



Empirical Study on Relations between Gender Differences and English Vocabulary and the Overall Proficiency Level

Hansheng Yan

School of Foreign Languages, Qingdao University of Science and Technology

Qingdao 266061, Shandong, China

E-mail: hanshengyan2006@yahoo.com.cn

Abstract

By comparing the cognitive CET-4 vocabulary, applied vocabulary and CET-4 scores of male group of English learners with that of female group of English learners in the university, the paper mainly discusses whether there exists gender difference in second language learning and whether female would keep their language advantages as time goes on. On the basis of analysis of the independent sample *t* in relevant data, the author of this article believes that there exist obvious differences in their CET-4 cognitive vocabulary, applied vocabulary and CET-4 scores, and that the female group always excels the male group in terms of both vocabulary and CET-4 scores. Therefore, we come to the conclusion that in second language learning in the university, there exist obvious gender differences among male and female and that female have language advantages. Finally, it is pointed out that these conclusions are just exploratory due to limitations of this study, and more research is required to make clear relations between gender and second language learning.

Keywords: Gender differences, Cognitive vocabulary, Applied vocabulary, Overall proficiency level

1. Theoretical background

Relevant studies in the last years indicate that differences exist between male and female in their learning of a language. Song Haiyan (1998) studied prototype of gender and its reflection in speech communication competence between the two genders. According to her, obvious differences exist between male and female in terms of language use, such as, females are inclined to passiveness, circumbendibus and not imposing their own will upon others, etc. Hou Songshan (1998) analyzed and studied influences of task and gender upon foreign communication strategies, and he pointed out that when male and female subjects finished a task, sometimes obvious discrepancy existed and sometimes didn't. From the above studies, it is not difficult to find out that, study on male and female in terms of language pays more attention to discrepancy exhibited in their use of a language, while rare study has been conducted on them in terms of their language learning, especially on whether significant difference exists in terms of second language learning, let alone study on that as a special subject. As a matter of fact, as for the problem whether difference exists between male and female in terms of learning a language, Larsen-Freeman (2000) believed that, in the process of first language acquisition, at least at the early stage, female excel male. Hu Zhuanglin (1989: 199) pointed out, it was generally believed that male and female are born with different linguistic advantages, such as, female learn to speak earlier than male, and female learn a foreign language faster and better than male, etc. Xu Liang (2001) pointed out more definitely, significance of gender difference in language acquisition at the early stage of language acquisition has been increasingly proved with development of study on linguistics in recent years and with breakthrough of neurolinguistics experiment. Then, he pointed out from the perspective of individual development that, effective integration of spatial skills and linguistic cognition provide favorable guarantee for advantages of female in their earlier language acquisition. All the above studies indicate theoretically that females are more adept at learning a language or a second language. Whether that inherent linguistic difference between male and female will continue to exist in the process of second language learning? In this article, the author is going to tentatively discuss and analyze these issues in terms of their CET-4 cognitive vocabulary, applied vocabulary and CET-4 scores in the university.

2. Experimental study

2.1 Respondents

Respondents are classified into two groups: control group and experimental group. The former group are constituted by 32 students randomly selected from Grade 98 of the Department of History, Qufu Normal University, which is mainly to verify reliability and validity of the two vocabulary questionnaires we have compiled. The experimental group is constituted by 318 students randomly selected from Grade 98 of the Department of Chinese Language and Literature, Department of Education, Department of Mathematics and Department of Chemistry, Qufu Normal University, among which are 144 male students and 174 female students.

2.2 Survey instrument

In survey of CET-4 vocabulary, the survey instrument we use is <<Questionnaire of CET-4 Vocabulary>> we have compiled. All respondents are non English majors in the university, so we choose as the target of the survey the vocabulary in the syllabus of College English and vocabulary in the Intensive Reading Course in <<College English>> (edited by Zhai Xiangjun et al, revived version of 1998) which is being studied by our respondents. First of all, we sort out words (including phrases) learned by students in the junior middle school and words (including phrases) in the word list at the end of the course book <<College English>>. Then, we classify them into five levels.

Vocabulary of Level A: altogether 1877 words learned during the junior middle school in the Appendix IV of <<College English>> (Intensive Reading Volume 1).

Vocabulary of Level B: altogether 900 words of CET-4 in Volume 1 of <<College English>>.

Vocabulary of Level C: altogether 797 words of CET-4 in Volume 2 of <<College English>>.

Vocabulary of Level D: altogether 767 words of CET-4 in Volume 3 of <<College English>>.

Vocabulary of Level E: altogether 624 words of CET-4 in Volume 4 of <<College English>>.

