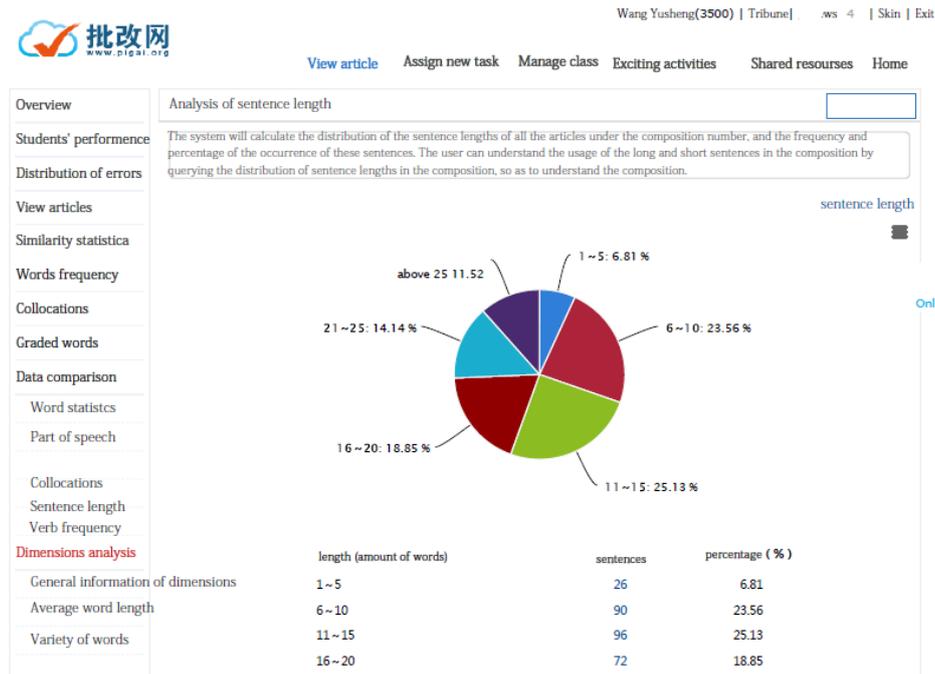

Fig. 4. Analysis of sentence length

The teacher may access any student's essay, click on "Student performance", and obtain the evaluation data of all the essays submitted by the student using the correction network, displayed in a histogram, including the error revision, the average score, the initial score and the final score of each essay. This allows the teacher to have a better understanding of the student's progress in a short time and provide appropriate assistance (see Figure 5).

Fig. 5. Students' performance

3.5 Excellent compositions sharing and forum

Sharing good compositions is an important means to improve students' writing. It helps students understand the gap between themselves and high scorers. Every registered teacher is provided with a teacher's webpage by Juku Correction Network which has three functions: "latest assignment", "excellent composition" and "shared document". Students can directly enter the writing platform by visiting the teacher's personal website, and students can also click on "Excellent composition" to check out the best essays on the same theme recommended by the teacher. Students can download the resources provided by the teacher through the "Shared document". The Correction Network has established its own writing forum (http://bbs.pigai.org/) for students to exchange their writing training experience in the forum. Teachers can interact with students here about English learning (see Figure 7) anytime and anywhere through the forum.

Fig. 6. Teacher's website

4 Experimental Study

4.1 Research objects

The experimental group (47 people) and the control group (38 people) in this study were all freshmen in Nanjing College of Information Technology. The school implements traditional teaching methods. Students were divided into five categories according to their previous qualifications. The experimental group and the control group were all students of the same category. The experiment time was the second semester of the first year.

4.2 Data analysis

This study used SPSS 21.0 to perform statistical analysis on the collected data. The researcher conducted an independent sample T test on the English writing ability and English proficiency test results before and after the experiment in the experimental group and the control group to determine the difference between the English writing level and the overall language level of the two groups of students.

