

Language Proficiency and Knowledge in Adjective-Noun Collocations: A Case Study of Vietnamese Learners of English

Thai Bao Ngoc Pham

Faculty of English Linguistics and Literature, University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh City, Vietnam

Abstract—Collocational development is of great significance to second language acquisition. Among different types of collocations, adjective-noun collocations are notoriously difficult to EFL learners, but there has been limited research, especially in Vietnam, on this type of collocations in the field of teaching and learning English. To address this issue, the current study, employing the quantitative approach, investigates Vietnamese university students' receptive and productive knowledge of adjective-noun collocations and the relationship between their language proficiency and their collocational knowledge. Results reveal that the difference between the students' receptive and productive knowledge was more significant when they reached higher levels of English, and even those at an advanced level had great difficulties in identifying erroneous collocations and using them correctly. The number of collocational errors appeared to increase with rising proficiency. Results also indicate a close relationship between language ability and collocational knowledge, thereby emphasizing the importance of teaching collocations explicitly in English classrooms.

Index Terms—adjective-noun collocations, collocational knowledge, language proficiency, receptive knowledge, productive knowledge

I. INTRODUCTION

Almost 70% of all spoken and written genres of discourse consist of formulaic sequences, among different types of which collocations are “the most powerful force in the creation and comprehension of all naturally-occurring text” (Lewis, 2000, p. 53). The mastery of collocations is considered as part of communicative competence and an important indicator of EFL learners' language proficiency (Celce-Murcia, 2008; Men, 2018). Thus, the lack of collocational knowledge can be a great hindrance to EFL learners in acquiring thorough understanding of both spoken and written discourse and developing native-like proficiency (Wray, 2002).

The problem is that collocations are notoriously difficult to learn due to the arbitrary nature of collocations, their rigid combination, and their variation in different languages. Because of these reasons, it appears inevitable for EFL learners, even those at advanced level, to make errors in using collocations (Nesselhauf, 2003; Laufer & Waldman, 2011).

Among different types of collocations, verb-noun and adjective-noun collocations are the most frequently used by native speakers and also the most problematic to EFL learners (Gui & Yang, 2003; Benson et al. 2010). In comparison with verb-noun collocations, there appears to be relatively little research focusing on EFL learners' errors in adjective-noun collocations and the relationship between their language proficiency and their knowledge of this type of collocations, especially in the context of Vietnam, where grammar instruction still dominates the EFL classes. Thus, this study is an attempt to fill in these gaps in the literature by addressing the following research questions:

- (1) Is there any statistically significant difference between Vietnamese learners' receptive and productive knowledge of adjective-noun collocations in each level of English?
- (2) Is there a relationship between learners' level of English proficiency and their attempts to use adjective-noun collocations?
- (3) Is there a relationship between Vietnamese learners' receptive knowledge of adjective-noun collocations and their level of English proficiency?
- (4) Is there a relationship between Vietnamese learners' productive knowledge of adjective-noun collocations and their level of English proficiency?
- (5) Is there any statistically significant difference in Vietnamese learners' knowledge of adjective-noun collocations between different levels of English proficiency?

Within the scope of this study, the term “receptive knowledge” is used to refer to the recognition of the right adjective-noun combinations in reading rather than listening, and “productive knowledge” is understood as the recall and use of adjective-noun collocations in writing, not in speaking.

II. LITERATURE REVIEW

A. *Definition of Collocation*

There are two main approaches frequently adopted by researchers when defining collocation, namely the frequency-based approach (Firth, 1957; Sinclair, 1991; Lewis, 2000) and the phraseological approach (Cowie, 1981; Howarth, 1998). Whereas the former approach examines the habitual co-occurrence of two or more lexical items, using quantitative evidence from corpora, the latter highlights the restriction in combination and the semantic transparency of a collocation, differentiating it with a free combination and an idiom. As both of these approaches have their own merits and demerits, it is suggested that a definition of collocation should embrace the essence of the two major approaches (Kurosaki, 2012).

However, even when this definition combines the ideas of the two approaches, without a clear-cut boundary between a collocation and other linguistic expressions, it is hard to determine whether a group of two or more words is a restricted collocation or, for example, a free combination. Two specific criteria which are often used in research on collocation are frequency (FREQ) and mutual information (MI).

According to Wolter and Yamashita (2015), corpus linguists tend to use “a minimum MI-score threshold of 3 as indicative of significant co-occurrence” to examine how closely related two words are (p. 1201). As regards frequency, while Hong et al. (2011) set the minimum FREQ of five tokens in BNC (the British National Corpus) as the standard threshold for a group of words to be considered a collocation, Shin and Chon (2019) points out that 20 repetitions prove sufficient as the minimum cut-off point for frequency in COCA due to its relatively considerable size (with more than one billion words of data, as compared to 100 million words in BNC) (p. 611). As COCA was used in the current study as a reference source for identifying collocation, the operational definition for the current study is formed as follows:

A collocation is an arbitrarily restricted lexeme combination which is syntactically fixed to a certain degree, is included in the collocation dictionary, presents a minimum MI score of 3.0, and has a raw FREQ of at least 20 tokens in COCA.

