# An Analysis of Collocations in Iranian ESP Materials

**Alipour Mohammad, Haghverdi Hamid Reza and Biria Reza**
Department of English, Khorasgan Branch, Islamic Azad University (IAU), Isfahan, IRAN

## Abstract

*Lexical Priming Theory claims that the collocational primings of a word depend on the genre or discipline in which it occurs. In this study, attempts were made to compare the collocates of two nouns (group, theory) and the positions of the collocates prior and subsequent to the node word in two specialized corpora of soft sciences and ESP textbooks in Iran to see if they are realized in the same or different ways. To serve this purpose, two corpora were compiled. Then, two nouns which enjoyed high frequency in the corpora were selected. The Word Smith Tools software was employed to detect the first 15 collocates of the nouns and the frequencies of their occurrences in 10 possible positions prior and subsequent to the node word. The findings demonstrated some remarkable discrepancies across the corpora. Nouns tend to collocate with different words, especially content words, in the Iranian ESP textbooks in comparison to the original texts. Furthermore, it was found that these collocates are positions specific. To put it differently, collocates of a word tend to occur in or avoid certain positions in different disciplines. The same also unfolded in the ESP corpus.*

## Introduction

Linguists have recognized for some time that words in natural language are neither randomly combined into phrases and sentences nor that they are only constrained by the rules of syntax. Curiously, this basic fact about collocations and, at the same time, their rather diverse and apparently idiosyncratic behavior, has been taken out of focus by a substantial part of contemporary mainstream linguistics which has been primarily concerned with examining language from a theoretical perspective. In particular, generative linguistics in the Chomskyan tradition demotes all lexical and syntactic idiosyncrasies safely into the realm of the lexicon.

Bartsch believes that by pointing out that "you shall know a word by the company it keeps!" [1], it was Firth who first introduced the notion of collocation into contemporary linguistics [2]. Still, as it is claimed, Firth used to be rather vague about a precise definition of the concept, and hence it is not surprising that there has been a rather enormous conceptual diversity surrounding the idea of collocation in linguistic research up to today [3,4]. Drawing a very rough dividing line, two lines of linguistic research may be identified in the last half-century. On the one hand, there is the structural-lexicographic approach which is mainly concerned with adequate representation forms of collocations within linguistic lexicons and dictionaries. On the other hand, there is the frequentist corpus-based approach to collocations which was initiated and significantly influenced by Firth's linguistic research and which is dedicated to an empirically grounded analysis of natural language. The present study follows the later approach and studies collocations through Hoey's Lexical Priming Theory [5].

Lexical Priming Theory is a new theory of both the lexicon and of language. Following this theory, lexis and grammar are given reversed roles so much so that lexis is complexly and systematically structured and that grammar is an offshoot of this lexical structure [5,6,7]. Hoey contends that "the theory grew out of an increasing awareness that traditional views of the vocabulary of English were out of kilter with the facts about lexical items that are routinely being thrown up by corpus investigations of text" [5].

Lexical priming theory attempts to show how naturalness is obtained and how naturalness might be dependant on explanations of what is possible. In this regard, collocation functions as a key factor in naturalness. Through this theory of priming, Hoey explains the existence of collocation: "the only explanation that seems to account for the existence of collocation is that each lexical item is primed for collocational use" [8]. The priming proposed by him draws on psycholinguistic arguments and it assumes that "as the word is learnt through encounters with it in speech and writing, it is loaded with the cumulative effects of those encounters such that it is part of our knowledge of the word that it co-occurs with other words" [8].

For Hoey, priming is the most appropriate psychological concept to explain collocation [5]. He holds that collocational naturalness would be inexplicable unless the words were stored in our minds separately or in sets. Hoey draws on the notion of semantic priming to elaborate on the way a priming word might trigger a given target word. As an instance, a listener will recognize the word *heart* more quickly if he or she is previously given the word *body* than if they had previously been given an

unrelated word such as *trick*; in this way, the listener is primed for *heart* by *body*. The word *body* establishes a word association with *heart*, which the word *trick* cannot.

