

A Study on Chinese EFL Learners' Vocabulary Usage in Writing

Lili Zhai

Foreign Languages School, Shanxi Normal University, Linfen, China

Abstract—Vocabulary acquisition is one of the hottest research fields in English learning, which has aroused researchers' great attention in recent years. However, their focus is on vocabulary size, vocabulary learning strategies and receptive lexical ability, seldom to productive lexical ability. Writing is an important productive ability for EFL learners, and a myriad of writing researches show that inappropriate vocabulary use leads to inferior writing quality. Therefore, research on learners' vocabulary proficiency, especially their vocabulary in English writing is quite profound. 66 subjects from a comprehensive university participating in this study, finished one composition for analyzing their vocabulary usage, i.e. lexical richness which includes lexical sophistication and lexical variation. All the data and writing papers were analyzed with RANGE and SPSS 17.0. The findings of the present study demonstrated that the subjects relied more on the first 1000 word level to express their meanings in productive tasks and the lexical sophistication and lexical variation are not high. Subjects with different writing ability have differences in vocabulary usage, but the two groups only have significant difference in lexical variation and not in lexical sophistication. The present study enriches the research on vocabulary acquisition in SLA and provides helpful implications for vocabulary teaching and learning to improve learners' vocabulary productive ability.

Index Terms—vocabulary usage, English writing, Chinese EFL learners

I. INTRODUCTION

Vocabulary, as the information carrier, plays an indispensable role of language. Likely, vocabulary is a significant part in English learning, so vocabulary acquisition becomes one of the hottest research fields. Wilkins (1972) states that without grammar very little can be conveyed; without vocabulary nothing can be conveyed, which shows that vocabulary is of crucial importance to language learners (cited from Chen Hui, 2001). Although the status and importance of vocabulary acquisition has been acknowledged by linguists and language teaching researchers, they have attached more attention to vocabulary size, vocabulary learning strategies and receptive lexical ability, seldom to depth of vocabulary knowledge and productive lexical ability.

Writing is an important productive ability for EFL learners, while it is a relatively ordinary phenomenon in China that university students are apt to use a host of high-frequency words to express their idea in English writing, and if trying to use low-frequency words, usually they use them in a wrong way. It indicates that university students have problems in vocabulary learning and use. Therefore, research on learners' vocabulary proficiency, especially their vocabulary in English writing is quite profound. The present study tries to find the characteristics of the subjects' vocabulary usage in English writing and the relationship between vocabulary usage and writing quality. The research mainly involves 3 questions as follows.

Q1: What are the characteristics of subjects' vocabulary usage in English writing?

Q2: What's the correlation between vocabulary usage and writing quality?

Q3: Does there exist significant difference on vocabulary usage between groups of different writing ability?

II. LITERATURE REVIEW

The assessment of vocabulary knowledge is an important field for Second Language Acquisition (SLA) research, which has aroused researchers' attention (Nation & Laufer, 1995; Read, 1993; Wesche & Paribakht, 1995). Bao Gui (2008) classifies two types of assessment according to the salience of words in the design. One is to design a host of vocabulary tests with target words on different word frequency, which contains receptive and productive vocabulary test. The other is to apply words to a communicative task (oral output or writing a composition), which aims to check learners' lexical usage during the process. The former has wide vocabulary coverage and is easy to design and test, while the latter pays more attention to communicative function of language and puts context into consideration. Researchers usually use 4 indexes, i.e. lexical variation, lexical sophistication, lexical density, lexical originality to describe learners' productive ability, which are called lexical richness in a general term. The present research placed an emphasis on the latter, and tried to find the relationship between vocabulary usage and writing quality.

A. Assessment of Vocabulary Usage

Vocabulary usage is to use vocabulary in productive tasks, and the assessment of vocabulary usage is measured by lexical richness which is calculated by four items, i.e. lexical originality (LO), lexical density (LD), lexical sophistication (LS); and lexical variation (LV).