B. *Language Proficiency and Collocational Competence*

A growing body of literature has investigated the relationship between English proficiency and collocational knowledge. According to Nesselhauf (2003) and Yan (2010), EFL students, even those at an advanced level, tend to make the most errors in verb-noun and adjective-noun collocations in comparison with other types. Until now there are several studies exploring the connection between verb-noun collocations and language proficiency (Hosseini & Akbarian, 2007; Bueraheng & Laohawiriyanon, 2014; Ebrahimi-Bazzaz et al., 2014, 2015a, 2015b; Nguyen & Webb, 2016; Men, 2018, etc.); however, few researchers, especially in Vietnam, have thoroughly examined the correlation between levels of English and adjective-noun collocational knowledge in terms of both reception and production. Following are the studies conducted by researchers from different countries in the world with EFL learners' adjective-noun collocational knowledge as part of their research aims.

Bueraheng and Laohawiriyanon (2014) investigated whether students with different degrees of exposure to English would achieve different levels of verb-noun and adjective-noun collocational knowledge. The participants were Thai fourth-year students, including 116 students in the International Program at the intermediate level of English and 80 students in the English Major Program at the lower intermediate level. Results indicated that all the participants irrespective of their levels had significantly higher results in terms of their receptive knowledge as compared to their productive knowledge, and that the intermediate students outperformed the lower intermediate students in both receptive and productive knowledge tests.

Sharing the same interest in EFL students' collocational competence at different levels of English, Talakoob and Koosha (2017) examined the possible difference in the receptive and productive collocational knowledge of 60 Iranian students at intermediate and advanced levels. The results showed that the intermediate students' test scores for receptive collocational knowledge were significantly higher than their productive knowledge scores, whereas no statistically significant difference was found in advanced learners' scores for receptive and productive collocational knowledge.

Similarly, Kamarudin (2020) examined whether there was a significant difference between Malaysian students' productive and receptive knowledge of collocations and between their performance on the verb-noun, adjective-noun, and verb-preposition collocation test scores. The participants were 21 university students at the intermediate level of English. The results of the study reveal that the students' receptive knowledge of collocation was significantly greater than their productive knowledge and that verb-preposition collocations were the most problematic to the students while verb-noun collocations caused the least problems.

The major theme running through these studies is that the students tend to have a greater passive knowledge of collocations than their active knowledge. Unlike these studies, Nguyen and Webb (2016) and Men (2018) only focused on the students' receptive knowledge or productive knowledge of collocations in their research.

Nguyen and Webb (2016) only focused on the learners' receptive knowledge of verb-noun and adjective-noun collocations at the first three 1,000 word frequency levels. The participants were 100 Vietnamese first-year students majoring in English, whose general level of English proficiency was estimated to be between pre-intermediate to intermediate. The results show that Vietnamese EFL learners had limited collocational knowledge at the first three 1,000 word frequency levels, for the participants' mean scores in both verb-noun and adjective-noun sections were below 50% across the three frequency levels (1,000, 2,000 and 3,000 words). Besides, there were statistically

significant differences in the mean scores across the three levels of each section of the collocation test. The findings also indicated a decrease in the students' collocational knowledge across the three levels.

Different from the aforementioned prior studies on collocations whose research instruments were mainly collocation tests, Men (2018) made use of the Chinese Learner English Corpus, a database of written texts by Chinese students at five different levels of English proficiency to investigate the relationship between the production of L2 collocation and vocabulary growth. The study only focused on the three most common types of collocations, including verb-noun, adjective-noun and noun-noun collocations. As regards adjective-noun collocations, the quantitative analyses of the study showed that the students at the higher levels of English produced significantly more correct adjective-noun combinations than the lower levels, and there was a significant decrease in collocation errors in the writings of advanced learners as compared with those at a basic level. It can be seen that their productive collocational knowledge develops with rising L2 proficiency. However, it is also noteworthy that the students in this study, even those at the basic level, did not have great difficulties with adjective-noun collocations.

Based on the findings from Nguyen and Webb (2016) and Men (2018), it can be concluded that EFL students tend to have a limited receptive knowledge of adjective-noun collocations, but surprisingly they did not encounter many problems in using this type of collocations. These findings appear to be in contradiction with Bueraheng and Laohawiriyanon (2014), Talakoob and Koosha (2017) and Kamarudin (2020) which highlighted that EFL students' receptive knowledge of adjective-noun collocations was generally greater than their productive knowledge.

In brief, the review of the literature shows that there have been an insufficient number of studies on the relationship between different levels of language proficiency and adjective-noun collocational knowledge. The few studies that did explore this issue have mostly been limited to their sample size and the number of English proficiency levels which were discussed and analyzed in correlation with L2 collocational knowledge. Besides, the conflicting findings among previous studies seemed to add to the general confusion about this issue rather than solve it adequately. Thus, this study attempts to fill the existing gap in the literature by investigating whether there is a statistically significant difference between EFL learners' receptive and productive knowledge of adjective-noun collocations and whether there is a relationship between the learners' different levels of English proficiency and different levels of collocational knowledge, i.e. recognition and production, in the context of Vietnam.

