Application and Future Development of Computer Aided Writing Software in Chinese Language and Literature

Ying Huang*
Xi'an Traffic Engineering Institute, Xi'an 710300 China

*Corresponding author e-mail: huangying@xjgyedu.cn

Abstract. In recent years, the use of computer technology to study Chinese language and literature has become a new trend. There are many resources for the introduction and research of Chinese multimedia textbooks, multimedia courseware, and Chinese language and literature websites. But at present, there are few researches on the application and practice of computer-aided writing software. In this regard, the purpose of this article is to study the application and future development of computer-aided writing software in Chinese language and literature. This article first discusses the status quo of my country's Chinese language and literature construction and resources, as well as the main problems currently existing. Then, through the method of questionnaire survey, to understand the Chinese language and literature researchers' understanding, like and cognition of computer-assisted writing software Chinese language, so as to more truly understand the help of computer-aided writing software to Chinese language and literature. After that, this article combines the survey results to compare traditional manual writing and computer-assisted writing software, and explores the effectiveness and necessity of computer-assisted writing in teaching Chinese as a foreign language. Through the comparison of real Chinese language writing and the test results, the problem is explored, and corresponding countermeasures are proposed to judge its future development trend. The experimental results show that computer-aided writing software is more recognized and loved by Chinese language and literature researchers. Computer-aided writing software can effectively search for the content of writing, and then integrate resources with reference value, and then perform auxiliary writing based on it, which greatly improves the writing efficiency of Chinese language and literature researchers has increased by about 20%. Computer-assisted writing software is of great significance to the development of Chinese language and literature, and it will be more perfect in the future to achieve more functions.

Keywords: Computer Technology, Computer-Aided Writing Software, Chinese Language and Literature, Questionnaire Survey
1. Introduction
With the promotion of China's international status, the promotion and dissemination of Chinese has become more frequent, and Chinese language and literature have developed rapidly [1-2]. As an important tool and cultural carrier for understanding China and interacting with China, Chinese language has received more and more attention from foreign governments, educational institutions, enterprises, and people [3], and its practical value is constantly improving. With the rapid development of computer technology and the Internet, more and more people learn through computers [4-5]. Using this learning method can find a lot of learning resources on the one hand, and on the other hand, you can communicate with other learners online and learn together [6]. The computer-assisted writing software will help to learn and use Chinese language and literature [7].

At present, with the continuous development of computer technology, various software used in literary research and writing have emerged [8]. In foreign countries, Kamrani A has achieved great success in domain engineering, component and component library standardization, component assembly technology, reuse-based software development process and reuse maturity model by using software reuse [9]. In China, Huang H has designed a Chinese language learning website. The content of this type of website contains both Chinese cultural knowledge and Chinese knowledge, providing users with personalized services that can be customized by themselves, and providing a good online Chinese learning system [10].

This paper studies the application and future development of computer-aided writing software in Chinese language and literature. This article first starts from the actual situation, discusses the current situation of Chinese language and literature in my country, and Chinese language and literature researchers' opinions and suggestions on computer-aided writing software. Then summarized the meaning of computer-aided writing software for Chinese language writing based on the survey results. Finally, this article explores its problems through real Chinese language writing comparison and test results, and proposes corresponding countermeasures to judge its future development trend. Chinese language-assisted writing software greatly improves the efficiency of Chinese language writing, and is of great significance to the development of Chinese language and literature.

2. Application of Computer Aided Writing Software in Chinese Language and Literature and Technical Research of Future Development

2.1 Application of Computer Technology to Assisted Writing
During the actual operation of the computer-assisted writing system, it was found that when users used the synonym recommendation function, they were mainly concerned with whether the first word recommended by the synonym is the word most similar to the query in the natural language processing field. Taking into account this user demand, the semantic dictionary generates a synonym candidate set of user query words in a general environment, and then uses the context vector-based similarity calculation method in a specific context to calculate the natural relationship between each word in the synonym candidate set and the target word. The similarity value in the context of language processing, and finally the synonym candidate set is displayed to the user in descending order of the similarity value.

