

Big Data Methodology and Teaching Innovation of English Writing

Liuchun Wen

Guangdong Technical College of Water Resources and Electric Engineering, China

Keywords: Big Data Methodology, Teaching Innovation of English Writing, Million English Writing Contest.

Abstract: In the context of the era of big data, higher vocational English writing courses should make full use of Internet technology and massive information data to build a new teaching model in terms of teaching concepts, teaching forms, teaching resources, and teaching evaluation. Employing big data methodology and statistics analysis, this paper explores the English compositions with the same topic for millions of students in 2016. The adopted instruments include SPSS software, Matlab, SAS, Python software. Different graphs, pictures and tables show the situation of students' participating in the competition in each area, modification and score change of students' writing, and dimensional changes of students' writing. The study shows that the continuous development of information technology provides new auxiliary means and tools for writing teaching. Big data brings opportunities and challenges to traditional English writing teaching.

1 INTRODUCTION

With the development and rapid popularization of cloud computing technology, the Internet has entered an era of rapid development. The Internet has penetrated all aspects of human social activities. Due to the rapid improvement in the ability of computers to process data, people have discovered laws from the seemingly disorganized massive data that could not be found in the pre-Internet era. As a result, we have entered the era of "big data". Under the background of the era of big data, the concept and teaching form of higher education have undergone profound changes. The intellectualization and informatization of English teaching reform is an important subject for higher vocational English education researchers and teachers (Chen, 2015). In the context of the era of big data, higher vocational English writing courses should make full use of Internet technology and massive information data to build a new teaching model in terms of teaching concepts, teaching forms, teaching resources, and teaching evaluation. We will use the big data method to study the English writing of millions of the same topic in 2022 to construct an English writing teaching model for higher vocational education.

The basic characteristics of big data can be summarized by five "Vs", namely Volume (large capacity), Variety (many types), Velocity (fast

speed), Value (high value), Visualization (visualization) (Eynon, 2013)

2 THE IMPACT OF BIG DATA ON ENGLISH WRITING

The arrival of the era of big data will surely bring about great changes in modern education. In the era of big data, information-based education has become one of its distinctive features. Faced with this new situation, how to realize teaching innovation through "technical delicacy" is a problem that every teacher should seriously think about. English writing is an important language skill, and big data has had a profound impact on English writing. Wang Haixiao believes that the era of big data is characterized by college English writing teaching, including writing teaching resources, writing purpose, writing content and organization, and writing aids (Wang, 2014).

The connotation of tools, writing assessment and writing ability has brought changes in many aspects from concept to behavior, also brought new opportunities and challenges to the reform of college English writing teaching (Yang & Dai, 2015).

The continuous development of information technology provides new auxiliary means and tools for writing teaching. The research on intelligent essay scoring system (automated essay scoring) at home

and abroad has begun to take shape. Juku Correction Network (hereinafter referred to as "Correction Network") developed by Beijing Language Intelligence Collaborative Research Institute is one of the most influential online writing intelligent platforms in China. The feature of the Correction Network is that students submit their compositions online, and the system immediately makes assessment and feedback, and students can revise their compositions unlimited times based on the feedback. The automatic scoring system operates quickly and accurately, saves a lot of labor, and enhances immediacy and interactivity. In addition, the intelligent automatic evaluation system realizes the individualization of evaluation, can evaluate from multiple dimensions such as vocabulary, grammar, text and content, and provides various feedback information for students' writing synchronously, eliminating students' disgust for evaluation feedback.

Emotions and anxiety in the writing process can effectively improve the efficiency of English writing teaching and promote the improvement of students' English writing ability.

In order to construct and mine the big data of Chinese students' English writing, since 2014, the Chinese College English Writing Teaching Alliance and Beijing Language Intelligence Collaborative Research Institute have carried out the "Millions of English Writing Activities with the same topic" for 8 consecutive years. After 8 years of development, the English writing activity on the same topic has attracted more than 9 million teachers and students from thousands of colleges and universities, and provided a large amount of real corpus data for Chinese English teaching and research. It has become a well-known English teaching brand competition in China.