Hence, we can account for collocation only if we believe that every word is primed mentally for collocational use. According to this view, primings are psychological phenomena and each person possesses inevitably unique experiences of any particular word. Then, it can be assumed that the primings for a word may differ from person to person, due to different encounters with the word in different contexts. The underlying assumption of lexical priming is that each time we encounter a word, we note and keep a record of the contexts in which it has occurred. As patterns of recurrence of context emerge, we become primed to associate the word with these recurrent contexts. Hoey says: As a word is acquired through encounters with it in speech and writing, it becomes cumulatively loaded with the contexts and co-texts in which it is encountered, and our knowledge of it includes the fact that it co-occurs with certain other words in certain kinds of context. The same applies to word sequences built out of these words; these too become loaded with the contexts and co-texts in which they occur[5].

On the basis of this view, when we come across with and acquire a word or phrase through repeated encounters (spoken or written), simultaneously we subconsciously note the contexts in which we are encountering the expression in question; the experience primes us. In a way, a concordance of the word or group is built up and processed in our heads. When called upon to use the word or phrase, we draw unconsciously upon this concordance knowledge and in so doing we at the same time strengthen our own primings and our listeners or readers' primings. Hoey refers to this property as nesting, where the product of a priming becomes itself primed in ways that do not apply to the individual words making up the combination. Nesting simplifies the memory's task. Necessarily the priming of word sequences is normally a second phase in the priming; occasionally, of course, a child acquires the primings of a combination first and the primings of the individual words later, e.g. *all gone*. There is no difference in principle between acquiring the word or word sequence and acquiring the knowledge of its collocations although, presumably, recognition of the word must notionally precede recognition of recurrent features in that the word has to have occurred twice at least for the latter process to begin.

According to Pecorari, the findings of English for Specific Purposes (ESP) research in various genres or disciplines must include not only what features happen, but also how and why they are employed[9]. The question of what linguistic features characterize a particular genre within a given subject area is redily answered by corpus methods. However, to realize the hows and whys, besides the rhetorical or discoursal functions, it is also crucial to examine the lexical items. Describing such an investigation is the overarching purpose of this study.

The genre approach in corpus work, unfortunately, has merely resorted to a macro-analysis perspective focusing on the rhetorical structures of texts and has ignored the micro-analysis of lexical patterns of texts which make a salient and indispensable part of any genre on the ground the same lexical items may be patterned and applied differently in different genres and disciplines. A great number of ESP studies have been conducted focusing on single grammatical categories such as hedging, negation, passivization and so on[10]. As far as we know, nonetheless, few analyses have been carried out on the distribution and patterning behavior of lexical items. As a result, this study attempted to compare two nouns in a corpus of psychology and sociology (soft sciences) with a corpus of Iranian ESP textbooks of the same disciplines in terms of their collocations. In other words, it attempted to find out whether these ESP textbooks corresponds to or differ from the conventional and acceptable collocational features of original soft sciences textbooks. Therefore, the present investigation aimed at investigating the question below: Do the ESP materials developed for Iranian universities conform to the collocational patterns of nouns in soft sciences texts?

## Material and Methods

**Software:** To conduct this study, we applied the fifth version of Oxford Word Smith Tools software[11]. This integrated suite of programs allows detecting words' behavior in texts. It enables you to use the tools to see how words are used in your own or others' texts. This software comprises three main programs: Concord, Word List, and Key Words. In this study, to select the nouns for analysis, the Word List function was run to identify two frequently used nouns in the Soft Sciences corpus (SS henceforth). The same procedure was also applied for the ESP corpus. Furthermore, we applied the Concord function to find the most frequently employed collocates of the nouns along with their positions regarding the node word in our corpora.

**Corpora:** For the purpose of the current project, two home-made specialized corpora were the most suitable resource. This decision was made because suitable specialized corpora for different disciplines are unavailable. Nevertheless, there are always a series of issues which ought to be considered when compiling corpora including size, content or composition, representativeness, and generalizeability[11-16]. Efforts were made to account for all these factors.