The similarity calculation of the auxiliary writing system is based on the following assumptions: It is assumed that two words are semantically similar if and only if they have similar contexts. In order to calculate the semantic similarity between two words, we need to count the context information of the words from the corpus. Construct a contextual word co-occurrence vector for each word whose semantic similarity needs to be calculated. Definition of contextual word co-occurrence vector: the context of a word contains rich semantic information about the word. In different contexts, a word generally has different semantics, and the context vector of each word can be determined by the method counts the co-occurring words and word frequency to obtain the word. After obtaining these related sentence resources, the auxiliary writing system will give corresponding writing instructions.
2.2 Computer-Aided Writing Software Related Algorithms

(1) Boolean model

The Boolean model is a main algorithm of computer-aided writing software. Boolean model is a retrieval model that combines set theory and Boolean algebra. In the Boolean model, the information retrieval four-tuple introduced earlier is: Document D represents a collection of index items. The user query Q indicates that the index items are connected with the Boolean symbol of "and or not", and the priority is indicated by brackets. And F is defined as: only a document can be retrieved when it satisfies the Boolean query, so the retrieval strategy of the Boolean model is a binary judgment criterion. Finally, the retrieval algorithm judges whether the document is relevant to the query according to F, and when relevant, returns the document. The characteristics of the Boolean model are: simple, fast and easy to structure expression. The disadvantages are: because it is a two-value judgment, the degree of document relevance cannot be measured, and the judgment method is strict, and it is likely that part of the content required by users cannot be retrieved.

Let s be the training sample data set, and the category identification attribute in s has m independent values, that is, m classes are defined, i=1, ..., m, $R_i$ is the subset of the data set s belonging to the $C_i$ class, and the number of tuples in the subset $R_i$ is expressed by $r_i$. The expected information amount of set s in classification can be given by the following formula.

$$I(r_1, r_2, \ldots, r_m) = -\sum_{i=1}^{m} P_i \log_2(P_i)$$

Where $P_i$ indicates the probability that any sample belongs to $C_i$ class, $P_i = r_i / |S|$, where $|S|$ is the number of tuples in the training sample data set.

If $S_j$ indicates the number of tuples belonging to $S_j$ class in subset $C_i$, the entropy of attribute A for classification $C_i (i=1,2,\ldots,m)$ can be calculated by the following formula.

$$E(A) = \sum_{j=1}^{m} \frac{S_j + \ldots + S_{mj}}{|S|} I(S_1,\ldots,S_m)$$

$$w_j = \frac{S_j + \ldots + S_{mj}}{|S|}$$

Let $w_j$ be the weight of $S_j$ subset, which indicates the proportion of $S_j$ subset in data set S, and the expected information amount of classification $C_j$ for each value of attribute A can be calculated by the following formula.

$$I(S_{ij},\ldots,S_{mj}) = -\sum_{i=1}^{m} P_{ij} \log_2(P_{ij})$$

$$\text{Gain}(A) = I(r_1,\ldots,r_m) - E(A)$$

Where $P_{ij} = S_{ij} / |S_i|$ indicates the proportion of $S_j$ belonging to $C_i$ class in the subset. $\text{Gain}(A)$ is the me.

(2) Vector space model

The characteristic of this model is that it has strong computability and maneuverability for the relevance of queries and documents. Therefore, it is very useful in the fields of information retrieval such as computer-aided writing software for text retrieval, automatic summarization and search.
engines. Application value, and the effect is good. Compared with the Boolean model, the vector space has the advantages that the index item weight greatly improves the retrieval effect; it uses a partial matching strategy to make the retrieval result closer to the user’s retrieval needs; according to the calculated similarity value sort the relevance of the retrieved documents.

