LEXICAL COLLOCATIONS (VERB + NOUN)
ACROSS WRITTEN ACADEMIC GENRES IN ENGLISH

Eser Ordem[1], Erdogan Bada[2]
[1]Adana Science and Technology University, Turkey
[2]Cukurova University, Turkey

Abstract:

The dominance of syntactic studies in linguistics has caused lexis and grammar to be perceived as two distinct categories. With introduction of the paradigm of cognitive linguistics, the studies in syntax have been replaced by those in lexis and concepts. Semantics has come to the fore through the studies in cognitive linguistics, and there has been a trend from syntactic studies to lexical ones. In addition to research in cognitive linguistics, construction grammar has also emphasized the continuum between lexis and grammar. With the emergence of corpus linguistics, the studies regarding the continuum between lexis and grammar have gained momentum, and thus studies of collocations have been theorized. Early studies of collocations have focused on only lexis and disregarded grammar. However, in the process, the studies have also incorporated grammar as well, and this view supports the idea that each word has its own grammatical properties. Therefore, lexis and grammar should be studied on the same continuum because there is a continuum between these two categories rather than a discontinuum. Within the framework of this paradigm, this study focused on verb+noun lexical collocations across the health, physical and social sciences in the written academic genre and analyzed these lexical collocations through the frequency and chi-square analysis. The study aimed to search for commonalities and differences between the verbs with their collocations. The results showed that there were more similarities and relationship between the health and physical sciences, while the social sciences indicated a significant difference compared to the other two. The study found 165 common verbs used across the three sciences. 12 verbs among the 165 verbs were found to be candidate's verb+noun lexical collocations as prototypes.

Keywords: Corpus linguistics, collocations, construction, prototype, written academic genre

Introduction

The realization of formulaic language, specifically collocations, can be traced back to pre-Chomskyan period (Saussere, 1916; Firth, 1951). However, the prevalence of syntactic studies downgraded the importance of lexis. Since computational linguistic and corpora studies were still immature, the importance of lexis and formulaic language had been long ignored. Collocation studies have been resuscitated with the emergence of corpora studies inasmuch as they provide a large amount of data empirically (Evert, 2007). New theories of formulaic language and lexicon have been prevalent with the contributions of construction grammar and corpus linguistics. Shifting from generative grammar to formulaic language has altered perspectives pertinent to domains of language (Wray, 2002). It has been often emphasized that language, whether spoken or written, is composed of prefabricated routines and fixed expressions. One of the subcategories of formulaic language is collocations, mainly made up of grammatical and lexical units. Howarth (1998) classifies collocations as lexical and grammatical units and explicates that ''lexical collocations consist of two open class words (verb + noun, adjective + noun), while collocations between one open and one closed word are grammatical'' (p.27). The studies of lexical collocations in particular have been prolific in recent decades resulting in approaching even the term 'collocation' from different perspectives and distinct definitions. However, it is still one of the most controversial topics in linguistics although it is often defined as 'a relationship between lexical items that regularly co-occur' (Carter, 1998, p.163). Even early linguists such as Saussere (1916), Bloomfield (1933) and Firth (1951) recognized the importance of collocations with similar approaches and definitions. In the same way, formulaic aspect of collocations was emphasized by other linguists as well (Hymes, 1962; Bolinger, 1976; Fillmore, 1979).

Subsequent to the diagnosis of importance of collocations, computational lexicographers (Sinclair, 1991, 1996) have empirically used collocations in their studies. These kinds of applications have led to the emergence of corpus based collocation dictionaries (Sinclair, 2004, 2005). However, it still remains a problem to determine which two words regularly co-occur in a text since one can encounter different kinds of collocations at different levels. It is quite important to make distinctions between collocations and apply the right statistical analysis while extracting collocations. Since different researchers have reached different conclusions even about the collocations of the same word, a closer look at the nature of collocations through the help of corpus linguistics is highly needed.