Before explaining the four items, another two items “token” and “type” need introducing firstly and the two are used to count words in a text. The number of *tokens* is the same as the total number of word forms, which means that individual words occurring more than once in the text are counted each time they are used. On the other hand, the number of *types* is the total number of the different word forms, so that a word which is repeated many times is counted only once (ibid).

Lexical Originality is the percentage of words in one learner’s text that were not used by any of the other learners.

$$LO = \frac{\text{Number of tokens unique to one writer} \times 100\%}{\text{Total number of tokens}}$$

The lexical origination index measures the number of words unique to one learner in the test group. If the group changes, the index changes too, so it is unreliable (Laufer, 1994).

Lexical Density is defined as the percentage of lexical words in the text, i.e. nouns, full verbs, adjectives, adverbs.

$$LD = \frac{\text{Number of lexical tokens} \times 100\%}{\text{Total number of tokens}}$$

Since lexical words are the words which primarily convey information, a text is considered ‘dense’ if it contains many lexical words relative to the total number of words, i.e. lexical and functional words. Since it depends on the syntactic and cohesive properties of the composition as reflected in the use function word, its validity is questionable (ibid).

Lexical Variation is the type/token ratio, i.e. the ratio in percent between the different words in the text and the total number of words.

$$LV = \frac{\text{Number of types} \times 100\%}{\text{Number of tokens}}$$

The type/token ratio is the most widely used measure and LV can show how well a learner can express himself with different vocabulary he knows.

Lexical Sophistication is the percentage of ‘advanced’ words in the text.

$$LS = \frac{\text{Number of advanced tokens} \times 100\%}{\text{Total number of lexical tokens}}$$

Because the definition of ‘advanced’ depends on the researcher, it is unreliable. Due to the limitation of this measurement, Laufer and Nation (1995) devised a new measure of lexical richness, that is, the Lexical Frequency Profile (LFP) and designed a word frequency program-VocabProfile. It is based on the relative frequency of words in the language and involves simply calculating the percentage of word families in the learner’s composition that belong to each of three or four frequency bands. In their original study, Laufer and Nation used a profile that comprised four levels: first 1000 most frequent words, second 1000 most frequent words, words in the University Word List, and any other (less frequent) words. Thus, if a learner wrote an essay containing 200 words family word families in total and these consisted of 150 from the first 1000 list, 20 from the second 1000 list, 20 from the University Word List, 10 other words, the profile would be 75%-10%-10%-5%. The authors’ argument is that the LFP provides a more objective and differentiated measure of the learners’ vocabulary use than the other statistics, because it draws on a range of established word lists to classify the words into categories (Read, 2000).

Qin Xiaoqing & Wen Qiufang (2007) described the advantages of using LFP to assess lexical sophistication. LFP depends on three word frequency to calculate the percentage of word families, so it has a high maneuverability. In addition, LFP can effectively differentiate the subjects who use the first 1000 most frequent words and second 1000 most frequent words etc. effectively. What’s more, it has a high reliability. At last but not least, LFP has a good forecast of language proficiency. Afterwards, Nation and Coxhead increased the function of VocabProfile and renamed RANGE which includes the first 1000 word family, the second 1000 word family, and 570 academic words in 2000.

RANGE is available at http://www.vuw.ac.nz/lals/staff/Paul_Nation. RANGE contains 3 word lists, named BASEWORD1.txt, BASEWORD2.txt, BASEWORD3.txt, representing the first 1000 word family, the second 1000 word family, and 570 academic words. RANGE can be used to compare a text against vocabulary lists to see what words in the text are and are not in the lists, and to see what percentage of the items in the text are covered by the lists. It can also be used to compare the vocabulary of two texts to see how much of the same vocabulary they use and where their vocabulary differs. The present study employed RANGE to assess lexical sophistication.

Wolfe-Quintero etc.(1998) summarize a host of literature and find that lexical richness, especially lexical variation, lexical sophistication have a significant relationship with language development (ibid). Furthermore, lexical density and lexical origination have their inherent weakness, so the present research only adopted the two items to analyze learners’ vocabulary usage in EFL writing.

B. Empirical Studies on Vocabulary Usage and Writing

Read (2000) assumed that good writing has the following lexical features.