(3) Probabilistic retrieval model

The probabilistic retrieval model estimates the relative probabilities of q and dj based on the probability model of a given user's query q and documents dj in the document collection. The probability model is a binary independent retrieval model, assuming that each feature is independent of each other and the document correlation is binary, that is, only relevant or not. The similarity sim(dj,q) between the document dj and the query formula q is defined as the ratio of the correlation probability and the irrelevant probability between the query and the document. The calculation formula is as follows:

\[
sim(d_j, q) = \frac{P(R|d_j)}{P(R|d_j^c)}
\]

(5)

The advantages of the probabilistic model are that it has a strict mathematical theoretical foundation, can adopt the principle of relevant feedback, does not use query techniques that require users, and is sorted in descending order according to the relevance of the documents. Its shortcomings are: the index items assumed in the model are all worth 2. It does not consider the different weights of different index items in the query and the document; the document needs to be divided into relevant and irrelevant before the search starts, so either it is more difficult or less accurate assurance of the attribute of decision classification.

3. Experimental Research on the Application and Future Development of Computer-Aided Writing Software in Chinese Language and Literature

3.1 Experimental Data

This article randomly selected 100 Chinese language and literature researchers from universities as volunteers. In order to ensure the effectiveness and representativeness of the research, five universities located in different regions of my country were selected. Considering the safety of experimental data, this article first conducts a questionnaire survey was conducted, and then the Chinese language writing data were collected and counted. Among them, a total of 100 questionnaires were distributed and 100 were recovered, of which 98 were valid questionnaires. This questionnaire investigates the use of computer-assisted writing software, the basic situation of users, and the characteristics of the software.

3.2 Experimental Process

This article retrieves the results of the questionnaires of all Chinese language and literature researchers, and counts the actual use of computer-assisted writing software. In addition, the Chinese language and literature works written by computer-aided writing software are compared with traditional manual writing, to understand the basic function of computer-aided writing software for Chinese language writing, and the application to users, and analyze computer-aided writing the practicality of the software, the real effect and the areas that need improvement.

4. Experimental Analysis of the Application and Future Development of Computer Aided Writing Software in Chinese Language and Literature

4.1 Analysis of Chinese Language Grammar Researchers' Views on Computer-Assisted Writing Software

In order to study the applicability of computer-assisted writing software, this paper conducts a questionnaire survey on Chinese language grammar researchers who use computer-aided writing
software, the use of the main software, the basic situation of users and the characteristics of the software. As shown in Table 1 and Figure 1.

**Table 1.** Questionnaire survey results of Chinese language and literature researchers

|              | efficient | Practical | Strong applicability | Very recognized |
|--------------|-----------|-----------|----------------------|-----------------|
| A University | 18        | 15        | 19                   | 16              |
| B University | 15        | 18        | 20                   | 18              |
| C University | 17        | 18        | 16                   | 17              |
| D University | 13        | 14        | 15                   | 13              |
| E University | 16        | 17        | 16                   | 18              |

*Figure 1. Questionnaire survey results of Chinese language and literature researchers*

It can be seen from the survey data that most Chinese language and literature researchers approve computer-assisted writing software. Each survey item in the questionnaire survey is supported by at least 15 people, and some even 20 people are satisfied. This proves that the computer-assisted writing software is applicable and is of great help to Chinese language and literature writing. However, computer-aided writing software is a new type of writing method. The reason why this writing method that keeps pace with the times has gained widespread attention lies in its scientific and accurate writing method, which provides convenience for Chinese language and literature writing. However, it is worth noting that the application of computer-assisted writing software also requires teaching students in accordance with their aptitude, and not all Chinese language and literature are suitable for using this software. In the survey results, the number of people who received support for some survey projects is still relatively small. Although this data appears small here due to the limited sample size, if the scope of the survey is expanded, this data will have certain changes. We should pay attention to such a small proportion, because it is also crucial for the future development of computer-aided writing software.