Altogether 6500 words are collected in << Glossary in College English Syllabus>>, among which approximately 1800 words are required to be fundamentally grasped by freshmen at their entrance into the university, and approximately 2400 words are required to be grasped from Grade 1 to Grade 4 (Huang Jianbin, 1999). There are altogether 4965 words we have sorted out, which can be said to have met the required number of 4200 words (1800+2400) from Grade 1 to Grade 4 in << Glossary in College English Syllabus >>. Therefore, the vocabulary questionnaire we have compiled is target-oriented with respect to investigation on CET-4 vocabulary and CET-4 score of students.

In sampling words, what we adopt is the method of equidistance of level classification. We draw at random 25 words in Level A, 24 in Level B, 23 in Level C, 24 in Level D and 23 in Level E. In such way, we altogether draw 119 words from Level A to Level E, and then we make a test paper (A). In order to guarantee and check the reliability and validity of test paper (A), we compile test paper (B) again. Words in test paper (B) are sampled as such: we draw at random 25 words in Level A, 24 in Level B, 23 in Level C, 24 in Level D and 24 in Level E, that is, altogether 120 words from Level A to Level E.

There are two methods to investigate vocabulary both at home and abroad. On one hand, the method of multiple-choice question is adopted, namely, to draw a sample of vocabulary, which is matched with a correct answer in the mother language and with three distracters. Gui Shichun (1985), The PLA University of Foreign Languages and Hazenberg & Hulstijn (1996) adopted this method in their investigation on vocabulary of language learners. On the other hand, respondents are required to write the most frequently used meaning in mother language for each word in the test paper or to explain one or two items for the word. Yu Aiju (1991), Gu & Johnson (1996) adopted the second method in their investigation on vocabulary of language learners. In our investigation on the CET-4 vocabulary of non English majors in our school, we adopt the latter, that is, to ask students to write the Chinese meaning of each word in the test paper, or to write down its synonym, or to explain the word so as to indicate that at least they understand one meaning of the word. Besides, we also require students to compose a phrase or make up a sentence with words given so as to indicate that they know how to apply the word.

2.3 Experimental process

The specific test is classified into steps. The first step is to evaluate reliability and validity of the two test papers. First of all, we use test paper (A) and (B) to test students in the control group within one morning. After the test, we immediately make a analysis in the result of these two test papers. We think that, if validity and difficulty of these two test papers are different, then scores of each students for each paper should be different. On the basis of examination on the independent sample T and through comparative analysis, we find out that, Value T of the cognitive vocabulary in these two test papers is .368, Value T of applied vocabulary is 1.699, and the Value P of both test papers is above 0.05, with no obvious difference, so it can be said that differences of these two test papers are not statistically significant. The Mean value, Root-Mean-Square Error (RMS error), Value T and Value P of cognitive vocabulary and applied

vocabulary in test paper A and test paper B are as follows.

Insert Table 1 Here

After getting the result that test paper A and test paper B had the same reliability and validity, we chose test paper B, and made the second test on students in the experimental group within the two weeks before the CET-4 (June 17, 2000). The reason for the choice of the time was that vocabulary of students at that time was almost finalized, and their vocabulary wouldn't change a lot in the following CET-4, so the correlation coefficient of vocabulary and CET-4 scores would be relatively reliable.

2.4 Correction of test paper and collection of data

After the test, we invited relevant teachers and students for correcting test papers. In order to ensure reliability of the scores of test papers, we unified the correction standard before correcting. As for cognitive vocabulary, as long as a student writes out correctly the meaning of one word, we assume that he has understood the word, whether the part of speech of the word is right or wrong. As for a polysemous word, it is considered to be correct so long as a student writes out one of its polysemous words. For example, for the word "awake", it is right that students translate it into the Chinese meaning "huan xiang, jiao xing" (Verb), and it's also OK that they translate it into "xing de, xing zhe de", (Adjective). For the word "sentence", either the translation of "ju zi" is right or the translation of "pan xing" is OK. As for applied vocabulary, it is generally believed that cognitive vocabulary of language learners is larger than their corresponding applied vocabulary. Therefore, when correcting applied vocabulary, we should take into consideration one principle, that is, applied vocabulary of students cannot be larger than their cognitive vocabulary. So long as a student does not know a word, it is assumed that he/she cannot use the word. For instance, a student does not know the word "loss", but when he uses the word, he writes out the phrase "at a loss", which we cannot count as being able to use the word. In addition, as long as a student uses a word correctly when taking an example or making up a sentence, we count him as being able to use the word, regardless of other mistakes (spelling or grammar mistakes) in the example or sentence. Such as, the sentence a student makes up for the word "crime" is: He was committed crime, and we count him as being able to use the word "crime".