III. RESEARCH METHODOLOGY

A. Participants

This study was conducted at the University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh City in the academic year 2020-2021. Due to the impact of COVID-19, it was almost impossible to opt for a probability sample, so the researcher used convenience sampling by choosing the sample from the Vietnamese students to whom she had easy access. Initially, there were 227 English-majored students, including freshmen, sophomores and juniors, who volunteered to participate in the research. However, at the time of data collection, 11 students missed one or two tests given in the research, so they were excluded from the study. As a result, the total number of participants in this study was 216. In order to evaluate the students' English proficiency, the Quick Placement Test (QPT) (UCLES, 2001) was administered to the students. The results show that the students' levels of English ranged from B1 (intermediate) to C1 (advanced). The description of the participants is summarized in Table I.

TABLE I
DEMOGRAPHIC CHARACTERISTICS OF THE PARTICIPANTS

		Freshmen		Sophomores		Juniors		Total	
		Count	%	Count	%	Count	%	Count	%
Gender	Male	12	14.8	11	15.1	9	14.5	32	14.8
	Female	69	85.2	62	84.9	53	85.5	184	85.2
Age	18-21	79	97.5	68	93.2	57	91.9	204	94.4
	>21	2	2.5	5	6.8	5	8.1	12	5.6
English Proficiency	B1	59	72.8	3	4.1	0	0.0	62	28.7
	B2	19	23.5	44	60.3	15	24.2	78	36.1
	C1	3	3.7	26	35.6	47	75.8	76	35.2
Number of students		81	100	73	100	62	100	216	100

B. Research Instruments

There were two instruments used in this study to measure the students' collocational knowledge, including the Receptive Knowledge Test and the Productive Knowledge Test. The Productive Knowledge Test was designed before the Receptive Knowledge Test as the students' answers to the questions in the former test would provide materials for the researcher to design the test items for the latter test.

In the Productive Knowledge Test, there are 20 gap-filling items, each of which requires the students to supply the missing adjective-noun collocation in each sentential context. As this is a collocation test, not a language proficiency test, the given contexts and the missing word combinations were all translated into Vietnamese to aid the students' comprehension, ensuring that any mistake that the students made was due to their limited knowledge of collocations,

not their misunderstanding of the contexts. For one Vietnamese translation of the target collocation, there may be different word combinations in English that express the same idea; therefore, the students were encouraged to provide as many adjective-noun collocations as possible to demonstrate their collocational knowledge. Following is an example of a test item in the Productive Knowledge Test.

5. He became totally toothless and had to wear _____ to eat.

(Ông ta mất to ần bộ răng và phải mang răng giả mới c ó thể ăn được.)

The Receptive Knowledge Test includes 20 multiple-choice items. As receptive knowledge refers to the ability to recognize not only the right words that can occur with a specified word but also the incorrect combinations (Gyllstad, 2007), the students were required to choose the best answer out of three given options. This is an example of test item in the Receptive Knowledge Test.

5. A. artificial teeth B. false teeth C. Both are correct.

The 20 target collocations in this test are the same as those in the Productive Test because one of the purposes of the receptive test was to find out whether the students were able to identify which word combinations are incorrect. Moreover, in the Productive Knowledge Test, the students might write only the collocations they were certain of and avoid providing the answers they felt uncertain, thus the inclusion of these test items in the Receptive Knowledge Test was expected to provide a more complete picture of their collocational knowledge. Specifically, one of the options in each test item is the target collocation, another option is either a wrong combination or another correct collocation with the same meaning as the first one, and the last option is *Both are correct*. The wrong word combinations were created by translating literally from the Vietnamese equivalents, substituting one component of the correct collocation with its synonym, or, in many cases, making use of the students' incorrect answers in the Productive Knowledge Test.

The validity of the tests was ensured by the researcher's senior colleagues and mentors. The tests were also piloted on 20 students of the same population, using the test-retest method, to assess their reliability. The correlation coefficients of the Receptive Knowledge Test and Productive Knowledge Test were .783 and .798 respectively, which confirmed the stability of the tests over time. In addition, the Cronbach's alpha values (α) for these tests were .787 and .802, reassuring the internal consistency of the collocation tests.

C. Data Collection and Analysis Procedure

The following steps were taken to collect the data. First, the QPT was distributed to the participants to evaluate their English proficiency. In the next meeting with the researcher, the students were given the Productive Knowledge Test and they had 30 minutes to complete the test. The Productive Knowledge Test was administered to the students before the Receptive Knowledge Test so that the students would not remember the options in the multiple-choice questions and wrote them in the gap-filling test items. After the first collocation test was completed, the researcher collected all the students' answer sheets before distributing the second test. The students continued to do the Receptive Test within only 15 minutes, for this test was less time-consuming than the previous one. The total time for the whole test was 45 minutes.

The maximum score of the Receptive Knowledge Test was 20. The answer to each item in this test was coded dichotomously as right (1) or wrong (0) in SPSS. As regards the Productive Knowledge Test, there was no maximum score decided beforehand. Since the students were encouraged to provide many answers for each test item, there may be both correct answers and incorrect answers in one test item. The students were given one point for every correct answer. This means the score for the Productive Knowledge Test was the total number of the correct answers they provided in the test. Besides, the attempts to use collocations in the Productive Knowledge Test are equal to the total numbers of both correct and incorrect collocations. COCA was employed to check the collocability of the adjective-noun combinations in the students' answers with the criteria set by the operational definition.