1. *A variety of different words rather than a limited number of words used repeatedly. It is reasonable to expect that more proficient writers have a larger vocabulary size that allow them to avoid repetition by using synonyms,*

superordinates and other kinds of related words. This can be reflected in the type-token ratio or lexical variation index.

2. *A selection of low-frequency words which are appropriate to the topic and style of the writing, rather than just general, everyday vocabulary. This is actually the lexical sophistication.*

3. *A relatively high percentage of lexical (content) words, as compared with grammatical (function) words. This is known as lexical density.*

4. *Few if any lexical errors in the use of words. Thus another measure of writing quality can be the number of lexical errors in the written text.*

The above description reflects the relationship between lexical richness and writing. The following part is the empirical studies in detail. Researches on lexical richness are limited, and these researches mainly from two aspects: one is the comparison of lexical richness on different writing abilities and the other is to investigate the relationship between lexical richness and writing quality.

Many researches consider that learners with higher writing ability use less repeated vocabulary than those with lower ability. Linnarud (1986) compares the compositions written by the Swedish English learners and by their native-speaking peers, and finds lexical variation of Swedish English learners is lower than that of native speakers (Bao Gui, 2008). Laufer (1991) finds that there is no difference of lexical variation of learners among different writing ability stages. Among researchers on the relationship between lexical variation and writing quality, Engber (1995) finds their significant relation, while others rarely get their positive correlation. Li Zhixue and Li Jingquan (2005) find Chinese learners with higher English proficiency produce less productive vocabulary than American students in EFL writing.

Linnarud (1986) finds native-speaking students' lexical density is higher than second language learners in writing, while Laufer (1991) and Engber (1995) find no obvious relationship between lexical density and writing quality.

For there are two kinds of definitions of lexical sophistication, there exist two results. Linnarud (1986) defines lexical sophistication is the ratio of sophisticated tokens and total tokens, and finds native speakers use more advanced words. Liu Donghong (2004) adopts LFP to measure lexical sophistication, and he defines words on UWL and not in the lists as sophisticated words. His result shows that there is no significant relationship of lexical sophistication between higher and lower writing groups.

Researches on lexical originality are rare, and the results are contradictory. Linnarud (1986) finds native speakers use more original vocabulary than second language learners, and lexical originality has a positive correlation with writing quality. Laufer (1991) finds no development of original vocabulary between different writing stages.

Recently, Wan Lifang (2012), Wang Haihua & Zhou Xiang (2012), Yang Yingying (2012) and Zhu Huimin & Wang Junju (2013) made researches on lexical richness in EFL writing, which shows that researchers increasingly pay attention to this field. Nonetheless, it is easy to find that researches on vocabulary usage are not enough and some related findings are inconsistent. Therefore, it is necessary to enrich such research. The present research investigates the characteristics of vocabulary usage as well as its correlation with EFL writing quality. This research can make up studies on productive vocabulary ability and give some implications for vocabulary teaching and learning.

In this present research, the vocabulary usage i.e. lexical richness is only assessed from lexical variation and lexical sophistication, and lexical sophistication is measured through RANGE.

III. METHODOLOGY

A. Subjects

Subjects of the present study were 66 (35 males and 31 females) sophomores of non-English majors, chosen from two parallel classes of a comprehensive university. They were 19-21 years old and had studied English for more than 8 years. All of them had passed CET4, so they were intermediate learners regarding their English proficiency level.

Engber (1995) states that English language learners who fall within the intermediate range can provide unique insight into the language learning and writing process. For intermediate learners, their interlanguage is unstable, and they are courageous to test hypotheses, so it is reasonable to survey this kind of learners.

B. Instruments

The present study involved a timed composition.

Timed composition was used to analyze characteristics of the subjects' vocabulary usage in English writing from the perspective of lexical sophistication, and lexical variation with the help of the software RANGE. Because RANGE is stable across writing with 200 words or even more (Laufer & Nation, 1995), the subjects were required to write more than 200 words within 40 minutes in class according to the following directions:

All of us would agree that in order to be successful in the present day society, we have to possess certain personal qualities that can enable us to realize our aim. What do you think is the most important personal quality of a successful person?