In the process of correction, we first of all calculate the number of cognitive vocabulary of students in all levels of vocabulary, and then based on the ratio of each level of vocabulary, convert that number into corresponding number in each level of vocabulary. In the conversion of each level of vocabulary, we use the following formula:

$$\text{Vocabulary of all levels} = \frac{\text{Total number of vocabulary} \times \text{Actual correct number of vocabulary}}{\text{Number of samples from all vocabulary in compilation of a test paper}}$$

Assuming that the correct cognitive vocabulary of a student in test paper A is 23, and applied vocabulary is 20, then the corresponding number of cognitive vocabulary and applied vocabulary after conversion in Level A is respectively as follows:

$$\text{Cognitive vocabulary of Level A} = \frac{1877 \times 23}{25} = 1726.84$$

$$\text{Applied vocabulary of Level A} = \frac{1877 \times 20}{25} = 1501.60$$

The survey instrument we have applied also includes the CET-4 score unifiedly formulated by the Examination Board of the National College English Test-4 and Test-6. Design of CET-4 and CET-6 is quite rational and the grading process is fair and precise, so the score has extremely high reliability and validity.

3. Data analysis

Data analysis is conducted in two steps.

The first step is to make a comparative analysis in cognitive and applied vocabulary among male and female students. First of all, we categorize the vocabulary test paper at hand into two groups based on the factor of gender, namely, male group and female group. Then, we input their cognitive vocabulary and applied vocabulary into the computer, and on

the basis of examination on the independent sample T, we get out the mean value of cognitive vocabulary and applied vocabulary, RMS error, Value T and Value P among male and female students.

Insert Table 2 Here

From Table 2, we can see that the Value T of cognitive vocabulary among male and female students is 2.166, and Value P is 0.031, below the critical value of 0.05, which indicates that obvious differences exist between male and female English learners in terms of their cognitive vocabulary. At the same time, Table also tells us that, Value T of applied vocabulary among male and female students is 2.941, and Value P is 0.004, below the critical value of 0.05, which indicates that obvious differences exist between male and female English learners in terms of their applied vocabulary. Considering the mean value, the average cognitive vocabulary of male students is 3953.82, and that of female students is 4071.95, and the RMS error of female students is 76.9 lower than that of male students. Also, the average applied vocabulary of male students is 3078.73, and that of female students is 3378.78, and the RMS error of female students is 176.83 lower than that of male students. Obviously, whether in terms of cognitive vocabulary or in terms of applied vocabulary, female students excel male students. Therefore, it can be said that significant differences exist between male and female students in terms of their vocabulary, and that female students excel male students.

The second step is to make a comparative analysis in the CET-4 scores of male and female students. We input of CET-4 scores of the two groups into the computer, and on the basis of examination on the independent sample T, we get out the mean value, RMS error, Value T and Value P of CET-4 scores of male and female students.

Insert Table 3 Here

Table 3 also tells us, Value T of CET-4 scores among male and female students is 2.429, and Value P is 0.016, below the critical value of 0.05, which indicates that obvious differences exist between male and female English learners in terms of their CET-4 scores. Considering the mean value, the average CET-4 score of male students is 61.67, that of female students is 64.74, and that the RMS error of female students is 1.45 lower than that of male students, so the CET-4 score of female students is higher than that of male students. Thereout, it can be said that, obvious differences exist between male and female students in terms of their overall English proficiency level, and that female students excel male students.

4. Conclusion and discussion

The previous relevant analysis indicates that obvious differences exist between male and female students in terms of their CET-4 cognitive vocabulary, applied vocabulary and CET-4 scores, and that female students excel male students. That inherent linguistic difference among male and female proposed by Mr. Hu Zhuanglin has been obviously exhibited in the vocabulary of college English and in the CET-4 scores. This indicates that obvious differences exist between male and female in terms of second language learning, and that as time goes on, the differences still exist even in the university. Furthermore, it is still such a case that male maintain a linguistic advantage.

