An Empirical Study on the Role of Vocabulary Knowledge in EFL Listening Comprehension

Suping Wang

Binzhou Medical College, Yantai, China

Abstract—**Existing literature supports the positive effects of vocabulary teaching on students' comprehensive English proficiency. However, there have been some imperfections in the research process such as the research framework being somewhat narrow and the statistical instruments being a little conservative, which makes it difficult to form differentiated instructional strategies aimed at different levels of students. Therefore, this paper, based on mature vocabulary testing instruments, incorporates vocabulary breadth, depth, listening comprehension into one study framework and conducts an empirical study by innovatively using quantile regression models, so as to analyse the relationship between the two vocabulary dimentions and listening conprehension at different levels of students. The study results show that both vocabulary breadth and depth influenced listening scores significantly, but the overall effect of vocabulary depth was significantly greater than that of vocabulary breadth; at three levels of listening scores, the enhancing effects of vocabulary breadth showed a gradual increasing trend, while the influence of vocabulary depth shows no significant increase.This study will offer important pedagogical implications to English vocabulary teaching strategies in EFL classes.**

Index Terms—**vocabulary breadth, vocabulary depth, listening comprehension**

I. INTRODUCTION

Virtually every scholar of language recognizes that vocabulary knowledge is the basis and precondition for studies in second language acquisition. Many studies have shown that vocabulary knowledge is an important factor for successful reading comprehension in EFL (e.g. Hu & Nation 2000; Mecartty 2000; Qian 2002).

However, few studies have explored the relationship between lexical knowledge and listening comprehension, especially the listening comprehension in EFL. Thus, teachers and students have been long puzzled about the exact role that vocabulary knowledge plays in L2 listening comprehension.

Many researchers have pointed out that vocabulary usually involves two dimensions of breadth and depth (e.g. Qian, 1999; Read, 1993). However, among the few studies about vocabulary and listening, more attention has been put on the effect of vocabulary breadth. Recently, more and more researchers tend to acknowledge the essential role of vocabulary depth in listening. Nevertheless, not a few empirical studies have been conducted on the exact roles of vocabulary breadth and depth in EFL listening comprehension

The emphasis in the study to be reported is on both dimensions of vocabulary knowledge. The analysis emphasis will be on the role of vocabulary knowledge related to listening comprehension.

A. Purpose of the Study

The purpose of the study is to provide a review of the research literature related to the role of vocabulary knowledge in listening comprehension and then to conduct an experimental study to test the respective effect of vocabulary breadth and depth on different parts of listening comprehension so as to provide an empirical support for vocabulary teaching in EFL listening class.

B. Significance of the Study

The study on the role of vocabulary knowledge in listening comprehension for EFL learners is important to the practice of vocabulary teaching in EFL listening classes. This empirical research will provide teachers and students with a relatively complete understanding about the correlation between vocabulary and listening, and thus enhance their awareness about the importance of vocabulary in listening and help improve their teaching and learning methods in terms of vocabulary. The study will also be of importance to others seeking information related to this topic.

C. Definition of Terms

EFL: English as a Foreign Language

Vocabulary Breadth: the size of a learner's vocabulary, that is, the number of words for which the learner has at least some knowledge of meaning (Staehr, 2009)

Vocabulary Depth: relates to how well one knows a word and is defined as the quality of lexical knowledge that reflects how well a learner knows individual words and how well words are organized in the learner's mental lexicon. (Staehr, 2009)

CET 4: College English Test held nationally twice a year for non- English majors in China

D.   Method of Approach

To conduct a brief literature review about the relationship between vocabulary knowledge and listening comprehension, and then to summarize and synthesize the recent relevant studies.

To design an empirical experiment consisting of three pen-and-paper tests: vocabulary levels test (Nation, 1990), productive levels test (Laufer & Nation, 1999) and CET-4 listening comprehension test. The quantitative research methods of Pearson Correlations analysis and a Multiple Regression Analysis programmed in SASS 9.3 will be employed as the primary analysis tools for the data collected.

