

# A Quantitative Study on the Function of Quantifiers in *Hanyu Shuiping Kaoshi* Text Difficulty Assessment

Linlin Zou, Yongquan Li

linlangyuxia@foxmail.com, lirata@foxmail.com

College of Chinese Language and Culture, Jinan University, Guangzhou, China

**Abstract**—At present, the country is vigorously promoting The Belt and Road, jointly building The Belt and Road, committed to the connectivity of Asia, Europe, Africa and the adjacent seas, establishing and strengthening the connectivity partnership of countries along the line, and achieving diversified, independent, balanced and sustainable development of countries along the line. Therefore, all kinds of projects cannot be carried out without language first. With the rise of the Chinese language craze, the importance of Hanyu Shuiping Kaoshi (HSK) test is self-evident. Therefore, it is necessary to deepen the research of HSK. Based on the methods of automatic composition correction and text difficulty rating, this paper conducts a quantitative study on quantifiers in HSK compositions. The research results show that the species of quantifiers have a greater impact on the difficulty of HSK compositions. Adding the characteristics of the quantity and species of quantifiers can't significantly improve the accuracy of HSK composition difficulty rating. However, this does not mean that quantifiers are not important in the HSK composition difficulty assessment. The acquisition of quantifiers by foreign students in composition writing cannot be ignored, because quantifiers are a special grammatical phenomenon that distinguishes modern Chinese from European and American languages.

**Keywords**-modern Chinese; Hanyu Shuiping Kaoshi (HSK); quantifier; difficulty; assessment

## 1 INTRODUCTION

In recent years, the development of computer technology has broken the shackles of traditional writing teaching mode. As a tool of computer aided language learning, Automated Writing Evaluation (AWE) system has been continuously applied to English teaching and Chinese teaching. Automatic essay grader is a program that is designed to grade an essay automatically. Some of automatic essay grader have been made using string kernel, word embedding and reinforced learning method [1].

### 1.1 Research on Automatic English Composition Correction System

The research on composition correction system began in the United States. The earliest research was the Project Essay Grade (PEG) scoring mechanism developed by Professor Ellis Page of Duke University in the 1960s. This scoring mechanism can analyze and evaluate the internal characteristics of composition, such as length of words, average sentence length, etc [2]. In addition to PEG system, foreign famous AES systems also include Intelligent Essay

Assessment (IEA) [3], Electronic Essay rater (E-Rater) [4], Intelli Metric [5], Criterion, etc. developed in the 1990s. These systems can not only automatically evaluate the scores of compositions, but also feedback and modify opinions. E-rater mainly analyzes compositions from syntax, text and theme. According to relevant research, the accuracy of E-rater evaluation is highly consistent with manual scoring, ranging from 85% to 91% [6]. At present, the above automatic composition correction system has been used in primary and secondary school classrooms in California, USA [7].

## **1.2 Research on Automatic Correcting System of Chinese Compositions**

Compared with the automatic correction system for English compositions, the research on the automatic correction system for Chinese is slightly lagging behind. From the perspective of composition scoring, Ren has established a set of objective criteria based on experience and empirical research, which has been applied to the scoring of HSK composition papers and verified [8].

In addition, Huang etc. research on the feature selection of HSK automatic composition scoring [9], Liu researched on the key technology of automatic scoring of college entrance examination compositions [10], Li's proposal of "the development of the Chinese ability evaluation system for primary and secondary students", and Liu Hua's commitment to research and development of the Chinese Second language learners' automatic composition correction system for Chinese compositions and compositions in primary and secondary schools.

In a word, the research on the evaluation of Chinese composition difficulty level is still in the ascendant. This paper selects the topic of automatic evaluation of foreign students' compositions.

## **2 LINGUISTIC BACKGROUND**

### **2.1 Modern Chinese**

Modern Chinese is not only a language, but also a discipline, which has broad and narrow meanings. In a broad sense, modern Chinese refers to the language used by the Han nationality after the May 4th Movement. It not only includes modern standard Chinese (Putonghua), but also includes Chinese dialects. In a narrow sense, modern Chinese only refers to the common language of modern Han nationality - modern standard Chinese Putonghua.

The foreign students' compositions studied in this paper are written in modern Chinese and simplified Chinese characters.

### **2.2 Quantifier**

Quantifiers in Chinese is a distinctive category of words, which reflects the language habits and characteristics of the Chinese nation. In the traditional grammar of Europe and America, quantifiers are not independent. Influenced by this, Chinese classifiers were set up late [11]. At first, people regarded classifiers as nouns. Until the middle of the 1950s, after a great discussion on parts of speech, people finally pointed out in the Tentative Chinese Teaching Grammar System that quantifiers are words that represent the quantitative units of things or actions, and

they belong to the notional category together with nouns, verbs, adjectives, numerals and pronouns. For example,

wǒ de wǎnfàn shì yī tiáo yú.

