

## **The Impact of Teaching Corpus-based Collocation on EFL Learners' Writing Ability**

Shabnam Ashouri, Ph.D Candidate, Islamic Azad University, Tonekabon Branch,  
Tonekabon, Iran

*arat\_8285@yahoo.com*

Davood Mashhadi Heidar, Assistant Professor, Islamic Azad University, Tonekabon Branch,  
Tonekabon, Iran

*davood.m\_tarbiatmodares@yahoo.com*

### **Abstract**

The present study explores the impact of corpus-based collocation instruction on intermediate Iranian EFL learners' writing ability. For this study, 84 Iranian learners, studying English as a foreign language in Bayan Institute, Iran, were selected and were randomly divided into two groups, experimental and control. Conventional methods of writing instruction were taught to the control group while the experimental group received corpus-based collocation instruction in writing essays for 15 sessions. The design of the research was based on pre- and posttest method. The tests were employed to measure the writing fluency of the two groups. Test results were scrutinized to answer a major question for correlation between the participants' variety of corpus-based collocations and their writing. The results illustrated that there was a significant difference between the mean scores of control and experimental groups in writing elements ( $p < 0.05$ ). Actually, lexical collocation instruction expanded the writing elements of vocabulary and mechanics rather than grammar and fluency in writing essays. The study findings indicated that there is a significantly positive correlation between the participants' use of various lexical collocations and their writing proficiency.

**Key words:** lexical collocation, instruction, writing skill, EFL learners, collocation pre-teaching

### **Introduction**

Numerous studies have been done in different areas of language such as grammar, lexis and collocation, but very few of them have explored the correlation between teaching corpus-based collocation and EFL learners' writing proficiency. First Palmer (1933) brought the field of theoretical linguistics up and later Firth (1957) introduced the most common definition of collocation, which is the tendency of words to co-occur with each other in a particular domain. According to McCarthy (1990) Knowledge of collocation is an important part of nativized linguistic competence of speakers, and can make difficulties for learners who are communicating in English. Nattinger and DeCarrico (1992) clarified that language is not individual words but in chunks that may have the form of multi-word items, sentence heads. A decade ago, the very first experimental study on collocations, done by Zhang (1993), fell into the area of EFL/ESL writing. After Zhang several scholars, investigated the relations between collocations and general English proficiency. It is about the right time to initiate looking into how the achievement of collocations may or may not affect each of EFL/ESL learners' language skills, writing, speaking, reading, and listening. Bahns and Eldaw (1993) quoted while students may have learnt a large number of vocabulary, they are still short of knowledge of collocation. Bahns (1993) also stated that lexical items and grammar are indivisible. We all know collocation is vital and recognize that what differentiates native and non-native speakers is collocational proficiency (McCarthy, 2004a; Koya, 2006). Koosha and Jafarpour (2006) claimed that phonological shift is most frequent in Second Language Acquisition. However lexical and collocational transfer looks to be main cause of lack of proficiency. Lewis (2008b) claimed that lexical approach is based on significant chunks that construct continuous coherent text.

### **Literature Review**

There have been a lot of research on the individual vocabulary items, but studies on corpus-based collocations have rarely been conducted. Collocation is not a new concept. Brown (1974) talked about students' obscurity in using correct collocations. Brown draw attention to the complexity of collocating verbs. She also emphasized the significance of giving collocation exercises to the learners and prepares a number of collocation activities for them. According to Marton (1977), collocations did not seem to cause comprehension problems since learners could understand and decode English sentences including collocations, but there was a fact that they could not produce the same collocations while speaking or writing in English revealed that collocations did make problems at the level of production. Channel (1981) declared that vocabulary instruction should include collocations. They believed that collocations are beneficial for learning new words. Starting from the 1990s, series of studies on the necessity of collocations draw the attention of second language researchers and EFL/ESL practitioners. Haswell (1991) claimed that the use of frequent set of expressions reveals a positive skillfulness in writing while the lack of these expressions is a sign of beginner writers. Another reason is that collocations are very hard to learn, especially because they are too many, senseless because nobody knows why some words collocate and others do not and memorizing them all is almost impossible. Empirical studies on collocations did not begin by Zhang (1993) who examined the correlations among the EFL students' knowledge of collocations, use of collocations, and English fluency, testing 60 English speakers at the University of Pennsylvania, with a collocation test to measure the students' knowledge of collocation and a writing task to collect their use of collocations and writing proficiency. Bahns and Eldaw (1993) also claimed in their study that English learners' general vocabulary knowledge exceeds their collocational knowledge.