Regarding the data analysis procedure, in order to find out whether there was a statistically significant difference between the learners' receptive and productive collocational knowledge, paired samples t-tests were conducted on the mean scores of the two collocation tests. Pearson product-moment correlation coefficients were also computed to examine the relationships between the students' language proficiency and their collocational knowledge in terms of both recognition and production and between their language competence and their attempts to use collocations. Finally, one-way ANOVA was run to find out whether there was any statistically significant difference in Vietnamese learners' knowledge of adjective-noun collocations between different levels of English proficiency.

IV. FINDINGS AND DISCUSSION

A. Receptive and Productive Knowledge of Collocations

The first research question focuses on whether the students' receptive knowledge of collocations was significantly greater than their productive knowledge in each English level. Paired-samples t-tests were used to compare the mean scores of the two collocation knowledge tests, the results of which are summarized in Tables II and III.

TABLE II
DESCRIPTIVE STATISTICS OF THE RECEPTIVE KNOWLEDGE TEST AND PRODUCTIVE KNOWLEDGE TEST

Level		Mean	N	Std. Deviation	Std. Error Mean
B1	Receptive Knowledge Test	4.53	62	1.067	.135
	Productive Knowledge Test	4.74	62	1.492	.190
B2	Receptive Knowledge Test	6.81	78	1.571	.178
	Productive Knowledge Test	7.46	78	1.535	.174
C1	Receptive Knowledge Test	7.18	76	1.449	.166
	Productive Knowledge Test	9.03	76	1.766	.203

TABLE III
RESULTS OF THE PAIRED SAMPLES T-TEST FOR THE SCORES OF THE RECEPTIVE KNOWLEDGE TEST AND THE PRODUCTIVE KNOWLEDGE TEST

Level	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
B1	-.210	1.611	.205	-.619	.199	-1.025	61	.309
B2	-.654	2.005	.227	-1.106	-.202	-2.880	77	.005
C1	-1.842	1.488	.171	-2.182	-1.502	-10.791	75	.000

As can be seen in Table III, the intermediate students' mean score of the Receptive Knowledge Test ($M = 4.53$, $SD = 1.067$) is not significantly lower ($t(61) = -1.025$, $p = 0.309 > 0.05$) than the score of the Productive Knowledge Test ($M = 4.74$, $SD = 1.492$). Thus, it could be concluded that there was no statistically significant difference between the intermediate students' receptive and productive knowledge of L2 collocations. In contrast, there was a significant difference ($t(77) = -2.880$, $p = 0.005 < 0.05$) between the receptive knowledge score of the upper-intermediate (B2) students ($M = 6.81$, $SD = 1.571$) and their productive mean score ($M = 7.46$, $SD = 1.535$). In a similar vein, the advanced students had a significantly lower mean score for their recognition of English collocations than for their production of these collocations ($t(75) = -10.791$, $p = 0.00 < 0.05$). These findings are in contradiction with the results in Bueraheng and Laohawiriyonon (2014), which reported that the intermediate learners' receptive knowledge was significantly higher than their productive knowledge of collocations, and they also differ from Talakoob and Koosha (2017), which argued that there was no significant difference between the advanced learners' receptive and productive collocational knowledge.

One of the most striking findings to emerge from the data is that the participants in this study generally performed significantly worse in the Receptive Knowledge Test. This result is in contrast with earlier evidence of EFL learners' tendency to have a better score in the receptive test than in the productive test of collocations (Bueraheng & Laohawiriyonon, 2014; Talakoob & Koosha, 2017; Kamarudin, 2020).

Possible explanations for these contradictions in research results might lie in the differences in the scope and the format of the receptive test in each study. Specifically, the Receptive Knowledge Test in this study focused only on adjective-noun collocations, one of the most problematic collocation type for EFL learners, whereas the tests in the previous studies included different types of collocations, which means adjective-noun collocations are just part of the test, and some target collocations were relatively easy even to B1 students such as *next week*, *take place*, etc. It is also noteworthy that the test items in the receptive test of the previous studies seem to be less challenging, for the participants were simply required to evaluate whether or not the given word combinations are acceptable or which collocation is correct between the two options. To put it differently, the students had a 50% probability to answer correctly even when they do not know the correct word combination. Besides, such a test format allows the respondents to know only one correct option and ignore the other one as it is clearly incorrect. In the current study, the third option, i.e. *Both are correct*, was added not only to increase the difficulty of the Receptive Knowledge Test but also to check the students' ability to recognize erroneous collocations as well as their knowledge of different adjectives that can collocate a certain noun. With limited knowledge about collocations, the students, especially those at the lower English level, might have difficulties in deciding whether an unfamiliar combination was a correct collocation or not. At B1 level, some students were even not aware of collocations; therefore, recognizing deviant collocations or producing well-formed collocations seemed to be impossible for these students. This explains the reason why the students at B1 level not only had very low mean score for the Productive Knowledge Test but also a similarly poor score for the Receptive Test.

Additionally, the occurrences of incongruent collocations in some items of the receptive test also caused a great problem for the participants in this study, even those at an advanced level. For example, the Vietnamese collocation *gió mạnh* (literally translated as 'strong wind') has an English congruent equivalent, i.e. *strong wind* and an incongruent one, i.e. *high wind*. Most Vietnamese students could easily identify *strong wind* as a correct collocation, but they did not consider *high wind* as an equivalent combination because of the inequivalent component *high*, which is why they failed to choose the correct answer, i.e. *Both are correct*. While in the Productive Knowledge Test the students could avoid using the word combinations of which they were uncertain to achieve error-free production and gain one point for every correct answer, the design in the Receptive Knowledge Test did not allow them to use such a risk-avoiding strategy; instead, their lack of knowledge was exposed and they lost the point if they did not know the unfamiliar collocations or could not recognize the erroneous collocations. All of these explain the reason why the students at higher levels of English proficiency had a significantly lower mean score in the receptive test than their productive test score.

