ISSN: 2959-6386 (Online), Volume 3, Issue 1, March 2024

Journal of Knowledge Learning and Science Technology

Journal homepage: https://jklst.org/index.php/home

The Role of Vocabulary Knowledge in Advanced Listening Comprehension in English as a Foreign Language and Some Suggested Strategies for HUIT Students to Improve Their Vocabulary

Nguyen Thi Kim Anh

Faculty of Foreign Languages, Ho Chi Minh City University of Industry and Trade, Ho Chi Minh City, Viet Nam

Abstract

This article aims at examining the impact of vocabulary knowledge on listening comprehension. The study revealed a substantial correlation between listening comprehension and the depth and breadth of vocabulary knowledge. The study used a quantitative method to collect data from 103 English-majored students at HUIT. The paper highlighted the particular methods for learning vocabulary which can assist students to learn more vocabulary in English and broaden their vocabulary at the same time. Additionally, it might increase students' motivation to learn the language and make the process simpler.

Keywords: V*ocabulary, listening comprehension, vocabulary learning strategies*

Article Information:
Article history: *Received:* 20/08/2023 *Accepted:* 05/01/2024 *Online:* 12/01/2024 *Published*: 29/03/2024
DOI: https://doi.org/10.60087/jklst.vol3.n1.p152
[i] **Correspondence author:** Nguyen Thi Kim Anh **Email:** anhntk@huit.edu.vn

1. Introduction

The majority of linguists acknowledge that research on second language acquisition is contingent upon having a sufficient vocabulary. Vocabulary knowledge is crucial for effective reading comprehension in EFL, stated Hu (2000), Mecartty (2000), etc. . However, there has not been much research done on the relationship between lexical knowledge and listening comprehension, especially EFL listening comprehension. For this reason, teachers and students have long been perplexed about the precise role that vocabulary knowledge plays in L2 listening comprehension. Qian (1998) and Read (1993) have noted that vocabulary often consists of two dimensions: depth and breadth. Nonetheless, the impact of vocabulary breadth has received more attention in the scant research on vocabulary and hearing. A growing number of academic studies agree that language depth plays a crucial role in listening. However, a great deal of empirical research has been done to determine the precise functions of vocabulary depth and breadth in EFL listening comprehension. Both vocabulary knowledge components are heavily weighted in the study.

The investigation will focus on the significance of vocabulary knowledge in relation to listening comprehension.

1.1 Purpose of the Study

In order to provide empirical support for vocabulary instruction in EFL listening classes, the study aims to review the research literature on the role that vocabulary knowledge plays in listening comprehension and it ascertains how vocabulary breadth and depth affect different aspects of listening comprehension. It also carries out an experimental investigation.

1.2 Significance of the Study

Research on the role that vocabulary knowledge plays in EFL learners' listening comprehension is very beneficial to the method of teaching vocabulary in listening classes. Teachers and students will gain a profound understanding of the relationship between vocabulary and listening from this empirical research, which will raise their awareness of the value of vocabulary in listening and help them develop more effective vocabulary-based teaching and learning strategies. The study will also be valuable to anyone looking for knowledge on this subject.

1.3 Definition of Terms

EFL: English as a Foreign Language

Vocabulary breadth is the number of terms in a student's vocabulary, or the number of words that the learner knows the definition of at least in part**.** (Staehr, 2008)

Vocabulary Depth: is the measure of a learner's proficiency in learning certain words and the organization of those words in their mental lexicon. It has to do with how proficient a term user is. (Staehr, 2009)

HUIT: Hochiminh City University of Industry and Trade

2. Literature Review

The fact that vocabulary knowledge has multiple aspects has long been acknowledged. Two of the main dimensions of vocabulary knowledge are depth and breadth, among the other dimensions (Nation, 2001). Thus, the association between the two lexical aspects and listening comprehension is the primary focus of the few studies that exist regarding the relationship between lexis and listening. According to Kelly (1991), vocabulary knowledge is the primary barrier preventing EFL learners from successfully completing listening comprehension tasks, yet actual data has not entirely corroborated this assertion. The relationship between gist comprehension of listening texts and lexis knowledge was examined by Bonk (2000). Tests of 49 Japanese university students with low-intermediate to advanced English proficiency were administered using dictation as a lexical familiarity measure and first-language recall procedures as comprehension measures.