To carry out research on writing corpus data, it is necessary to master the methods and technologies of big data. To this end, higher vocational English teachers need to master various tools and software, keep up with the development of modern information technology, and learn the most advanced data analysis and processing software, such as SPSS software, Matlab, SAS, Python software, etc. Especially worth mentioning is Python software, which is an object-oriented interpreted computer programming language, which is very useful in data processing. Because it is a programming language, many college teachers born in the liberal arts have a deep sense of fear of it, but once mastered, it will greatly help improve the ability to process information.

3 EXPLORATION INTO THE ENGLISH COMPOSITIONS WITH THE SAME TOPIC FOR MILLIONS OF STUDENTS IN 2016 BY WAY OF BIG DATA METHODOLOGY

The 2016 English Writing Contest with the same topic for millions of students is jointly sponsored by the National Institute of Foreign Language Teaching in Colleges and Universities and the China College English Writing Teaching Alliance, and organized by the Correction Network (www.pigai.org).

3.1 Situation

This topic is provided by Peking University. It focuses on the impact of AI on human beings and guides Chinese students to think about AI. From April 6 to May 31, 2016, 22532 teachers from 9384 schools in 32 provinces and cities across the country participated in the activity. The number of student essays submitted reached 1739660, covering junior high school (7.9%), high school (13.89%), higher vocational (7.12%) and undergraduate (71.09%). Table 1 shows The situation of students' participating in the competition in each area.

Table 1. The situation of students' participating in the competition in each area

Provinces/municipalities/autonomous regions	The number of compositions handed in	Provinces/municipalities / autonomous regions	The number of compositions handed in
Beijing	72,232	Chongqing	2,672
Sichuan	39,272	Tianjin	2,527
Shandong	23,694	Anhui	1,818
Guangdong	18,293	Liaoning	1,728
Jiangsu	11,246	Fujian	1,722
Shanxi	8,839	Guangxi	1,415
Jiangxi	5,793	Henan	1,318
Hebei	5,126	Yunnan	1,149
Zhejiang	3,065	Hubei	1,034

3.2 Modification and Score Change

As shown in Fig. 1, in this activity, the average number of revisions of more than 200000 students' compositions is 6.1, which means that each student has revised more than 6 compositions on average with

the help of machine correction under the condition of completely independent learning during the winter vacation.

After an average of 6.1 revisions per composition, the original machine score of students' compositions

increased from 67.22 in the first edition to 70.48 in the final edition, with an average increase of 3.26 points.

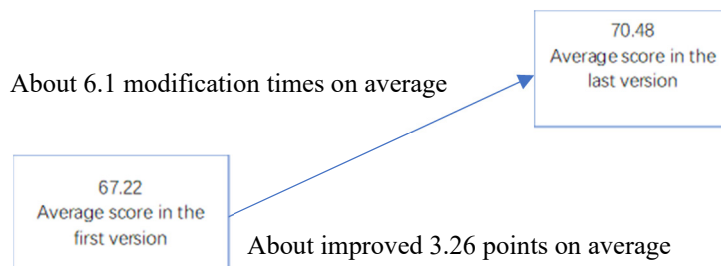


Figure 1. The change of the score of students' composition in different editions

3.3 Dimensional Changes

Fig. 1 shows the change of the diction in different editions. It can be seen that the measured values of

vocabulary richness, average word length and average vocabulary difficulty have all improved, which proves that students use more words in the final version of the composition than in the first version, and the difficulty of using words is slightly improved.

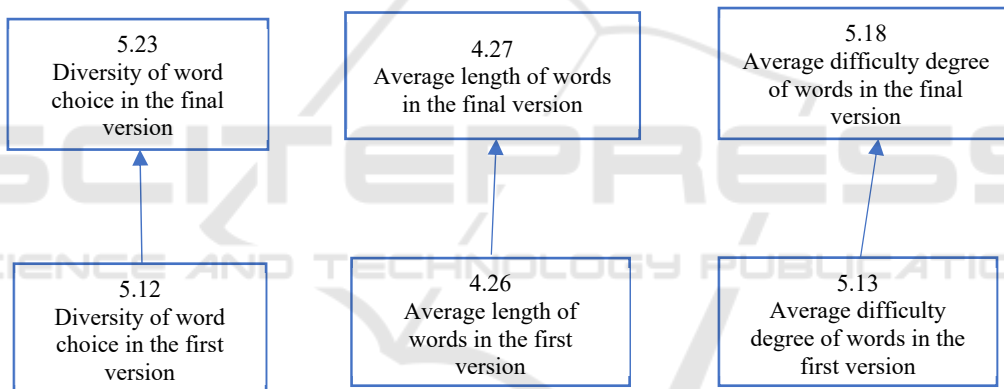


Figure 2. The change of the diction in different editions

Fig. 3 shows the change of the sentences in different editions. It can be seen that the average sentence length and clause density have slightly