In theory, the present study is corresponding with comments by such experts as Hu Zhuanglin (1989) and Larsen-Freeman (2000). In practice, the present study is also consistent with conclusions by some other researchers, for example, Wu Yi'an, Liu Runqing, etc. (1996) investigated English majors, and believed that female students obviously excel male students in terms of language learning. However, at the same time, the present study somewhat differs from investigation results by other scholars. For instance, Wang Xuemei (2005) make an investigation on differences in their linguistic skills of listening, speaking, reading and writing among English undergraduate majors came to the conclusion that there is no gender difference between male and female students in terms of grammar vocabulary. This might be correlated with different respondents and survey instruments, etc. Therefore, some conclusions of the present study are just exploratory. We testify that obvious differences exist between male and female in terms of their vocabulary and their overall proficiency level, but that cannot mean obvious differences also exist in other aspects, such as listening, writing, and grammar. Difference exists between male and female of non English majors, which also does not mean that difference exists between male and female of English majors. Obvious differences exist between male and female in the university, which does not mean differences exist between male and female in the primary school and in the middle school. All these problems are open for further discussion to give a correct conclusion on relationship between gender and second language learning.

References

- Ding, Guodong & Wu, Yaxin. (2002). Language and Gender: A Linguistic Research Focusing on Both Differences and Similarities. *Foreign Language Education*, No. 2.
- Gu, Yongqi and Robert Keith Johnson. (1996). Vocabulary Learning Strategies and Language Learning Outcomes. *Language Learning*, Vol. 46/4:643-79.
- Gui, Shichun. (1985). Survey and Analysis on English Vocabulary in China. *Modern Foreign Languages*, No. 1.
- Hazenber, Suzanne & Janh. Hulstijn. (1996). Defining a Minimal Receptive Secong-Language Vocabulary for Non -

- native University Students: An Empirical Investigation. *Applied Linguistics*, Vol. 17, No. 2.
- Hou, Songshan. (1998). Influences of Task and Gender on Foreign Communication Strategies. *Journal of PLA University of Foreign Languages*, No. 6.
- Hu, Zhuanglin. (1989). *Linguistics: An Introduction*. Peking University Press.
- Huang, Jianbin. (1999). On Glossary in "College English Syllabus". *Foreign Language World*, No. 4.
- Larsen-Freeman, D. & Michael H. Long. (2000). *An Introduction to Second Language Acquisition Research*. Beijing: Foreign Language teaching and Research Press.
- Li Yue'e. (2001). A Review of Women, Men and Politeness. *Contemporary Linguistics*, No. 1.
- Song, Haiyan. (1998). Prototype of Gender and Its Reflection in Speech Communication Competence between the Two Genders. *Foreign Language*, No. 1.
- Wang, Gaiyan. (1999). Exploration in Differences in Linguistic Behaviors among Men and Women. *Foreign Language Education*, No. 1.
- Wang, Xuemei. (2005). Study on Gender Differences in Linguistic Skills among EFL Learners. *English Teaching and Research in Normal Colleges and Universities*, No. 1.
- Wu, Jihong. (2001). On Theoretical Perspective of Community of Practice in Language and Gender Research. *Foreign Language Education*, No. 2.
- Wu, Yi'an, Liu, Runqing, et al. (1996). *Learner Factors and Learning Achievement --- A Study of the Effect of Factors Affecting English Language Learning*. Beijing: Foreign Language teaching and Research Press.
- Xu, Liang. (2001). The Neurolinguistic Significance of Gender Differences to the Establishment of Optimized SLA Strategies. *Foreign Languages and Their Teaching*, No. 6.
- Yu, Aiju. (1991). Survey and Analysis in English Vocabulary among Trainees. *Foreign Language Teaching and Research*, No. 1.
- Zhai, Xiangjun et al. (1998). *College English* (Volume 1 to Volume 4, revised version). Shanghai Foreign Language Education Press.

Notes

Note 1. Words in the present study include phrases, not including proper nouns, names of persons and places.

Note 2. All data in the present study are calculated based on the statistical software of SPSS (9.0).

Table 1. Table of comparison between Test paper A and B

Test paper	NO of students	Cognitive vocabulary				Applied vocabulary			
		Mean value	RMS error	Value T	Value P	Mean value	RMS error	Value T	Value P
A	32	4174.90	615.50	.368	.721	3958.71	763.47	1.699	.124
B	32	4149.67	456.02			3690.08	604.10		

Table 2. Mean value, RMS error, Value T and Value P among male and female students

Gender	NO. of students	Cognitive vocabulary				Applied vocabulary			
		Mean value	RMS error	Value T	Value P	Mean value	RMS error	Value T	Value P
Male	144	3953.82	524.74	2.166	0.031	3078.73	997.27	2.941	0.004
Female	170	4071.95	447.84			3378.78	820.44		

Table 3. Mean value, RMS error, Value T and Value P of CET-4 scores among male and female students

Gender	NO. of students	Score of CET-4			
		Mean value	RMS error	Value T	Value P
Male	144	61.67	12.04	2.429	0.016
Female	170	64.74	10.59		