II.   LITERATURE REVIEW

It has long been recognized that there are many dimensions of vocabulary knowledge. Among the varying dimensions, breadth and depth are two primary dimensions of vocabulary knowledge.(Nation, 2001.) So the few studies about the relationship between lexis and listening mainly focus on the correlation of the two lexical dimensions and listening comprehension.

It was claimed that vocabulary knowledge is the main obstacle to successful listening comprehension for EFL learners (Kelly, 1991), but this claim has not been fully supported by empirical results. Bonk (2000) investigated the relationship between knowledge of lexis in listening texts and gist comprehension of the texts. Fifty-nine Japanese university students of low-intermediate to advanced English ability were tested using first-language recall protocols as comprehension measures, and dictation as measures of lexical familiarity. The research was based on four texts of increasing amounts of low-frequency lexical words. The two variables were found to be only moderately associated (45%). Good comprehension frequently occurred with text-lexis familiarity levels at 90+ percent levels. But the study results also show that some learners obtained good comprehension although they knew less than 75% of the word types in the text and other learners knew more than 90% of the word types but did not obtain good comprehension. Bonk relates this finding to learners' ability with comprehension strategies, but "part of the explanation may be attributed to the measures used to assess word knowledge and listening comprehension" (Staehr, 2008, p.140). So this study by Bonk offered an equivocal result about the relationship between lexis and listening comprehension.

A slightly clearer picture is found by Milton, Wade and Hopkins (forthcoming), who investigated the relationship between two receptive vocabulary size measures (the X_Lex and the A_Lex) and the IELTS test with 29 EFL learners. The study produced modest, significant Spearman correlations between vocabulary size (the X_Lex) and the reading and listening components of IELTS (0.54 and 0.52, respectively).

"An important issue in studies of how much vocabulary is needed to read a text or listen to a movie is what amount of text coverage is needed for adequate comprehension to be likely to occur"(Nation, 2006, p.61). Hu and Nation (2000) examined the relationship between text coverage and reading comprehension for non-native speakers of English with a fiction text, finding that a 98% text coverage would be needed for most learners to gain adequate comprehension. Then Nation (2006) reported on the trialling of fourteen 1,000 word-family lists made from the British National Corpus, and used these lists to see what vocabulary size is needed for unassisted comprehension. In this study Nation transferred this 98% lexical coverage found in reading to listenng comprehenshon, suggesting that a vocabulary of as much as 6000–7000 word families is needed to gain adequate comprehension in spoken discourse. These figures are to some extent supported by an analysis of the spoken CANCODE corpus conducted by Adolphs and Schmitt (2003).

A study by Staehr (2008) about vocabulary size and the skills of listening, reading and writing similarly found a substantial correlation between vocabulary size and a listening comprehension test. Vocabulary size could pridict 39% of the variance in the listening scores and 6000-7000 word families have been suggesed for adequate listening. Staehr's another study (2009) got the result that a lexical coverage of 98% is needed for coping with the spoken texts that constitute the listening test.

A recent study by Zeeland and Schmitt (2012) directly investigated the lexical coverage in regard to listening comprehension, to determine whether it is reasonable to transfer the 98% text coverage got in studies about reading to listening. Results showed that most native and non-native participants could adequately comprehend the spoken texts with only 90% coverage, but the non-natives showed considerable variation at this level. At 95% coverage, non-native participants also demonstrated relatively good comprehension, but with much less variation. Based on a 95% coverage figure, language users would need to know between 2,000 and 3,000 word families for adequate listening comprehension. The figures show a little difference from Nation's (2006) calculation of 6,000–7,000 word families based on a 98% coverage.

Although a small number of studies have provided empirical research for the relationship between vocabulary knowledge and listening comprehension in L2, few of them have focused on the depth of vocabulary knowledge.