Considering size, following Sinclair, although a specialized corpus of around one million words is appropriate in ESP studies, we endeavored to expand the corpus as large as possible so long as the criterion of manageability remained untouched[17]. Hence, in order to compile the corpus, 50 books on psychology and sociology were used. Some of these books included less than 100,000 words and some of them consisted of more words. But, generally, each book comprised roughly 100,000 words. Therefore, the corpus consisted of approximately 5 million words which is thought to be a big number for a specialized corpus of this type. With respect to the content, our corpus also

represented these two disciplines since a vast variety of different books by various authors were included to ensure that no author's style would dominate and impact on the collocational patterns of the nouns.

Regarding the ESP corpus, to investigate the appropriate use of lexical items in ESP textbooks written by Iranian authors, we made a corpus of soft sciences comprising two commonly taught books on psychology and two commonly taught books on sociology[18-23]. This corpus included approximately 300,000 words.

**Procedure:** After the required books on psychology and sociology were collected, we developed a specialized corpus of soft sciences and an ESP corpus. Afterward, two nouns (*group, theory*) were selected to juxtapose their collocations in the corpora. Some certain criteria had to be met for the selections of the nouns. First of all, they must have been non-special words so that each given discipline would not influnce their collocational primings. Secondly, they had to be used with a high frequency rate across both corpora to guarantee the reliably of detecting their collocational patterns to scrutinize any possible similarities or differences. We did this selection by means of the fifth version of the Oxford WordSmith Tools software [11]. In the end, by employing the software, we tried to scrutinize the ways by which the nouns were primed for their collocates across the SS and ESP corpora.

## Results and Discussion

In the present study, the corpora were analyzed in the same line with Hoey [5]. In other words, we employed the same method to analyze the data. Therefore, the top 15 collocates of the nouns occurring immediately prior and subsequent to the node word on each side in order of decreasing were detected. The reason behind the selection of the first 15 collocates across the corpora,

as you can see in the following tables, is that the first 10 collocates were function words in most cases, such as *'the', 'is'* and *'with'*. Accordingly, the five collocates which were selected next were intentionally content words, not specific to the disciplines, such as *'protecting', 'individual'* and *'special'*.

According to table-1, except for the seventh collocate which was *members*, all the first 10 collocates of the word *group* in the soft sciences corpus were function words. Here, they all occurred in specific positions. In this corpus, the least favorite position for the collocates of the word *group* was R1 which represented 6 of the lowest occurrences, while the most favorite position was L1 which represented 5 of the highest occurrences. As an example, the first collocate of *group* was the article *the* which transpired 1718 times in the corpus, 1200 occurrences before and 518 ones after the node word. It is should be noted that, in the same vein, the main landing site for *the* was L2, and the least common position was R1. This is also indicative of the much higher likelihood of its occurrence in L2 position than its R1 position. By way of contrast, the fourth most recurrent collocate, that is, *and*, showed a strong tendency to happen rather equally on both sides of the node word, 427 and 463 instances respectively. Nevertheless, its highest and lowest occurrences both occurred before the node word; R2 included 132 and R3 comprised 34 instances.

According to table-2, all the first 10 collocates of the word *'group'* in the ESP corpus were function words too, including articles, prepositions and conjunctions, except for the 8th and 10th collocates, i.e. *'control'* and *'counseling'* respectively. Here, all of them occurred in specific positions too. As you can see, the least favorite positions for the collocates of the word *'group'* in ESP texts were L5 and L2, each accounting for 7 of the lowest occurrences.

**Table-1**
**The top 15 collocates of the noun 'group' occurring immediately prior and subsequent to the node word in the Soft Sciences corpus**

