The Impact of Collocations On Iranian EFL Learners Interlanguage

Dr. E. Faghih[1]
M. Sharafi[2]

Abstract

Knowing a word implies the knowledge of the possible combinations – or collocations into which a given item can enter. Collocations tend to vary among languages and topic domains. While attaching due significance to the inclusion of lexical items in EFL syllabus, highlighting the importance of collocations is a necessity. The present study was undertaken with the belief that more work on collocations would enhance the practices of English learning/teaching. Specifically, this study is the result of an attempt to investigate an error pattern in the vocabulary of Iranian EFL learners i.e. the confusion of collocations. It was assumed that such an error

1. Professor of TEFL, Alzahra University, email:esfaghih@alzahra.ac.ir
2. MA in TEFL

pattern reflected a difficulty in vocabulary acquisition and that it was a feature of interlanguage. The analysis of the data indicated that: collocation confusion is indeed a common error in Iranian EFL learners' interlanguage; there is a positive correlation between learners' overall proficiency and their knowledge of collocations; all kinds of collocations are not equally hard for Iranian EFL learners, and adjective plus noun type of collocations will pose the largest amount of difficulty to Iranian learners.

Key words*: collocation, vocabulary, interlanguage*

1. Introduction

The mastery of vocabulary is an essential component of second language acquisition. In Laufer's words, "vocabulary learning is not only absolutely necessary, it is also never ending. Long after the acquisition of phonology, morphology and syntax has been completed, the learner will still be encountering new word and expanding his lexicon" (Laufer, 1991: 1). It is only common sense that no communication can take place without words. Madsen (1983) admits that, "basic communicative competence is largely concerned with the strategies that learners use to solicit the vocabulary they need in order to get meaning across" (1983: 74). Krashen is also often quoted to have said that "learners don't carry grammar books around in their pockets. They carry dictionaries" (in Laufer 1991: 1) As it has been well established, "knowing a word implies the knowledge of possible combinations into which a given item can enter. Such combinations are called collocations (Laufer, 1991: 19). As Laufer admits,

"collocations are problematic when their meaning is apparent at first glance but their constituent elements can not be given their translation equivalents". (ibid)

Combinations of words will differ from language to language. For example the verb 'have' collocates with the word 'dream' in English therefore, it is said "I had a dream". However, in Persian this combination would not be used. Instead, one would say "I saw a dream". EFL learners probably make mistakes because they collocate English words on the basis of their native language principles. Lexical collocations are frequently occurring word pairs in natural languages. These collocations are used by native speakers of a language almost without thought, yet they must be acquired by non-native speaker of that language. Collocations tend to vary among languages and topic domains. Despite the obvious necessity for more research in this aspect of language proficiency, it has not been sufficiently explored yet.

Some researches such as Zughoul's study have been conducted with the assumption that the result of their study on collocations would contribute to a better understanding of interlanguage and would give better insights into the strategies employed by language learners for lexical choice. Zughoul believes that "the study of lexical choice, an area that may be classified under interlanguage semantics has not received as much emphasis as the other two components of interlanguage, namely phonology and syntax in language learning – teaching research (1991: 51).

The present study has been undertaken with the belief that intuition and experience of students and instructors shows more work on collocations is a necessity in English learning settings. It was also an

attempt to argue that highlighting the collocational aspects of lexical items is as important as teaching them individually. This, it was believed, may be a way of nurturing the active use of language and helping the English language learner construct lexically as well as grammatically acceptable sentences. This study had a triple purpose. First, it attempted to stress the significance that collocations are an integral part of any language and therefore must be incorporated into the EFL syllabus in order to be taught formally and systematically. Second, it attempted to indicate which type of collocation was the most difficult one for Iranian EFL learners and third, it tried to propose and substantiate that there was a positive correlation between learners overall proficiency in English language and their knowledge of collocations. Even though, the primary focus was to understand which type of collocations posed the most difficulty to students, it also tried to indicate which aspects of interlanguage affected the learners, recognition of the required collocations.

2. Research Questions and hypotheses

As already mentioned, collocations are among multi-word combinations in English language which are believed to believed to create problems for EFL student. Five types of collocations are error provoking and have been included in the present study as follows.

1. Verb – Noun
2. Adjective + Noun
3. Count Noun + of + Mass Noun
4. Collective Noun + Count Noun
5. Subject – Verb

To achieve the purpose of the study, the following research questions were proposed:

1) Is collocational confusion indeed a common error in Iranian EFL learner's interlanguage?

2) Is the learners overall proficiency systematically related to their susceptibility to collocational error?

3) Are all kinds of collocations equally hard for Iranian EFL learners.

4) Do adjective plus noun type of collocation pose the largest amount of difficulty for Iranian learners?