The topic of composition was chosen considering the criterion that it was fit to the subjects' cognition and familiar to them. They can easily expand the content to 200 words.

C. Data Collection and Processing

Composition writing was conducted during regular class time. The researcher, also the subjects' English teacher,

informed that composition score would be integrated into their final score. Thus the subjects were motivated to finish the composition carefully.

During the process composition writing, dictionaries, peer help were unavailable, and writing should be within 40 minutes. The experiment moved smoothly with the subjects' active cooperation.

All the 66 pieces of writing were typed into computer with 2 copies, one for teachers' scoring rating, and the other for statistical analysis. The subjects' writing was the mean score of 2 college English teachers according to the writing rating scales of CET4 with total score of 15 points (see Appendix).

Finally, SPSS 17.0 was used to do statistical analysis and RANGE for exploring characteristics of vocabulary usage in writing.

IV. RESULTS AND DISCUSSION

A. The Description of Vocabulary Usage in Writing

Writing paper is a kind of form to embody how learners use words in communicative tasks. In order to get more information of the subjects' vocabulary knowledge, their writings were analyzed by the software RANGE. Two variables, lexical sophistication and lexical variation were considered to examine the characteristics of their vocabulary usage.

To calculate RANGE, there exist three kinds of data in the form of percentage according to token, type and family. Because writing papers in the present research are small texts (about 200 words) and the concept of type and family are too broad, the research adopts the data in the item of token.

TABLE 4.1
THE DESCRIPTIVE STATISTICS OF VOCABULARY USAGE

| | Minimum | Maximum | Mean | Std. Deviation |
|-------------------------|---------|---------|--------|----------------|
| LS 1 st 1000 | 80.20% | 98.83% | 89.13% | 3.79 |
| 2 nd 1000 | .58% | 10.89% | 4.96% | 2.20 |
| UWL | .00% | 6.28% | 1.84% | 1.360 |
| Beyond list | .00% | 10.84% | 3.91% | 2.40 |
| tokens | 135.00 | 342.00 | 217 | 38.460 |
| types | 76.00 | 160.00 | 113.7 | 18.010 |
| LV | 34.62% | 64.85% | 52.79% | 6.21 |

Note: 1st 1000= the first 1000 vocabulary
 2nd 1000= the second vocabulary
 Beyond list= vocabulary not in the list
 LV= lexical variation

Table 4.1 is the descriptive statistics of vocabulary usage and the two items, lexical sophistication and lexical variation, were described. From this table, it is easy to find the subjects can write enough long words to express their thought, for the mean score is 217, but they rely more on the first 1000 word level to express their meaning (89.13%>4.96%>3.91%>1.84%), and the highest one is 98.83%. There are two reasons to explain this phenomenon. Firstly, the subjects have only mastered limited productive vocabulary, so they have to use repeated words to express the same or similar meaning. Secondly, they may be scared to make mistakes with complicated words which influence their writing quality. The learners avoid using new words consciously, so they depend on the high-frequency words. The avoidance is a serious impediment to the increase of productive vocabulary. Therefore, the subjects should use the other three word levels to product in communicative tasks.

The mean ratio of lexical variation is 52.97%, and it is relatively low, which shows that the subjects lack productive vocabulary. Furthermore, they rely on the 1st 1000 words in writing. Due to these limitations, they are doomed to lack vocabulary to express their meanings but repeat words.

B. The Correlation between Vocabulary Usage and Writing Quality

Vocabulary is the foundation of writing, and vocabulary is one of the most important features that determine writing quality. Vocabulary usage influences writing quality, and lexical richness is one of distinct characters of good writing; however, researches on vocabulary usage and writing are not only limited, but also inconsistent, so the present study made a correlation analysis on vocabulary usage and writing quality and the results were shown in Table 4.2.