Lennon (1996) worked on a group of German EFL learners disclosed that even the advanced learners employed high frequency verbs erroneously, which designated their deficiency in collocational knowledge. It was said that we need proficient ways to present vocabulary to EFL learners (Zimmerman, 1997). Another research was conducted later by Al-Zahrani (1998). In this study it was found that native writers performed considerably better than non-native writers in the collocation test and in writing. Al-Zahrani (1998) claimed that there is a significant difference in students' knowledge of lexical collocations among the different academic years. To be more detailed, the knowledge of lexical collocations enhanced along with the subjects' studied. According to Lewis (2000), the problem of not having collocational competence often makes learners form longer utterances since they are lack of knowledge of collocations which convey in particular what they want to say. According to Nation (2001) there is a classification of eight features of word knowledge, which has to comprise knowledge of lexical collocations. According to Sung (2003) there is a significant strong correlation between EFL students' knowledge of lexical collocations and their speaking fluency. According to Siik (2006) some students endure from the lack of knowledge of collocation. Although they may have good knowledge of vocabulary, they still make deviant language. According to Hyland (2008) multi-word structures are elements of fluent linguistic production and main components in language learning. Lewis (2008a) considered that, it is the collection of chunks which shapes the data by which learners initiate to recognize patterns, morphology and other features of linguistic. Ghonsooli, Pishghadam, and Mahjoobi (2008) stated that Iranian EFL learners may have a superior knowledge of English vocabulary and grammar but problems with having the ability to use collocations. McIntosh, Ben and Richard (2009) argued that collocations are the way words come together in a language to create natural sounding communication and writing.

Although a number of studies have been completed in the area of grammar and collocation, a few of them have examined the relationship between teaching lexical

collocation and English learners' writing ability (Ghonsooli et al., 2008; Mounya, 2010). According to Hsu (2010) collocational studies display the effect of lexical collocation instruction on other features of general English skill. The present study endeavors to fill this space, i.e. to examine the impact of corpus-based collocation instruction on developing writing proficiency of EFL learners. Zahedi & Mirzadeh (2010) have highlighted the significance of drawing second language learners' attention to collocations and idioms. The language awareness framework was opted to assist students in learning collocations, since students did seem to have difficulties in not perceiving collocated expressions, and it seemed required to increase this awareness before collocation learning could occur. Hashemi (2012) in his research state that EFL College students, high school students, and professors have little knowledge of collocation because collocation has been neglected in EFL classrooms. He believes that a good way to promote the fundamental mastery of collocation would be to work on an exercise that raises learners' awareness of collocations and then help them with tokens in their output. The present study, therefore, aimed to investigate the effect of pre-teaching vocabulary and collocations on writing development of advanced students. It actually sought to find answer to the following research question:

### **Research Question**

Does corpus-based collocation instruction have any significant effect on EFL learners' writing proficiency in terms of fluency, cohesion, coherence and grammar?

### **Hypothesis**

H<sub>0</sub>: Corpus-based collocation instruction does not have any effect on EFL learners' writing proficiency.

## **Methodology**

### **Participants**

Participants in the study consisted of 84 intermediate students who were studying in Bayan institute for two and a half years. They were selected through administering an OPT exam to get sure of their homogeneity. The participants were randomly assigned to two groups-- experimental and control-- each consisting of 42 students. They had all passed Pre-intermediate courses, studying *American File*.

### **Instruments**

To fulfill the aims of the current study, the following instruments were employed: 1) An OPT exam to ensure the participants' homogeneity in terms of language proficiency level, 2) Two writing tasks to check the participants' progress in writing essays.