Four texts with increasing numbers of low-frequency lexical words served as the basis for the study. The degree of association between the two variables was quite moderate (45%). At 90% and above, there

was often good comprehension with text-lexis familiarity levels. The study's results do, however, also show that even though some students were able to comprehend the content effectively, they did not know more than 75% of the word kinds in the text, and while some pupils could understand it effectively even if they knew more than 90% of the word kinds. While some of the rationale might be ascribed to the word knowledge and listening comprehension assessments, Bonk connects this discovery to students' mastery of comprehension techniques. (Bonk, 2000)Therefore, the conclusion drawn from this study by Bonk about the connection between lexis and listening comprehension was ambiguous. Researchers found a slightly clearer picture after looking at the relationship between the IELTS exam and two measures of receptive vocabulary size (the Lex and the Lex) with 29 EFL learners. The study found weak but significant Spearman correlations between the reading and listening sections of the IELTS and vocabulary size (the Lex) (0.54 and 0.52, respectively). One major question in research on vocabulary requirements for reading texts or watching movies is how much text coverage is necessary for appropriate comprehension to be likely to occur. They discovered that most learners would require 98% text coverage to achieve satisfactory comprehension. Nation (2006) then reported on the testing of fourteen lists of 1,000 words from the British National Corpus, which were used to determine the size of vocabulary required for understanding without assistance. According to Staehr (2008), there was a significant relationship between vocabulary size and listening comprehension tests as well as vocabulary size and reading, writing, and listening comprehension skills. A vocabulary of 6000–7000 word families has been suggested for sufficient listening, and vocabulary size has been found to predict 39% of the variance in the listening scores. A lexical coverage of 98% is required to comprehend the spoken sentences that make up the hearing exam, according to the findings of another study conducted by (Staehr, 2009). To determine whether it is suitable to generalize the 98% text coverage found in studies pertaining to reading to listening, Zeeland and Schmitt (2012) conducted a recent study in which they directly examined the lexical coverage with reference to hearing comprehension. The majority of both native and non-native participants were able to understand the spoken texts with just 90% coverage, according to the results, yet there was a significant difference in the non-natives' performance at this level. Non-native individuals showed substantially less heterogeneity in their relatively high comprehension at 95% coverage. For sufficient listening comprehension, language users would need to be familiar with between 2,000 and 3,000 word families, based on a 95% coverage number. The results indicate a slight deviation from Nation's (2006) estimate of 6,000–7,000 word families using a 98% coverage. Few studies have examined the depth of vocabulary knowledge, despite the fact that a small number of them have offered empirical evidence for the link between vocabulary knowledge and listening comprehension in L2. In order to better understand the impact of vocabulary knowledge on listening comprehension, Staehr (2009) conducted an empirical study including 115 advanced Danish language learners of English as a second language. The study revealed a substantial correlation between listening comprehension and the

characteristics of vocabulary knowledge, namely depth and breadth. 49% of the variation in the hearing scores could be predicted by vocabulary breadth. However, listening comprehension variabilities could only reach 51% after the regression model took the vocabulary depth into account. The association between foreign language learners of Spanish's lexical and grammatical knowledge and their reading and listening comprehension was investigated in another research (Mecartty, 2000). She found that while lexical and grammatical knowledge were significantly correlated to listening, only lexical knowledge explained the variance in listening (13%). The main causes of the inconsistent research results are the various participant levels and measurement tools. Forty percent of the 115 advanced Danish language learners who took part in Staehr's study had previously spent at least five months in an English-speaking nation. The Vocabulary Levels Test assessed both the breadth and depth of vocabulary knowledge by Schmitt et al (2001) and the Word Associates Test by Read (1993). Zhang (2011) studied the relationship between lexical knowledge and listening comprehension of the TEM4 (Test for English Majors) using sophomore English majors as subjects. The results showed that lexical depth explained 2% of the variance in dictation and listening comprehension, while lexical breadth accounted for 27%, 24%, and different variance of other parts. The findings of this study on vocabulary depth align with those of Staehr's (2009) investigation. The aforementioned studies demonstrate that, despite small variations in particular numbers, researchers generally agree on the importance of vocabulary breadth in listening comprehension. Study findings regarding vocabulary depth, however, have been sparse and erratic. Additionally, there have been flaws in the research process, such as a rather restrictive study framework and somewhat conservative statistical tools, which makes it challenging to develop differentiated instructional strategies for students at different skill levels.