| Collocates | Total | Left | Right | L5 | L4 | L3 | L2 | L1 | R1 | R2 | R3 | R4 | R5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1-THE | 2950 | 2075 | 875 | 300 | 317 | 303 | **687** | 468 | **55** | 192 | 233 | 208 | 187 |
| 2-OF | 2124 | 1287 | 837 | 199 | 202 | 352 | 363 | 171 | **414** | **77** | 84 | 119 | 143 |
| 3-AND | 1632 | 776 | 856 | 143 | 134 | 126 | 213 | 160 | 233 | **239** | **100** | 138 | 146 |
| 4-A | 1498 | 1109 | 389 | 106 | 107 | 175 | 299 | **422** | **23** | 94 | 101 | 91 | 80 |
| 5-IN | 1412 | 921 | 491 | 129 | 146 | 159 | 168 | **319** | **81** | 126 | 99 | 99 | 86 |
| 6-TO | 1173 | 619 | 554 | 141 | 125 | **171** | 141 | **41** | 62 | 107 | 125 | 130 | 130 |
| 7- THAT | 512 | 299 | 213 | 67 | 43 | **78** | 57 | 54 | 48 | 43 | 42 | **39** | 41 |
| 8- MEMBERS | 487 | 135 | 352 | 31 | 55 | 46 | 3 | **0** | **288** | 8 | 14 | 31 | 11 |
| 9-FOR | 471 | 265 | 206 | 46 | 54 | 62 | **65** | 28 | 38 | 41 | 38 | 45 | 44 |
| 10-AS | 458 | 199 | 259 | 26 | 26 | 31 | 42 | **1** | **88** | 26 | 21 | 33 | 22 |
| 11- MINORITY | 174 | 153 | 21 | 11 | 5 | 15 | 4 | **118** | 3 | 2 | 8 | **1** | **7** |
| 12INDIVIDUAL | 154 | 105 | 49 | 7 | 14 | 22 | **50** | 12 | **1** | 13 | 13 | 9 | 13 |
| 13-IDENTITY | 133 | 52 | 81 | 14 | 12 | 16 | 9 | **2** | **58** | **2** | 10 | 4 | 7 |
| 14- CONTROL | 133 | 104 | 29 | 3 | 1 | 4 | 1 | **95** | **1** | 8 | 7 | 8 | 5 |
| 15-MAJORITY | 125 | 112 | 13 | 5 | 3 | 11 | 1 | **92** | **0** | **0** | 5 | 6 | 2 |

On the other hand, the most favorite position was L1 which comprised 6 of the highest occurrences. As an instance, the most frequently used collocate of *'group'* was the article *'the'* which occurred 71 times in our ESP corpus, with 45 occurrences prior to and 26 ones subsequent to the node word. It goes without saying that this article has a stronger tendency for the left hand side position of the word *'group'* as it occurred roughly twice more here compared to the right hand side position. Moreover, L2 was the main landing site for *'the'*, with 21 occurrences, while it only occurred 2 times in L3 position. In contrast, the second most frequently used collocate, i.e. *'of'*, tended to occur remarkably differently from *'the'* on both sides of the node word, with 16 ones on the left and 31 ones on the right. Nevertheless, this preposition was mainly associated with the R1 position of *'group'* which embraced 25 occurrences. The lowest ones belonged to L1 position incorporating 0 occurrences.

A simple comparison of tables 1 and 2 reveals some intriguing points. First of all, only 8 of the collocates of the word *'group'* were same across the two corpora, which are 6 function words and 2 content words. The rest were unique to each corpus. Secondly, these shared collocates were not in the same descending order between the corpora, except the first three ones, namely, *'the', 'of',* and *'a'*. For instance, while *'and'* was the 4th collocate, *'in'* the 5th collocate, and *'to'* the 6th collocate in the SS corpus, they were the 5th, 7th, and 4th collocates in the ESP corpus respectively. Moreover, the L5-to-R5 positions for the highest and the lowest occurrences of the shared collocates differed remarkably in the corpora. To illustrate this point, although the least favorite landing sites for the collocates *'the'*, *'of'*, and *'a'* were R1, R2, and R1 respectively in the SS corpus, they were L3, L1, and L5 in the ESP corpus. Conversely, the positions with the biggest number of occurrences for *'and'*, *'in'*, and *'to'* in the SS corpus were R2, L1, and L5 respectively; they were R1, L4, and L2 in the ESP corpus. Another point to mention is that merely 2 content-word collocates were the same between the corpora, i.e. *'control'*, and

*'work'*. Nevertheless, they ranked quite differently across our corpora. For example, whereas *'control'* ranked 12th in the SS corpus, it was the 8th collocate in the ESP corpus.