Staehr (2009) presents an empirical study to investigate the role of vocabulary knowledge in listening comprehension with 115 advanced Danish learners of English as a foreign language. The dimensions of depth and breadth of vocabulary knowledge were found to be significantly correlated with listening comprehension. Vocabulary breadth could predict 49% of the varience in the listening scores. However, after the vocabulary depth was included in the regression model, only 51% varience in listening comprehension could be achieved. Another study by Frances H. Mecartty (2000) of Denver University examined the relationship of lexical and grammatical knowledge with reading and listening comprehension by foreign language learners of Spanish, finding that lexical and grammatical knowledge

were significantly correlated to listening, but only lexical knowledge explained the variance in listening (13%). The inconsistency in research results mainly comes from the different levels of participants and the measurement instruments. In Staehr's research, the participants were 115 advanced Danish learners of English as a foreign language, 40% of whom once stayed in English-speaking countries for at least 5 months. The vocabulary breadth and depth knowledge were respectively measured by the Vocabulary Levels Test by Schmitt (2001) and the Word Associates Test by Read (1993). While in Mecartty's (2000) study, the 154 participants were "from a large Midwestern university in their final semester of a four-semester basic Spanish language sequence" (p. 328) and the test score was based on word-association tasks and word-antonym tasks. Taking sophomore English majors as the subjects, Zhang (2011) examined the relationship between lexical knowledge and listening comprehension of TEM4 (Test for English Majors) and found that lexical breadth accounted for 27% variance of listening comprehension, 24% variance of dictation and different variance of other parts, while the depth explained 2% variance of listening comprehension and dictation respectively. This study result about vocabulary depth is consistent with that in Staehr's (2009) research.

From above studies, we can see that, despite the minor differences in specific figures, researchers are close to agreement on the role of vocabulary breadth in listening comprehension. However, the study results about vocabulary depth have been few and inconsistent. What's more, there have been some imperfections in the research process, such as the research framework being somewhat narrow and the statistical instruments being a little conservative, which makes it difficult to form differentiated instructional strategies aimed at different levels of students. Therefore, this paper, based on mature vocabulary testing instruments, incorporates vocabulary breadth, depth, listening comprehension into one study framework and conducts an empirical study by innovatively using quantile regression models, so as to analyse the relationship between the two vocabulary dimentions and listening conprehension at different levels of students. This study will offer important pedagogical implications to English vocabulary teaching strategies in EFL classes.

This paper will try to answer two questions:

1). How are vocabulary breadth and depth correlated with listening comprehension in general.

2).To what extent can vocabulary breadth and depth predict the varience of' listening comprehension at different English levels.

III. TEACHING IMPROVEMENT PROJECT

A. Participants

120 non-English majors (80 females and 40 males respectively) participating in this study were chosen randomly from 3 natural classes in a medical university, with some extreme cases of especially being poor in English ticked out according to their academic performances in English exams. All the participants were sophomores and from different majors ranging from clinical medicine, nursing, psychology and laboratory medicine.They used the same English textbooks and attended the same English classes twice a week taught by the same teacher and they shared similar educational background. Moreover, all of them have learnt English for at least 8 years and have prepared for the coming CET 4 in December for several months. All of them were quite serious about their performances in the CET4, which would be linked to their bachelor's degree.

B. Instruments

1. Test for breadth of vocabulary knowledge. Tests measuring vocabulary breadth in this study are based on Nation's Vocabulary Level Test (the 1990), which has been widely used in vocabulary research context and can provide a valid estimate of leamers'vocabulary knowledge at the different frequency levels (Staehr, 2009). This kind of test is divided into five levels: 2000-word level, 3000-word level, 5000-word level, the university word level and 1 0000-word level. The 2000-and 3000-word level contain high-frequency words; the university word level represents one type of specialized vocabulary; and the 5000-word level is on the boundary of high-and low-frequency words. Finally, the 10000-word level contains low-frequency words (Nation, 1990, P. 261) and few EFL learners can achieve this level. In accordance with participants' reading vocabulary level and the requirement of Chinese New College English Syllabus (To master the vocabulary of about 4795 words and 700 phrases, of which about 2000 words are active vocabulary), the first four word frequency levels in Vocabulary Levels Test were chosen to test the participants. At each level, there are 10 groups comprising six words and three definitions.The participants are required to match the three definitions on the right with three of the six words on the left.