B. Language Proficiency and Attempts to Use Collocations

Pearson product-moment correlations were run to determine the relationship between the Vietnamese students' levels of English and their attempts to use adjective-noun collocations. These attempts could result in well-formed collocations or deviant collocations. Following are the statistics of correct and incorrect collocations and the SPSS output of the Pearson correlations.

TABLE IV
DESCRIPTIVE STATISTICS OF WELL-FORMED AND DEVIANT COLLOCATIONS IN THE PRODUCTIVE KNOWLEDGE TEST

Level	Productive Knowledge Test					
	Well-formed collocations		Deviant collocations		Total number of collocations (or total number of attempts to use collocations)	
	Count	%	Count	%	Count	%
B1	294	27.0	793	73.0	1087	100
B2	582	34.9	1087	65.1	1669	100
C1	686	38.9	1079	61.1	1765	100

As shown in Table IV, the students at the three levels of L2 proficiency produced a particularly high number of deviant collocations, which constituted well over half of all the adjective-noun collocations they attempted to produce. The results presented in Table VI reveal that there was a weak relationship between the learners' language proficiency and the number of deviant collocations they produced ($r = 0.272$, $p = 0.000 < 0.05$). The direction of the relationship was positive (i.e., the proficiency level and the number of incorrect collocations were positively correlated), meaning that these variables tended to increase together, though the magnitude, or strength, of the association was weak ($0.20 < r < 0.40$).

TABLE V
DESCRIPTIVE STATISTICS OF THE SCORES OF THE QUICK PLACEMENT TEST AND OTHER DATA RELATED TO THE PRODUCTIVE KNOWLEDGE TEST

		Mean	Std. Deviation	N
Productive Knowledge Test	Quick Placement Test	43.86	6.351	216
	Well-formed collocations	7.23	2.347	216
	Deviant collocations	13.70	2.057	216
	Attempts to use collocations	20.93	3.254	216

TABLE VI
RESULTS OF THE PEARSON PRODUCT-MOMENT CORRELATIONS FOR THE SCORES OF THE QUICK PLACEMENT TEST AND THE ATTEMPTS TO USE COLLOCATIONS

		Quick Placement Test	Well-formed collocations	Deviant collocations	Attempts to use collocations
Quick Placement Test	Pearson Correlation	1	.772**	.272**	.729**
	Sig. (2-tailed)		.000	.000	.000
	N	216	216	216	216
Well-formed collocations	Pearson Correlation	.772**	1	.088	.777**
	Sig. (2-tailed)	.000		.199	.000
	N	216	216	216	216
Deviant collocations	Pearson Correlation	.272**	.088	1	.695**
	Sig. (2-tailed)	.000	.199		.000
	N	216	216	216	216
Attempts to use collocations	Pearson Correlation	.729**	.777**	.695**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	216	216	216	216

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

It can be seen from Table VI that there were strong correlations between the students' language proficiency and the accuracy in their production of adjective-noun collocations ($r = 0.772$, $p = 0.000 < 0.05$) and between their English proficiency and their attempts to use collocations ($r = 0.729$, $p = 0.000 < 0.05$). Besides, the learners' attempts to use collocations and the number of incorrect collocations were moderately positively correlated ($r = 0.695$, $p = 0.000 < 0.05$). These findings reveal that the students at higher levels of English tend to use more adjective-noun collocations and produce more well-formed collocations but at the same time make relatively more mistakes.

C. Language Proficiency and Collocational Knowledge

The third and the fourth research questions ask about the relationship between language proficiency and receptive collocational knowledge and between language proficiency and productive collocational knowledge. Two Pearson product-moment correlations were run with the results shown below.

TABLE VII
DESCRIPTIVE STATISTICS OF THE QUICK PLACEMENT TEST AND THE COLLOCATION TESTS

	Mean	Std. Deviation	N
Quick Placement Test	43.86	6.351	216
Receptive Knowledge Test	6.29	1.792	216
Productive Knowledge Test	7.23	2.347	216

TABLE VIII
RESULTS OF THE PEARSON PRODUCT-MOMENT CORRELATIONS FOR THE SCORES OF THE QUICK PLACEMENT TEST AND THE COLLOCATION TESTS

Quick Placement Test	Pearson Correlation	Quick Placement Test	Receptive Knowledge Test	Productive Knowledge Test
	Sig. (2-tailed)	1	.646**	.772**
	N	216	216	216
Receptive Knowledge Test	Pearson Correlation	.646**	1	.630**
	Sig. (2-tailed)	.000		.000
	N	216	216	216
Productive Knowledge Test	Pearson Correlation	.772**	.630**	1
	Sig. (2-tailed)	.000	.000	
	N	216	216	216

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

Table VIII reveals that there was a strong, positive correlation between the mean score of the Quick Placement Test and that of the Receptive Knowledge Test, which was statistically significant ($r = 0.646$, $p = 0.00 < 0.05$). Similarly, the score of the Quick Placement Test and that of the Productive Knowledge Test had a statistically significant linear relationship ($r = 0.772$, $p = 0.00 < 0.05$), and the magnitude of the association was also strong ($0.59 < r < 0.80$). Thus, it can be inferred that there was a close relationship between learners' levels of English proficiency and their collocational knowledge, which means increases in English proficiency were consistently correlated with increases in their knowledge of English adjective-noun collocations. In other words, when a learner's general knowledge and skills in English improve, his/her knowledge and use of collocations will increase as well. This finding is consistent with previous results (Namvar, 2012; Men, 2018), supporting the view that collocational knowledge is an important indicator of language proficiency.

