A Comparative Study of the Use of Collocation in Iranian High School Textbooks and American English File Books

Mohsen Shahrokhi

English Department, Shahreza Branch, Islamic Azad University

Isfahan, Iran

Tel: +98-3213292013 E-mail: shahrokhi1651@yahoo.com

Samira Moradmand (Corresponding author)

English Department, Shahreza Branch, Islamic Azad University

Isfahan, Iran

Tel: +98-9132295915 E-mail: sa_moradmand2007@yahoo.com

Abstract

The present study investigates the extent to which lexical and grammatical collocations are used in Iranian high school English textbooks, compared with the American English File books. To achieve the purposes of this study, this study had to be carried out in two phases. In the first phase, the content of the instructional textbooks, that is, American English File book series, Book 2 and Iranian high school English Book 3, were analyzed to find the frequencies and proportions of the collocations used in the textbooks. Since the instructional textbooks used in the two teaching environments (i.e., Iranian high schools and language institutes) were not equal with regard to the density of texts, from each textbook just the first 6000 words, content words as well as function words, were considered. Then, the frequencies of the collocations among the first 6000 words in high school English Book 3 and American English File Book 2 were determined. The results of the statistical analyses revealed that the two text book series differ marginally in terms of frequency and type of collocations. Major difference existed between them when it came to lexical collocations in American English File book 2.

Keywords: Grammatical and lexical collocation, Textbook, Public school, Language institute.

1. Introduction

After grammar to which a lot of attention has been paid by different teaching approaches, the other language area which has been considered prominently during the last decades is vocabulary learning and teaching. This reorientation from grammar to vocabulary is definitely a right step which has been taken in the realm of language teaching and learning. However, there is one crucial aspect of vocabulary learning and teaching which deserves attention receiving even more than the amount it has received so far, and this is the problem of formulaic language and in fact word combinations (Bahn, 1991). Rudzka, Channel, Putseys, and Ostyn (1981, p. 5 cited in Bahn, 1993) maintain that, "one of the main difficulties students encounter in relation to new items of vocabulary is knowing what their collocational properties are". It can be claimed that in some languages (like Persian and English), words do not co-occur completely freely; instead, they follow different levels of limitation in their co-occurrence (Allerton, 1984 cited in Bahn, 1991) and, therefore, learners in these languages have some restrictions in combining words and they should be aware of them. Conklin and Schmitt (2007) referred to word combination and formulaic language as one entity, but Benson, Benson and Ilson (1986) mentioned that formulaic language has different categories namely: free word combination, idioms and collocation. Each one of these categories is discussed below. As it is inferred from the phrase 'free word combination' it can be understood that in free word combination, the meaning of a whole is obtained by summing the meaning of the constituent words (Benson et al., 1986). Of course even in such kinds of combinations it is argued that there are some semantic constraints for compatibility (Gelbukh, Sidorov, Han & Rubio, 2004). This means that for example, for the infinitive 'to sew' it is expected that its object will be a kind of clothing. Moreover, Gelbukh et al. (2004) mentioned that compared to other words, some words may have a broader compatibility, for example, 'to see' can be combined with practically any physical object; while 'to read' only with something that contains written material. Finally, they summarized that some combination of words have high conditional probability, and therefore when one encounters one of the words in the text, then the probability to encounter the other word is relatively high. On the other hand, they discussed most of randomly chosen word pairs with low conditional probabilities which are called mutual information. The definition of collocation mentioned by Firth (1951) was as the company which a word keeps. Therefore, according to Gabriella (1994), in this way Firth distinguished "meaning by collocation" from both the "conceptual" and the "contextual" meaning of words .

As mentioned earlier, different scholars treated collocations differently. For example for Benson et al. (1986) collocations are transitional between idioms and free combinations but Cowie and Howarth (1996) introduced a model

based on which there can be a continuum containing four categories of collocations:

- Free combinations: whose meaning is inferred from the dictionary meaning of individual components (open the door).
- Restricted collocations: which are more restricted in the choice of compositional elements and these collocations usually have one component used in a particular context (meet the demand).
- Figurative idioms: which have a metaphorical meaning as a whole that can tell its literal interpretation to some extent (call the shot).
- Pure idioms: which are single units whose meanings are completely unpredictable from the meaning of its components (hit the hays).