increased, indicating that students constantly adjust the rationality of their sentence structure, and use clauses to make sentences in the text more abundant.

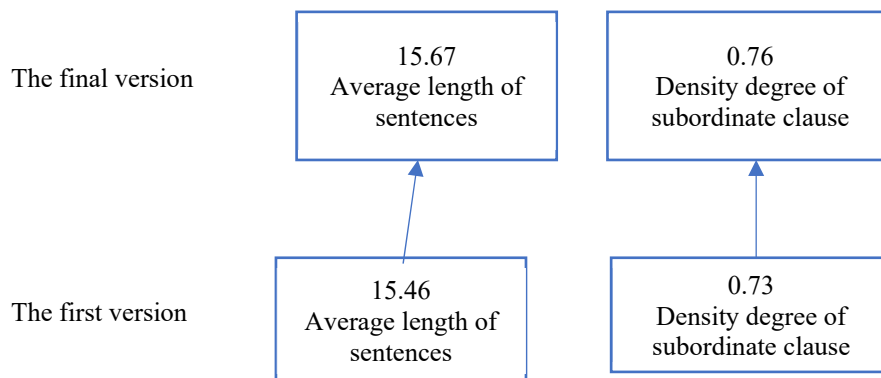


Figure 3. The change of the sentences in different editions

The following Table 2 shows the change of the text in different editions. It can be seen that the length of the article increases, and the number of conjunctions students use increases, and students noticed that the use of cohesive words in writing can improve the coherence of the structure of the article, the number of paragraphs increases, and students' sense of segmentation increases.

Table 2. The change of the text in different editions

Index Version	Length of composition	The number of conjunctions	The average number of paragraphs
The final version	172.41	11.38	3.19
The first version	161.51	10.51	3.1

3.4 Big Data Method and Innovation in English Writing Teaching

English writing teaching in the era of big data must first change the educational concept and the role of teachers (Wang, 2014). In traditional writing classrooms, a teacher-centered teaching model based on experience is often used. Teachers play a central role in the teaching process, teaching writing skills and assigning writing tasks. Students passively accept knowledge and complete homework, and it is difficult to get timely and targeted feedback in the teaching process, so they cannot revise and improve their writing in time. This kind of teaching mode is monotonous, lacks pertinence, and cannot mobilize students' interest and enthusiasm, so it is difficult to achieve ideal teaching results. In the era of information technology and big data, learning is more of a self-organizing behavior of students. Students are the center of learning activities and the main body of the teaching process. Teachers are more likely to provide guidance, support and services for learners. With the rapid development of network technology, network resources are readily available, and students can obtain and utilize no less learning resources than teachers. At this time, the main role of teachers is no longer the transmitter of knowledge resources, but the integrator of autonomous learning resources. The teacher's responsibility is to adhere to the "student-centered" educational philosophy, integrate excellent learning resources on the Internet, make full use of diverse resources and teaching methods, stimulate students' continuous thinking, cultivate students' interest in writing, guide students to learn independently, and timely. As far as English writing

teaching is concerned, teachers' role is to become an integrator of writing learning resources, a digger of student writing data, a designer of data-driven precision teaching, and a professional assessor of students' writing level. The data analysis of millions of writing on the same topic can provide specific guidelines for the design of precise teaching. On the whole, Chinese students lack in-depth English reading, as the proportion of students mentioning literary works in the composition is only 1.33%.

In traditional writing teaching, students "write for the sake of writing" and "practice for the sake of practice" are inevitably boring, students lack interest, and it is inevitable that they will be perfunctory. Students' writing for practice ignores the expressive function of writing, which may cause students to pay too much attention to grammar, vocabulary, structure and other formal problems and ignore the core part of the article, that is, content and ideas.