In the last few decades, studies of collocations have advanced in four main courses. Firstly, Firthian perspective of collocations refers to the predictability of co-

occurrences with the mutual expectancy of collocations (Evert, 2007). As Firth (1957) puts it, ''you shall know a word by the company it keeps'' (p.179). Sinclair (1991, 1995) also depicts *collocations* as a combinaton of two words that are biased to act together using the term in Firthian sense. Sinclair (1991) refers to *collocation* as "the occurrence of two or more words within a short space of each other in a text"(p.170). In this sense, Sinclair's studies are often referred to as Neo-Firthian (Evert, 2007). Secondly, the term '*collocation'* has also been used as phraseological or semi-compositional (Cowie, 1981, 1994; Hausmann, 1989). In the phraseological approach, *collocations* are syntactically related. Thirdly, Hoey (2005) regards collocations as psychological associations between words. Lastly, an eclectic view related to collocations is shared by Bartsch (2004) who describes collocations as "lexically and/or pragmatically constrained recurrent co-occurrences of at least two lexical items which are in a direct syntactic relation with each other" (p.76).

Although collocation studies are on the dramatic increase, Hoey (2005) still views definitions and properties of collocations as vague and rather complex, and therefore adds that ''lexis is complexly and systematically structured and that grammar is an outcome of this lexical structure'' (p.1). New perspectives to collocations studies still remain to be developed. Since the term 'collocation' covers a very heterogeneous collection of combinations of lexical items, it is important to produce alternatives and solutions to this problem. Another problem with collocations is that this is taken into consideration without the analysis of current linguistic theories. The previous research did not focus on lexical collocations from the perspective of certain linguistic theories. Rather, those studies viewed collocations as only a subpart of vocabulary. In the current studies except a few, collocations are still, to a large extent, related to only lexis.

Grammar and lexis were not handled on the same continuum in language studies for a long time. It is only in recent times that studies on grammar and lexis have begun to appear. Therefore, it seems quite crucial to approach the issue under the umbrella of a certain linguistic theory. As most of the studies are based on frequency effect, it is also necessary to employ other statistical tools in order to be able to reach sound results.

The main aim of this study is to investigate verb+noun lexical collocations across written academic genres under the theory of construction grammar and collostruction analysis by using corpora. We aimed to use Evert's (2007) empirical definition of collocations and statistical analysis in order to extract verb+noun collocations, and Gries' analysis (2012) of collostructions. Another aim of this was to analyze lexical collocations closely with the analysis of construction grammar. By applying construction grammar analysis, we intended to reach prototypes of some certain lexical

collocations and collostructs across the written academic genres. Thus, we sought responses to the following questions:

1. What verb+noun lexical collocations (collostructions) can be observed across academic genres?
2. How are these lexical collocations (collostructions) constructed from a constructionist grammar view?
3. Is it possible to discover prototypical lexical collocations (collostructions) according to the academic genre?

The first question intends to seek an answer to the types of verb+noun collocations across different written academic disciplines. The second question aims to analyze these collocations from constructionist and collostructionist perspective. The last question purports to find out whether prototypical lexical collocations can be extracted and elicited from the distinct academic disciplines.

Methodology

Since this study aims to describe and analyze lexically-initiated verb + noun lexical collocations across written academic genres, and intends to discover whether there are prototypical verb+noun collocations and collostructions across and within three different sciences, health, physical and social sciences and to examine these collocations in terms of construction grammar, thus the scope necessitates establishing a design considering corpus-driven approach together with theoretical analysis. Hunston and Francis (2000) refer to this kind of methodological design as below:

Potentially, then, we have two competing (or complementary) sets of generalizations arising from a corpus, one that depends on entirely on frequency of co-occurrence and is able to be generated by computer software alone, and one that is more interpretative and demands the input of a human researcher. How these sets of generalizations might differ from each other, and the implications of such difference are topics that have yet to be explored (Hunston & Francis, 2000, p.27).