In their study, Alemi and Mesbah (2013) aimed to evaluate Top Notch series, which was widely used in EFL contexts. Fifty EFL teachers evaluated the series based on ACTFL standards by Cisar's (2000) checklist. The findings indicated that the series enjoyed some benefits for language learners such as encouraging the students to communicate successfully by offering lots of opportunities for interaction, and demonstrating culture-based aspects through lively and authentic unbiased visual images. The series also let students connect their foreign language learning with other disciplines such as Medical Science, Art, Geography, etc. Moreover, the series offered vast opportunities for EFL learners to compare their own language with English as a foreign language.

Peacock (2012) described a corpus-based analysis of the distribution of the high frequency collocates of abstract nouns in 320 research articles across eight disciplines: Chemistry, Computer Science, Materials Science, Neuroscience, Economics, Language and Linguistics, Management, and Psychology. Conclusions revealed that there were disciplinary norms and some standard terminology within the disciplines arising from the topics discussed, research methods, and content of discussions. It was also concluded that the collocations were important parts of the meanings and functions of the nouns, and that this evidence of sharp discipline differences underlined the importance of discipline-specific collocation research.

Finally, it can be concluded that the main characteristics of collocations are that their meanings reflect the meaning of their constituent parts (in contrast to idioms), they are used frequently, they spring to mind readily, and they are psychologically salient (in contrast to free combinations) (Bahn, 1991).

2. Statement of the problem

EFL learners' knowledge of collocations has been investigated in several studies (Bahns & Eldaw, 1993; Koya, 2005; Shehata, 2008; Shei & Pain, 2000). Lots of these studies have revealed lack of collocation knowledge among EFL learners and the results of these studies have confirmed that collocations are still controversial issues for language learners especially in EFL settings. In EFL settings, Learners are mostly exposed to individual words rather than collocations and therefore they usually find it difficult to encounter collocations in different settings (Farghal & Obiedant, 1995). Obviously, it is necessary to understand collocations both for language production and comprehension (McIntosh et al., 2009), but unfortunately according to the results of lots of studies learners' collocation proficiency is restricted and they have lots of problems producing and comprehending collocations (Farghal & Obiedat, 1995; Koosha & Jafarpour, 2006; Shokouhi & Mirsalari, 2010; Zareie & Koosha, 2002). As EFL teachers and learners, we all have frequently observed students' collocation errors after leaving high school and even institutes. So as a learner and a teacher of English in an EFL environment (Iran), the researchers, have come to this point that perhaps the collocational content of the text books is not enough or satisfying. The researchers believe that it is very crucial to find a suitable way to solve the learners' problems with collocations because collocation proficiency is very important for a person who is learning a second language, as it increases native likeness and decreases foreignness. Since collocation learning and teaching are very important and they occupy a noticeable part of language learning and proficiency, this study is conducted to evaluate the collocation content of English text books. As part of this study, the collocation contents in Iranian high school and American English File textbooks are evaluated. In this study, the aim is to consider the type (lexical or grammatical) and the amount of collocations used in high school textbooks and American English Files Series.

3. Purpose of the study

According to the problem stated above, the purposes of this study are to evaluate Iranian high school English textbooks with regard to their type and amount of collocations available in the content, and to compare these books with American English File textbooks that have been developed by native authors of English. To this end, three research questions are provided.

4. Research questions and hypotheses

There are three research questions supporting the objectives of the present study:

- To what extent are the lexical and grammatical collocations used in Iranian high school English textbook 3, and American English File Book 2?
- Is there a significant difference in the frequency of lexical and grammatical collocations between Iranian high school English textbook 3, on the one hand, and American English File Book 2, on the other hand?
- Do the EFL learners who study high school English textbook 3 and those who study the American English File Book 2 differ significantly in their performance on collocation tests?

This research focuses on the following two research null hypotheses:

H01: There is no significant difference in the frequency of lexical and grammatical collocations between Iranian high school English Textbook 3, on the one hand, and American English File Book 2, on the other hand.

H02: There is no significant difference between the performance of the EFL learners who study Iranian high school English Textbook 3 and those who study the American English File Book 2 on collocation tests. It is hoped that the outcomes of our study can shed light on collocation pedagogy and presentation among Iranian language learners, teachers, material developers and curriculum designers.

5. Methodology

5.1 Participants

In order to select the required participants at lower-intermediate level of proficiency, two groups of students (totally 120, 72 from high school and 48 from institute) were tested through administering the Quick Placement Test (Oxford University Press and University of Cambridge Local Examinations Syndicate), one group high school students and the other group language institute students. Then, 60 of those who scored between 28 and 36 (according to the test level chart) were selected, 30 from high school and 30 from institute. They were all females, and their age range was between 16 and 23, with the mean age of 16.82.