TABLE 4.2
THE CORRELATION BETWEEN VOCABULARY USAGE AND WRITING QUALITY

| Correlations | | | | | | | |
|----------------------|----------------------|----------------------|---------|--------------|---------|---------|--------|
| | LS | | | | | | |
| | 1 st 1000 | 2 nd 1000 | UWL | Beyond lists | tokens | LV | Score |
| 1 st 1000 | 1 | -.578** | -.511** | -.736** | -.105 | -.582** | -.175 |
| 2 nd 1000 | -.578** | 1 | .054 | .081 | .210 | .423** | .408* |
| UWL | -.511** | .054 | 1 | .280* | .208 | .409** | .473* |
| Beyond list | -.736** | .081 | .280* | 1 | -.010 | .404* | .545* |
| Tokens | -.105 | .210 | .208 | -.010 | 1 | -.493** | -.060 |
| LV | -.582** | .423** | .409** | .404** | -.493** | 1 | .552** |
| Score | -.175 | .408* | .473* | .545* | -.060 | .552** | 1 |

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Note: Score= the writing score

Table 4.2 reports the correlation between lexical sophistication, lexical variation, tokens and writing quality. It shows that the use of 1st 1000 words has a highly negative relationship with the use of 2nd 1000 words, the academic words as well as the words not in the lists (-0.578, -0.511, and -0.736). Lexical variation has a negative relationship with the use of 1st 1000 words and tokens (-.582, -.493), but a highly positive relationship with the 2nd 1000 words, UWL and the words not in the list (0.423, 0.409, 0.404). It indicates that the more 1st 1000 words learners use, the smaller lexical variation is; in reverse, the learners are apt to use the words beyond 1st 1000 words, while the lexical variation is high. Besides, lexical variation has a strong negative relationship with tokens (-.493), for it is the ratio of types and tokens. Bao Gui (2008) also considers that the longer learners write, the smaller lexical variation is, for they are apt to repeat words; therefore, length has an influence on lexical variation. In the present study, nearly all the subjects write a little more than 200 words, and the length has no great difference, so the length of writing papers has no great influence on calculation of LV in this way. However, it can not be denied that there exists a disadvantage to calculate LV in this way. Learners who have good command of productive vocabulary knowledge can write longer compositions but their lexical variation are low, so this calculation needs revising.

Writing quality has a negative relationship with the use of 1st 1000 words (-0.175), but a high positive correlation with the use of words beyond 1st 1000 words (0.408, 0.473 and 0.545), which means that when learners use ordinary and high-frequency words to express their meanings, they usually get low marks. Besides, writing quality has a significant positive correlation with lexical variation (0.552). When learners choose different and changeable words to express the same meaning, they can get high marks, for various kinds of expression make their writings interesting and attractive. Ma Guihua and Shi Yongzhen (2006) also find that writing quality has a significant negative correlation with 1st 1000 words, but a positive correlation with words beyond 1st 1000 words. What's more, for learners of lower language proficiency, their writing quality has no significant correlation with lexical variation, while only for learners of the higher language proficiency, their writing quality has a significant correlation with lexical variation. As they mentioned, this is probably because learners of lower language proficiency master limited vocabulary and their lexical variation is pretty low. In their research, language proficiency to some extent influences the correlation between lexical richness and writing quality. In the present study, language proficiency is not considered, but these two studies have got some similar results.

C. T-test of Vocabulary Usage between Groups of Writing Ability

Table 4.2 has shown that writing quality has a certain relationship with lexical sophistication and lexical variation. In detail, writing quality has a negative relationship with the use of 1st 1000 words, but a high positive correlation with the use of words beyond 1st 1000 words and lexical variation. Whether there exist significant differences in vocabulary usage between groups of different writing ability? The current study made a descriptive statistics of vocabulary usage and T-test between groups. The results were presented in Table 4.3 and Table 4.4. The subjects were reordered according to their writing score and divided into three groups. The first 25% was the higher group (Group 1) and the last 25% was the lower group (Group 2).