As illustrated in table-3, except for *social*, which was the sixth collocate of the word *theory*, all the first 10 collocates in the SS corpus were content words. It is very astonishing to note that, in this corpus, L1 functioned both as the most and the least favorite positions for the collocates, and it included 5 of the highest and 8 of the lowest occurrences. The first collocate of *'theory'* was *'of'* which occurred somehow evenly on both sides of the node word, with 1949 occurrences prior to and 2202 ones subsequent to the node word. For this collocate, R1 was the main landing site and R2 was the least common position. However, with 1182 instances on the left hand side, the collocate *social* tended to pile up there than the right hand side position which comprised only 287 instances out of the total 1469 instances. This position specificity characteristic also applied to the content-word collocates, whereby the collocates *research* and *sociological*, for example, leaned toward different positions of *theory*.

Following table-4, function words made up all the first 10 collocates of the noun *theory* in the SS ESP corpus, except for the third and tenth ones, *control* and *counseling*. Here, all of them occurred in specific positions too. As you can see, the inclusion of 9 of the lowest occurrences made R1 the least favorite positions for the collocates in the SS ESP texts. On the other hand, the most favorite position was R2 which comprised 4 of the highest occurrences. As an instance, the most frequently used collocate of *theory* was the article *the* which occurred 50 times in the SS ESP corpus; it was equally distributed on both sides of the node word. Moreover, L4 was the major position for *the*, with 10 occurrences, while it only occurred once in L5 position. In contrast, the third most frequently used collocate, i.e. *organizational*, tended to occur remarkably differently from *the* on both sides of the node word, with 17 ones on the left and 7 ones on the right. This adjective was chiefly associated with the L1 position which embraced 15 occurrences.

**Table-2**
**The top 15 collocates of the noun 'group' occurring immediately prior and subsequent to the node word in the ESP corpus**

| Collocates | Total | Left | Right | L5 | L4 | L3 | L2 | L1 | R1 | R2 | R3 | R4 | R5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1-THE | 71 | 45 | 26 | 9 | 3 | **2** | **21** | 10 | 3 | 5 | 6 | 6 | 6 |
| 2-OF | 47 | 16 | 31 | 3 | 2 | 5 | 6 | **0** | 25 | 2 | 1 | 2 | 1 |
| 3-A | 45 | 33 | 12 | **0** | 1 | 2 | 7 | 23 | 1 | 3 | 3 | 4 | 1 |
| 4-TO | 25 | 14 | 11 | 3 | 1 | 2 | **5** | 3 | 1 | 2 | 4 | 4 | **0** |
| 5-AND | 23 | 10 | 13 | 1 | 2 | 3 | 4 | **0** | **5** | **0** | 1 | 4 | 3 |
| 6-IS | 22 | 8 | 14 | 1 | 3 | 3 | 1 | **0** | **5** | 3 | 1 | 3 | 2 |
| 7-IN | 19 | 14 | 5 | **0** | **5** | 5 | 2 | 2 | 1 | 1 | 1 | 2 | **0** |
| 8-CONTROL | 16 | 11 | 5 | **0** | 1 | 1 | **0** | 9 | 0 | 0 | 2 | **0** | 3 |
| 9-OR | 14 | 8 | 6 | **0** | **0** | 1 | 2 | 5 | 3 | **0** | 1 | 1 | 1 |
| 10COUNSELING | 12 | 3 | 9 | 1 | **0** | 1 | **0** | 1 | 8 | 0 | **0** | **0** | 1 |
| 11EXPERIMENT | 11 | 7 | 4 | 1 | **0** | **0** | **0** | 6 | 0 | 1 | **0** | 2 | 1 |
| 12-WORK | 8 | 8 | **0** | **0** | **0** | **0** | **0** | 8 | **0** | **0** | **0** | **0** | **0** |
| 13-CALLWD | 8 | 4 | 4 | 1 | **0** | 3 | **0** | **0** | **0** | 2 | 1 | 1 | **0** |
| 14- ONE | 7 | 5 | 2 | 0 | 1 | 1 | **0** | **3** | **0** | 1 | **0** | **0** | 1 |
| 15-POWER | 6 | 1 | 5 | **0** | **0** | 1 | **0** | **0** | 4 | **0** | 1 | **0** | **0** |