This definition of the methodology in language studies can be regarded as a referential point of this study. Therefore, the bidirectional perspective of the methodology enables the study to look into lexical collocations in a richer context (Biber et al., 1998; Fillmore, 1992; Leech, 1992). The database of this study was formed from 249 research articles from 44 journals of health, physical and social sciences. Table 1 shows the steps followed in conducting the study.

Table 1: Research Type and Stages of the Study

Stages	Process	Research type
Stage 1	Selection of articles from journals	Corpus-based approach
Stage 2	Formation of corpora from 249 research articles from 44 journals of health, physical and social sciences.	Corpus-based approach
Stage 3	Conversion of corpus into text format	Corpus-based approach
Stage 4	Automatic generation of frequency lists	Descriptive analysis
Stage 5	Selection of meaningful lexical collocations manually	Corpus-based approach
Stage 6	Application of statistical analyses across and within corpora (Fisher`s exact test)	Quantitative corpus analysis
Stage 7	Checking prototypical lexical collocations	Descriptive and Interpretative
Stage 8	Analysis of prototypical lexical collocations through construction grammar	Interpretative

Three different types of software were used in order to reach reliable results. The first software used in this study was concordance that provided the basic results (Watt, 2012). This software does not carry out detailed inferential statistics but offers basic descriptive statistics. Counting words, making word lists and word frequency lists, full concordances, choosing pick lists, using multiple input files are among the functions of this software. The second software utilized was Antconc that offers a better service because Antconc provides multi-layered results composed of clusters, concordance plot and basic statistical measurement. The basic statistical tools in Antconc are log-likelihood, average value and clustering. Although it does not present a detailed statistical measurement, it was used for the basic statistical results. The third software was Wordsmith, a relatively sophisticated and integrated corpus software program used for text processing and extracting verb+noun lexical collocations descriptively and inferentially (Scott, 2010).

Coding of Verb+Noun Collocations

In this study, only verb+noun collocations were handled. However, even verbs and nouns alone denote intricate variations. Although a wide range of criteria could be used owing to the complexity of verbs and nouns, only certain criteria were determined in this study:

    a.   Verb+noun collocations do not necessarily have to be transitive verbs. Not all verbs that take nouns have to be transitive.

        Ex 1:   *It seems a problem.*

    b.   Verbs that are used in passive constructions and take infinitive forms were ignored.

Ex 2:   Reflection has been found to compromise.

c.   Verbs that have to take some prepositions when they co-occur with a noun were added.

Ex 3:   Women can deal with stress more easily.

d.   Verbs that are combined with a particle and form phrasal verbs were ignored.

Ex 4:   We found out that men resort to violence more.

e.   Nouns that are followed by infinitive and gerund were removed.

Ex 5:   *We intended to focus on different criteria.*

f.   Nouns that are followed by noun clause were classified as verb+noun collocations since noun clauses are extensions of nouns, and it is important to determine noun clauses in construction grammar. Thus, noun constructions were added to verb+noun collocations.

Ex 6:   Previous research has found that people tend to consume fast food.

g.   Verbs that are transitive but have not taken any kind of collocations or collostructs were ignored.

Ex 7:   People know.

Written Academic Corpora

The corpora for this study were retrieved from internationally recognized, electronic journals research articles (RA). A corpus of 249 research articles (116 for health: 84 for physical, and 49 for social sciences) included 1,217.197 words. Each science type was planned to have the similar number of words. Therefore, the number of articles varied but the number of the words for each science remained similar (see Table 2). Recent articles published between 2009 and 2011 were chosen.

Only professional texts were chosen from the journals of three mainstream sciences to gain an insight into the analysis of across and within disciplines.

Table 2: The Overall Data of the Texts

Science type	Number of disciplines	Number of research articles	Years	Total words
Health science	20	116	2009-2011	405,753
Physical science	14	84	2009-2011	405,751
Social science	10	49	2009-2011	405,693
Totals	44	249	2009-2011	1,217.197

Table 2 indicates that the number of the words in each genre was rendered almost equal so that results that are more reliable could be obtained between and across the genres. The number of the disciplines varied because each discipline has a different number of pages and words. However, the number of the words remained similar. The disciplines for each genre are shown in Table 3.