TABLE 4.3
THE GROUP STATISTICS OF VOCABULARY USAGE

| | Group | Mean | Std. Deviation | Std. Error Mean |
|----------------------|---------|-------|----------------|-----------------|
| 1 st 1000 | Group 1 | 87.79 | 3.98 | .913 |
| | Group 2 | 90.23 | 3.89 | .892 |
| 2 nd 1000 | Group 1 | 5.42 | 1.99 | .456 |
| | Group 2 | 4.74 | 2.21 | .506 |
| UWL | Group 1 | 2.16 | 1.23 | .282 |
| | Group 2 | 1.60 | 1.53 | .351 |
| Beyond lists | Group 1 | 4.62 | 2.60 | .596 |
| | Group 2 | 3.42 | 2.68 | .614 |
| LV | Group 1 | 56.68 | 5.37 | 1.231 |
| | Group 2 | 50.46 | 6.24 | 1.432 |

Note: Group 1= Group of higher writing ability
Group 2= Group of lower writing ability

Table 4.3 is a descriptive statistics of vocabulary usage between groups. It shows that learners with higher writing ability can use more words on the 2nd 1000 word list, university word level and words not in the lists than the lower group (5.42>4.74; 2.16>1.60; 4.62>3.42), but use less words on the 1st 1000 word list (87.79<90.23). In addition, it also shows that the higher group’s lexical variation is higher than the lower group (56.68>50.46). It means that the higher group can use more different words to express their meanings than the lower group. The reasons leading to the above results are that writing quality has a negative relationship with the use of 1st 1000 words, but a high positive correlation with the use of words beyond 1st 1000 words and lexical variation in Table 4.2. Whether there exist significant differences of vocabulary usage between the higher and lower group? T-test was needed. The results of T-test were presented in Table 4.4.

TABLE 4.4
T-TEST OF VOCABULARY USAGE OF TWO GROUPS

| | t | df | Sig. (2-tailed) | 95% Confidence Interval of the Difference | |
|----|----------------------|--------|-----------------|---|----------|
| | | | | Lower | Upper |
| LS | 1 st 1000 | -1.914 | .064 | -5.03059 | .14533 |
| | 2 nd 1000 | .998 | .325 | -.70234 | 2.06234 |
| | UWL | 1.248 | .220 | -.35138 | 1.47560 |
| | Beyond lists | 1.403 | .169 | -.53523 | 2.93628 |
| LV | 3.291 | 36 | .002 | 2.38481 | 10.04572 |

Table 4.4 shows that there is no significant difference of lexical sophistication between two groups, because the four Sig. (2-tailed) are larger than .05 and 0 is between the 95% Confidence Interval of the Difference. However, it shows that the difference of lexical variation is significant, for the Sig (2-tailed) are smaller than .05 and 0 is excluded in the 95% Confidence Interval of the Difference. Although the two groups have differences in lexical sophistication, but the differences are not significant. The reason is that the subjects lack low-frequency words but rely on 1st 1000 words, so all the subjects need to improve their vocabulary knowledge, especially productive vocabulary. However, learners with different writing ability have a significant difference of lexical variation, so higher lexical variation leads to higher writing scores in this research. Table 4.2 has shown that lexical variation has the highest correlation with writing quality. The two results agree with each other.

V. CONCLUSION

The current research explores the correlation between vocabulary usage and writing quality. The major findings are summarized in the following parts, together with the implications from pedagogical perspectives and the limitations as well as suggestions for further study.

A. Major Findings

The research found that the subjects relied more on the first 1000 word to express their meanings in productive tasks. Although they can write enough long papers, their lexical variation was not high. Besides, it also found that the use of 1st 1000 words had a highly negative relationship with the use of 2nd 1000 words, the academic words as well as the words not in the lists. Lexical variation had a negative relationship with the use of 1st 1000 words and tokens, but a highly positive relationship with the second 1000 words, UWL and the words not in the list.

Writing quality had a negative relationship with the use of 1st 1000 words, but a high positive correlation with the use of words beyond 1st 1000 words as well as lexical variation. Subjects with different writing ability had differences in word use. The higher group used fewer words on the 1st 1000 word level but inclined to the words beyond 1st 1000, and

the difference was not significant. What's more, the higher group used different words to express the same or similar meaning, so that their lexical variation was high. The lower group was in reverse. The two groups had significant difference in lexical variation.