**Table-3**
**The top 15 collocates of the noun 'theory' occurring immediately prior and subsequent to the node word in the Soft Sciences corpus**

| Collocates | Total | Left | Right | L5 | L4 | L3 | L2 | L1 | R1 | R2 | R3 | R4 | R5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1-OF | 4151 | 1949 | 2202 | 242 | 255 | 438 | 839 | 157 | **1483** | **40** | 157 | 266 | 256 |
| 2-THE | 3265 | 1845 | 1420 | 411 | 470 | 218 | 241 | **505** | **108** | 487 | 247 | 303 | 275 |
| 3-AND | 2798 | 986 | 1812 | 191 | 159 | 189 | 323 | **124** | **941** | 227 | 240 | 221 | 183 |
| 4-A | 1621 | 1119 | 502 | 138 | 120 | 182 | 313 | **366** | **46** | 127 | 109 | 114 | 106 |
| 5-IN | 1551 | 825 | 726 | 141 | 138 | 172 | **293** | 81 | 265 | **70** | 151 | 125 | 115 |
| 6-SOCIAL | 1469 | 1182 | 287 | 55 | 45 | 29 | 228 | **825** | **14** | 160 | 39 | 40 | 34 |
| 7-TO | 1200 | 602 | 598 | 125 | 120 | 150 | **173** | 34 | 131 | 91 | 119 | 137 | 120 |
| 8-THAT | 761 | 278 | 483 | 58 | 58 | 47 | 81 | **34** | 113 | **142** | 89 | 73 | 66 |
| 9-IS | 705 | 213 | 492 | 55 | 65 | 51 | 38 | **4** | **249** | 54 | 54 | 82 | 53 |
| 10-AS | 515 | 216 | 299 | 65 | 40 | 58 | 45 | **8** | **105** | 24 | 66 | 53 | 51 |
| 11-RESEARCH | 312 | 78 | 234 | 14 | 10 | 20 | 32 | **2** | 38 | **151** | 25 | 11 | 9 |
| 12PERSONALITY | 196 | 97 | 99 | 4 | 5 | 17 | **1** | 70 | 2 | 69 | 15 | 5 | 8 |
| 13SOCIOLOGICAL | 183 | 172 | 11 | 9 | 5 | 5 | 2 | **151** | **1** | 3 | 3 | 2 | 2 |
| 14-PSYCHOLOGY | 161 | 78 | 83 | 17 | 22 | 22 | 16 | **1** | 21 | 6 | **28** | 15 | 13 |
| 15- PRACTICE | 145 | 27 | 118 | 4 | 8 | 7 | 8 | **0** | 18 | **65** | 25 | 8 | 2 |

**Table-4**
**The top 15 collocates of the noun 'theory' occurring immediately prior and subsequent to the node word in the ESP corpus**