To obtain a logical answer to the aforementioned research questions the following null hypotheses were formulated:

Ho_1) Collocational confusion is not indeed a common error in Iranian EFL learners interlanguage.

Ho_2) The learners overall proficiency is not systematically related to their susceptibility to collocational error.

Ho_3) All kinds of collocations are equally hard for Iranian EFL learners.

Ho_4) Adjective plus noun type of collocations will not pose the greatest level of difficulty to Iranian learners.

3. Method

3. 1. Subjects

Over one hundred EFL students participated in this study. The students were selected from both male and female juniors and seniors studying English as their major field of study at: a). AlZahra University and b) the Islamic Azad University of Torbat-Heydarieh. After administering the Michigan test to the entire group of the

students, they took a multiple choice test on collocations. The objective of the elicitation test was to measure learner's confusion on collocational aspects of lexical knowledge viewed from the perspective of recognition.

After scoring the Michigan test papers, the testees were ranked in terms of the scores they obtaind. The top half were grouped as the high intermediate and the bottom half as the low intermediate. The subjects who, for one reason or another did not take one of the tests were ruled out from the list. The remaining subjects of the sample added up to 100 students.

3.2. Insturmentation

1. Michigan Proficiency Test. As a means for the estimation of the subjects' proficiency level, the subjects were given a Michigan test, consisting of 100 items: 40 items on structure, 40 on vocabulary and 20 on reading comprehension.

2. Elicitation test. The elicitation test which was especially prepared by the researchers was a multiple choice test on collocations. This test included five categories of collocations as follows.

2. 1. Verb + Noun	15 Items
2. 2. Adjective + Noun	15 items
2. 3. Count Noun + of + Mass Noun	10 items
2. 4. Collective Noun + Count Noun	10 items
2. 5. Subject + Verb 10 items	10 items

In all categories the learners were required to select the correct response from among four given alternatives. In the first category, the

verb is the collocate of a noun as its object and the subjects were expected to recognize the correct verb, e. g.:

Please -------- us a visit next time you're in Birmingham.
A) give   B) pay   C) make   D) provide

In the second category the adjective was missing and the subjects had to choose the correct adjective, e. g.:

To avoid gaining weight, most women eat light rather than ---------- -- food.
A) heavy   B) oily   C) rich   D) greasy

In the third category the appropriate count noun should have been selected by the learners, e. g.:

He had to eat a ------------- of crusty bread.
A) loaf   B) slice   C) circle   D) roll

In the fourth category, it was the collective noun which had to be recognized by the testees, e. g.:

He had left a huge ------------ of flowers in her hotel room.
A) bunch   B) pack   C) bundle   D) collection

And finally the last category, i. e. subject + verb category required learners to select the correct verb which suited the subject, e. g:
Don't let the dog ------------.
A) bark   B) scream   C) shout   D) yell

To construct this test, some of the simplest examples given in the related literature, and Collins Cobuild Dictionary of English Language were used. The test, having a multiple choice format, was intended to be a recognition task, attempting to simulate the situation in which various competing words were simultaneously activated. The items provided four alternative words one of which would collocate with the word given before or after the blank and three distracters. One of the distracters was given on the basis of word for word translation from Persian into English with the intention of observing the possibility of interference from learner's mother tongue. In this test each correct response had one score.

3.3. Procedure

The Michigan proficiency test was first administered to determine the subjects' level of proficiency. After this test, the subjects were categorized as 'high' and 'low' intermediate learners in terms of their proficiency scores. Since all of the subjects belonged to intermediate level, a collocation test consisting of sixty items was administered to all of the subjects after a time interval of 1 – 2 weeks. In order to find out whether we can reject the first null hypothesis, i. e. "Collocational confusion is not a common error in Iranian EFL learners interlanguage", the mean and standard deviation of the whole group and also of high and low groups on both tests were computed. The second and third null hypotheses state that all kinds of collocations are equally hard for Iranian learners and that adjective plus noun type of collocation will not pose the greatest level of difficulty for them. In order to reject these null hypothese, the mean of each category of collocations (e. g. verb + noun) was computed separately. Then each

two obtained means (e. g. verb + noun and adjective + noun) were compared through a t-test to see which type of collocations has posed the highest level of difficulty for learners. Finally, using Pearson correlation formula, the means of Michigan test and the collocation test were compared in order to find out if the last null hypothesis can be rejected. This would have meant that there is no correlation between the learners, overall level of proficiency in English language and their knowledge of collocations.

3.4. Statistical analysis

As a subset of an ex post facto design, the study followed a correctional design in which no treatment was involved, nor did the researchers have any control over what had already happened to the subjects. The scores obtained on the elicitation test were divided into sub-scores of different categories of collocations which were then compared to decide about the most difficult type of collocation. In addition, the scores obtained in the elicitation test were compared with the scores on Michigan proficiency test in order to investigate the relationship between the two sets of scores and eventually to portray the degree of relationship, if any, in terms of a correlation coefficient. Finally, it should be mentioned that the statistical package for social sciences (SPSS) computer program was used to analyze the data.