B. Pedagogical Implications

Pedagogically, the findings indicate that L2 vocabulary instructions should pay more attention to productive vocabulary. Teachers should master vocabulary teaching theories systematically and understand the rules of vocabulary development. According to Henriksen (1999), vocabulary development is a continuum, from partial to precise, from receptive to productive and the gradual increase of depth of knowledge. Abided by the rules, teachers should help learners build lexical networks, e.g. the syntagmatic and paradigmatic lexical relations in their minds. Thus, learners can eliminate the anxiety of lacking vocabulary but use different words to express the same or similar meaning. When learners learn new words, teachers should provide more opportunity for them to use the words they have learnt. Teachers can adopt different kinds of activities to motivate learners' interest of vocabulary learning, such as guessing words, making stories with selected words. Moreover, because the class time is limited, teachers can arrange some writing tasks to learners after class. In the writing instruction, teachers can help learners get accustomed to thinking of vocabulary as part of the writing process. Writing is a kind of productive task, and during this process, learners can use words to express their meanings.

Learners take the major responsibility of expanding and deepening vocabulary knowledge, so they should pay attention to the following things. Learners should use the new words boldly, and do not be afraid about lexical errors especially in productive tasks. Lexical errors give some hints of weakness in vocabulary learning. Of course, when learners learn new words, they should master the vocabulary knowledge consciously. They should know its meaning, spelling, pronunciation, collocation and grammatical usage etc. In the present research, the subjects relied more on the 1st 1000 words, and they used easy and simple words to express the same and similar meaning, so that their writings were in low lexical sophistication and lexical variation. Therefore, learners should not repeat the same words and try to use the low-frequency words. Furthermore, they need to master some vocabulary learning strategies. For example, they should actively make word association to build syntagmatic and paradigmatic relations, which is a good way to build lexical networks.

C. Limitations of the Study

Obviously the present study is far from adequate due to the limitations of the author's ability and practical difficulties. There surely exist some problems in the present research.

Firstly, the sample is very small, only 66. If the study covered a larger sample, the results could be more convincing.

Secondly, the current study explores vocabulary usage from two variables, lexical sophistication and lexical variation. Lexical variation is the ratio of type and token, which is influenced by the length of writing papers. Although this calculating method has a good reliability and validity, it needs revising. Besides, the research doesn't take lexical errors into consideration. It also plays an important significance in writing rating and embodies learners' productive ability.

Thirdly, the present research only chooses one writing style of the composition. Different writing styles may influence learners' choice in vocabulary use, so learners' lexical richness can change. For example, learners can use many descriptive words in narrative papers, so lexical variation is high, while learners may be inclined to use explanatory words in argumentative papers, but this kind of papers may have less variation in vocabulary choice. Ni Lan (2000) makes a study on characteristics of vocabulary use frequency in English writing, with sophomores of English majors as the subjects. She finds that there exists a significant difference on vocabulary use frequency between descriptive essays and expository papers. Therefore, writing style is a variable in the research, which needs a certain consideration.

D. Suggestions for Further Study

The present research provides an initial empirical study for further research. Going with limitations, suggestions for further study are mentioned as follows.

Firstly, further investigations can cover a larger sample or involve English majors or postgraduates in the study, so that the results could be more convincing and comprehensive.

Secondly, further studies should take lexical errors into consideration and eliminate the influence of length of writing papers on lexical variation.

Thirdly, further studies should adopt several kinds of writing styles to find the characteristics of lexical richness, and the relationship between vocabulary knowledge and vocabulary usage.

Fourthly, vocabulary learning is a kind of psychological phenomenon, so further studies can use psycholinguistic theory to explain the data, and instruct vocabulary teaching and learning.

Finally, writing is a complex psychological activity, and too many factors may influence vocabulary usage in EFL writing, such as learners' language proficiency, grammatical knowledge and affective factors, so further studies can involve these factors and explore their relationship with vocabulary usage in writing.

