Integrated teaching of college English reading and writing under the background of big data

Yangzi Sun and Chen Liu

1 School of Foreign Languages, Dalian Neusoft University of Information, Dalian, Liaoning 116023, China
2 School of the English Language and Culture, Xiamen University Tan Kah Kee College, Zhangzhou, Fujian 363105, China

Abstract. At present, reading and writing are divided into two independent courses in English teaching at most universities. According to the views of many scholars at home and abroad, it is more efficient to learn a language through integration of input and output, which advocates the integrated teaching of English reading and writing. The purpose of this paper is to verify whether the integration of reading and writing is helpful to the improvement of college students' ability through comparative experiment and random sampling method, and put forward an integrated learning mode of online reading and writing based on data algorithm. The main content of this paper is to elaborate the current situation of mobile reading and mobile writing in the era of big data, and then extract two groups of college students with similar learning conditions from two universities for empirical analysis, to study the differences in students' performance through comparison in the traditional reading and writing teaching mode and the new mode of integration of reading and writing teaching, and finally put forward the application of clustering algorithm based on K-Modes under the background of big data English learning scoring system, so that students can learn independently and efficiently on the Internet. The experiment shows that the average score of students in the integration of reading and writing is generally higher than that in traditional teaching. Under the integration of reading and writing, the students' reading accuracy is as high as 90%, which is 10% higher than that in the traditional way, and the score of writing is 5 points higher than that in the traditional way.

1. Introduction
Reading is generally acknowledged as the process of English learning input, while writing the process of English learning output. Through reading, learners can enrich lexical resources and acquire sentence patterns, as well as syntactical knowledge. Reading also provokes ideas for writing. Writing needs to read the output of this input, and writing is the performance of knowledge internalization [1-2]. However, at present, reading and writing are regarded as two unrelated branches of teaching and learning in many universities, which leads to the common phenomenon that people with good reading ability may show...
poor writing. But there is another phenomenon, that is, people with good reading ability will also have outstanding writing ability [3-4]. Therefore, it is necessary to study the integration of reading and writing. Over the past decade, scholars have been exploring the integration of reading and writing. They have theoretically discussed the feasibility of the combination of reading and writing, and they have also tried to combine reading and writing in practical teaching [5-6]. For example, Xuan believes that the input and output should be combined in the learning of a language, because only when the input is output in time can it leave a deeper impression [7-8]. Zhao, W. and others believe that only when language output and information are combined with other inputs can they promote students' language learning. Through the deep absorption and integration of other input information materials, and through the reconstruction and flow of basic language knowledge, students can better achieve output [9].

This paper attempts to analyze the internal attributes of reading and writing, explore the close relationship between reading and writing and the feasibility of integration, and then find out the obvious advantages of integrated course of reading and writing, compared with separate teaching of reading course and writing course. Finally, based on the second language acquisition, input and output theory and constructivism, it proposed an innovative reading and writing integrated classroom teaching mode in the era of big data [11-12].

2. The Integration of Reading and Writing in the Era of Big Data

2.1. Mobile Reading
Mobile library is the focus of research on mobile reading mode. Mobile library means that students can read all books through mobile phones or computers without going to the library. In the era of big data, every university has its own library website and some universities have even established their own learning platform, which stores a large number of books. In addition, mobile reading is also reflected in that people can read news, e-books and other electronic materials anytime and anywhere.

2.2. Mobile Writing
In the past, the common way of teaching writing in college classroom is that teachers first explain the methods and requirements of writing, let students write in exercise books, and then students give the written composition to the teacher for correction and evaluation. This process is time-consuming and inefficient. But today's big data era has triggered a new way of writing, that is, mobile writing. There are various correcting websites, writing websites and English electronic dictionaries that have launched the mode of intelligent composition modification. The website database will store correct grammar and writing requirements, which can realize the function of machine or manual composition modification, saving a lot of time for college students and teachers.