### 4.2 Comparative Analysis of Traditional Manual Writing and Computer-Aided Writing Software

In order to further verify the effect of computer-aided writing software, we compare traditional manual writing and computer-aided writing software. The Chinese language content of the two writings is the same, and as the amount of writing content increases, the writing efficiency and quality of the two are recorded. As shown in Table 2 and Figure 2.
Table 2. Writing efficiency and quality

| Amount of writing | 20  | 40  | 60  | 80  | 100 |
|-------------------|-----|-----|-----|-----|-----|
| Traditional manual writing | effectiveness | 80% | 76% | 70% | 65% | 60% |
|                     | quality       | 90% | 84% | 79% | 72% | 63% |
| Computer-aided writing software | effectiveness | 87% | 85% | 83% | 82% | 80% |
|                     | quality       | 93% | 92% | 95% | 94% | 94% |

Figure 2. Writing efficiency and quality

It can be seen from the experimental data that as the amount of writing content increases, the efficiency and quality of traditional manual writing will gradually decrease. This is because people are prone to fatigue if they write too much at a time. If tired, the efficiency and quality of writing will be both slowly decrease, so that it will affect the Chinese language and literature writing. But when computer-aided writing software encounters this situation, it will not be affected by this. The efficiency and quality of computer-assisted writing software have always been maintained at a high level, which is of great help to the writing of Chinese language and literature. This has once again confirmed that the application of computer-assisted writing software in Chinese language and literature is very extensive and effective.

5. Conclusions

This paper studies the application and future development of computer-aided writing software in Chinese language and literature. In this article, first, based on the actual situation, discuss the current situation of Chinese Chinese language and literature, as well as the use of computer-aided writing software and related opinions by Chinese language and literature researchers. Based on these investigations, this article summarizes the application effects of computer-assisted writing software on Chinese language and literature writing. Finally, this article analyzes the efficiency and quality of computer-aided writing software by comparing traditional manual writing and computer-aided writing software, and proposes corresponding measures and future development trends. Computer-aided writing software has greatly improved the efficiency and quality of Chinese language writing, and the application effect is relatively good. It is very important for the development of Chinese language and literature.
References

[1] Peker M, En B, Delen D. Computer-Aided Diagnosis of Parkinson's Disease Using Complex-Valued Neural Networks and mRMR Feature Selection Algorithm [J]. Journal of Healthcare Engineering, 2015, 6(3):281-302.

[2] Patuk I, Borowski P F. Computer aided engineering design in the development of agricultural implements: a case study for a DPFA [J]. Journal of Physics: Conference Series, 2020, 1679(5): 052005.

[3] Shi X, Geng L. The application of computer technology in Chinese language and literature teaching [J]. Agro Food Industry Hi Tech, 2017, 28(1):1147-1151.

[4] Mudawe O M N. Ramping the Future of Translation Studies through Technology-based Translation [J]. International Journal of Comparative Literature and Translation Studies, 2019, 7(3):74.

[5] Yeh Y F, Lin T C, Hsu Y S, et al. Science Teachers' Proficiency Levels and Patterns of TPACK in a Practical Context [J]. Journal of Science Education & Technology, 2015, 24(1):78-90.

[6] Mudawe O M N. Ramping the Future of Translation Studies through Technology-based Translation [J]. International Journal of Comparative Literature and Translation Studies, 2019, 7(3):74.

[7] Liao K H. The abilities of understanding spatial relations, spatial orientation, and spatial visualization affect 3D product design performance: using carton box design as an example [J]. International Journal of Technology & Design Education, 2015, 27(1):1-17.

[8] Yeh Y F, Lin T C, Hsu Y S, et al. Science Teachers' Proficiency Levels and Patterns of TPACK in a Practical Context [J]. Journal of Science Education & Technology, 2015, 24(1):78-90.

[9] Kamrani A, Abouel Nasr E, Al-Ahmari A, et al. Feature-based design approach for integrated CAD and computer-aided inspection planning[J]. International Journal of Advanced Manufacturing Technology, 2015, 76(9-12):2159-2183.

[10] Huang H, Yin L. Creating sustainable urban built environments: An application of hedonic house price models in Wuhan, China [J]. Journal of Housing and the Built Environment, 2015, 30(2):219-235.