Thus, this paper, which is based on sophisticated vocabulary testing instruments, combines vocabulary breadth, depth, and listening comprehension into one study framework and conducts an empirical study by inventively using quantile regression models in order to analyze the relationship between the two vocabulary dimensions and listening comprehension at various student levels. This work will have important pedagogical ramifications for teaching vocabulary in English to foreign language learners.

This essay will attempt to respond to two queries:

1. *How general listening comprehension is connected with vocabulary depth and breadth?*
2. *To what extent may vocabulary breadth and depth be used to predict variations in listening comprehension among English skill levels?*

3. Methodology

3.1 Participants

For this study, 103 sophomores from the Hochiminh City University of Industry and Trade (HUIT) enrolled in course 0101100822 were chosen. My place of employment, HUIT served as the research location is situated in Ho Chi Minh City's Tan Phu District , this is one of the largest colleges in Vietnam.

3.2 Methods

To create a pen-and-paper empirical experiment with three tests: the CET-4 listening comprehension test, the productive levels test (Laufer & Nation, 1999), and the vocabulary levels test (Nation, 1990), the key analysis techniques for the data acquired will be a Multiple Regression Analysis developed in SASS 9.3 and Pearson Correlations analysis, both of which are quantitative research methods.

3.2.1 Test your vocabulary's depth and breadth

The Vocabulary Level, which has been used extensively in a vocabulary research context and can offer an accurate estimate of learners' vocabulary knowledge at the different frequency levels, is the basis for the tests measuring vocabulary breadth in this study (Staehr, 2009). There are five different levels for this type of test: 2000 words, 3000-words, 5000 words, university words, and 10,000 words. High-frequency words can be found in word levels 2000 and 3000-word; one sort of specialist vocabulary is represented by the university word level; and words in the 5000-word level fall between high- and low-frequency word boundaries. Finally, because of its low-frequency phrasing, only few EFL learners are able to attain the 10000-word level. The first four-word frequency levels in the Vocabulary Levels Test were selected to test the participants based on their reading vocabulary level. Ten groups, each with six terms and three definitions, make up each level. It is necessary for the participants to match three of the six terms on the left with the three definitions on the right.

This test had a 120-point maximum score, with one point deducted for each blank.

3.2.2 Assess the level of vocabulary depth

To gauge pupils' depth of vocabulary, Laufer and Nation (2001) developed the Productive Levels Test (Version A). The Productive Levels Test uses sample words from the same five word frequency levels as the Vocabulary Levels Test. Eighteen sentences make up each word level; the target words are omitted, but the first letter or letters are provided as hints to put the word down. According to the meaning of each sentence, the students must fill in the underlined words with the initial letter or letters as hints. Based on the participants' reading vocabulary level and the requirements of the HUIT English Syllabus, the first four-word frequency levels in the Productive Vocabulary Levels assessment were chosen. Each right response on the vocabulary test was worth one point; grammatical errors were not taken into account. This test had a total score of 72 points.

3.2.3 Test of listening comprehension

The listening comprehension exam results were taken from September, 2022 and the results were released in January 2023. It is intended to assess language proficiency of university students. Two primary sub-sections comprise the thirty-minute Listening Comprehension portion which is read at a speed of one hundred and thirty words per minute: Section A comprises fifteen multiple-choice questions for both short and long conversations from the Listening Conversations section; Section B has three short passages presented as multiple-choice questions; and Section C is compound dictation that calls for both productive and receptive knowledge.

With the exception of the compound dictation, which was played three times, each text is only played once. Furthermore, the recordings have native-sounding English accents, either British or American. The tasks vary in sentence completion and multiple choice (refer to Table). The listening material selection is genuine and comes from periodicals and newspapers written in native English. It covers a wide range of topics, from social science and natural science to humanities and sciences, and includes daily discussions, academic lectures, radio programs, and interviews.

