A Comparative Study of Vocabulary Richness in English Compositions between Chinese Non-English Majors and International Students from Non-Native English-Speaking Countries

Xuefei Peng[1], Mengyu Xu[2,a]

[1]International Students' Education Center, Dalian Maritime University, Dalian, 116000, China

[2]English Language and Literature, Korea Maritime and Ocean University, Busan, 612022, Korea

[a]xumengyu@g.kmou.ac.kr

Keywords: Vocabulary richness, Chinese virtue, English compositions

Abstract: According to the "accessing vocabulary" proposed by Read (2000), this paper analyzes the vocabulary richness in English writing. Moreover, the results show that both Chinese and international students have understood and experienced Chinese culture, and their writing styles are different. The results can give teachers a guideline when they teach, especially about Chinese culture.

1. Introduction

With the promotion of the Belt and Road Initiative, China's cultural soft power and influence are gradually increasing, and China has gradually become a choice for more and more overseas students to study. In China, international students experience Chinese culture, making them rich in professional ability. While studying in China, international students will also learn English with local students. Therefore, studying the English learning of the two types of students can effectively analyze students' acceptance of knowledge, which is conducive to teachers' timely adjustment in future teaching. Second-language acquisition (SLA) is to study how people learn a second language. SLA includes the type and amount of input, guidance and feedback, and the interactive environment for learning (Larsen-Freeman&Long, 1991). Vocabulary is an effective indicator to measure the language proficiency of second language learners (Linnarud, 1986). Currently, the study of lexical richness has attracted the attention of many scholars. So far, most lexical studies have taken lexical richness as the primary method of study and measurement.

2. Previous Studies

Gui and Yang (2003) collected the corpus, built the CLEC corpus and then observed the English characteristics and the situation of all kinds of students by marking speech errors. Lu (2012) used automatic calculation to detect three indicators of lexical richness, namely, lexical density, lexical complexity and lexical diversity, to seek the relationship of lexical richness to the quality of ESL learners' oral narratives. Wang (2017) selected 90 compositions from the HSK Dynamic Composition Corpus of Beijing Language and Culture University as research samples and investigated the factors affecting Chinese learners' composition scores from the three aspects.

3. Data and Methodology

3.1 Data

This paper randomly selected 48 English writings from non-native English speaking students (postgraduates) and Chinese non-native English majors (postgraduates) studying at Dalian Maritime University. Nationality of international students come from 15 countries, such as Pakistan, Tanzania, Mongolia and so on. Table 1 shows the situation of textual studies.

Table 1 the Situation Of Textual Studies.

	Chinese students	Non-native English language students
Number of essay submissions	49	59
Number of valid essays	48	56
The number of essays chosen at random	48	48
Average essay word number	316.10	368.06
Composition writing time	October 2021	October 2021

The topic of this essay is: There are so many virtues (such as loyalty, piety, affection, etc.) cherished by most Chinese people. Please give examples to illustrate which virtue has affected you deeply. Please use your own words and write at least 300 words.

3.2 Methodology

3.2.1 Lexical Diversity

TTR (Tokens/Types) is a typical indicator of vocabulary diversity. However, TTR lacks comparability between compositions of different lengths (Wan Lifang, 2010). Jarvis (2000) believes that Uber Index is a more appropriate way to calculate vocabulary diversity because Uber Index is more suitable for studies with a smaller corpus (Wan Lifang, 2010). The formula of the Uber Index is: $Log^2N/(logN-logV)$. N means tokens, and V means types. This paper adopts TTR and Uber Index dual detection methods to calculate the lexical diversity.

3.2.2 Lexical Complexity

Lexical complexity is also known as lexical sophistication. Laufer and Nation (1995) proposed a lexical frequency profile, which was used to calculate the low-frequency words in the selected corpus, and divided written words into four groups. The first group and the second group contain 1000 words, respectively. The third batch is the academic word list, and the fourth group is the words not listed above (other words). Now there are much software for analyzing the word complexity. This paper will use AntWordProfiler.

3.2.3 Lexical Density

Lexical density is used to estimate the linguistic complexity in a written or spoken composition from the functional words and content words. The formula of lexical density is the number of content words/ tokens in the text.

3.2.4 Lexical Error

This paper will from 13 aspects, Add an article, Add the particle, Change preposition, Change the adjective, Change the pronoun, and so on to analyze common mistakes made by students.

4. Results and Discussion

4.1 A Comparison of Lexical Diversity between Two Types of Texts

Table 2 Values of TTR and Uber Index in Two Kinds of Types.

	Types	Tokens	TTR%	Uber Index
Chinese students	2480	15, 173	16.38	0.011
Non-native English language students	2947	17,667	16.76	0.012

According to Table 2, the value of TTR and Uber Index of international students' writings are all higher than these of Chinese students' writings. It can be concluded that compared with Chinese students, international students are better at using rich vocabulary to express their ideas in writing.

4.2 A Comparison of Lexical Complexity between Two Types of Texts

Table 3 Lexical Complexity Of Two Kinds of Texts by Antwordprofiler Software.

	Band One	Band Two	Band Three	Not in the list
Chinese students	77.6	5.3	5.0	12

Non-native English language students	79.8	5.4	4.3	10.5

From this table, it can be summarized that Chinese students prefer to use advanced vocabulary because the statistic of brand three is higher than that of non-native English language students.

This paper also analyzes the frequency of words of different lengths in the two types of texts. It found the ratio of 1-letter and 2-letter words in the two types of texts is relatively consistent. Chinese students are better at using 3-letter words, while international students are better at using 4-letter words and words with more than four words.

4.3 A Comparison of Lexical Density between Two Types of Texts

Through WordSmith8.0, the vocabulary lists were obtained. Then, function words were selected by manual deletion, and finally content words, function words and lexical density were calculated.