In order to answer the last research question, one-way ANOVA and Tukey post hoc tests were conducted to find out whether there was any statistically significant difference in Vietnamese learners' knowledge of adjective-noun collocations between different levels of English proficiency.

TABLE IX
DESCRIPTIVE STATISTICS ON THE SCORES OF THE STUDENTS AT THREE ENGLISH LEVELS IN THE COLLOCATION TESTS

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	
						Lower Bound	Upper Bound			
Receptive Knowledge Test		B1	62	4.53	1.067	.135	4.26	4.80	2	8
		B2	78	6.81	1.571	.178	6.45	7.16	2	10
		C1	76	7.18	1.449	.166	6.85	7.52	4	10
		Total	216	6.29	1.792	.122	6.05	6.53	2	10
Productive Knowledge Test	Well-formed collocations	B1	62	4.74	1.492	.190	4.36	5.12	2	9
		B2	78	7.46	1.535	.174	7.12	7.81	4	11
		C1	76	9.03	1.766	.203	8.62	9.43	6	14
		Total	216	7.23	2.347	.160	6.92	7.55	2	14
	Deviant collocations	B1	62	12.79	2.490	.316	12.16	13.42	7	17
		B2	78	13.94	1.630	.185	13.57	14.30	11	19
		C1	76	14.20	1.833	.210	13.78	14.62	9	18
		Total	216	13.70	2.057	.140	13.42	13.97	7	19

TABLE X
RESULTS OF ONE-WAY ANOVA FOR THE SCORES OF THE STUDENTS AT DIFFERENT LEVELS OF ENGLISH IN THE COLLOCATION TESTS

		Sum of Squares	df	Mean Square	F	Sig.	
Receptive Knowledge Test		Between Groups	273.232	2	136.616	69.787	.000
		Within Groups	416.972	213	1.958		
		Total	690.204	215			
Productive Knowledge Test	Well-formed collocations	Between Groups	633.223	2	316.611	122.347	.000
		Within Groups	551.203	213	2.588		
		Total	1184.426	215			
	Deviant collocations	Between Groups	74.447	2	37.223	9.495	.000
		Within Groups	834.993	213	3.920		
		Total	909.440	215			

TABLE XI
MULTIPLE COMPARISONS BETWEEN DIFFERENT LANGUAGE LEVELS IN THE NUMBER OF WELL-FORMED AND DEVIANT COLLOCATIONS
Tukey HSD

Dependent Variable		(I) Level	(J) Level	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Receptive Knowledge Test		B1	B2	-2.275*	.238	.000	-2.84	-1.71
			C1	-2.652*	.239	.000	-3.22	-2.09
		B2	B1	2.275*	.238	.000	1.71	2.84
			C1	-.377	.226	.219	-.91	.16
		C1	B1	2.652*	.239	.000	2.09	3.22
			B2	.377	.226	.219	-.16	.91
Productive Knowledge Test	Well-formed collocations	B1	B2	-2.720*	.274	.000	-3.37	-2.07
			C1	-4.284*	.275	.000	-4.93	-3.63
		B2	B1	2.720*	.274	.000	2.07	3.37
			C1	-1.565*	.259	.000	-2.18	-.95
		C1	B1	4.284*	.275	.000	3.63	4.93
			B2	1.565*	.259	.000	.95	2.18
	Deviant collocations	B1	B2	-1.146*	.337	.002	-1.94	-.35
			C1	-1.407*	.339	.000	-2.21	-.61
		B2	B1	1.146*	.337	.002	.35	1.94
			C1	-.261	.319	.691	-1.01	.49
		C1	B1	1.407*	.339	.000	.61	2.21
			B2	.261	.319	.691	-.49	1.01

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

As shown in Table X, a significant difference was found in the average numbers of the well-formed collocations produced by the students in each pair of English levels ($p < 0.05$). As regards the number of deviant collocations, only the results of the students at B2 level ($M = 13.94$, $SD = 1.630$) and C1 level ($M = 14.20$, $SD = 1.833$) were not significantly different ($p = 0.691$), which is similar to the results of the Receptive Knowledge Test. These data suggest that there is generally an increase in the students' reception and production of adjective-noun collocations with rising L2 proficiency; however, higher English proficiency does not guarantee a significant decrease in deviant collocations, and the increase in their collocational knowledge appears to be slower when they reach an advanced level.