2.3. Input and Output in English Reading and Writing
According to the input-output theory, language learning is the best when input and output are carried out simultaneously. Therefore, in English learning, we should combine input and output, and pay attention to output, with students as the center of teaching.

3. Experimental Research on the Integration of Reading and Writing

3.1. Subjects
In order to make the experiment more convincing, this experiment takes first-year English majors from two universities with the same grade in CET-4 test: the first group of research objects are 20 English Majors University X; the second group of research objects are 20 English Majors University Y. It should be noted that there is no significant difference in English proficiency between the two groups.
3.2. Experimental Purpose
When college students write English, especially first-year students, they prefer setting up templates and have nothing to say for lack of proficient training in academic writing. The goal of this experiment is to collect and integrate a large number of experimental data through a large number of case studies to verify that under the support and guidance of the theory of language input and output, the teaching of the new English reading and writing course with the integration of reading and writing can better cultivate and improve college students' English reading and writing skills, and we can observe in practice whether the teaching mode of this kind of reading and writing course can effectively mobilize and cultivate students' enthusiasm and self-confidence in writing activities.

3.3. Experimental Design
In this experiment, two groups of subjects were conducted simultaneously in a controlled way. Group from University X used the traditional teaching method for English teaching, and the other group from University Y used the integration of reading and writing.

3.3.1. Reading Test. The experiment lasted for 3 months, a total of 12 weeks of teaching practice, and 20 English Majors in University X were taught English reading. In the process, let everyone read a certain amount every day: two articles in the 21st century English newspaper (one long and one short, with a total of about 700 words) or one newspaper article plus one article in the western media (with a total of about 800 words), and make reading notes. The way of reading is the combination of intensive reading and extensive reading. Take the following teaching methods:

1) Check students' understanding through reading comprehension multiple choice questions, or in class, adopt question-and-answer teaching method according to the text content, or understand students' understanding through reading comprehension multiple choice questions.

2) Outline the key words, idioms and sentence patterns of each article, and explain some knowledge points in depth. Present sample sentences, and deepen the understanding of knowledge through relevant fill in the blanks, translation, or multiple choice exercises.

3) Consolidate input by spot check and dictation.

3.3.2. Writing Experiment Test. In the same way, the same students in University X are trained separately from reading. In the classroom, the teacher designed tasks to help students analyze the selected language input materials to strengthen students' discourse awareness, and help students grasp the writing tasks and make layout. At the same time, it can improve the students' awareness of the coherence and unity of the article, form good logical thinking habits over time, and make the expression more fluent.

3.3.3. Integration of Reading and Writing. Twenty students from University Y were experimented in different spaces at the same time, but in different ways. In addition to keeping the same teaching progress with the control class, maintaining the same input of classroom language and textbook reading materials, on the basis of relatively reducing the capacity of dictation, the teacher strengthened the language output training of students' composition form in the classroom. Each class has an average of 5 minutes, mainly in the form of speech, story continuation or debate based on the related topics involved in the articles read this week. In the weekly tutoring class, the traditional multiple-choice exercises focusing on reading comprehension are abandoned, and the writing training of reading and writing is carried out by selecting an article from "college students learning English newspaper" or selected according to hot topics.

3.4. Research on Integrated Teaching of Reading and Writing Based on Big Data
In the information age, there are many English learning software for college students' autonomous learning, such as: silent word memorizing software, daily English listening software, Youdao dictionary. These apps can provide the integrated learning function of reading and writing, so that college students can improve their writing ability while reading.
3.5. Outlier Mining Algorithm Based on Frequency

3.5.1. Selection of clustering algorithm. When learning English on software, it is necessary to involve the requirement of scoring. A large number of students answer questions through the Internet and transmit data about English reading and writing. Different data need to be classified in order to correspond to the corresponding level of scoring standards and give appropriate ratings. K-Modes algorithm is a commonly used classification data clustering algorithm, which uses dissimilarity to measure the dissimilarity between two data. The smaller the dissimilarity between two data, the smaller the distance between them. This algorithm can be used to classify students’ data. In this paper, we use the data of teaching scores of an English course. In the experiment, the scores (5 excellent, 4 good, 3 medium, 2 pass and 1 fail) are simplified to (5 4 3 2 1). K-Modes algorithm is used to cluster the teaching evaluation data used in the experiment.