Table 3: Disciplines Chosen for the Corpora in Three Distinct Sciences

Health science	Physical science	Social science
Anatomy	Agriculture-Plant sciences	Literature
Anesthesiology	Astronomy	Anthropology
Bacteriology	Bioengineering	Education
Brain-Neuroscience	Botany	Gay and lesbian studies
Cardiology	Chemistry	Law
Cell Biology	Chemical and Materials	Philosophy
Dentistry	engineering	Political Science
Dermatology	Civil Engineering	Psychology
Endocrinology	Environmental Sciences	Recreation and Sports
Gastroenterology	Geology	Sociology
Genetics	Marine Science	
Geriatrics	Mechanical Engineering	
Immunology	Meteorology and Climatology	
Internal medicine	Physical Geography	
Nephrology	Physics	
Ophthalmology		
Pediatric		
Physiology		
Psychiatry		
Radiology		

These texts were transformed into text format in order to create an electronic corpus of 1,217.197 words. Lexical collocations, specifically verbs, were extracted from the corpus. Since the aim of this study was to analyze verb+noun lexical collocations, other word classes were excluded. The classification of the collocations was done in accordance with the operational definition.

Results and Discussion

The overall descriptive results and the summary statistics of the three disciplines are shown in Figures (1.1.), (1.2.) and (1.3.)
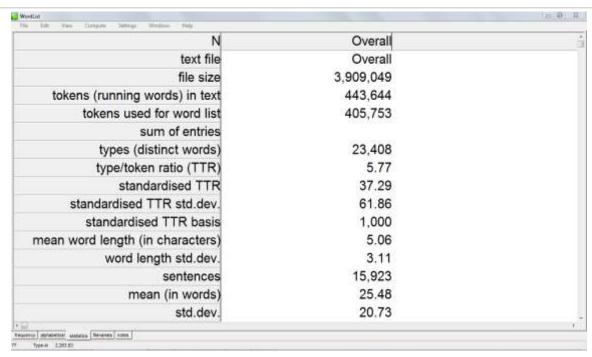
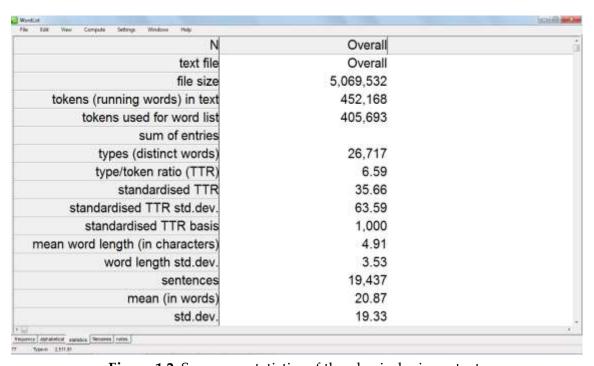**Figure 1.1:** Summary statistics of the health science texts

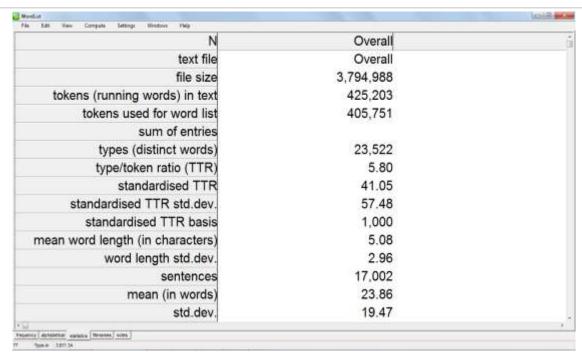**Figure 1.2:** Summary statistics of the physical science texts

Figure 1.3: Summary statistics of the social sciences texts

As indicated in Figures (1.1.), (1.2.) and (1.3.), the number and ratio of types and tokens across the three genres were intended to be balanced so that the sampling could be reliable. The ratios of types and tokens were 5.77% in health sciences, 5.80% in physical sciences and 6.59% in social sciences. The relatively slight ratio difference in social science stems from the nature of the interpretative aspect of social sciences. After the summary statistics of the words was given, the frequency of the verbs that co-occurred with nouns was given in Table 4.