5.2 Instruments and Materials

The Instruments of this research included the Quick Placement Test (QPT), and the collocation pretest/posttest. The QPT was a multiple-choice test with 50 items testing situations (five items), cloze passages (25 items), and grammar and vocabulary (20 items). Of course, there were 10 more items for the advanced level which were not used in this study because the participants of this study were up to lower-intermediate level.

As for the collocation pretest/posttest, it was prepared by the researchers, and included 30 multiple-choice items. The test was administered once at the beginning and once at the end of the semester. The test incorporated both types of grammatical collocations and lexical collocations. After consulting two university professors about the test items and making the necessary revisions, the test was piloted on 20 participants (10 from high school and 10 from institute) who were similar to the main participants of this study to check the wordings, instructions, timing, and scoring procedure. After the results of the pilot test were determined, the researchers determined the reliability of the test by using Cronbach's alpha, and it turned out to be .794, which was high enough to claim the test was reliable.

The material of the study included high school English textbooks and American English Files. Regarding the high school English textbooks, it was decided to investigate English Book 3 which has more reading texts, and, therefore, more collocations could be found. As for the American English File book series, Book 2 was selected because it corresponds to high school English Book 3. Moreover, the scores obtained through the administration of collocation pre- and post-test from among the participants of the study set up another part of the material to be analyzed for the purpose of answering the third research questions.

5.3 Procedure

With regard to the nature of the study and the research questions, this study had to be carried out in two phases. In the first phase, the content of the instructional textbooks, that is, American English File book series, Book 2 and high school English Books 3, were analyzed to find the frequencies and proportions of the collocations used in the textbooks. Since the instructional textbooks used in the two teaching environments (i.e., high schools and language institutes) were not equal with regard to the density of texts, from each textbook just the first 6000 words, content words as well as function words, were considered. Then, the frequencies of the collocations among the first 6000 words in high school English Books 3 and American English Files Book 2 were determined.

In order to analyze the content of the textbooks, De Carrico's (2001, cited in Celce-Murcia, 2001) classification of grammatical and lexical collocations was used. Six major types of grammatical collocations were considered for the analysis, designated here by G1 through G5, as follows:

- G1. Noun+ to (infinitive): e.g., a pleasure to do;
- G2. Preposition + noun combinations: e.g., by accident;
- G3. Adjective + preposition combinations: e.g., angry at everyone;
- G4. Predicate adjective + to (infinitive): e.g., It was necessary to work;
- G5. Adjective + that clause: e.g., afraid that;

Moreover, five combinations of lexical collocations were considered in this study, which are designated by G1 through G5, as follows:

- G1. Verb (creation/activation) + noun: e.g., She does the laundry every day;
- G2. Adjective + noun/noun + noun: e.g., The room has a sour smell;
- G3. Noun (unit) + of + noun: e.g., David gave Elisa a bouquet of flowers;
- G4. Adverb + adjective: e.g., They are closely acquainted;
- G5. Verb + adverb: e.g., They argued heatedly in that debate (Benson et al., 1997).

To collect the targeted collocations in the textbooks under investigation, two research assistants who hold BA in TEFL were trained as two raters to identify both types of collocations. Once the research assistants were briefed enough and

the researchers made sure the research assistants are proficient enough to do the task, they analyzed the textbooks manually and identified instances of both grammatical and lexical collocation. The raters were asked to annotate the type of identified collocations. In order to check the inter-rater reliability of the identification of collocation types, the researchers checked the accuracy of raters' identifications. The inter-rater reliability was done first to find similar identifications, and second to find instances of difference between the raters' collocation type identifications. There were few differences between the raters' collocation type identifications that were further discussed by the researchers and raters and a consensus was made among the researchers and raters on the type of collocation.

For the second stage, 120 students were tested using QPT and 60 students (30 from high school and 30 from institute) who scored between 28 and 36 on the test were selected as the lower-intermediate students to take part in the study. Then, at the beginning of the semester, the collocation pretest was administered to both groups of participants to make sure that they were not familiar with the collocations which were to be taught during the semester. After that, the books were taught to the students. The high school students had 5 hours of English class a week and the institute students had three hours of English class in two sessions every week. Finally, at the end of the semester, the same collocation test which was given to the participants in the beginning was administered as the posttest to find out if there were any differences between the two groups with regard to collocations. Then the collected data was submitted to SPSS version 16 for further analysis.