Testing contents		**Task kinds**	**Score**
Talks	Brief conversations	Multiple choice	9%
	extended talks	Multiple choice	8%
Sections	Sections	Multiple choice	11%
	Complex voice	Sentence completion	11%

3.3 Data Collection

The vocabulary exams were used to gather data throughout the week of September 2022. Regular English classrooms were used to administer the survey from researcher's. Prior to moving on to the Productive Levels Test, the students were instructed to complete the Vocabulary Levels Test. The examiners took the time to explain the requirements in detail and provide precise examples before beginning the examination, ensuring that each test taker was prepared for success.

The participants in the vocabulary tests were instructed not to consult dictionaries or ask their classmates for assistance when completing any portion of the examination on their own. The researcher oversaw the entire survey session. The Nation's Vocabulary Level Test took up 30 minutes of the 60-minute testing period, followed by the Test of Productive Levels (Version A) for the remaining 30 minutes.

The vocabulary breadth assessment papers were collected after the first thirty minutes so that students might focus on the second test. The participants were made aware that every exam would be documented and assessed as a component of their daily evaluation, which would be added to their final semester score.

As a result, the validity and effectiveness of the experiment were guaranteed. Furthermore, the listening test data was gathered in early January 2022. Following the collection of all the data, the principal analysis tools for the data were a Multiple Regression Analysis designed in SASS 9.3 and Pearson correlations analysis, both quantitative research methodologies.

3.4 Data analysis

It is simple to make a mistake in a hypothesis test because of the increasing parameter estimate interval caused by the multi-collinearity between the independent variables during model design. As a result, we ought to start by determining if the independent variables are multi-collinear, usually with the use of the Spearman rank correlation coefficient index. Table 1's null hypothesis is rejected since the correlation test P value is less than 0.05 and the correlation coefficient between the depth score (CHSD) and breadth score (CHGD) is 0.758. Since the two independent variables have a substantial correlation with each other, They cannot be utilized to build simultaneous regression models using the listening score (TLCJ) as the dependent variable. Establishing the regression models is necessary in each case. Additionally, table 1's correlation test findings between the two independent variables, CHGD and CHSD, and the dependent variable, TLCJ, demonstrate a significant correlation that is appropriate for the creation of regression models.

Table 1. Results of a correlation test between variables

Variables	**Score for breadth**	**Score for depth**	**Score for listening**
Score for breadth	1.000	0.757	0.369
Score for depth	0.757	1.000	0.428
Score for listening	0.369	0.428	1000

The traditional least-squares linear regression method is typically utilized in the current study literature to evaluate the correlation between language performance and vocabulary knowledge. This is due to three main reasons: A general technique for estimating conditional mean value function is provided by the least squares method. When the random disturbance items are uncorrelated with independent variables and follow a same-variance normal distribution with a zero mean, it is easy to compute and exhibits impartiality and efficacy. However, it is an indisputable fact that when data show Leptokurtosis or contain outliers and heteroscedastic conditions, it is impossible to meet the normality, independence, and variance homogeneity assumptions of least squares, which results in poor stability. Furthermore, the least squares regression fails to provide a more comprehensive representation of the conditional distribution or to elucidate its

fundamental features, assuming that the independent factors have an impact exclusively on the dependent variables' conditional distribution position. The quantile regression concept was first put forth by (Koenker and Bassett 1978) as a way to get around the drawbacks of least squares. The fundamental idea is to create regression models on various quantiles and then perform regression to independent variables based on the dependent variables' conditional quantiles. As a result, this method may be able to precisely describe the impacts of independent variables on the variation range and conditional distribution of the dependent variables, as well as effectively capture the features at tail distribution. In order to achieve the transverse and longitudinal quantitative analysis evaluation about the relationship between vocabulary mastery and students' English listening proficiency at different English levels, this paper thus pioneered the introduction of the concept of fractional regression to data mining and model design. Formula 1 shows the two model forms that were finally decided upon:

The above algorithm uses TLCJ for listening scores, CHSD for vocabulary depth, and CHGD for vocabulary breadth presents the TLCJ conditional quantile, which in a given CHGD instance corresponds to quantile theta (0<theta <1) (Liu, 2008). Furthermore, the coefficient vector corresponding to the approach of reducing the absolute deviation is as follows in order to estimate the regression coefficient: By using continuous sampling with replacement to acquire the samples' confidence interval and infer the coefficient, TLCJ-TLCJU eliminates Bootstrap intensive method technology.