Table 4: Descriptive Statistics of Verbs According to the Tokens

Academic genre	Total words	Verbs with collocates	%
Health science	405,753	8740	2.15
Physical science	405,751	7298	1.79
Social science	405,693	12206	3.00
Total	**1,217.197**	**28244**	**2.32**

The percentage of the verbs in Table 4 shows a similar variation. The percentage of the verbs in social science is the highest (3.00%), while physical science forms the lowest percentage (1.79%). Health science accounts for only 2.15% of the verbs. Table 5 exhibits the ratio of verbs considering the types.

Table 5: The Overall Statistical Results of Verbs According to the Types

Academic genre	Total words	Collocational verbs	%
Health science	23.408	724	3.09
Physical science	26.717	556	2.08
Social science	23.522	920	3.91
Total	**73.647**	**2190**	**2.98**

It can clearly be seen from Table 5 that the collocational verbs in the social sciences account for the highest percentage (3.91%), whereas the verbs in the physical sciences constitute the lowest percentage (2.08%). The percentage of the verbs in the health sciences is only 3.09%. The total percentage of the verbs in terms of types is 2.98%. The academic genres were compared considering the common verbs extracted from the corpora. The results were obtained through descriptive statistics. The number of the common verbs across the three genres was 165. The percentages and ratio of the common verbs in proportion to total verbs between and across academic genres are given in Table 6.

Table 6: Descriptive Statistics of Common Verbs across the Three Genres

Academic genres	Collocational verbs	Common verbs	%
Health sciences	714	165	23.10
Physical sciences	556	165	29.67
Social sciences	920	165	17.93
Total	**2190**	**495**	**22.60**

Table 6 shows that common verbs used in the three genres composing 22.60% of total verbs, which imply that 77.40% of all are different from each other. The ratio of common verbs in the health sciences is 23.10%; physical sciences, 29.67%, and social sciences, 17.93%. Since the verbs and their collocates in the social sciences showed more variations, the ratio of common verbs was lower than those of health and physical sciences. In order to look into the relationship and significance level of the 165 common verbs across the three genres, a chi square analysis was conducted.

Table 7: Overall Results of the Chi-Square Analysis of the Common Verbs in Three Academic Genres

Academic Genres	n	value	df	p
Health-physical sciences	165	628	164	.000
Health–social sciences	165	1307	164	.000
Physical-social sciences	165	912	164	.000

Table 7 shows that there is a significant difference between the 165 verbs across the health, physical and social sciences (p<0.00). Hyland (2004) argues that each academic discipline is unique in that it uses different textual collocations. In line with this observation, our results support Hyland's argument, although similar verbs are used across the three genres. The similarity is, however, more striking between the health and physical sciences. The social sciences showed more variation in all the data.

Considering the research questions, it can be said that the findings regarding the first research question showed that similar verbs were used across the three academic genres: health, physical and social. However, these verbs showed some variation in terms of the collocates they attracted. Collocates in the social sciences showed more variation compared to those in the health and physical sciences. The number of verbs taking collocates was more limited in the health and physical sciences. As for the possible answer to the second question, the results showed that the verbs in written academic genres tended to occur with constructions besides only simply co-occurring words. Almost each verb was seen to have its collostructional properties. It was found that there were no pure verb+noun collocations in their pure and naïve form. The most frequently used verbs with their collostructions showed a similar result in several studies as well (Thompson & Ye, 1991; Hyland, 1999, 2000; Hyland & Tse, 2005). In the three academic genres, one of the strongest collostruct was found to be *that-clause* collostruct. This finding is important in that Goldberg (2006) stresses the importance of the frequency and entrenchment of a specific construction. The last question in the study intended to find out whether the data could produce some prototypicalities similar to those in linguistics and psychology.