For example: (From 2000-word level)

1. copy
2. event   (6) end or highest point
3. motor   (3) This moves a car
4. pity   (1) Thing made to be like another
5. profit
6. tip

The total score for this test was 120 points, with each blank one point.

2. Test for depth of vocabulary knowledge. Productive Levels Test (Version A) by Laufer and Nation (2001) is used to test students' vocabulary depth. Like the Vocabulary Levels Test, the Productive Levels Test involves the sample words from the same five word frequency levels. Each word level consists of eighteen sentences in which the target words are removed but the initial letter(s) given as clues to write it down. The students are required to complete the underlined words with initial letter(s) as clues according to the meaning of each sentence. Given their reading vocabulary level and the requirementof Chinese New College English Syllabus, the first four word frequency levels in Productive Vocabulary Levels Test were chosen to test the participants. In scoring the vocabulary test, each correct answer was given one point (grammar mistakes were ignored). The total score for this test was 72 points.

For example: (From 2000-word level)

I'm glad we had this op<u>portunity</u> to talk.

3. Listening comprehension test. In the case of listening comprehension test, the scores were extracted from CET4 held on December 19th, 2014. All the subjects took part in the test and the exam results came out in March, 2015. CET is the most formal and large-scale nationwide English proficiency test in China with relative high validity and reliability.It is designed for all the college non-English majors who have finished the required English course to check their language proficiency after approximately one and half years' study. The Listening Comprehension part in CET-4, which takes about 30 minutes to complete and is read at the speed of 130 words per minute, contains three main sub-sections: Section A includes 15 multiple choices for short and long dialogues belonging to Listening Conversations part; Section B contains 3 short passages in the form of also multiple choices; and Section C is compound dictation requiring not only receptive knowledge but also productive knowledge. And each text is played only once, except the compound dictation which was played three times. Additionally, the accents of the recordings are native-like English, either British English or American English. The task types are ranged in multiple choice and sentence completion(see Table). As for the listening material selection, it is authentic and sourced from native English magazines and newspapers ranging from daily conversations, academic lectures, broadcast programs to interviews, concerning the subjects that vary from natural science and social science to humanity science, etc.

TABLE:
STRUCTURE OF CET-4 LISTENING COMPREHENSION

Testing contents		Task type s	Score
Conversations	Short dialogues	Multiple choice	8%
	Long conversations	Multiple choice	7%
Passages	Passages	Multiple choice	10%
	Compound dictation	sentence completion	10%

The total score for listening comprehension is 35 points.

C.   Procedures and Data Collection

Date collection was conducted through the vocabulary tests in the second week of December, 2014, the very week before the nationwide CET4. With the help of the fellow teachers, the survey was administered in regular English classes. The students were asked to finish the Vocabulary Levels Test first and then the Productive Levels Test. Before they went down with the test, the examiners spent a span of time instructing the requirements specifically with the examples illustrated explicitly, making sure every testee was well equipped during the testing. When doing the vocabulary tests, the participants were asked to finish each part on their own without looking up dictionaries or turn to their classmates for help. The whole survey session was under the supervision of the English fellow teachers. The testing time for vocabulary was 60 minutes, the first 30 minutes for Nation's Vocabulary Level Test and the second 30 minutes for Productive Levels Test (Version A). After the first 30 minutes, the testing papers for vocabulary breadth were collected so that students could focus on the second test. For the sake of the effectiveness and authenticity of the experiment, the participants were told that all the tests would be documented and scored as one part of their daily evaluation which would contribute to the final score in the end of the semester.

And data for the listening test were collected at the beginning of March, 2015 after the exam results for CET4 2014 were released to public.

After all the data were collected, the quantitative research methods of Pearson Correlations analysis and a Multiple Regression Analysis programmed in SASS 9.3 were employed as the primary analysis tools for the data collected.