The criteria includes 5 ranks: 14 points, 11 points, 8 points, 5 points and 2 points. One point can be added or subtracted considering the paper's quality.

14 points: The content of the article is coherent and complete and is closely correlated with the title. The article is well structured and grammatically correct. The language used is fluent and appropriate while the sentences are variable. There should be some "bright spots" and there should be no language errors.

11 points: The content of the article is coherent and complete and is correlated with the title. The article is well organized and grammatically correct. The language is appropriate with only a few grammatical mistakes.

8 points: The content of the article is relatively coherent and correlated with the title. Some parts of the article can't express thought clearly. There exist some grammatical mistakes, some of which are serious.

5 points: The content of the article is complete and relatively correlated with the title but not coherent. The article can't express thought clearly, and there are many grammatical mistakes.

2 points: The content of the article is neither complete nor correlated with the title. The structure is in a mess and there exist serious mistakes in most sentences.

REFERENCES

- [1] Bao, G. (2008). A Multi-dimensional Study on the Development of Lexical Richness in EFL Writing, *Computer-assisted Foreign Language Education*, 123: 38-43.
- [2] Chen, H. (2001). Vocabulary Learning Strategies of Chinese Non-English Majors. *Foreign Language Teaching*, 6:46-51.
- [3] Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*, 4(2), 139-155.
- [4] Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21, 303-317.
- [5] Laufer, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal*, 25, 21-33.
- [6] Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16 (3), 307-322.
- [7] Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Applied Linguistics*, 19 (2), 255-271.
- [8] Liu, D. H. (2004). The Influence of Writing Strategies and Productive Vocabulary Size on Writing Quality, *Modern Foreign Language*, 2:180-187.
- [9] Li, Z. X., & Li, J. Q. (2005). A study of Productive Vocabulary Use by Chinese EFL Learners with High Proficiency—Based on a Comparative Research on English Compositions of Chinese and American College Students. *Shandong Foreign Language Teaching*, 5:56-59.
- [10] Ma, G. H., & Shi, Y. Z. (2006). A Correlation Research on Lexical Richness and English Writing Quality, *Sino-US English Teaching*, 10: 56-60.
- [11] Nation, P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.
- [12] Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- [13] Ni, L. (2000). A Study on Vocabulary in Writing of Sophomores in English Major, *Foreign Language Teaching*, 2:38-41.
- [14] Qin, X. Q., & Wen, Q. F. (2007). *A study on English Writing Proficiency Development Rules and Characteristics of Non-English Majors*. Beijing, China Social Science Press.
- [15] Read, J. (1993). The development of a new measure of L2 vocabulary knowledge, *Language Testing*, 10, 355-371.
- [16] Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- [17] Wang, H. H., & Zhou, X. (2012). A Longitudinal Study on the variation of Lexical Richness in Non-English Major's Writing, *Foreign languages and their teaching*, 2: 40-44.
- [18] Wan, L. F. (2012). A study of Lexical Richness in EFL writing of English Majors. *Foreign Language World*, 1:40-46.
- [19] Wesche, M., & Paribakht, T. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *The Canadian Modern Language Review*, 53 (1), 13-40.
- [20] Yang, Y. Y. (2012). On English Major's Vocabulary Use in Writing, *Journal of University of Shanghai for Science and Technology*, 1: 46-51,84.
- [21] Zhu, H. M., & Wang, J. J. (2013). A Study on the Characteristics of Lexical Richness in English Writing- Based on a Self-built Linguistic Corpus, *Foreign Language World*, 6:77-86.

Lili Zhai was born in Linfen, Shanxi Province, China in 1984. She received her Master degree in linguistics and applied linguistics from Chongqing University, China in 2010.

She is currently a teacher in the school of Foreign Languages, Shanxi Normal University, Linfen, Shanxi Province, China. Her major research interests are issues of Second Language Acquisition and English Teaching.

Ms Zhai got a reward in English Teaching Contest held by Foreign Language Teaching and Research Press in 2014. She continues her teaching and researching and tries to make more progress at present.