5.4 Data Analysis

In order to analyze the collected data, both descriptive and inferential statistics were incorporated in this study through the use of Statistical Package for Social Sciences (SPSS). For the first research question, the proportion of the collocations used in each textbook to the total number of words in that book was determined. For the second research question, Chi-square test was employed. And for the third research question, t-test was decided to be the best test to compare the two groups of students. All these tests are accompanied by the related tables and graphs wherever deemed necessary.

6. Results

- Investigating the First Research Question

In order to answer the first research question of the study the number of the lexical as well as the grammatical collocations in the two textbooks was determined. Table 1 presents the frequency and the percentage of each type of collocation for each textbook separately (in the first 6000 words), and Figure 1 shows the frequencies graphically.

Table 1. The Frequency and the Percentage of Collocations for Each Textbook

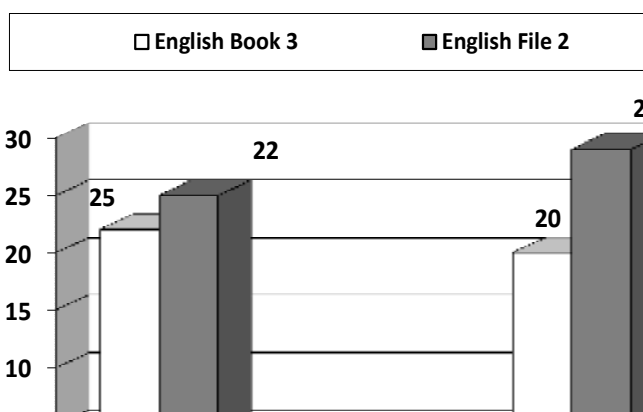| Book	Grammatical		Lexical		Total
	Frequency	Percent	Frequency	Percent	
English Book 3	22	.58	20	.53	42
English File 2	25	.74	29	.91	54

Figure 1. The graphical representation of the frequency of the collocations

According to the data in Table 1, the number of grammatical collocations in English Book 3 was 22 (about 47.6% of the total grammatical collocations) and that of American English File Book 2 was 25 (about 46.3% of the total grammatical collocations). As the figures in the table suggest, in English Book 3 the percentage of the grammatical collocations in the first 6000 words was .58 and in American English File Book 2 was .74, which means that the percentage of grammatical colocations used in American English File Book 2 was higher than the percentage of the grammatical collocations in English Book 3.

According to the same table, that is, Table 1, the number of lexical collocations in English Book 3 was 20 (about 52.4% of the total lexical collocations) and that of American English File Book 2 was 29 (about 53.7% of the total lexical collocations). As the figures in the table suggest, in English Book 3 the percentage of the lexical collocations in the first 6000 words was .53 and in American English File Book 2 was .91, which means, once again, that the percentage of lexical colocations used in American English File Book 2 was higher than the percentage of the lexical collocations in English Book 3.

- Investigating the Second Research Question

The second research question was raised to find out if the frequency of the collocations in the two books under investigation differed significantly. For this purpose, Chi-square calculation was employed to determine the issue. Table 2 depicts the results of Chi-square.

Table 2. The Results of the Chi-square Analysis

Books	χ^2	df	Sig
English Book 3 vs. English File Book 2	.350	1	.554

As it can be observed in the above table, that is, Table 2, the amount of chi-square is not statistically significant ($X^2 = .350$, $p > .05$). In other words, the difference between the frequency of the lexical and grammatical collocations in the two books cannot be considered significant. Therefore, the first null hypothesis which states that "there is no significant difference in the frequency of lexical and grammatical collocations between high school English Textbook 3, on the one hand, and American English File Book 2, on the other hand" can be retained.

- Investigating the Third Research Question

To answer the third research question, the performances of the two groups (that is, high school students and institute EFL learners) on the test of collocations were compared to see if there were any differences between the two groups. Table 3 reports the descriptive statistics for this comparison, and Figure 2 illustrates the means graphically.