4.3 Teaching experiment process

The author applied for the account in Juku Correction Network, began to assign the first essay and obtained the composition number 1049862. In the correction network, each essay assigned by a teacher has its own unique essay number. Thus the student can enter the number to find the writing task promptly. Thereafter, the students were requested to register on the network and find the essay number, and start submitting the essay online as required.

The research was conducted for one semester, which is of 15-week duration. The experimental class used Juku Correction Network to assist the English writing assessment. The teaching process included assigning the writing task, completing one essay every two weeks, submitting the essay, modifying it online and resubmitting it until they felt satisfied (the number of submissions is unlimited). The teacher browsed essays online and added manual review comments in a timely manner. The control class continued to learn English writing in the traditional teaching method. The writing instruction and writing tasks of the two classes were identical. The difference is that the experimental class submitted the essay online, and made multiple revisions and submissions according to the network correction tips while the control class could only submit the written essay once, and was marked by the teacher.

At the end of the semester, the teacher conducted a questionnaire survey on the network evaluation of the network in the experimental class and control class to obtain information about students' attitudes and opinions on the teaching of writing. The teacher also interviewed some students.

4.4 Research results

SPSS 22.0 was used to analyze the pre-experiment and post-experiment scores, and compared the differences between the two classes before and after the experiment. The purpose of the pre-experiment writing was to understand the students' writing level and provide some reference for our teaching experiment. The average score of the control class was 64.07, which was higher than the average score of the experimental class of 61.80. The standard deviation of the experimental class and the control class were 9.90 and 7.96 respectively. The independent sample T test results showed no significant difference in the pre-experiment scores of the two classes (t=-0.61, df=77, p=0.543>0.05). This indicated that the students in the experimental class and the control class had the same level of writing before the experiment, which meets the experimental requirements.

Table 1. Comparison of pre-experiment writing scores between experimental group and control group

Pre-test Comparison	Test Group		Control Group		MD	T(77)
	M	*SD*	*M*	*SD*		
	61.80	9.90	64.07	7.96	-1.21	-0.61

After one semester, the two classes were tested again. It can be seen from Table 2 that the post-test scores of the experimental class students have significantly improved, with an increase of 8.43 points. The results of the independent sample T test (see Table 2) showed significant differences in the writing performance of students using different models (t=3.13, df=78, p=0.002<0.05). The experimental class using the teaching model based on the Juku Correction Network got significantly higher scores than the students using the traditional model.

Table 2. Comparison of post-experiment writing scores between experimental group and control group

Post-test Comparison	Test Group		Control Group		MD	T(78)
	M	*SD*	*M*	*SD*		
	70.23	9.04	63.25	7.30	5.82	3.13

At the end of the semester, questionnaires and semi-open interviews were conducted for the two classes on students' opinions on English writing instructions and arrangement over the last semester. The results of the questionnaire are shown in Table 3. On the question whether the writing arrangement helps to improve the writing level, 71% of the experimental class students chose "Yes", while only 40% of the control class chose "Yes". In the experimental class 82% of students indicated an increase in their interest in writing in English while only 40% of the students in the control class shared the same sentiments. The proportion of students of "lowered anxiety in writing" reached more than 88% in the experimental class, while in the control class only 31% students reported so.

Table 3. Students' feedback on teaching of writing

Items / Groups	Improved writing skills			Increased interest in writing			Lowered anxiety in writing		
	Yes	*Hard to judge*	*No*	*Yes*	*Hard to judge*	*No*	*Yes*	*Hard to judge*	*No*
Test Group	71%	29%	0	82%	8%	10%	88%	6%	6%
Control Group	40%	60%	0	43%	7%	55%	31%	11%	58%

Students who used Juku Correction Network have greatly improved their sense of accomplishment and satisfaction in English writing. In the process of online multiple revisions, students play the role of both the reader and the teacher, and have a certain understanding of their own English writing level and problems in writing. Using the network and the provided corpus to continuously search, select, and modify the errors in the composition, the students' English writing potential is developed fully. When the grade changes from a lower score to a higher score, students feel an unprecedented sense of accomplishment and satisfaction in this process-oriented writing.