Network technology based on big data can help solve the problem of writing purpose from two aspects. On the one hand, writing software and online writing systems can correct most of the formal errors, freeing students to spend more time and energy to conceive the content of the essay. On the other hand, open web platforms make every writing a "share". For example, teachers can let students write e-mails to others, or share their comments or comments on the Internet, or adopt the method of students' mutual evaluation of compositions, so that every writing has a "reader", so that each student is both an author and a reader. In this way, writing is no longer a dry exercise, but a real transmission of information and emotion, which returns to the most authentic purpose of writing (Tang & Wu, 2012).

4 CONCLUSION

Big data brings opportunities and challenges to traditional English writing teaching (Liu, 2014). How teachers and students adjust their learning concepts and methods in this context will determine the future results of college English teaching. Teachers should have a clear understanding of this trend, and should actively participate in and adapt to this trend, big data is used to analyze students' learning characteristics, learning ability and learning motivation, so as to make adjustments in teaching strategies and teaching concepts, in order to achieve the best teaching effect. Further research on English writing in the era of big data involves not only college English teaching itself.

In addition, it can control and intervene in the entire English learning process of students, such as

primary school and middle school, to form an integrated teaching of the entire English learning process, optimally allocate learning resources, and complete English learning goals.

ACKNOWLEDGEMENTS

2022 Guangdong Education Science Planning Project (Higher Education Project) "Research on Higher Vocational English Teaching Based on 'Output oriented Approach' from the Perspective of National Foreign Language Competence" (No. 2022GXJK506); The 2021 Guangdong Higher Vocational Education Teaching Reform Research and Practice Project "Exploration and Practice of the Teaching Reform of the 'Comprehensive English' Course for Higher Vocational English Majors under the Theory of Curriculum Ideology and Politics".

REFERENCES

- Chen Jianlin. Research on MOOC and foreign language teaching in the era of big data [J]. Foreign Language Electronic Teaching, 2015(1).
- Eynon, R. The rise of Big Data: What does it mean for education, technology and media research[J]. Learning, Media and Technology, 2013(3).
- Gao Yujuan, Zhao Xiaodong. Corpus-based English majors in China Quantitative analysis of the use of adjectives and adjectives [J]. Foreign Language Studies, 2020, 37(2): 24 - 31.
- Hu Xuewen. The influence of online composition self-correction on college students' English writing results [J]. Foreign Language Electronic Teaching, 2015(3).
- Liu Runqing. Foreign language education and scientific research in the era of big data [J]. contemporary foreign language studies Research, 2014 (7).
- Tang Jinlan, Wu Yi'an. Writing automatic evaluation system in college English teaching research on the application of Chinese language [J]. Foreign Language and Foreign Language Teaching, 2012 (4): 53-59.
- Victor Meier-Schoenberg, Kenneth Cooke. when big data Generation: The Great Change of Life, Work and Thinking [M]. Hangzhou: The People of Zhejiang Press, 2013.
- Wang Haixiao. The reform of college English writing teaching in the era of big data [J]. Modern Online Education Research, 2014(3): 66-72.
- Wang Tsunami. College English writing teaching reform in the era of big data [J]. Modern Distance Education Research, 2014(3).
- Wang Ying, Li Zhen Yang. A Research Review on Electronic Feedback Models in Foreign Second Language Writing[J]. Foreign Language Teaching, 2012 (4): 11-16.
- Wang Zhe, Li Junjun. Exploration on the reform of foreign language general education in colleges [J]. foreign language Chemical Teaching, 2010(5).
- Yang Xianmin. The connotation and characteristics of smart education in the information age [J]. China Electronic Education, 2014, 33(1): 29-34.
- Yang Xiaoqiong, Dai Yuncai. College English Autonomous Writing Teaching Based on Correction Net Model practice research [J]. Foreign Language Electronic Teaching, 2015(2).
- Yang Yonglin. English teaching in the era of globalization, informatization and digitalization Learning—a research on the construction of training system based on “experience English writing” [J]. Foreign Language and Foreign Language Teaching, 2008(5).