4. Data analysis and Results

As mentioned earlier the subjects were given two tests: a Michigan proficiency test and an elicitation test on collocations. The data analyzed was limited to the scores gathered from 100 subjects since 16 subjects had not completed both tests. Each correct choice was

5. Discussion

This study had a triple purpose of: a) stressing the point that collocations are an integral part of any language and therefore, must be incorporated into EFL syllabus to be formally and systematically taught; b) indicating which type of collocations is the most difficult for Iranian EFL learners and c) substantiating that there is a positive correlation between the learners overall proficiency in English language and their knowledge of collocations. In order to achieve these purpose four research questions were proposed and four null hypothesis were formulated.

Research question: 1) Is collocational confusion indeed a common error in Iranian EFL learns' interlanguage?

Null hypothesis: 1) Collocational confusion is not a common error in Iranian EFL learners interlanguage.

In this study, the basis for classifying an error as a common one was decided to be the Mean Score. Since generally in Iran, half of the total score is the minimum score required to pass the course, a score below thirty in our collocation test which consisted of 60 items will be an indication of poor performance. In order to be able to reject the first null hypothesis, we therefore decided to take into account the first part of the computations regarding the test of collocations. In this test the highest score was 40 (out of 60) and the lowest score was 1. The mean score was 24.14 and the standard deviation was 6.1759. It is clearly obvious that the subjects performances was very low and the number of collocational errors was higher than that of Michigan test. Among one hundred subjects, the highest score was 40 which was obtained by only one subject. The mean score of 24.14 was also a poor one and thus the first null hypothesis can be rejected. Comparing this mean

score with the learners mean score on Michigan proficiency test also indicates that the learners performance on collocation test, is poorer that of Michigan test.

Research question: 2) Is the learners overall proficiency systematically related to their susceptibility to collocational error?

Null hypothesis: 2) The learners overall proficiency is not systematically related to their susceptibility to collocational error.

In the order to be able to reject the second null hypothesis the correlation formula was adopted. The mean and standard deviation of each test, was used to compute the correlation between the learners scores on the Michigan proficiency test and their scores on the collocation test. The result indicated a positive correlation of 0.6949 which substantiates the second hypothesis and indicates that there is a high relationship between the learner's proficiency and their knowledge of collocations. It means that as a learner's general knowledge in English language increases, his / her knowledge on collocations will increase too.

Research question: 3) Are all kinds of collocations equally hard for Iranian EFL learners?

Null hypothesis: 3) All kinds of collocations are equally hard for Iranian EFL learners.

In order to be able to reject the third null hypothesis, the lowest, the highest, the mean and the standard deviation of each category of collocations was separately computed and the means were compared. Comparing the mean of all five categories (Table 4) resulted in the following rank order of the different categories of collocations on the basis of their level of difficulty.

Table 4. Level of Difficulty of Collocations

Variable	Mean	SD	Min.	Max.	N	# of items
Adj + Noun	5.23	1.90	1	9	100	15
Count Noun + of + Mass Noun	3.84	1.57	0	8	100	10
Verb + Noun	5.98	2.37	0	10	100	15
Subj + verb	4.72	1.97	1	9	100	10
Collective noun + Count noun	4.94	1.87	0	9	100	10

The results indicate that all types of collocations were not equally hard for Iranian EFL learners. The above table shows the ranking of scores on the basis of their level of difficulty. Obviously, the category 'adjective plus noun' which consisted of 15 items has been the most difficult type for learners. Thus, the fourth null hypothesis which presumed that "Adjective plus noun type of collocation will not pose the greatest level of difficulty to Iranian learners" was also rejected. The analysis of different categories of collocations indicated that the second most difficult one was count noun + of + mass noun, the third was verb + noun, the fourth was subject plus verb and finally the last and in fact the easiest for the learners was collective noun plus count noun category.

The analysis of the obtained data rejected all four null hypotheses and substantiated the authenticity of the research hypotheses. Some interesting patterns were also found in the data which are worth mentioning. As already pointed on, in the construction of elicitation instrument, one of the distractors was chosen with an assumption that learners collocate words on the basis of collocability of words in their first language. Therefore, such distractors were in a sense the word for

word translation of the target word, e. g. Please …. us a visit next time you're in Birmingham.

A) give   B) pay   C) make   D) provide

Alternative C was presented on the basis of this assumption that learners may choose the word "make" as a collocation for "visit" because in Persian people "make a visit", but in English they "pay a visit" The analysis of the data indicated that such distractors were so attractive to the learners that they overruled the other alternatives and this was an indication of transfer of collocations from mother tongue to English language. Following examples are the most frequently selected ones which clearly exemplify this strategy of the learners in the selection of collocations.

