Comparative Analysis on Lexical Richness in Chinese-English Translation with Help of E-Dictionary Among Medium and Low Level Students

Wanji Xie

Xi'an Jiaotong University City College

Xi'an, China

Abstract—With the widespread of Mobile Phone E-Dictionary in English class, its impact should not be overlooked. The research is set out to find out the influence of E-Dictionary on Chinese-English translation based on a comparative analysis of lexical richness in texts before and after the use of E-Dictionary. The research finds: on one hand, using E-Dictionary will lengthen the texts of medium and low level students, reflecting a certain increase in their translating willingness and confidence; on the other hand, using E-Dictionary will help increase the lexical variation and their lexical frequency profile, indicating that students intentionally use relatively higher-end words with the help of E-Dictionary.

Keywords—E-Dictionary; lexical richness; Chinese-English translation; CET4

I. INTRODUCTION

Dictionary has always been one of the most common tools for language learners. With the advance in technology, portable electronic dictionaries are becoming more widespread. According to Zhang, 91% of the students often use electronic dictionaries, while 82% indicate that they rely on dictionary APP to look up words. Such a high rate demonstrates that students rely heavily on the dictionary APP.

The research subjects are from third tier college, mostly having a low intermediate English level and heavy dependence on the dictionary APP. When assigned in-class exercises, almost all students will use their phones.

The research on e-dictionary is mainly divided into 2 categories. One is to develop and improve the e-dictionary and its derivative functions through text analysis or technology advance. The other is interdisciplinary research between dictionaries and language learning, focusing on status survey or vocabulary acquisition during readings with the help of a dictionary.

On the one hand, the interdisciplinary research topic is often quite narrow and partial; on the other hand, the articles are not often supported by data. This article focuses on in-class CET-4 translation exercise, and analyzes the use of e-dictionary in the actual teaching by using text analysis and statistics.

II. LITERATURE REVIEW

A. Theory

Text analysis in the language teaching field often refers to a descriptive analysis of learners' vocabulary use. Lexical knowledge is often divided into range and depth. According to Read, It is not very common for researchers to elicit learner's deep knowledge about a particular word in large scale text analysis; most often, researchers will emphasize on the lexical range (richness) of the text.

Laufer & Nation defines that lexical richness includes four aspects: complexity, density, variety, and novelty. Read has made some improvements: complexity is measured by words beyond the first 2000 frequent word list, variety is measured by Type-Token Ratio and density is measured by the percentage of notional words in a text. While Laufer & Nation also claim that Vocabulary Level can also be used to assess learners' lexical richness. This article will be based on the above theories to study the lexical richness in students' translation.

B. Literature Review

1) E-dictionary and language teaching:

Yao (2003) introduced the function and prospect of e-dictionary, and predict that it will play an important role in the language classroom. This article has been published in the early 2000s and didn't have any data support.

Liu (2013) has highly recommended the use of e-dictionary in language skill teaching and showed how to use it to improve learning efficiency. Chen (2010) conducted a survey and interviews to discuss the use of e-dictionary in high school extensive reading class. She discovered a positive effect on raising students' reading efficiency and motivation. Lu (2016) carried out a survey of e-dictionary use among students in the Japanese major and mentioned "helpful to translation" as part of the conclusion. Since most of these articles are qualitative research without data, it is not easy to determine in which aspect or how much the e-dictionary helped students.

2) Research on lexical richness:

Research on lexical richness started over two decades ago in China. However, almost all related research is focused on writing texts. Some research focuses on the parallel comparison of texts differed by learners, genre, tasks, etc. Some research emphasizes on vertical comparison of similar texts produced by the same learners at different times. These researches paid attention to learners' acquisition of words and their lexical development. With slight nuisance in their conclusion, mostly they all state that lexical richness is important in assessing learners' texts.

Compared to a large quantity of writing text research, few have paid attention to the lexical richness in translation texts. One of the reasons may be confined to the nature of translation and contents, not much variance can be demonstrated across texts. This may hold true to high-level learners, but for learners with a low intermediate level, using e-dictionary can help them get rid of simple words and try using higher level words, and also reflect their choice of words which can also indicate their mindset in learning vocabulary.