Table 3. Descriptive Statistics for Research Question Three

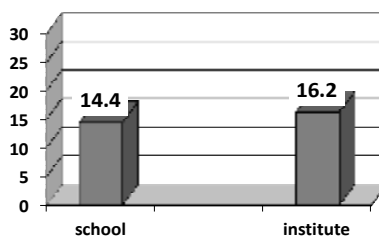|	N	Min	Max	Mean	SD	SEM
School	30	11	18	14.40	1.632	.298
Institute	30	13	23	16.20	2.203	.402

Figure 2. Graphical representation of the means for research question three

Figure 2 and Table 3 reveal that there is a difference between the two groups of students, but it does not tell whether or not the difference is statistically significant. An independent-sample t-test was employed in order to find this out. Table 4 presents the results of this t-test.

Table 4. The Results of the Independent-sample t-test for Research Question Three

Mean Difference	Sig.	df	t
1.80	.001	58	-3.596

According to the figures in Table 4, the amount of t-observed ($t_{(55)}$= 3.596) is high enough to be statistically significant (p= .001), and since the amount of t-observed is negative, it can be said that the institute participants answered questions related to collocations better than the high school participants did. Therefore, the second null hypothesis which states that "there is no significant difference between the performance of the EFL learners who study high school English Textbook 3 and those who study the American English File Book 2 on collocation tests" can safely be rejected. In other words, the institute participants outperformed the high school participants on the test of collocations.

7. Discussion

The findings of this study showed that both grammatical and lexical collocations were presented more in American English File Book 2 than high school English Text Book 3, and the participants from institute outperformed the participants from high school in collocation test. Ahangari and ZununiVahed (2013) aimed at comparing two sets of language books namely, Interchanges books and Iranian high school books in order to investigate the types of collocations used in each and the differences between the two books regarding these types. In order to investigate the collocation types employed in Interchange series and English books for high school students, the four books of Interchange books and the four high school English books were studied in depth and phrases or the word combinations were written down according to what collocation type they were. Then, the collocations used in the eight books were organized into six types according to their classification by McCarthy (2005). As the results revealed, the four Interchange books were abounded with collocations specially the verb- noun pairs considering their frequency. Among the six collocation types, the verb- noun pairs had the highest frequency and the verb preposition- noun rated the last in all four books. The adjective- noun pairs were also prevalent in the four books, especially Interchange book three. Then the noun-noun pairs rated the third in the series and the other three types, namely verb-adverb, adverb-adjective and verb preposition- noun pairs seemed to be meagerly employed in the Interchange series. On the whole, from this part of their study, it could be concluded that more collocations were used in interchange book series than high school text books and most of these collocations were lexical collocations. This result is exactly in line with the result of our study because in our study it was also found that the percentage lexical collocations was more in American English File Book 2 than in high school English Text Book 3. In another study, Roohani (2011) investigated the extent to which lexical and grammatical collocations were used in high school and pre-university English textbooks, compared to the New Interchange book series. To achieve the purposes of the study, a content analysis was first carried out in the corpora obtained from the textbooks to see the frequency and proportion of lexical and grammatical collocations. Second, textbook-based collocation tests were developed and administered to the participants of the study. The results showed that the frequency and proportion of collocations in the high school and pre-university English textbooks were generally lower than those in the New Interchange book series, and it was also indicated that the learners studying the New Interchange book series in the private language institutes had a better performance on both lexical and grammatical collocation tests. These results are also congruent with the results of our study.

The researchers of this study did not found any studies investigating the percentage of grammatical and lexical collocations in American English file book 2 and English high school text book 3. Through comparing these two books it was revealed that lexical collocations were used more than grammatical collocations and the percentage of both grammatical and lexical collocations was more in American English File Book 2 than in high school English Text Book 3. It was also found that the institute participants outperformed the high school participants on the collocation test.

8. Conclusion

The use of appropriate collocations in a spoken or written text increases native likeness. Based on the importance of collocation in language proficiency and native likeness, different studies, including this one, have made attempts to investigate collocations from different aspects in different books. In the present study the researchers considered the frequency and type (lexical and grammatical) of collocations which were presented in two books namely, American English File Book 2 which has been prepared for teaching English to foreigners and English Iranian High School Text Book 3 which has been produced by the Ministry of Education and taught in all state and non-state schools in Iran. In Iranian high schools the emphasis in English classes is only on the text books and in fact these text books are regarded as the core materials which set the English class directions in Iran .As far as the researchers studied, there has been no study concerning with the comparison of collocations in American English file book 2 and Iranian high school text book 3, and therefore, this study aimed to ascertain the frequency and types of collocations in these two books. The results of the statistical analyses revealed that the two text book series differ marginally in terms of frequency and type of collocations. Major difference existed between them when it came to lexical collocations in American English file book 2. To conclude, according to the observation made in the present study, it was revealed that collocations occupy the content of American English File Book 2 more than English high school text book 3 and it was also proven that most of these collocations were lexical collocations compared with grammatical collocations. Moreover, it was shown that the participants who took part in the language institute and studied American English file book 2 outperformed the typical high school participants who studied and relied on English high school text book 3 in the collection test.