IV.   DATA ANALYSIS AND RESEARCH RESULTS

When designing the model, the multicollinearity between the independent variables results in greatening parameter estimation interval, which is easy to make wrong judgment in hypothesis testing. Therefore, we should be first to check the presence of multicollinearity among the independent variables, generally using Spearman rank correlation coefficient index. In table 1, the correlation coefficient between breadth score (CHGD) and depth score (CHSD) is 0.758, and correlation test P value is less than 0.05, that is, reject the null hypothesis. There is significant correlation between the two independent variables, so, they are not suitable for us to establish simultaneous regression models with the dependent variable, namely listening score (TLCJ). The regression models should be established respectively.

What's more, from table 1 we can see the correlation test results between the dependent varible(TLCJ) and the two independent variables(CHGD and CHSD) show significant correlation, sutible for the establishment of regression models.

TABLE 1
CORRELATION TEST RESULTS BETWEEN VARIABLES

Variables	breadth score	depth score	Listening score
Breadth score	1.000	0.758	0.368
depth score	0.758	1.000	0.408
Listening score	0.368	0.408	1.000

The existing research literature, when assessing the relationship between vocabulary mastery and language performance, usually employ classical least-squares linear regression method, for three reasons: linear estimate method is easy to calculate; when the random disturbance items are uncorrelated with independent variables and comply with a same-variance normal distribution whose mean is zero, the least square method shows unbiasedness, effectiveness and some other good properties; least squares provides a general method of estimating conditional mean value function. However, the inescapable fact is that, when data show Leptokurtosis or have outliers and heteroscedastic cases, the normality, independence and variance homogeneity assumption of least squares are difficult to meet, which results in poor stability. In addition, the least squares regression assumes that the independent variables can only affect the dependent variables' conditional distribution position, but fail to depict the conditional distribution in greater detail and to describe the basic characteristics of the conditional distribution. In order to overcome the disadvantages of least squares, Koenker and Bassett in 1978 proposed the quantile regression (Quantile regression) idea. The basic thought is to perform regression to independent variables according to the conditional quantiles of dependent variables, then regression models on different quantiles are established. So, this method can accurately describe the influence of independent variables on the variation range and conditional distribution of the dependent variables, and at the same time capture the characteristics at tail distribution.

Therefore, to realize the transverse and longitudinal quantitative analysis evaluation about the relationship between vocabulary mastery and students' English listening proficiency at different English levels, this paper pioneered to introduce the concept of fractional regression to data mining and model design. The 2 model forms eventually established are shown in formula 1:

$$\text{TLCJ} = \alpha_0 + \alpha_1 \text{CHGD} + \varepsilon \qquad \text{Quant}_\theta(\text{TLCJ}|\text{CHGD}) = \beta^\theta_{\text{TLCJ}} \text{CHGD}$$

$$\text{TLCJ} = \alpha_0 + \alpha_1 \text{CHSD} + \varepsilon \qquad \text{Quant}_\theta(\text{TLCJ}|\text{CHSD}) = \beta^\theta_{\text{TLCJ}} \text{CHSD}$$

In the formula above, CHGD stands for the breadth of vocabulary, CHSD for the depth of vocabulary, and TLCJ for listening scores. $\text{Quant}_\theta(\text{TLCJ}|\text{CHGD})$ reprasents the conditional quantile of TLCJ that corresponds to quantile theta (0< theta <1) in a given case of CHGD (Liu Shenglong, 2008). And the coefficient vector $\beta^\theta_{\text{TLCJ}}$ corresponding to θ is estimated by minimizing the absolute deviation , i.e:

$$\beta^\theta_{\text{TLCJ}} = \text{argmin}\{\sum \theta \mid_{\text{TLCJ-}}\text{CHGD}\beta\mid + \sum(1-\theta)\mid_{\text{TLCJ-}}\text{CHGD}\beta\mid\}$$