C. Research Question

Based on the above literature review, this article mainly studies the short term vertical development on translations texts using e-dictionary. Subjects are non-English major students doing CET-4 translation exercises in class. By comparing their two translation texts, the article seeks to answer the following questions:

- What is the difference of lexical richness between two texts in the aspect of text length, lexical variety and level?

- Is the change in lexical richness related to students' level? If so, what relationship is it?

III. RESEARCH METHOD

A. Research Subject

Research subjects come from non-English major sophomore students, with low intermediate or even elementary levels and very few intermediate levels.

B. Research Procedure

Firstly, students finish a CET-4 translation exercise about Jade in class within 20 minutes. Students can neither use e-dictionary nor ask their classmates. The first translation is named Text 1.

Secondly, after finishing Text 1, students are asked to do the same exercise immediately and are expected to finish within 15 minutes. The students can use e-dictionary this time but they cannot discuss with classmates. The second translation is named Text 2.

Thirdly, both texts are collected.

Finally, students are divided into high, middle and low-level groups based on their score of the 17.06 CET-4 tests. High-level groups are made up of 8 students scoring higher than 480, the middle-level group contains 24 students scoring between 420 and 450, and the low-level group consists of 43 students scoring between 360 and 400. The total number of texts for analysis is (8+24+43)*2=150.

C. Data Collection and Tools

- Microsoft Word 2010. Word numbers are counted using this software after the texts have been typed into a computer. All mistakes or errors are kept original.

- Complete Lexical Tutor v8.3. This web-based tool is developed by Canadian researcher Tom Cobb based on many lexical research, including various lexical tools such as Range 3.2 for lexical frequency measurement and Vocab Profile for mistake ratio, lexical variety, and vocabulary level analysis.

- IBM SPSS Statistics 19.0. This software is used for statistical analysis, such as the Wilcoxon test and paired t-test.

IV. RESULT AND DISCUSSION

A. Text Length Between Different Groups

After all the texts are typed into Word 2010, English word counts are as followed.

TABLE I.   TEXT LENGTH COMPARISON

	Overall	AVG	High Overall	High AVG	Mid Overall	Mid AVG	Low Overall	Low AVG
Text 1	5639	75.1	721	90.1	1885	78.5	3033	70.5
Text 2	6373	84.9	676	84.5	2060	85.8	3637	84.6

a. * misspelled words are included.

TABLE II. GROUP DIFFERENCE IN VARIANCE ANALYSIS OF TEXT 1

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
High	Mid	11.66667	5.67941	.145	-3.6310	26.9643
	Low	19.38081*	5.86801	.015	3.8550	34.9066
Mid	High	-11.66667	5.67941	.145	-26.9643	3.6310
	Low	7.71415	4.03743	.144	-1.9760	17.4043
Low	High	-19.38081*	5.86801	.015	-34.9066	-3.8550
	Mid	-7.71415	4.03743	.144	-17.4043	1.9760

*. The mean difference is significant at the 0.05 level.

It can be seen from "Table I" that the number of words in the low and middle-level group has risen while the number dropped in the high-level group. After variance analysis (see "Table II"), there is no statistical difference among the high and middle group and middle-low groups, but a distinct difference between the high and low-level group. Confined by the length of the original text, text 2 of all groups reaches a similar number of words without any significant variation.

TABLE III. WILCOXON TEST OF TEXT LENGTH OF THE HIGH-LEVEL GROUP

Test Statistics[b]	
	VAR00004 — VAR00003
Z	-1.192a
Asymp. Sig. (2-tailed)	.233
a. Based on positive ranks.	
b. Wilcoxon Signed Ranks Test	

TABLE IV. PAIRED SAMPLE T-TEST BETWEEN TEXTS AMONG THE MIDDLE AND LOW-LEVEL GROUPS

		Paired Differences		t	df	Sig. (2-tailed)
		95% Confidence Interval of the Difference				
		Lower	Upper			
Pair 1	midtext1 - midtext2	-13.03452	-1.71548	-2.696	23	.013
Pair 2	lowtext1 - lowtext2	-19.32139	-8.30652	-5.062	42	.000

A further analysis results in "Table III" and "Table IV". The number of subjects in the high-level group is too small and do not follow a normal distribution. After the Wilcoxon test, the data do not reach a statistical significance. Therefore, it can't be concluded that e-dictionary had any effect on this group. However, paired t-test of the text length of the middle and low-level group showed a statistically significant difference, which suggests e-dictionary does affect students' translation text length. One of the reasons may be students have more confidence and tools to complete a longer text with help of e-dictionary.