3.5.2. Description of clustering algorithm. Hypothesis \( A_i \) is an attribute in the dataset, and \( x \) and \( y \) are two values of \( A_i \). \( A_j \) is another attribute in the dataset, \( w_1 \) is \( A_j \) subset of the range of \( J \), \( w_2 \) is the complement of \( P_i(w_1 | x) \) represents attribute \( A_i \). When \( I \) is \( x \), \( A_j \) is the conditional probability of \( \text{W set}_{i} P_i(w_2 | x) \), when \( A_i \) is \( y \), \( A_j \) is the conditional probability of \( \text{W set} \). Then the two values \( x \) and \( y \) under \( A_i \) are relative to the distance of \( A_j \) is:

\[
\delta^{ij}(x,y) = P_i(w_1 | x) + P_i(w_2 | x)
\]  

(1)

Where \( W \) is the maximum value of \( \delta^{ij}(x,y) \). Due to the \( P_i(w_1 | x) \) and \( P_i(w_2 | x) \) has a range of \([0,1]\). In order to limit the value of \( \delta^{ij}(x,y) \) between 0 and 1, we define \( \delta^{ij}(x,y) \) as:

\[
\delta^{ij}(x,y) = P_i(w_1 | x) + P_i(w_2 | x) - 1
\]  

(2)

Assuming that there are \( m \) attributes in a dataset, the distance between any two different values can be expressed as:

\[
\delta(x,y) = \frac{1}{m} \sum_{j=1...m,j\neq j} \delta^{ij}(x,y)
\]  

(3)

In the process of scoring students, the algorithm is used to analyze the similarity of data among students, and cluster scoring is carried out.

4. Analysis of Experimental Results

After the end of the experiment, 40 students from two different universities were tested and assessed. Finally, the average score of each class was compared to analyze the experimental effect of the integration of reading and writing.

4.1. Experimental Results of Asynchronous Teaching of Reading and Writing

Three months later, each student in the University X was given an English test paper to test. The difficulty of the test paper was based on the CET-6 level, with 50 points for reading and 50 points for writing. Submit the test paper for correction after the specified time. The results of reading test are shown in Table 1.

| No.1 | No.2 | No.3 | No.4 | No.5 | No.6 | No.7 | No.8 | No.9 | No.10 |
|------|------|------|------|------|------|------|------|------|-------|
| Reading | 42   | 44   | 38   | 30   | 36   | 32   | 48   | 40   | 42    | 38    |
| Writing | 36   | 32   | 35   | 33   | 38   | 32   | 40   | 40   | 41    | 37    |
| No.11 | No.12| No.13| No.14| No.15| No.16| No.17| No.18| No.19| No.20 |
| Reading | 34   | 30   | 32   | 36   | 42   | 40   | 50   | 38   | 36    | 34    |
| Writing | 30   | 28   | 32   | 34   | 41   | 43   | 45   | 35   | 31    | 31    |

Table 1. Reading and writing scores (group 1)
As can be seen from Figure 1, the students with excellent reading performance also have high writing performance, and the students with average reading performance also have reached average writing performance. But most of the students’ writing performance is slightly lower than that of reading. And it can be seen that the number of top students in this group is small, most of the reading scores are about 40 points, and the writing scores are about 35 points.