Using Bootstrap intensive algorithm technology to achieve the estimation of regression coefficient, means obtaining samples' confidence interval by continuous sampling with replacement, thus to infer the coefficient. The estimation results from the 2 models are shown in Figure 1, Figure 2, Table 2 and Table 3.
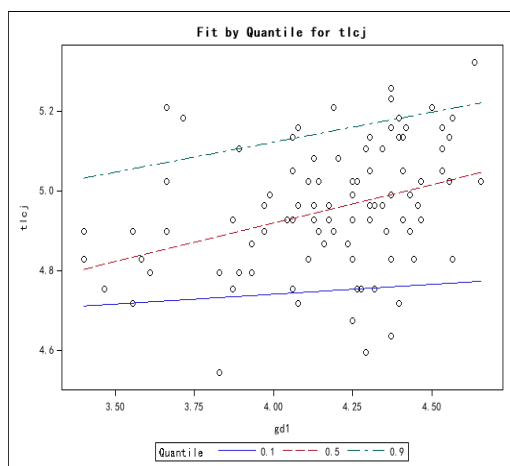

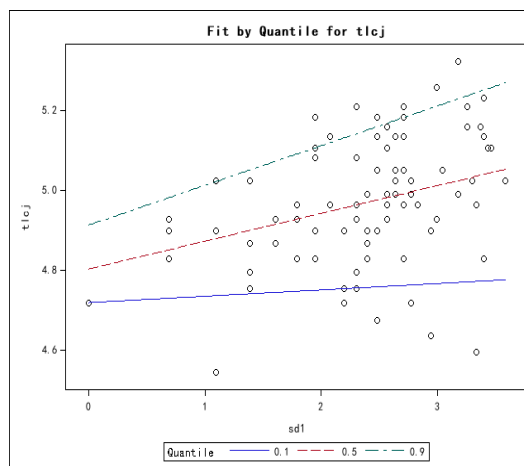Figure 1   Vocabulary breadth and listening scores

Figure 2   Vocabulary depth and listening scores

TABLE 2
QUANTILE REGRESSION RESULTS OF VOCABULARY BREADTH AND LISTENING SCORES

quantile	variable	coefficient	S.E	the lower limit of the 95% confidence interval	the upper limit of the 95% confidence interval	P_ value
0.2	constant term	102.800	11.646	79.693	125.907	0.000
0.2	breath	0.371	0.184	0.006	0.737	0.046
0.5	constant term	111.071	9.524	92.158	129.985	0.000
0.5	breath	0.464	0.145	0.177	0.752	0.002
0.8	constant term	117.929	15.520	87.108	148.749	0.000
0.8	breath	0.643	0.203	0.240	1.046	0.002

TABLE 3
QUANTILE REGRESSION RESULTS OF VOCABULARY DEPTH AND LISTENING SCORES

quantile	variable	coefficient	S.E	the lower limit of the 95% confidence interval	the upper limit of the 95% confidence interval	P_ value
0.2	constant term	111.500	6.422	98.748	124.252	0.000
0.2	depth	1.125	0.512	0.108	2.142	0.031
0.5	constant term	130.793	4.071	122.708	138.878	0.000
0.5	depth	1.069	0.292	0.488	1.650	0.000
0.8	constant term	148.000	9.055	130.019	165.981	0.000
0.8	depth	1.300	0.624	0.060	2.540	0.040

As Table 2and Table 3 show, the dependent variable fits the independent variables well, and in cases of 0.2, 0.5, 0.8 quantiles, the regression coefficients of the 2 models all show significicance at 0.1 significance level. If the quantiles 0.2, 0.5 and 0.8 are respectively regarded as the low, the middle and the high grade levels, following conclusions can be drawn:

a. Conclusion one: From the transverse perspective, both vocabulary breadth and depth influenced listening scores significantly, but the overall effect of vocabulary depth was significantly greater than that of vocabulary breadth.

b. Conclusion two: From a vertical view, at the lower, middle, and higher levels, the enhancing effects of vocabulary breadth showed a gradual increasing trend, i.e. vocabulary breadth teaching can greatly improve the listening scores of high level students more than the scores of other two level students; while the influence of vocabulary depth shows no significant difference at three levels of listening scores.