TABLE V. MISTAKE RATIO IN TEXT 1 AND TEXT 2

High Level		Mid Level		Low Level	
Text 1	Text 2	Text 1	Text 2	Text 1	Text 2
0.017%	0.013%	0.045%	0.012%	0.056%	0.013%

Nevertheless, the increase in word counts doesn't necessarily mean an increase in correctness. Vocab Profile is used to calculate the misspelled words in both texts, as shown in "Table V". General mistake ratio of text 1 increases as students' English level lowers. After introducing the e-dictionary, this ratio dropped in all three groups and came to a very similar level. One explanation may be under the constriction of time, students are bound to make a certain percentage of mistakes; the other reason may be some students are very confident in their spelling and do not check their writings, so that the misspellings remained wrong in the second text.

B. TTR Comparison Between Texts Among Different Groups

Type-Token Ratio, one of the measurements of lexical richness, is often used to measure lexical density within a given text. Restricted by the original text of the paragraph, the words that can be used are largely confined without much variance. Therefore, this research does not focus on the TTR on text 1 or text 2, but on the difference of TTR between text 1 and 2 among within groups and among groups. All misspelled words are excluded.

TABLE VI. TTR OF DIFFERENT GROUPS

High Level		Mid Level		Low Level	
Text 1	Text 2	Text 1	Text 2	Text 1	Text 2
0.627	0.685	0.642	0.686	0.654	0.686

It can be seen from "Table VI" that TTR of text 1 of all groups are lower than that of text 2, which means students tend to use repetitive words when translating on their own. For example, one of the texts had a total number of 105, but the TTR is only 0.549, indicating that almost half of the words in this text are repeated. This student doesn't know the word ancestor, so he chose to use ancient people, making them unnecessarily reused.

Although most students' TTR rises during the two translations, there are students whose TTR dropped significantly. This phenomenon, centered on students whose total number of words in text 1 is less than 60, is resulted

from the change in length. One student wrote 58 words in text 1 with a TTR 0.709 (the actual type is 41.2). While the same students wrote 90 words in text 2 with a lower TTR 0.656, the actual type rose to 59.4. According to Nation & Warring, the first 1000 words in any text will account for 70% of the word numbers. As long as students write a text of reasonable length, the ratio of the first 1000 words is hard to change. Only when students try to use more words beyond the first 1000 words to write a longer text, will the TTR drop.

TABLE VII.   WILCOXON TEST OF TTR DIFFERENCE IN THE HIGH-LEVEL GROUP

Test Statistics[b]	
	hightext2 — hightext1
Z	-1.820[a]
Asymp. Sig. (2-tailed)	.069
a. Based on negative ranks.	
b. Wilcoxon Signed Ranks Test	

TABLE VIII.   PAIRED T-TEST OF TTR DIFFERENCE IN THE MIDDLE AND LOW-LEVEL GROUPS

Paired Samples Test					
	Paired Differences				
	95% Confidence Interval of the Difference				
	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1 midtext1 - midtext2	-.06878	-.01922	-3.673	23	.001
Pair 2 lowtext1 - lowtext2	-.06194	-.00150	-2.118	42	.040

According to "Table VII", the high-level group is not statistically significant on two-tailed Wilcoxon test but is significant in the one-tailed test. Therefore, it can't be definitely concluded. From "Table VIII", the paired t-test showed a significant difference in the middle and low-level

groups. Using e-dictionary can help students improve their TTR and increase lexical density.