The findings in this study indicate that there was no significant error decrease when comparing the results of the students at B2 and C1 levels in terms of both reception and production of adjective-noun collocations. This is in good agreement with the finding of Nguyen and Webb (2016) about a decline in the development of collocational knowledge of students at high levels of language proficiency. It is also unexpected that the number of erroneous adjective-noun collocations produced by B2 and C1 learners in this study turned out to be far greater than those at a lower level (Table 4). This concurs well with Laufer and Waldman (2011)'s finding when they investigated the relationship between language proficiency and verb-noun collocations. Their study revealed that the advanced and the intermediate learners produced significantly more incorrect verb-noun collocations than those at the basic level. The reason for such a decrease in the development of collocational knowledge and an increase in incorrect collocations produced by students at higher English levels may lie in what Men (2018) called as "collocation lag". Specifically, collocational knowledge lags far behind the development of grammar and vocabulary. It appears that the increase in single-word vocabulary does not ensure the accuracy of word combinations. Take the target collocation "răng giả" (*false teeth*) in the Productive Knowledge Test as an example. Several students at the intermediate level did not give any answer to this test item or gave a wrong collocations (e.g. **not real teeth* or **unreal teeth*), the components of which are mostly from A1 to B1 level, whereas the students at B2 or C1 level provided more correct answers (i.e. *false teeth*), but at the same time introduced several erroneous collocations (e.g. **fake teeth*, **man-made teeth*, **artificial teeth*, **counterfeit teeth*, **ceramic teeth*, etc.). As can be seen in Table IV, the students at higher levels tend to make more attempts to produce different collocations for each test item possibly due to their confidence in their language proficiency and their wide range of single-word vocabulary that allows them to use a variety of synonyms to create several expressions with the same meaning, though many of them are incorrect combinations.

According to Gitsaki (1999), adjective-noun collocations are an easy and early acquired type of collocations. Men (2018) also proved that Chinese students did not find this type of collocation particularly problematic. However, the data in this study proved that the students, even those at high levels of English are very likely to encounter great difficulties in using adjective-noun collocations. The mean scores of the Receptive Knowledge Test and the Productive Knowledge Tests were well below half of the total scores. The cause for such a sharp contrast in the research results might lie in the source of data. In order to collect well-formed and deviant collocations, Men (2018) analysed several writings by Chinese learners at different English levels and pointed out that there were not so many instances of erroneous adjective-noun collocations in the corpus. However, the relatively small number of incorrect collocations of this type does not necessarily mean that the students do not struggle to use them. They may simply avoid using the expressions or word combinations which they are unsure of in order to feel safe from errors by using more simple phrases, avoiding the topic or concept that pose difficulties or even leaving the message unfinished. This phenomenon was termed as "avoidance behavior" or "avoidance strategy" in second language acquisition (Schachter, 1974;

Kleinmann, 1977; Brown, 2000). Unlike Men (2018), this study collected the data from the two collocations tests which required the respondents to identify the correct collocations as well as the incorrect collocations and produce the adjective-noun collocations to complete the sentences, thereby minimizing their use of avoidance strategies. It is in these tests that the students' collocational errors were exposed and analyzed. Thanks to the results of the collocation test, this study has provided a revealing insight into the Vietnamese learners' knowledge of adjective-noun collocations.

V. CONCLUSIONS

This article investigates Vietnamese students' knowledge of adjective-noun collocations in terms of both reception and production and explored the relationship between knowledge of collocations and language proficiency among English-majors at a university in Vietnam. The results show that the students, even those at advanced levels, had a very limited knowledge of adjective-noun collocations. It is noteworthy that the students had greater difficulties in identifying erroneous collocations than recognizing the correct combinations. One interesting finding of the study is that the students at higher levels of language proficiency were more likely to produce deviant adjective-noun collocations, for they tended to make more attempts to use collocations and they tried to combine unfamiliar words to create new expressions. Besides, this study confirms the results of prior research suggesting that language ability is a good predictor of collocational knowledge. It is also important to note that the increase in the students' collocational knowledge appears to be slower when they reach an advanced level.

Due to the importance of adjective-noun collocations in second language acquisition and the difficulties they cause for EFL learners irrespective of their levels, this type of collocations deserves more attention from researchers, teachers and learners. It is advisable that teachers raise the students' awareness of the significance of collocations by instructing them to discover the effectiveness of collocation use in both spoken and written discourse. Teachers also need to teach the students how to search for well-formed collocations by using the collocation dictionaries and corpuses. Their continual encouragement is also of great importance in helping the students develop the habit of using collocations in both speaking and writing.

Owing to the constraints of time and resources, this work clearly has some limitations in the sample size, the sampling method as well as the number of adjective-noun collocations included in the tests. It is hoped that the results of this study can be further validated with a larger sample size selected by probability sampling for stronger generalizability. The findings of this study also suggest the following directions for future research: (1) analyzing errors related to English adjective-noun collocations, (2) exploring the role of L1 in collocational development, and (3) examining the development of collocations over time.