| Collocates | Total | Left | Right | L5 | L4 | L3 | L2 | L1 | R1 | R2 | R3 | R4 | R5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1-THE | 50 | 25 | 25 | **1** | **10** | 6 | 2 | 6 | 4 | 9 | 4 | 6 | 2 |
| 2-OF | 39 | 16 | 23 | 2 | **0** | 2 | **11** | 1 | 8 | **0** | 3 | 8 | 4 |
| 3-ORGANIZATIONAL | 24 | 17 | 7 | **0** | 2 | **0** | **0** | 15 | **0** | 4 | **0** | 3 | **0** |
| 4-A | 19 | 9 | 10 | 2 | 1 | **0** | 1 | 5 | **0** | 6 | 1 | 1 | 2 |
| 5-TO | 19 | 7 | 12 | 2 | 2 | 3 | **0** | **0** | 3 | **4** | 3 | 2 | **0** |
| 6-AND | 18 | 3 | 15 | **0** | **0** | 1 | 1 | 1 | **10** | 1 | 2 | 2 | **0** |
| 7-IS | 17 | 6 | 11 | 4 | 2 | **0** | **0** | **0** | **7** | 1 | 1 | 2 | **0** |
| 8-IN | 17 | 12 | 5 | 2 | 2 | 1 | **5** | 2 | **0** | 2 | **0** | 1 | 2 |
| 9-THAT | 12 | 2 | 10 | **0** | 1 | **0** | **0** | 1 | **3** | 2 | 2 | 2 | 1 |
| 10-BEHAVIOR | 14 | 6 | 8 | 1 | 1 | 4 | **0** | **0** | **0** | **0** | 5 | **0** | 3 |
| 11-PERSONALITY | 8 | 1 | 7 | **0** | **0** | **0** | **0** | 1 | **0** | 6 | **0** | **0** | 1 |
| 12-PROBLEMS | 7 | 4 | 3 | 2 | 1 | **0** | 1 | **0** | **0** | **0** | 3 | **0** | **0** |
| 13-BASED | 6 | 3 | 3 | **0** | 1 | 1 | 1 | **0** | **0** | 2 | **0** | **0** | 1 |
| 14- RESPONSE | 5 | 4 | 1 | **0** | **0** | **0** | **0** | 4 | **0** | **0** | **0** | **0** | 1 |
| 15-PSYCHOLOGICAL | 5 | 5 | 0 | **0** | **0** | **0** | **0** | 5 | **0** | **0** | **0** | **0** | **0** |

One can come up with some fascinating points when he or she compares tables-3 and 4. First of all, only 9 of the collocates of the word *theory* were the same across the SS and SS ESP corpora, 8 of which were function words. The rest were unique to each corpus. The only content-word collocate which was shared between the corpora was *personality*, and the others were absolutely different. Secondly, apart from the fourth collocate, namely, *in*, the other shared collocates were in varied descending orders between the corpora. For instance, while *the* was the second collocate, *and* the third collocate, and *personality* the twelfth collocate in the SS corpus, they were the first, sixth, and eleventh collocates in the SS ESP corpus respectively. What is more, there were remarkably different L5-to-R5 positions for the highest and the lowest occurrences of the shared collocates in the corpora. To illustrate this point,

although the least attractive positions for the collocates *of* and *the* were R2 and L1 respectively in the SS corpus, they were L4 and L5 in the SS ESP corpus. However, the positions with the biggest number of occurrences for *a*, *to*, and *personality* in the SS corpus were L1, L2, and L1 respectively; it was R2 for the three of them in the SS ESP corpus. These difference go to reveal that Iranian SS ESP did not match the conventions of original texts concerning the appropriate use of the noun *theory* with its collocates.

## Conclusion

This study made endeavors to juxtapose the collocates of two nouns and their collocates positions prior and subsequent to the node words in two specialized corpora of soft sciences and ESP materials to scrutinize whether or not these collocates are used

in similar or different ways in ESP textbooks which are taught in Iranian universities compared to original textbooks. Having counted the frequencies of the collocates in 10 different possible positions immediately prior and subsequent to the word, we could obtain solid evidence to claim that the nouns *group* and *theory* possess different positions for their collocates in the SS and ESP corpora. The findings of the present study lend great support to the lexical priming theory put forth by Hoey[5]. According to this theory, a word becomes cumulatively loaded with the contexts and co-texts in which it is encountered, and our knowledge of it includes the fact that it co-occurs with certain other words in certain kinds of context. Therefore, a word is primed in the context of its occurrence for its collocates and the positions of those collocates. That is why the nouns *group* and *theory* have a number of different collocates in original psychology and sociology texts compared to Iranian ESP textbooks. The same also applies to the positions of occurrences for those similar collocates which occupied different positions in our corpora.