C. Word Level Comparison in Two Texts Among Groups

TABLE IX.   WORD LEVEL RATIO OF TWO TEXTS AMONG GROUPS

		High Level		Mid Level		Low Level	
		Text 1	Text 2	Text 1	Text 2	Text 1	Text 2
K-1	**Type**	71.49%	62.74%	72.53%	63.04%	73.35%	62.04%
	Token	64.95%	50.86%	51.93%	51.78%	52.71%	50.16%
	Family	99	68	140	96	187	115
K-2	**Type**	9.60%	11.93%	7.42%	11.57%	6.21%	11.75%
	Token	14.43%	17.14%	13.81%	13.44%	11.46%	11.58%
	Family	20	21	38	23	46	23
K-3	**Type**	8.21%	12.37%	6.02%	12.63%	4.87%	12.54%
	Token	8.25%	15.43%	6.63%	13.44%	7.08%	12.54%
	Family	13	18	20	25	29	26
K-4	**Type**	0.70%	2.50%	0.59%	1.98%	0.47%	1.94%
	Token	1.55%	5.14%	2.76%	5.53%	1.88%	5.14%
	Family	3	5	8	10	8	9
K-5	**Type**	0.42%	0.88%	0.38%	0.87%	0.40%	1.20%
	Token	1.55%	1.71%	1.38%	1.98%	1.25%	2.25%
	Family	3	3	5	4	5	6
K-6	**Type**	0.14%	0.44%	0.05%	0.87%	0.07%	0.79%
	Token	0.52%	1.71%	0.28%	2.77%	0.42%	2.57%
	Family	1	2	1	6	2	4

Web Vocab Profile uses the 1-25K word level of BNC-COCA to grade the lexical level of a given text. Excluding proper noun (place, name, numbers, etc), other English words are divided into 25 levels, each containing 1000 words. K-1 includes the most frequent words; k-2 includes the next frequent words and so on. Nation & Warring (2010) believes the first 6K words should cover 90 percent of all contents based on the Brown Corpus; therefore, this study focuses on the K1-K6 level. The given topic word jade belongs to K-7

and is excluded, so are proper names such as China and Chinese and all misspelled words. The results are shown in "Table VII".

It can be seen from the "Table IX" that although all groups increased their K-1 word numbers (both type and token), the percentage drops. The K-2 to K-6 level words are increasing both in numbers and percentages. It represents that students had consciously started to use more high-level

words and avoid the repetitive words in K-1 with the help of e-dictionary.

Occasionally, there are words beyond K-7 level in middle and low-level groups, but mostly is a coincidence of misspelled words that is the same with some high-level words. For instance, a student wrote Chinese as chines and was classified as K-16 level word. At the same time, some students wrote words related to the Chinese meaning in the original text but rarely used. The word fete belongs to K-10 level, and it means festival or religious celebration or banquets. However, the students didn't confirm the Chinese meaning and used it to translate sacrifice. It demonstrates that middle and low-level students' entire dependence the e-dictionary without any confirmation of the real meaning or usage of the word they checked.

V. CONCLUSION

Based on the text analysis of students' two same translation exercises with and without the help of e-dictionary, it is discovered that:

- With the help of e-dictionary, the translation text length increases, especially a significant increase among middle and low-level groups; the increase in length does not necessarily mean an improvement in correctness. From the aspect of spelling, the problem cannot be completely eradicated.

- The type-token ratio also increases after using the e-dictionary, showing that students consciously substitute repetitive words with more variety.

- Word level of the translation texts diversifies with the help of e-dictionary. Under the precondition that K-1 word will account for most of the text, the use of K-2 and higher level words increases to make translation more precise rather than a repetition of general words. Also, among the middle and low-level students are situations where they will use words they found without a second check of the meaning.

Above conclusions statistically proved that using e-dictionary will provide help for students of low and mid English level when doing translation exercises. It can be further inferred from the data that students will have a higher self-confidence with the help of e-dictionary to complete the task, and teachers should encourage students to use e-dictionary when doing similar tasks. Only when the confidence is enhanced, can their acquisition of news contents and methods be obtained.

There are some improvements that can be further made into subsequent researches. Analyzing the lexical richness of texts does not reveal the quality of the whole translation, for example, the correctness of words does not equal to correctness in lexical usage. In order to further research into the development of translational texts, qualitative study into the correctness of words and sentences should be incorporated into further research.