REFERENCES

- [1] Benson, M., Benson, E., & Ilson, R. (2010). *The BBI combinatory dictionary of English: Your guide to collocations and grammar* (3rd ed.). Amsterdam: Benjamins.
- [2] Brown, H. (2000). *Principles of language learning and teaching* (pp. 122-130). San Francisco: Longman.
- [3] Bueraheng, N., & Laohawiriyanon, C. (2014). Does learners' degree of exposure to English language influence their collocational knowledge. *Psychology*, 4(3), 1-10.
- [4] Celce-Murcia, M. (2008). Rethinking the role of communicative competence in language teaching. In E. A. Soler & M. P. S. Jordà (Eds.), *Intercultural language use and language learning* (pp. 41-57). Dordrecht: Springer.
- [5] Cowie, A. P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics*, 2(3), 223-235. doi: 10.1093/applin/II.3.223
- [6] Ebrahimi-Bazzaz, F., Abd Samad, A., bin Ismail, I. A., & Noordin, N. (2014). Verb-noun collocation proficiency and academic years. *International Journal of Applied Linguistics and English Literature*, 3(1), 152-162.
- [7] Ebrahimi-Bazzaz, F., Abd Samad, A., bin Ismail, I. A., & Noordin, N. (2015a). Verb-noun collocations in written discourse of Iranian EFL learners. *International journal of Applied Linguistics and English literature*, 4(4), 186-191.
- [8] Ebrahimi-Bazzaz, F., Samad, A. A., bin Ismail, I. A., & Noordin, N. (2015b). Verb-noun collocations in spoken discourse of Iranian EFL learners. *International Journal on Studies in English Language and Literature (IJSELL)*, 3(3), 41-50.
- [9] Firth, J. R. (1957). *Papers in Linguistics 1934-1951* (pp. 190-215). London, England: Oxford University Press.
- [10] Gitsaki, C. (1999). *Second language lexical acquisition: A study of the development of collocational knowledge*. San Francisco: International Scholars Publications.
- [11] Gui, S.C. & Yang, H.Z. (2003). *Chinese learner English corpus*. Shanghai: Shanghai Foreign Language Education Press
- [12] Gyllstad, H. (2007). *Testing English collocations: Developing receptive tests for use with advanced Swedish learners*. Lund University. Retrieved July 03, 2021 from <https://lup.lub.lu.se/search/ws/files/5893676/2172422.pdf>
- [13] Hosseini, S. M. B., & Akabarian, I. (2007). Language proficiency and collocational competence. *Journal of Asia TEFL*, 4(4), 35-58.
- [14] Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19(1), 24-44. doi: 10.1093/applin/19.1.24
- [15] Hong, A. L., Rahim, H. A., Hua, T. K., & Salehuddin, K. (2011). Collocations in Malaysian English learners' writing: A corpus-based error analysis. *3L: The Southeast Asian Journal of English Studies*, 17 (Special Issue), 31-44.
- [16] Kamarudin, R., Abdullah, S., & Aziz, R. A. (2020). Examining ESL learners' knowledge of collocations. *International Journal of Applied Linguistics and English Literature*, 9(1), 1-6.
- [17] Kleinmann, H. H. (1977). Avoidance behavior in adult second language acquisition. *Language learning*, 27(1), 93-107.

- [18] Kurosaki, S. (2012). *An analysis of the knowledge and use of English collocations by French and Japanese learners* (Doctoral dissertation, University of London, Institute in Paris). Retrieved July 01, 2021, from <https://core.ac.uk/download/pdf/28902137.pdf>
- [19] Laufer, B., & Waldman, T. (2011). Verb-Noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), 647–672. doi:10.1111/j.1467-9922.2010.00621.x
- [20] Lewis, M. (2000). *Implementing the Lexical Approach: Putting theory into practice*. Hove: Language Teaching Publications.
- [21] Men, H. (2018). *Vocabulary increase and collocation learning: A corpus-based cross-sectional study of Chinese learners of English*. Singapore: Springer.
- [22] Namvar, F. (2012). The relationship between language proficiency and use of collocation by Iranian EFL students. *3L: The Southeast Asian Journal of English Language Studies*, 18(3), 41-52.
- [23] Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223–242. doi:10.1093/applin/24.2.223
- [24] Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, 21(3), 298-320. doi: 10.1177/1362168816639619
- [25] Shin, D. & Chon, Y. V. (2019). A Multiword Unit Analysis: COCA Multiword Unit List 20 and ColloGram. *The Journal of Asia TEFL*, 16(2), 608-623. doi: 10.18823/asiatefl.2019.16.2.11.608
- [26] Schachter, J. (1974). An error in error analysis. *Language learning*, 24(2), 205-214.
- [27] Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- [28] Talakoob, F., & Koosha, M. (2017). Productive and receptive collocational knowledge of Iranian EFL learners at different proficiency levels. *International Journal of Applied Linguistics and English Literature*, 6(7), 11-16.
- [29] Wolter, B. & Yamashita, J. (2015). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics* 36, 1193-1221. doi:10.1017/S0142716414000113
- [30] Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- [31] Yan, H. (2010). Study on the causes and countermeasures of the lexical collocation mistakes in College English. *English Language Teaching*, 3(1). doi:10.5539/elt.v3n1p162

Thai Bao Ngoc Pham received a Master's degree in TESOL from the University of Social Sciences and Humanities, Ho Chi Minh City, Viet Nam in 2016 and is currently a lecturer at the Faculty of English Linguistics and Literature at the same university. Her research interests include English language teaching, English linguistics, and comparative linguistics. Her recent publications are *The Application of Conceptual Metaphors to Teaching English Idioms to English-major Students in Viet Nam*, *Theory and Practice in Language Studies*, 9(6), 2019, pp. 610-619, and *Teaching English Idioms of Happiness and Sadness through Conceptual Metaphors in Vietnamese Context*, *Ho Chi Minh City Open University Journal of Science*, 7(1), 2017, pp. 94-102.