V.　TEACHING IMPLICATIONS

The study results above have great implications to vocabulary teaching in EFL listening classes:

Firstly, both vocabulary breadth and depth influenced listening scores significantly, so via these study results, we could help students realize the importance of vocabulary study for their listening proficiency. Secondly, in actual English classes, vocabulary teaching should appropriately tilt to vocabulary depth, requiring students not only to remember the literal meaning of words, but give enough attention to lexical collocation, semantic field, etc. Teachers should lead students to lay stress on the accumulation and mastery of active words that are listed in textbooks. Thirdly, with the advance of students' English level, teachers should strengthen the teaching of vocabulary breadth by increasing students' reading quantity and diversifying their reading materials.

The factors affecting listening comprehension are various, such as cognitive strategies, emotion, motivation, attention, etc. This paper just investigated the influence of vocabulary at two dimentions. And, limited by the the testing instruments, we can only test the relationship between reading vocabulary and listening comprehension. If conditions permitting, a study can be conducted to investigate the correlations between listening vocabulary and listening comprehension, and that would be of more practical significance to EFL listening.

REFERENCES

[1] Adolphs, S. & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics, 24(4)*, 425-438.
[2] Bonk, W. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening, 14*, 14-31.
[3] Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*, 403-430.
[4] Kelly, P. (1991). Lexical ignorance: The main obstacle to listening comprehension with advanced foreign language learners. *IRAL, 29(2)*, 135-149.
[5] Koenker, R. & Bassett, G. (1978) Regression quantiles. *Econometrica*, 46(1), 33-50.
[6] Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud, & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp.126–32). London: Macmillan.
[7] Liu, S. L. (2008). Effects of education and experience of Chinese residents' income -- An Empirical Study Based on Quantile Regression and censored quantile regression. *Technology and economy research on Quantitative Economics*, 4, 29-25.
[8] Mecartty, F. (2000). Lexical and grammatical knowlwdge in reading and listening comprehension by foreign language learners of Spanish. *Applied Language Learning, 11*, 323-348.

[9]   Milton, J., Wade, J., & Hopkins, N. (Forthcoming). Aural word recognition and oral competence in a foreign language. In R. Chacon-Beltran, C. Abello-Contesse, M. Torreblanca-Lopez, & M. Lopez-Jimenez (Eds.), *Further insights into non-native vocabulary teaching and learning.*

[10]   Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review, 63(1)*, 59-82.

[11]   Qian, D. D. (1998). Depth of vocabulary knowledge: Assessing its role in adults' reading comprehension in English as a second language [Unpublished doctoral dissertation], University of Toronto, Toronto..

[12]   Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance. *Language Learning, 52*, 513-536.

[13]   Read, J. (1993). A development of a new measure of L2 vocabulary knowledge. *Language Testing, 10*, 355-371.

[14]   Schmitt, N., & McCarthy, M. (1997). Editors' comments - pedagogy section. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy.* Cambridge: Cambridge University Press.

[15]   Schmitt, N., Schmitt, D. & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the vocabulary levels test. *Language Testing 18(1)*, 55–89.

[16]   Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing, *The Language Learning Journal, 36(2)*, 139-152.

[17]   Staehr, L. S. (2009). Vocabulary knowledge and advanced listening comprehenson in English as a foreign language. *Studies in Second Language Acquisition, 31*, 577-607.

[18]   Zeeland, H.V. & Schmitt, N. (2012). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34(4), 457-479.

[19]   Zhang, X. D. (2011). On the relationship of lexical knowledge and listening comprehension. *Foreign Language World, 143*, 36-42.

Suping Wang was born in Binzhou, China in 1977. She received her master's degree in English Education from University of Wisconsin-Platteville, America in 2014.

She is currently a lecturer in Foreign Language Department of Binzhou Medical College, Yantai, China. Her research interests include applied linguistics and English Education.