REFERENCES

[1] Laufer, B. and Nation, I.S.P. Lexical richness in L2 written production: Can it be measured? [J]. Applied Linguistics, 1995(3).

[2] Read, J. Assessing Vocabulary [M]. Beijing: Foreign Language Teaching and Research Press, 2010.

[3] Nation, P., & Waring, R. Vocabulary size, text coverage, and word lists. In Schmitt, N. & McCarthy, M. (Eds.) Vocabulary Description, acquisition, pedagogy [M]. New York: Cambridge University Press.

[4] Bao, G. A Survey of the Development of Lexical Richness in L2 Compositions from a Multidimensional Perspective [J]. Computer-Assisted Foreign Language Education in China, 2008(5). (in Chinese)

[5] Bao, G A comparative study of EFL learners' linguistic complexity [J]. Modern Foreign Languages, 2010(2). (in Chinese)

[6] Bao, G The Effects of School Type and Writing Proficiency on EFL Learners' Lexical Complexity Across Course [J]. Journal of PLA University of Foreign Languages, 2011(4). (in Chinese)

[7] Chen F, Wang Q, Chen ZH. An Empirical Study of Electronic Dictionary Application in High School English Teaching [J]. China Educational Technology, 2010(10) (in Chinese)

[8] Dai Y, Xu H. Electronic Lexicography: Status Quo and Prospects [J]. Lexicographical Studies, 2014(4). (in Chinese)

[9] Li JH, Qin J. Effects of the Electronic-Dictionary Use Caused by Reading Tasks on Incidental Vocabulary Acquisition [J]. Theory and Practice of Education, 2014(21). (in Chinese)

[10] Liao HH. Electronic Dictionary and College English Reading Teaching [J]. Lexicographical Studies, 2004(3). (in Chinese)

[11] Liu CQ, Yuan KL. On the Use of Electronic Dictionaries in English Learning [J]. Lexicographical Studies, 2013(2). (in Chinese)

[12] Liu CY. Language output and second language development through second language curriculum — A case study of second language writing in the curriculum of a Hong Kong university [J]. Foreign Language World, 2014(3). (in Chinese)

[13] Liu XL, Lin LH. A study on the Chinese lexical attrition of advanced English learners in China [J]. Foreign Language Teaching and Research, 2013(5). (in Chinese)

[14] Lu MJ, Li R. Analysis of Electronic Dictionary Use for Japanese Language Teaching in Colleges [J]. China Educational Technology and Equipment, 2016(10). (in Chinese)

[15] Qin L. Lexical Development of Non-English Major Students — A Study Based on Written English Corpus [J]. Foreign Language Research, 2013(2). (in Chinese)

[16] Shi XL, Wu SJ. A Study of Improving HVC Students' Productive Skills Through Online Asynchronous Interaction [J]. Journal of Xi'an International Studies University, 2015(3). (in Chinese)

[17] Wan LF. An empirical investigation into lexical diversity of Chinese English majors' TEM writings [J]. Foreign Language World, 2010(1). (in Chinese)

[18] Wang HH, Zhou X. A Longitudinal Study on the Features of Lexical Richness in Writing by University Non-English Majors [J]. Foreign Languages and Their Teaching, 2012(2). (in Chinese)

[19] Wang HH, Li BB, Xu L. A dynamic case study on the development of Chinese EFL learners' written language proficiency [J]. Foreign Language Teaching and Research, 2015(1). (in Chinese)

[20] Wang YD, Zhang XM. Retrieval Functions of Electronic Dictionaries and Their Future Development [J]. Lexicographical Studies, 2016(3). (in Chinese)

[21] Yao XM, Liu SL. E-Dictionaries and English Teaching [J]. Media in Foreign Language Instruction, 2003(2). (in Chinese)

[22] Zhang H, Zhang J, Jia XY. Investigation and Analysis on the Current Situation of Electronic Dictionaries' Use among College Students [J]. Journal of Heilongjiang Vocational Institute of Ecological Engineering, 2017(3). (in Chinese)

[23] Zhu HM, Wang JJ. Developmental features of lexical richness in English writing:A self-built corpus-based longitudinal study [J]. Foreign Language World, 2013(6). (in Chinese)