References

Ahangari, S., & Zununi Vahed, F. (2013). Collocation types in Interchange Series and high school books. *Basic and Applied Scientific Research, 3*(2), 367-374.

Bahns, J. (1993). Lexical collocations: a contrastive view. *ELT Journal, 47,* 56-63.

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations. *System, 21*(1), 101–114.

Benson, M. (1985). Collocations and idioms. In R. Ilson, (Ed.) .*Dictionaries, lexicography and language learning* (pp. 61-68).

Benson, M., Benson, E., & Ilson, R. (1986). *The BBI combinatory dictionary of English.* Amsterdam: John Benjamin's Publishing Company.

Conklin, K & Schmitt, N. (2007). Formulaic Sequences: Are They Processed More Quickly than Nonformulaic Language by Native and Nonnative Speakers? *Applied Linguistics, 29*(1), 72-89.

Cowie, A. P. (1981). The Treatment of Collocations and Idioms in Learners' Dictionaries'. *Applied Linguistics 2,* 223-235.

Cowie, A. P. (1994). Phraseology. In R. E. Asher & J. Simpson (Eds.), *The encyclopedia of language and linguistics*, (pp. 3168-71). Oxford: Pergamon Press.

Firth, J. R. (1951). Modes of meaning. In J. R. Firth (Ed.), *Papers in linguistics,* (pp. 115-197). Oxford: Oxford University Press .

Gitsaki, C. (1996). *The development of ESL collocational knowledge.* Doctoral thesis, The University of Queensland, Australia. Retrieved from http://www.Cltr.uq.ozau.

Halliday, M. A. K. (1966). Lexis as a linguistic level. In C. E. Bazell, J. C. Caford, M. A. K. Halliday, & R. H. Robinso (Eds.), *In memory of J. R. Firth*, (pp.148-162). London: Longman.

McCarthy, M. (2005). *English collocations in use*. Cambridge: Cambridge University Press.

Roohani, A., (2011). *Collocations in high school and pre-university English textbooks versus New Interchange Book Series. The Journal of Asia TEFL, 8*(3), 55-81.

Rudzka, B., Channel, P., Putseys, Y., &Ostyn, P. (1981). *The words you need. Teacher's book*. London: Macmillan.

Searle, J. R. (1976). *Speech acts: An essay in the philosophy of language*. London: Cambridge University Press.

Shahedi, S. (2001). *Constructing an analytical framework for the analysis of Persian language texts for foreign learners*. Unpublished master's thesis, Shiraz University, Shiraz. Iran.

Shehata, A. (2008). *L1 influence on the reception and production of collocations by advanced ESL/EFL Arabic learners of English.* Retrieved from: http://etd.ohiolink.edu

Shokouhi, H., &Mirsalari, GH. (2010). *Collocational knowledge versus general linguistic knowledge among Iranian EFL learners. Teaching English as a second or foreign language.* Retrieved from: http://tesl-ej.org

Sinclair, J. M. (1966). Beginning the study of lexis. In C. E. Basell, J. C. Catford, M. A. K. Halliday, & R. H. Robins (Eds.), *In memory of J. R. Firth* (pp. 410-430). London: Longman.

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.

Sinclair, J. M. (2003). *Reading concordances*. London: Longman.

Soleimani, H., & Dabaghi, A. (2010). Textbook evaluation: A reflection on the New Interchange Series. *International Journal of Research Studies in Language Learning*, *1*(2), 19-32.

Wang, J., & Good, R. (2007). *The repetition of collocations in EFL textbooks: A corpus study.* Paper presented at the fourth corpus linguistics conference held at the University of Birmingham.

Woolard, G. (2000). Collocation encouraging learner independence. In M. Lewis (Ed.), Teaching collocation*: Further developments in the lexical approach,* (pp. 28-46). London: Heinle Cengage Learning.

Wu, W. C. (2007). *Lexical collocations*: *One way to make passive vocabulary active*. Retrieved from: http://: web.chu.edu.tw.