### **Procedures**

The current study aimed to scrutinize the impact of the independent variable, i.e. corpus-based collocations instruction, on the participants' writing development. The homogenized participants (through OPT) were therefore first assigned randomly to two groups, experimental and control. A pretest posttest design was employed. The pretest was first administered to both groups to check their sameness in terms of collocation knowledge and general writing ability. Then, the posttest was run to find out the effect of the treatment. The treatment comprised a wide circle of corpus-based collocations to equip the experimental group with relevant background to facilitate their writing. Simultaneously relevant collocations were written on the board and taught to draw their attention to collocations and the restrictions on how words can collocate with each other. To assess the participants' performance in exploiting the input they received, they were required to do some collocation

exercises, prior to giving them some topics to write about. The exercises planned to activate their individual autonomous employment of collocations. They also helped them to explore how they are utilized and put them into practice. Having followed the above-mentioned procedures within 15 sessions, the experimental group was assigned two different writing tasks to see their possibly different performance in comparison with the control group on the same tasks. Finally, the writings of all participants were analyzed in terms of such writing elements as vocabulary, fluency, cohesion, coherence and grammar, as well as the frequency of the learned collocations, new vocabulary, and variety of vocabulary and collocations use. Lastly, two scorers were requested to correct the papers. The results of the statistical analysis of scores are shown in tables of the next section.

### Analysis of the data

**Table1.** Correlations between First and Second Administration of Writing Test

		writing test second administration
writing test first administration	Pearson Correlation	.867 <sup>**</sup>
	Sig. (2-tailed)	.000
	N	15

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

Table 1 above shows the results of the Pearson correlation run for estimating the reliability index of the writing test through test- retest method. The value obtained was .867, which indicates a considerably high correlation between the results of the writing tests on two different administrations.

**Table 2.** Correlations for the raters

		Rater B
Rater A	Pearson Correlation	.601 <sup>*</sup>
	Sig. (2-tailed)	.018
	N	15

As the above table shows, there was a significant correlation between the two raters' scores for the writing tests ( $p .018 \leq 0.05$ ). After examining the correlation between the two raters' scores, the total writing score for each individual was the mean of the scores given by the two raters.

### Results of NELSON Test (sampling phase)

To select the main sample and to make certain that the participants were at similar level of general foreign language proficiency at the beginning of the study, NELSON test was given to the whole population. The test included fifty multiple-choice items with the maximum possible score being 50 points. Eighty-four students whose score fell  $\pm 1SD$  from the mean score were selected as the main sample for the present study. The results of NELSON test for 145 students are presented in the following table:

**Table 3,**Statistics for NELSON Test

NELSON		
N	Valid	145
	Missing	0

Mean	31.0828
Median	29.0000
Mode	19.00
Std. Deviation	10.16564
Variance	103.340
Skewness	.441
Std. Error of Skewness	.201
Kurtosis	-1.079
Std. Error of Kurtosis	.400
Range	35.00
Minimum	15.00
Maximum	50.00
Sum	4507.00

The main participants of the study were selected from among those who scored within the range of  $\pm 1SD$  from the mean score (31.0828+ 10.16564, 31.0828- 10.16564). Therefore, a cut-point of (20.91716) to (41.24844) was set and 84 learners (two groups each consisting of 42 students) whose general foreign language proficiency scores were within this range (20.91716 to 41.24844) were selected as the main sample for the present study and were randomly divided into two groups (one control and one experimental).

#### **Descriptive Statistics of the Pretest Scores of the writing test**

To examine the null hypothesis and to find if corpus- based collocation instruction affects Iranian EFL learners' writing proficiency, a writing pretest was administered to the both groups to examine their possible differences at the beginning of the study. Tables 4 and 5 show the results of an independent samples t- test used to analyze students' scores in the pre-writing test.

Based on Table 4 below, there was no significant difference between the mean scores of the two groups in pretest of writing proficiency ( $p>0.05$ ), that is, the control and experimental groups were almost at the same level of proficiency in terms of their writing ability in the administered writing test at the beginning of the study.