Table 4 Lexical Density In Two Kinds of Texts.

	Types	function words	content words	lexical density%
Chinese students	2480	27	2453	98.91
Non-native English language students	2947	19	2928	99.36%

It can be seen the number of function words used by non-native English language students is less than that of Chinese students, and they have a higher vocabulary density in writing.

4.4 A Comparison of Lexical Error between Two Types of Texts

4.4.1 Add an Article

(1) Bai Juyi, great poet ...(Chinese student) → Bai Juyi, a great poet ...

(2)...it has deeper meaning ... (Chinese student) →...it has a deeper meaning ...

4.4.2 Add the Particle

(3) He must stand firm himself ... (Chinese student) →He must stand to firm himself ...

(4)They begin make efforts to ...(Chinese student) →They begin to make efforts to ...

4.4.3 Change Preposition

(5) ...has been put on a more important position. (Chinese student) →...has been put in a more important position.

(6)When I did something to boast of and ... (Non-native English language students) → When I did something to boast about and ...

4.4.4 Change the Adjective

(7) ...we want to destroy evil and promote good. (Chinese student) → ...we want to destroy evil and promote well.

4.4.5 Change the Pronoun

(8) And the one who you should be... → And the one which you should be...

4.4.6 Change the Word

(9) ...if everyone has kindness in the heart.(Chinese student) → ...if everyone has kindness in their heart.

4.4.7 Correct Article Usage

(10) Liu Heng, emperor Wen of Han Dynasty... (Non-native English language students) → Liu Heng, emperor wen of the Han Dynasty...

4.4.8 Correct Word Choice

(11) Kindness is tolerance, love, compassion. (Chinese student) → Kindness is tolerance, love, and compassion.

4.4.9 Fix the Infinitive

(12) It is that fuel that drives us to do things well, to be patient, to improve day by day and to look ... (Chinese student) →It is that fuel that drives us to do things well, ~~to~~ be patient, ~~to~~ improve day by day and ~~to~~ look ...

4.4.10 Replace the Word

(13) ...Ren is the base of a harmony society. (Chinese student) → ...Ren is the base of a harmonious society.

4.4.11 Singular and Plural Problem

(14) ...to achieve higher prize. (Non-native English language students) → ...to achieve higher prizes.

(15) Both cultural idea and etiquette system. (Non-native English language students) → Both cultural ideas and etiquette systems.

4.4.12 Wrong Verb Form

(16) His stepmother said she wants to... (Chinese student) → His stepmother said she wanted to ...

(17) ...until some people helps me. (Chinese student) → ...until some people helped me.

(18) ...two craps jumping out of the river. (Non-native English language students) → ...two craps jumped out of the river.

(19) ...making him lost his father's love. (Non-native English language students) → ...making him lose his father's love.

4.4.13 Wrong Verb Tense

(20) ...I read a story that makes me know... (Chinese student) →...I read a story that made me know...

(21) ...there are many loyal people... (Chinese student) → ...there have been many loyal people...

From the above 13 vocabulary problems commonly encountered by students in writing, it can be concluded that Chinese students encounter more vocabulary problems than non-native English language students. Both Chinese and non-native English language students have many problems with verb forms. Chinese students have fewer problems than non-native English language students in terms of singular and plural vocabulary.

4.5 A Comparison of Writing Theme and Examples between Two Types of Texts

In order to better study the understanding degree of Chinese and international students to Chinese culture, the writing themes of 96 Chinese and international students' compositions were counted by manual retrieval. It can be concluded that compared with Chinese students, non-native English language students have a more comprehensive understanding of traditional Chinese qualities. Respect, peace and love, faithfulness, friendliness, loyalty and filial piety were demonstrated in the compositions of non-native English language students. Among all traditional virtues, harmony is reflected in Chinese and international students' compositions. It accounts for a large proportion, accounting for 13% of all Chinese students' compositions and 15% of non-native English language students' compositions.

5. Conclusion

This paper from four aspects: lexical diversity, lexical complexity, lexical density and lexical errors analyzes students' English compositions from China and non-native English-speaking countries to study vocabulary richness. Besides the four perspectives, this paper also researched the

composition theme and found that international students from non-native English-speaking countries are more willing to explore traditional Chinese virtues.

However, there are several limitations, such as only taking 96 writings to analyze vocabulary richness, focusing on one university to collect and select students' writing. All in all, the results shown in this paper deserve other teachers' attention during their teaching process.

Acknowledgement

References

[1] Gui, S.C.; Yang, H.Z. Chinese Learner *English* Corpus. Shanghai: Shanghai Foreign Language Education Press, 2003, pp.11-13.

[2] Larsen-Freeman, D.; Long, M. An Introduction to Second Language Acquisition Research. London: Longman, 1991, pp.155.

[3] Laufer, B.; Nation, P. Vocabulary Size and Use: Lexical Richness in L2 Written Production. Applied Linguistics, no.16, pp.307-322, 1995.

[4] Linnarud, M. Lexis in Composition: A Performance Analysis of Swedish Learners' Written English. LundSweden:CWK Gleerup, 1986, pp.138.

[5] Lu, X. F. The Relationship of Lexical Richness to the Quality of ESL Leaners' Oral Narrative. The Modern Language Journal, vol.96, no.2, pp. 190-208, 2012.

[6] Read, J. Assessing Vocabulary. Cambridge: Cambridge University Press, 2000, pp.168.

[7] Wan, L.F. A study on vocabulary richness in Second Language studies of Chinese College students majoring in English. Foreign language world, no.1, pp.41, 2010.

[8] Wang, Y.X. Correlation between HSK composition score and sentence length, complexity and grammatical errors. Research on Chinese Applied Linguistics, no.12, pp. 139-148, 2017.