My supper is a fish.

我的晚饭是一条鱼。

### **2.3 Hanyu Shuiping Kaoshi (HSK)**

The Chinese Proficiency Test (HSK) is an international standardized test that focuses on the ability of candidates who are not the first language of Chinese to communicate in Chinese in their life, study and work. HSK is a national standardized examination set up to test the Chinese language proficiency of native speakers (including foreigners, overseas Chinese and Chinese national minorities). The examination grades usually include HSK (Level 1), HSK (Level 2), HSK (Level 3), HSK (Level 4), HSK (Level 5) and HSK (Level 6).

HSK composition is the writing part of the Chinese Proficiency Test (HSK) (Advanced) launched in 1993. It plays an important role in improving the writing level of the tester and in understanding the expression and logical structure of Chinese deeply.

## **3 EXPERIMENT DESIGN AND ANALYSIS**

In this section, we will discuss the influence of quantifiers on the difficulty of HSK composition through quantitative methods.

### **3.1 Introduction to HSK Compositions Corpus**

In July 2021, the Ministry of Education released the “国际中文教育中文水平等级标准” (Chinese Proficiency Grading Standards for International Chinese Language Education, CPGSICLE). In short, it is the new version of HSK. As we all know, the previous version of HSK exam syllabus is divided into 6 levels. In contrast, the new version for HSK has greatly improved in terms of the number of Chinese characters and vocabulary, which shows that the difficulty of the HSK test has gradually increased. The old HSK standard has been used for nearly 10 years. Over the past decade, the number of people taking part in the Chinese language level test has continued to increase, and the standards for measuring various language skills have also improved, which also reflects the gradual increase in the international influence of Chinese.

However, since HSK (Level 7-9) will take the first global exam on November 26, 2022, it is also believed that the data on the classification of HSK composition level 7-9 is very lacking. Therefore, when assessing the difficulty of HSK compositions, a large number of HSK compositions accumulated previously are still classified according to the level 1-6 standard. Therefore, this paper still divides HSK compositions according to the old HSK standards.

In order to carry out HSK composition automatic difficulty evaluation smoothly, we extracted 5498 compositions from the HSK examination papers over the years, and then rated them by 5 experts according to the differences of difficulty levels. The vertical axis in the table I represents 1-6 grades of HSK, and the horizontal axis 0-3 indicates that each HSK grade is subdivided into 4 sub grades. Table 1 below shows the details of the corpus.

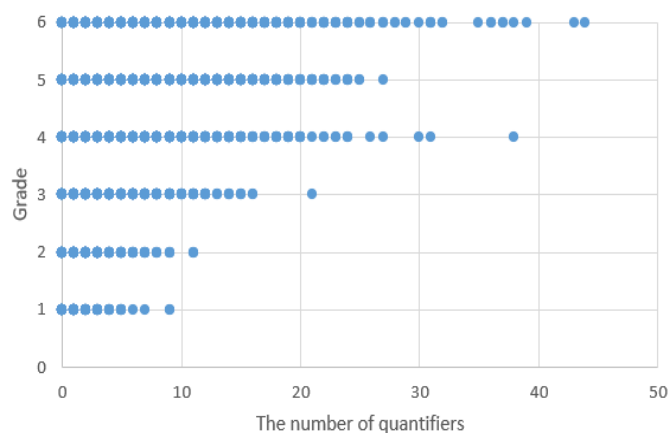
**Table 1** Details of HSK Compositions Corpus

Grade	0	1	2	3	Sum
1	49	23	57	22	151
2	28	120	80	72	300
3	159	337	135	176	807
4	313	540	435	261	1549
5	245	600	502	259	1606
6	92	281	372	340	1085
Sum	886	1901	1581	1130	5498

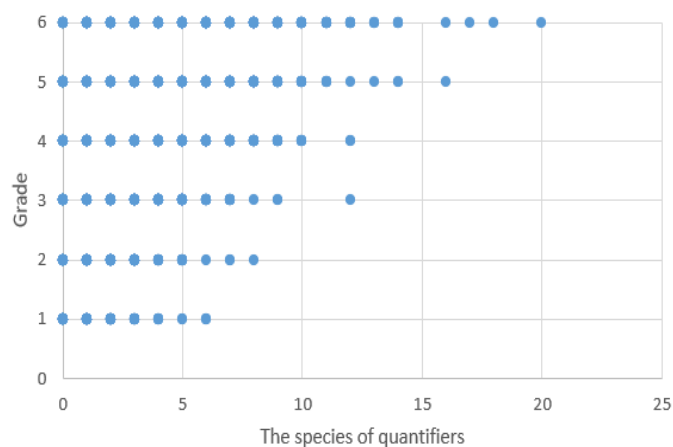
a. This corpus is the same as Li [12].