**Table 4, Group Statistics for the Writing Test (Pretest)**

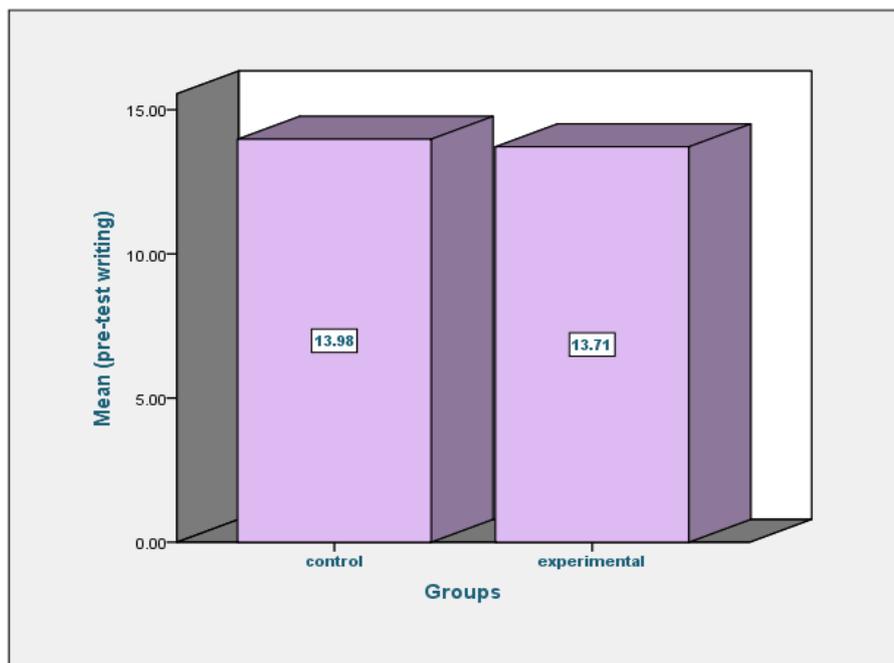
	Groups	N	Mean	Std. Deviation	Std. Error Mean
Pretest writing	Control	42	13.9762	1.78733	.27579
	Experimental	42	13.7143	1.92906	.29766

The above descriptive table illustrates the sample size, mean, standard deviation, and standard error for both control and experimental groups at the beginning of the study. The control and experimental groups mean score were (13.9762) and (13.7143), respectively. In addition, the degree of scatteredness of the scores for the experimental group ( $SD= 1.92906$ ) was somewhat higher than the degree of deviation of scores around the mean score for the control group ( $SD= 1.78733$ ). The two groups differed somewhat around their average score. The mean score of the experimental group was .2619 points higher than that of the control group.

**Table 5.** Independent Samples Test for the writing test (pretest)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Pretest writing	Equal variances assumed	.75	.38	.64	82	.52	.2619	.4057	-.545	1.069
	Equal variances not assumed			.64	81.5	.52	.2619	.4057	-.545	1.069

The independent sample T-test procedure introduced two tests of the difference between the control and experimental groups. Levene statistics above tested the equality of the variances, and it was reported that the significance value was equal to (sig= .38). Since this approximation was higher than (0.05), it could be assumed that the two groups had equal variances and consequently the first test was considered. As the above tables show that there were no significant differences between the group's performances in terms of their writing proficiency at the beginning of the study. The sig. value was equal to (.52) indicating that this mean difference was not significantly important.



**Figure 1,** Comparison between the two groups on writing test (pretest)

## Restatement of the research question and hypothesis

### Research Question

Does corpus-based collocation instruction have any significant effect on EFL learners' writing proficiency in terms of fluency, cohesion, coherence and grammar?

### Hypothesis

H<sub>0</sub>: Corpus-based collocation instruction does not have any effect on EFL learners' writing proficiency.

As far as the research question is concerned, i.e. whether Corpus- based collocation instruction affects the learners' writing proficiency, an independent t-test was run to check the results of the writing posttest and compare the experimental and control groups. Table 6 below show the results:

**Table 6, Group statistics for the writing test (Posttest)**

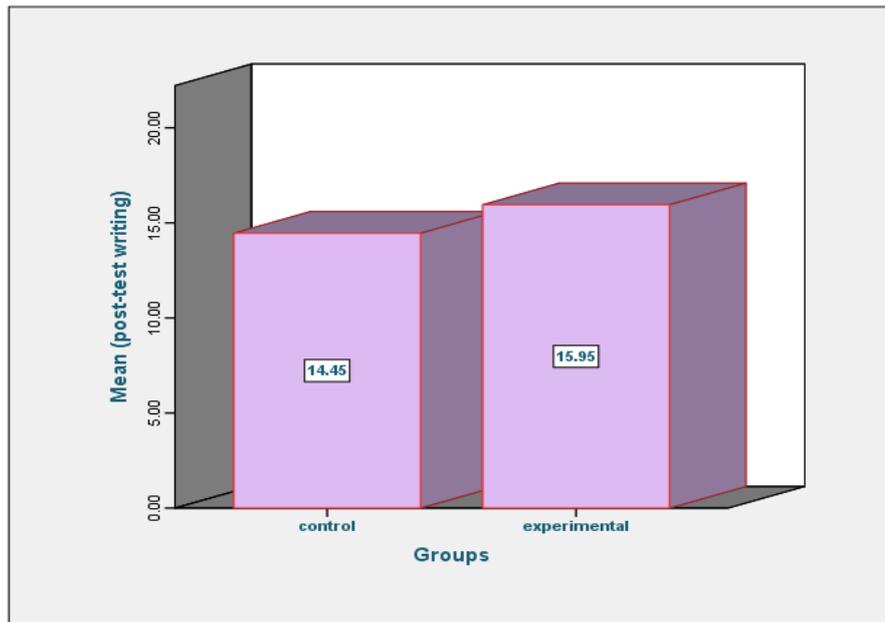
	Groups	N	Mean	Std. Deviation	Std. Error Mean
posttest writing	Control	42	14.4524	1.69939	.26222
	experimental	42	15.9524	1.69580	.26167

Based on the above results, the participants' performance in the experimental group (Mean =15.9524) weighed more than that of the control group (Mean =14.4524) in posttest of writing. In fact, the results revealed that Corpus- based collocation instruction affects the writing proficiency of the two groups differently ( $t=4.04$ ,  $0.00 < .05$ ).

**Table 7, Independent Samples Test for the Writing Test (Posttest)**

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
posttest writing	Equal variances assumed	.90	.34	-4.04	82	.000	-1.50	.370	-2.23	-.76
	Equal variances not assumed			-4.04	82.00	.000	-1.50	.370	-2.23	-.76

Since the participants were homogenous at the beginning of the study and had been randomly assigned to two groups, the difference found in their posttest was not due to chance. Actually, it could be related to the specific treatment of corpus- based collocation instruction on the experimental group. As a result, the null hypothesis was rejected implying that teaching corpus-based collocation affects Iranian EFL learners' writing proficiency.



**Figure 2**, Comparison between the two groups on writing test (posttest)

### Results and discussion

The major purpose of the present research was to inspect the possible effects of corpus-based collocation instruction on Iranian EFL learners' writing proficiency. A total number of 84 homogeneous EFL learners were selected based on their performance on NELSON test. The main sample was selected from among those who scored within the range of  $\pm 1SD$  from the mean score on NELSON test. Then, they were randomly divided into two groups each including forty-two participants (one control and one experimental group). At the beginning of the study, they were given a writing test. The purpose was to ascertain that there was no significant difference between the two groups in terms of their writing proficiency prior to the implementation of the specific treatment. Afterwards, Corpus-based collocation was practiced for the experimental group while traditional teaching was worked on for the control group. The data was collected through the pre- and post- writing tests administered to both groups.

In the analysis phase of this research, the data collected from the pretest and posttest were summarized, and the procedures of descriptive statistics (including frequencies, means, standard deviations, etc.) along with inferential statistics; namely, independent samples T-tests were run. The parametric Independent samples t-test was run in order to find out if there was any significant difference between the control and experimental groups in terms of their vocabulary size and writing proficiency both at the beginning and at the end of the study. Besides, the measure of inter-rater reliability for the two raters of the writing test, which was the Pearson product-moment correlation, was estimated. Before running the main statistical analysis of the present study, normality that is the main assumption of parametric tests was examined for all of the distributions (control group pretests, control group posttests, experimental group pretests, and experimental group posttests). In addition, the reliability of the tools employed in the study was estimated through a pilot study on 15 EFL students who were representative of the main group in terms of their general language proficiency.

Finally, the findings of this study demonstrated that there was a significant difference between the scores of the participants in the control and experimental groups. Actually, the experimental group has become aware of the fact that using corpus-based collocations, can advance their writing and make it more natural. Corpus-based collocation instruction can, in

fact, considerably influence EFL intermediate language learners' writing ability. It can thus be asserted that this type of instruction should be included in the writing syllabus while teaching writing at English institutes so that collocational competence could be improved and result in an increase in students' communicative competence. The study can be wider in scope, to include all the possible corpus-based collocations, i.e. going further than the two- word collocations, so that the results could represent the learners' knowledge of collocations further. Another area of development over the current study could be the choice of texts in which the phrases could be located. Further research might also scrutinize grammatical collocations and their possible effects in improving written fluency.