Further, the experimental class students mentioned that automated software correction can provide personalized feedback, but the content is too general, and sometimes too mechanized. This result is consistent with the findings of Chen et al. [10] on English majors. Although there is a quick evaluation of the automated software, the students still agree that they also want to be corrected by the teacher, and that the teacher's more specific guidance and help is also very important [11-13].

5 Conclusion

Through this research, it is found that the application of Juku Correction Network in college English writing teaching has positive significance for improving students' English writing ability, stimulating students' learning motivation and improving their

self-efficacy. The college English writing teaching mode based on Juku Correction Network has made up for the drawbacks in traditional writing teaching to a certain extent. Teachers can focus more on guiding students to analyze, synthesize, compare and develop critical thinking. The correction feedback and automatic error correction function of the correction network is conducive to the improvement of students' expressive ability, leading to improvement of the quality of college English writing teaching. It is worthy of promotion in college English teaching.

The assessment system has become a useful teaching tool for teachers. However, the system cannot replace the teacher. It still requires teachers to combine classroom teaching and online writing training closely to get satisfactory results.

6   References

[1] Yeh, Y., H. C. Liou & Y. T. Yu. The Influence of Automated Essay Evaluation and Bilingual Concordancing on EFL Students' Writing [J]. English Teaching & Learning, 2007, (1).

[2] Shermis, M. D. & Burstein J. Automated essay scoring: Across disciplinary perspective ［M］. Mahwah，NJ: Lawrence Erlbaum Associates，2003: 13. https://doi.org/10.1162/coli.2004.30.2.245

[3] Valenti，S．，Neri，F．& Cucchiarelli，A． An Overview of Current Research on Automated Essay Grading ［J］. Journal of Information Technology Education，2003，24 (2)

[4] J. Burstein，M. Chodorow，C. Leacock． Criterion Online Essay Evaluation: An Application for Automated Evaluation of Student Essays ［C］． In Proceedings of the Fifteenth Annual Conference on Innovative Applications of Artificial Intelligence，Acapulco，Mexico，2003: 3－10

[5] Li, Jinhui． Using latent semantic analysis for automated essay scoring in the Chinese EFL context. ［D］. Guangzhou：Guangdong Foreign Language Study University，2009

[6] Liang, Maocheng． Constructing model for automated writing evaluation for Chinese students. ［M］. Beijing: Foreign Language Research and Study Press, 2010.

[7] Jiang, Y. & Ma, W.L. Chinese Automated Scoring Systems in English Writing: Achievements and Challenges [J]. E- Education Research, 2013, (3): 76-81

[8] Flower H． A cognitive process theory of writing ［J］． College Composition and Communication，1981 (4): 365-387

[9] Zamel V． Teaching composition in the ESL classroom: What we can learn from research in the teaching of English ［J］. TESOL Quarterly，1976, (10): 67-76 https://doi.org/10.2307/3585940

[10] Chen，E．& E． Cheng． Beyond the design of automated writing evaluation: pedagogical practices and perceived learning effectiveness in EFL writing classes ［J］. Language Learning & Technology, 2008, (12): 94-112.

[11] Lai, Yi-hsiu. Which Do Students Prefer to Evaluate Their Essays: Peers or Computer Program [J]. British Journal of Educational Technology, 2010, (3).

[12] Goldberg, A., Russell, M., & Cook, A. (2003).The effect of computers on student writing: A meta-analysis of studies from 1992 to 2002． Journal of Technology, Learning, and Assessment, 2(1), 1-51

[13] Denton, P.et al. Students' Response to Traditional and Computer-assisted Formative Feedback: A Comparative Case Study [J]. British Journal of Educational Technology, 2008, (3). https://doi.org/10.1111/j.1467-8535.2007.00745.x

7   Author

Yusheng Wang is an associate professor at Nanjing College of Information Technology, China. His research interests include applied linguistics and cross-cultural studies.