Bearing in mind the findings of the comparisons, we can draw a number of conclusions about the differences between collocations in original SS textbooks and Iranian ESP texts. First of all, it is revealed that in Iranian ESP textbooks of psychology and sociology compared to their original texts, a word such as *group* and *theory '*tend to collocate with some different and special words in general, and some different and special function and content words in particular. Secondly, even though some of the function words are shared, almost all of the content words are not the same across the corpora. Additionally, this study concludes that the collocates of a noun may not occur in the same descending order with respect to the node word in ESP texts of those disciplines. Still another variation we came up with is that collocates of a word can occupy different most and least favorite positions in ESP materials of soft sciences disciplines. Probably, the major conclusion this study arrives at is that the occurrence of collocates of a word might be position specific. In other words, a collocate tends to occur in or avoid some specific positions. This difference was conspicuously observed in the comparisons made throughout the current study.

## References

1. Bartsch S., *Structural and functional properties of collocations in English: A corpus study on lexical and pragmatic constraints on lexical co-occurrence.* Tubingen: Gunter Narr Verlag **(2004)**

2. Firth J.R., *Papers in linguistics.* London: Oxford University Press **(1957)**

3. Hoey M., *Textual interaction*, London : Routledge **(2001)**

4. Genji P. and Yaghoubi Doust M., Examining the effect of the social confidence. *Research Journal of Recent Sciences,* **2(9),** 25-28 **(2013)**

5. Hoey M., *Lexical priming: A new theory of words and language.* London: Routledge **(2005)**

6. Hoey M., Why grammar is beyond belief, In J.P. Van Noppen, C. Den Tandt and I. Tudor (Eds.), *Beyond: New perspectives in language, Literature and ELT. Special issue* of *Belgian Journal of English Language and Literatures*, **1**, 183–96 **(2003)**

7. Hoey M., Lexical priming and the properties of text. In A. Partington, J. Morley and L. Haarman (Eds.), *Corpora and discourse.*, 385–412, Bern : Peter Lang **(2004)**

8. Pecorari D., Formulaic language in biology: A topic-specific investigation. In M. Charles, D. Pecorari, and S. Hunston (Eds). (91-104). London. Coninuum, **(2009)**

9. Biber D. and Finegan E., *Sociolinguistic perspectives on register*, Oxford : Oxford University Press **(1994)**

10. Atkins S., Clear J. and Ostler, N. Corpus design criteria. *Literary and Linguistic Computing,* **7(1)**, 1–16 **(1992)**

11. Scott M.R., *WordSmith* : *Software tools for Windows*, Oxford : Oxford University Press **(2010)**

12. Bianchi F. and Pazzaglia R., Student writing of research articles in a foreign language : Metacognition and corpora, In R., Facchinetti (Ed.), *Corpus linguistics : 25 years on.*, (259-287), New York : Rodopi **(2007)**

13. Sadia K. and Muhammad Z., Humanitarian Intervention : A New Perspective, *Research Journal of Recent Sciences,* **3(1),** 97-102 **(2014)**

14. Biber D., Corpus-based analyses of discourse: Dimensions of variation in conversation. In V. K. Bhatia, J. Flowerdew and R. H. Jones (Eds.), *Advances in discourse studies,* (100-114), London : Routledge **(2008)**

15. Biber D., Conrad S. and Reppen R., *Corpus linguistics: Investigating language structure and use*, Cambridge: Cambridge University Press **(1998)**

16. Bora A., Science Communication through Mass Media, *Research Journal of Recent Sciences,* **1(1),** 10-15 **(2012)**

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

18. Koosha M., *English for the students of guidance and counseling*,Tehran: SAMT Publications **(2010)**

19. Rastegarpour H., *English for the students of psychology*, Tehran : SAMT Publications **(2009)**

20. Pooya A., Barfoei H.R., Kargozar N. and Maleki F., Relationship between Emotional Intelligence and Conflict Management Strategies, *Research Journal of Recent Sciences,* **2(7),** 37-42 **(2013)**

21. Moshfeghi F., *English for the students of social sciences*, Tehran : SAMT Publications **(2009)**

22. Moshfeghi F. and Afghari A., *English for the students of social sciences*, Tehran : SAMT Publications **(2010)**

**23.**   Khan Y.D., Ahmad F. and Khan S.A., A Survey on use of Neuro-Cognitive and Probabilistic Paradigms in Pattern Recognition. *Research Journal of Recent Sciences,* **2(4),** 74-79 **(2013)**