### References

- Al-Zahrani, M. S. (1998). *Knowledge of English lexical collocations among male Saudi college students majoring in English at a Saudi university*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Pennsylvania.
- Bahns, J. (1993). Lexical collocations - a contrastive view. *ELT Journal*, 47(1),56-63.
- Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21(1), 101-114.
- Brown, D. (1974). Advanced vocabulary teaching: The problem of collocation. *RELC Journal*, 5(2), 1 – 11.
- Channell, J. (1981). Applying semantic theory to vocabulary teaching. *English Language Teaching Journal*, 35, 115-122.
- Firth, J. R. (1957). Modes of meaning. In J. R. Firth, *Papers in linguistics 1934 -1951* (pp. 190-215). Oxford: Oxford University Press.
- Ghonsooli, B, Pishghadam, R, & Mahjoobi, F. (2008). The impact of collocational instruction on the writing skill of Iranian EFL learners: A case of product and process study. *Iranian EFL Journal*, 2, 36-59.
- Hashemi, M., Azizinezhad, M. & Dravishi, S. 2012. Collocation a neglected aspect in teaching and learning EFL. *Procedia-Social and Behavioral Sciences* 31: 522-525.
- Haswell, R. (1991). *Gaining ground in college writing: Tales of development and interpretation*. Dallas: Southern Methodist University Press.
- Heaton, J. B. (1990). *Writing English language tests*. New York: Longman Group UK Limited.
- Hsu, J-Y. (2010). The effect of collocation instruction on reading comprehension and vocabulary learning of Taiwanese college English majors. *Asian EFL Journal*, 12(1), 47-87.
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27, 4-21.
- Koosha, M., & Jafarpour, A. A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL Journal*, 8(8), 1-14.
- Koya, T. (2006). What is the reality of collocation use by native speakers of English? *Dialogue*, 5: 1-18.
- Lennon, P. (1996). Getting 'easy' verbs wrong at the advanced level. *IRAL*, 34(1), 23-36.
- Lewis, Michael. (2000). Learning in the lexical approach. In Michael Lewis (Ed.), *Teaching collocation* (pp. 155-184). Hove, England: Language Teaching Publications.
- Lewis, Morgan. (2000). There is nothing as practical as a good theory. In Michael Lewis (Ed.), *Teaching collocation* (pp. 10-27). Hove, England: Language Teaching Publications.
- Lewis, M. (2008a). *The lexical approach: Putting theory into practice*. England: Language Teaching Publication.

- Lewis, M. (2008b). *Implementing the lexical approach: Putting theory into practice*. England: Language Teaching Publication
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- McCarthy, M. J. (2004a, August). *Collocation in vocabulary teaching and learning*. Lecture given at the meeting of JACET summer seminar program, Gunma, Japan.
- McIntosh, Colin, Francis, B., & Poole, R. (2009). *Oxford collocations dictionary for students of English (OCD)*. Oxford: Oxford University Press.
- Marton, W. (1977). Foreign vocabulary learning as problem no. 1 of language teaching at the advanced level. *Interlanguage Studies Bulletin*, 2(1),33-57.
- Mounya, A. (2010). *Teaching lexical collocations to raise proficiency in foreign language writing*. Unpublished MA Thesis, Guelma University, Guelma, Algeria.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Palmer, H. E. (1933). *Second interim report on English collocations*. Tokyo: Kaitakusha.
- Siik, S. (2006). The teaching of lexical collocations and its effects on the quality of essays and knowledge of collocations among students of program Persedian Ijazah Sarjana Muda of Instut Perguruan Batu Lintang, Kuching. Unpublished MA Thesis, Faculty of Education University Teknologi Malasia, Malasia.
- Sung, J. (2003). *English lexical collocations and their relation to spoken fluency of adult non-native speakers*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Pennsylvania.
- Zahedi, H., & Mirzadeh, M. (2010). Collocational cloze: The effect of deletion procedure and gender. *Iranian EFL Journal*, 6(1), 143-157.
- Zhang, X. (1993). *English collocations and their effect on the writing of native and non-native college freshmen*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Pennsylvania.
- Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 5-19). Cambridge: Cambridge University Press.