### 3.2 Quantitative Statistics of Quantifiers in HSK Compositions Corpus

In order to better understand the use of quantifiers in HSK compositions corpus, we have made statistics on their quantity and species, and the results are shown in the following figure 1 and figure 2.



**Figure 1.** Statistics of the number of quantifiers in HSK compositions corpus.



**Figure 2.** Statistics of the species of quantifiers in HSK compositions corpus.

The richness of vocabulary use is closely related to the grade of HSK composition. HSK vocabulary learning is a process of accumulation. Generally speaking, a low level composition cannot contain a high level vocabulary, while a high level composition can contain a low level vocabulary. From figure 1 and 2, it can be clearly found that with the increase of HSK level, the quantity and species of quantifiers in compositions are gradually increasing. Therefore, the use of quantifiers in Table 1 and Table 2 also conforms to our cognition of HSK learners' learning process.

### 3.3 The Influence of Quantifier Characteristics on Text Difficulty of HSK Compositions

#### 3.3.1 The contribution of quantifiers to measuring the difficulty of HSK composition.

In order to evaluate the contribution of quantifiers to measuring the difficulty of HSK's compositions, we respectively made correlation coefficient analysis between the number and species of quantifiers and the difficulty grades of HSK's compositions. The correlation coefficient between the number of quantifiers and the difficulty grades of HSK compositions is 0.38538407. In the meanwhile, the correlation coefficient between the species of quantifiers and the difficulty grades of HSK compositions is 0.392583051. 0.38538407 and 0.392583051 in the interval [0.2, 0.4), this also means that the quantity and types of quantifiers are weakly related to the difficulty level of HSK composition.

The latter is greater than the former, which means that the species of quantifiers have a greater impact on the difficulty of HSK compositions. Both values are positive, which means they are positively related to HSK compositions' difficulty. The higher the value, the greater the difficulty, and the smaller the value, the smaller the difficulty.

#### 3.3.2 Display the difficulty formula of HSK compositions.

To some extent, automatic composition correction can also be transformed into composition difficulty rating, which is one of the research fields of readability. Therefore, automatic composition evaluation can draw on the corresponding achievements of readability research.

Our research team has constructed a parameter system of 131 characteristics in three categories -- fluent features, complex features and cohesive features to measure the difficulty grades of HSK compositions. Through stepwise

regression, the following formula (1) is summarized by [12].

HSK compositions' readability grade  $Y = 0.0024833 * \text{the quantity of Chinese characters in } SGCCP - 0.9962586 * \text{the quantity of characters not in } SGCCP + 0.0690185 * \text{the average species of Chinese characters in sentences} + 1.3112404 * \text{the ratio of the species of words to the quantity of words} - 0.3901110 * \text{entropy (word)} + 0.0108346 * \text{the species of Chinese words in } SGWCP - 0.0364927 * \text{the species of words of grades 1} - 0.0312857 * \text{the species of words of grades 3} - 0.0043296 * \text{the quantity of characters} + 0.2541129 * \text{the square root value of the quantity of characters} - 0.0119805 * \text{the species of nouns} - 0.0697599 * \text{the ratio of the species of parts of speech to the quantity of sentences} - 0.0250014 * \text{the quantity of typos} - 1.6593806$  (1)

Since  $Y$  is a positive decimal, HSK composition grade is a discrete value within 1-6. Therefore, it is necessary to round  $Y$  to an integer as the readability grade of a HSK composition. The accuracy rate (ACR) of the formula (1) is 39.45%, and the residual interval of the formula is [-4.7917, 3.9641].

Formula (1) is a formula used by the research team to evaluate the difficulty of HSK compositions, which is convenient and practical. Of course, the accuracy rate of Formula (1) is certainly not the best among all kinds of machine evaluations, but the formula method can explain what features are important, which can bring great convenience to our teaching work.

*SGCCP* and *SGWCP* in Formula (1) are abbreviations of *the Syllabus of Graded Characters for Chinese Proficiency* and *the Syllabus of Graded Words for Chinese Proficiency* respectively. *SGCCP* divides Chinese characters into 4 grades: A, B, C and D. In the meantime, *SGWCP* divides Chinese words into 6 grades: 1, 2, 3, 4, 5 and 6. *SGCCP* and *SGWCP* are the normative guidelines for teaching Chinese as a foreign language. Therefore, they also guide the HSK examination, and it is self-evident that they are important to evaluate the difficulty of HSK compositions.