Correct Collocation	Incorrect collocation
pay a visit	make a visit
set my alarm clock	fix my alarm clock
adopt a child	accept a child
rich food	oily food
fast color	stable color
weak tea	light tea
pat of butter	block of butter

Even a quick look at the examples shows traces of students' mother tongue. Translation from the native tongue to the target language could be regarded as the main cause of error. Another reason could be lack of extensive reading of contemporary English prose whereby the students may have acquired and built up the necessary competence to use the lexicon of the target language effectively.

6. Conclusions and implications

The present study was an attempt to define, illustrate, classify, and validate collocation errors. It was necessary to draw the attention of learners, teachers and translators to the existence of this phenomenon and some of its general characteristics, like the relative difficulty of collocations and the relationship between collocational errors and the learner's level of proficiency. The main aim of the empirical part of the study was to validate the existence of the problem, i. e. to find out whether the confusion of collocations was a common error made by learners and whether the learners overall proficiency is systematically related to their susceptibility to collocational errors. The study also compared various categories of collocations based on their difficulty. The results of the study indicate that confusion of collocations is indeed evident in the performance of language learners and that the task of correctly identifying lexical collocations, as admitted by many researchers seems to be very difficult. The study also indicated and substantiated a positive correlation between the learners' overall proficiency in English and their knowledge of collocations.

Finally, the confusion of collocations as evidenced by the results of the study lends support to the contrastive analysis and interlanguage studies. The first language effect is reflected in the results of the learner's performance. Since collocations have been shown to be a pattern of difficulty in vocabulary learning and collocational errors would probably rate high, it is reasonable to claim that collocations should receive pedagogical treatments, in the form of exercises, materials and tests. The findings of the study can provide guidelines as to which categories of collocations are more problematic and should therefore be selected for practice. Such practice can take the form of

various exercises, such as filling in the blanks, word family building, controlled writing, multiple choice etc. These practices can be supplemented by carefully prepared materials on vocabulary, including special dictionaries, etc. In conclusion, collocational treatment should be included in testing. Collocations could serve as a criterion for the selection of items to be tested in vocabulary tests. The general principle behind the suggested practices is the development of the learner's awareness of the collocational problems. This principle is based on the belief that a conscious and systematic analysis of language contributes to language learning. It should be reiterated that "language is not made up of a large number of words which can be used together in free variation" (Baker, 1992: 63). Words have a certain tolerance of compatibility. Like individual words, collocational patterns carry meaning and can be cultural specific. This gives rise to numerous pitfalls in vocabulary learning and requires appropriate attention and systematic teaching. Learners, teachers and translators have much to gain from the study of collocations. It is hoped that the results of this study, will have an impact on English textbook writing, particularly textbooks written for Persian speakers learning English as a foreign language.

Dictionary writers may also benefit from studies on collocations and when writing dictionaries they will hopefully take into account the collocational aspects of words. Translators should demonstrate high lexical competence in the foreign language they are translation from or into. This competence comprises a good knowledge of collocations as well. The translators awareness of collocation restriction is a fundamental prerequisite for improving translation. It is very important, therefore, that translators be made aware of the

collocational nature of English language through extensive training that views collocations as a central concern. Thus translators are another group who will hopefully benefit from the study of collocations.

References

Baker, Mona, (1992), *In other words: A course book on translation*, London: Routledge.

Collins, *Cobuild English Dictionary*, (1995), London, Harper Collins Publishers.

Farghal Mohammad and Obiedat Hussein, (1995), Collocations: A Neglected Variable In EFL, *IRAL*, Vol. xxxIII.

Greenbaum, S. (1970), *Verb-Intestifier Collocation in English*, The Haugue: Mouton.

Jones, S. and Sinclair, J. M. (1974), English Lexical collocations, *Cahiers de Lexicologie*, 30: 15-61.

Larson, Mildred L. (1984), *Meaning – based Translation: A Guide to Cross – Language Equivalence*, New York: University press of America.

Laufer – Dvorkin, Batia, (1991), *Similar lexical Forms in Interlanguage*, Tubingen: Gunter Narr Verlag.

Madsen, Harold S., (1983), *Techniques in Testing*, Oxford: Oxford University Press.

Newman, Aryeh, (1988), The contrastive analysis of Hebrew and English Dress and Cooking Collocations: Some Linguistic and Pedagigic Parameters, *Aplied linguidtics*, vol. 9, No. 30, Oxford University Press.

Stubbs, Michael, (1995). *Corpus Evidence for Norms of Lexical Collocation* in Guy Cook and Barbara Seidlhofer (eds.)

Zughoul, Mohammad Raji: (1991), Lexical Choice : Towards writing Problematic Word Lists, *IRAL*, Vol. xxxlx.