Table 2. Vocabulary breadth and listening scores as quantitative regression results

Quantile	variable	coefficient	S.E
0.2	constant term	102.801	11.656
0.2	breath	0.361	0.174
0.5	constant term	112.071	9.534
0.5	breath	0.454	0.146
0.7	constant term	117.829	15.510
0.8	breath	0.653	0.213

Table 2 demonstrates how well the dependent variable fits the independent factors. The regression coefficients of the two models all demonstrate significance at the 0.2 significance level for quantiles of 0.2, 0.5, and 0.8. If the quantiles 0.2, 0.5, and 0.8 are used to represent the low, middle, and high grade levels, respectively, then the following conclusions can be drawn:

Students' listening scores more than it can other two level students' scores. However, at three listening score levels, there is no discernible difference in the influence of vocabulary depth.

4. Results and Discussion

4.1 First finding

In L2 learning in general and listening comprehension in particular, vocabulary knowledge is crucial. The study's findings clearly imply that listening comprehension, vocabulary depth, and breadth are closely related to one another. They both had a considerable impact on hearing scores; but, altogether, vocabulary depth had a significantly bigger impact than vocabulary breadth.

4.2 Second finding

When looking at listening scores from a vertical perspective, the enhancing effects of vocabulary breadth at the lower, middle, and higher levels showed a gradually increasing trend. This means that teaching vocabulary breadth can significantly improve high level. Certain skills that can impact comprehension may have varying degrees of influence on the respective effects of vocabulary breadth and vocabulary depth on various listening measures. With more sophisticated measures of listening comprehension, future research can reexamine and clarify the impacts of variety.

5. Conclusion

The aforementioned study findings have significant ramifications for vocabulary instruction in EFL listening classes:

First off, the results of this study showed that vocabulary depth and breadth both had a substantial impact on hearing scores. This means that by understanding the significance of vocabulary study for listening proficiency, we may better assist students.

Second, vocabulary instruction in practical English classes should appropriately emphasize vocabulary depth, asking students to pay attention to semantic field, lexical collocation, and other aspects of words in addition to their literal meaning. Lecturers should guide their students to place a strong emphasis on learning and mastering the active words included in textbooks.

Thirdly, as students' English proficiency advances, teachers should make a stronger effort to teach vocabulary breadth by having students read more and a wider variety of books.

Numerous elements, including cognitive processes, emotion, motivation, attention, and so on, influence listening comprehension.

This essay only looked into vocabulary's impact in two dimensions. Furthermore, we are only able to examine the connection between vocabulary reading and listening comprehension due to the limitations of the testing tools. More practically speaking, a study examining the relationships between acquiring rich vocabulary and listening comprehension might be carried out. This study would focus on EFL listening.

References

Bonk, W. (. 2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening, 14*, , 21-28.

Hu, M. &. (2000). Unknown vocabulary density and reading comprehension. *Reading in a foreign language*, 358-388.

Hu, M. &. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*, 358-388.

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*, 33-39.

Mecartty, F. (2000). Lexical and grammatical knowledge in reading and listening comprehension by foreign language learners. *Reading in a foreign language*, 217-219.

Mecartty, F. (2000). Lexical and grammatical knowlwdge in reading and listening comprehension by foreign language learners of Spanish . *Applied language learning (11)*, 237-263.

Nation, I. S. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review, 63(1)*, 77-100.

Qian, D. D. (1998). Depth of vocabulary knowledge: Assessing its role in adults' reading comprehension in English as a second language. *Doctoral dissertation*, 163-175.

Read, J. (1993). A development of a new measure of L2 vocabulary knowledge. *Language Testing, 10*, 237-253.

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)*, 45-69.

Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing,. *The Language Learning Journal, 36(2)*, , 148-165.

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

Zeeland, H. &. (2012). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? . *Applied Linguistics, 34(4)*, , 363-392.

Zhang, X. D. (2011). On the relationship of lexical knowledge and listening comprehension. *Foreign Language World, 143*, 54-60.