There are characteristic measures of the number and species of quantifiers in 131 parameters system. But through stepwise regression, we have eliminated the two characteristics of the number and the species of quantifiers. This does not mean that the characteristics of quantifiers do not contribute to measuring the difficulty of HSK compositions. Because the features that can enter the formula (1) must be statistically significant (in R language, the p value is very close to 0 and the form is \*\* \*), and all pass the hypothesis test. In other words, quantifiers may not be as important as we expected to measure the difficulty of HSK compositions.

### **3.3.3 The Influence of Quantifier Characteristics on Text Difficulty of HSK Compositions**

However, in order to quantitatively measure the contribution of quantifiers to the difficulty level of HSK compositions. We combine the quantity and species features of quantifiers with the 13 features of formula (1) to conduct multiple linear regression calculation.

- The quantity of quantifiers is combined with the 13 features of formula (1), and after multiple linear regression calculation, the weight value of it is 0.0111637. The accuracy rate of the whole formula is 39.87%.
- The species of quantifiers is combined with the 13 features of formula (1), and after multiple linear regression calculation, the weight value of it is 0.0125505. The accuracy rate of the whole formula is 39.91%.

0.0111637 is lower than 0.0125505, which, to some extent, again shows that the species of quantifiers have a greater impact on the difficulty level of HSK compositions. And the accuracy of 39.87% and 39.91% is slightly lower than that of Formula (1) 39.98% [12]. Adding the characteristics of the quantity and species of quantifiers can't significantly improve the accuracy of HSK composition difficulty rating.

#### 4 CONCLUSION

At present, the country is vigorously promoting *The Belt and Road*, jointly building *The Belt and Road*, committed to the connectivity of Asia, Europe, Africa and the adjacent seas, establishing and strengthening the connectivity partnership of countries along the line, and achieving diversified, independent, balanced and sustainable development of countries along the line. Therefore, all kinds of projects cannot be carried out without language first. With the rise of the Chinese language craze, the importance of HSK test is self-evident. Therefore, it is necessary to deepen the research of HSK. Based on the methods of automatic composition correction and text difficulty rating, this paper conducts a quantitative study on quantifiers in HSK compositions. The research results show that the species of quantifiers have a greater impact on the difficulty of HSK compositions. Adding the characteristics of the quantity and species of quantifiers can't significantly improve the accuracy of HSK composition difficulty rating. However, this does not mean that quantifiers are not important in the HSK composition difficulty assessment. The acquisition of quantifiers by foreign students in composition writing cannot be ignored, because quantifiers are a special grammatical phenomenon that distinguishes modern Chinese from European and American languages.

#### REFERENCES

- [1] Y. Salim, V. Stevanus, E. Barlian, A. Cempaka Sari, D. Suhartono, "Automated English Digital Essay Grader Using Machine Learning," IEEE Intelligent Systems, October 2020.
- [2] M. D. Shermis, J. Burstein, *Automated Essay Scoring: A Cross- disciplinary Perspective*, Mahwah, NJ: Lawrence Erlbaum, 2003, pp.43–54.
- [3] M. Hearst, "The debate on automated essay grading," IEEE Intelligent Systems, 2000, pp. 15:22–37.
- [4] S. Valenti, F. Neri, A. Cucchiarelli, "An overview of current research on automated essay grading," *Journal of Information Technology Education*, 2003, pp.3193–30.
- [5] M. D. Shermis, J. Burstei, "Automated Essay Scoring: A Cross- disciplinary Perspective," Mahwah, NJ:Lawrence Erlbaum, 2003, pp.113–122.

- [6] S. Valenti, F. Neri, A. Cucchiarelli, "An overview of current research on automated essay grading," *Journal of Information Technology Education*, 2003, pp.319–330.
- [7] S. Nobles, L. Paganucci, "Do digital writing tools deliver? Student perceptions of writing quality using digital tools and online writing environments," *Computers and Composition*, 2015, pp. 16–31.
- [8] C. Ren, "Discussion on Objectivity of HSK Composition Scoring," *Journal of Chinese Language Learning*, 2004(06). (in Chinese)
- [9] Z. Huang, J. Xie, and E. Xun, "Research on Feature Selection of HSK Automatic Composition Scoring," *Journal of Computer Engineering and Application*, 2014(06). (in Chinese)
- [10] M. Liu, "Research on the Key Techniques of the automatic scoring in college entrance examination essay," (Master Thesis) Harbin Institute of Technology, Harbin, China, 2015. (in Chinese)
- [11] J. Zhou, "A Study on the Combination of Verbs and Verbal Classifiers in Mandarin Chinese," (Doctor Thesis) Jinan University, Guangzhou, China, 2007. (in Chinese)
- [12] Y. Li, "Research on Readability Grade Formula Based on HSK Compositions," 2022 4th International Conference on Natural Language Processing (ICNLP), 2022, pp. 439–443, doi: 10.1109/ICNLP55136.2022.00081.