4.2. Reading and Writing Combined With Experimental Results

In order to reduce the error, 20 students from University Y were given the same test. The test results are shown in Table 2.

|         | No.1 | No.2 | No.3 | No.4 | No.5 | No.6 | No.7 | No.8 | No.9 | No.10 |
|---------|------|------|------|------|------|------|------|------|------|-------|
| Reading | 46   | 42   | 40   | 42   | 48   | 38   | 42   | 40   | 44   | 36    |
| Writing | 40   | 44   | 35   | 41   | 45   | 39   | 43   | 40   | 41   | 38    |
|         | No.11| No.12| No.13| No.14| No.15| No.16| No.17| No.18| No.19| No.20 |
| Reading | 45   | 50   | 48   | 40   | 44   | 46   | 36   | 40   | 34   | 42    |
| Writing | 45   | 48   | 47   | 39   | 41   | 45   | 45   | 38   | 32   | 39    |
Figure 2. Reading and writing scores (group 2)

From Figure 2, the overall test results of the second group of students in the integrated reading and writing experiment are higher. The number of reading scores is more than that of the first group, and there are also many cases of writing scores higher than reading scores. The reading scores are about 45, and the writing scores are about 40. Therefore, the teaching effect of integrating reading and writing is better than that of separate teaching.

5. Conclusion
This paper mainly studies the integrated teaching of reading and writing for college students under big data background. A group of controlled experiments shows that the integrated teaching of reading and writing can effectively improve college students' reading and writing abilities at the same time, so this teaching mode can be carried out in other colleges and universities. We should use the data advantage of the current network to implement the integrated learning mode of reading and writing on the network to improve students' learning efficiency. Although this paper concludes the advantages of integration of reading and writing, there are still many shortcomings in the experiment. It is impossible for the two groups of experimental students to be completely consistent in their learning situation, and there are also differences in the level of the teachers who teach the experimental subjects. Therefore, there are great random errors in the experimental data.

References
[1] Ho-jung Yoo. (2016) College Students' Perceptions about the Connection between English Reading and Writing in L2 Reading-Based Writing[J]. Studies in English Language & Literature, 42(4):245-262.
[2] Vasudevan L, Kerr K R.(2016) "Unflattening" Our Ways of Seeing, Reading, and Writing[J]. Journal of Adolescent & Adult Literacy, 60(1):103-105.
[3] Anyang Zhao. (2016) "Research of Integration Method of Interpersonal Teaching and Man-machine Teaching in College English." International Technology Management 2:39-41.
[4] Wang Haiyan. (2016). The study and practice of ESP and EGP integration pattern on College English. Journal of Heihe University, 007 (004), 61-63
[5] Chen Xiaoqin, Tang Mingfeng. (2019) Teaching practice of College English reading and writing based on output oriented approach [J]. Journal of Wuzhou University, 029 (002): 106-112

[6] Zhang Dan. (2018) The application of "Mosoc teach" in College English reading and writing course under the mixed teaching mode [J]. Heilongjiang Science, 009 (021): 56-57

[7] Xuan G. (2016) Action Research on College English Writing Based on Information Technology from the Perspective of Moocs[J]. English Language Teaching, 9(11):48.

[8] Zhao, W., D. Xu, and H. Wang. (2019) "A Study on the Integration of Chinese Excellent Traditional Culture and College English Teaching." The frontier of educational research 009(003):142-146.

[9] Chen Chunmei. (2017) An evaluation of New Horizon College English (reading and writing, book I, Second Edition) [J]. Overseas English, 000 (012): 247-248

[10] Meng, Z. L., T. N. Wydell and H. Y. Bi. (2018) "Visual-motor integration and reading Chinese in children with/without dyslexia." Reading & Writing, pp.1-18.

[11] Wang Aiqin. (2018) An Integrated Model of Teaching College English Reading and Writing with the Aim of Cultivating Students' Critical Thinking Abilities[J]. Journal of Shangqiu Polytechnic, 017(006):68-72.

[12] Chunmei Chen. (2017) An Evaluation of New Horizon College English(Reading and Writing, Book I , Second Edition)[J]. Overseas English, 000(012):247-248.