The results of the study showed that prototypes existed in the social context of written academic prose. In general, 165 common possible prototypical verbs were detected, although statistically there seemed no significant relationship between the genres. Out of these 165 common verbs, 12 most frequent and most common verbs across the three genres were seen to have prototypical features at high frequencies. As the degree of the frequency decreased, the variation of the verbs increased. This result is also supported by Hyland and Tse (2007) stressing that only 8% and 10% of the words show similar frequencies across different genres, and in terms of technical vocabulary, only 5% of the running words indicate similarities implying that genres show 'discursive variability'(p.251). It is not surprising that only a small percentage of the data show similarities because each sub-discipline produces different combinations. Therefore, Hyland and Tse (2007) approach academic vocabulary list with caution by insistently stating that these kinds of results may refer to the misrepresentation of academic literacy. Psychological explanation of conceptual combinations and linguistics explanation of collocations have shown that it is a thorny issue to find prototypes at the

level of collocations (Murphy, 2004; Hyland & Tse, 2007). Hyland and Tse (2007) in their study conclude that it may be pedagogically misleading for learners to direct them to 'overarching, universally appropriate teaching items' (p.251).

Recommendations

This study has revealed that similar verbs with their collocations across written academic genres might be followed by advanced foreign and second language users so that their academic writing and publication goals can be accomplished. Since each genre requires certain conventions that each member of this genre is supposed to comply with, learners are also expected to attend this community with full competence. Teachers should help learners gain awareness of the fact that knowledge is socially constructed within particular domains, and thus this line of thinking is reflected to academic writing as well. This basic theoretical background in the minds of teachers can motivate learners to pay attention to certain constructions in a certain genre.

More practically, learners need to be aware of not only common verbs used but more importantly of the collocates each verb attracts because the main competence in writing a professional article in a specific genre requires noticing certain collostructs in this very particular discipline (Hoey, 2001, 2004; Hyland, 2008). Hyland (2008, p.561) suggests that each learner should be trained in a 'genre approach' by teachers who are supposed to regard texts as a dynamic 'social interaction' rather than only a sequence of verbs given in a list. In parallel with this explanation, this study recommends teachers to show the similarities and difference in using collocates. Teachers can direct their attention to specific genres so that they can help learners notice lexico-grammatical patterns in academic writing rather than present a list of verbs or nouns.

This study showed that teaching writing is beyond listing only similar content words because each genre is specially and socially constructed and compromised (Hyland, 2007). It is important for both teachers and learners to discover and develop genre-specific corpora for themselves elaborately, and work on these constructions together. Thus, Hyland (2007, p.251) stresses the fact that 'discursive' similarity as well as variability should be noticed and detected by learners. In this sense, teacher educators should introduce and guide teachers and learners into genre-oriented theory and pedagogy presupposing that learners shall write only in socially constructed domains, and learners should bear in mind that they are liberal and can be creative only within constraints in order to attend the world of socially determined and constructed meanings in academic writing because each genre refers to a particular social world with certain patterns of language (Hoey, 2005; Hyland, 2007). The study has got significant implications for English language learning and teaching, particularly specific

to academic writing in that while introducing academic texts to learners, teachers have a reservoir of available data of collocations which they can put into the utilization of language users while producing an academic text. This availability is bound to facilitate the process of writing in general.

In terms of classroom application, teachers and learners have new roles in language teaching and learning because they can constitute their own corpus in the classroom so that they can extract their own collocations and reach reliable generalizations over examples and exemplars. Before learners are asked to write about an academic topic, as a warm-up activity forming a corpus in a two-three week period might prepare learners to use the target language according to the specific topic or genre they are supposed to write. Unless learners are entrenched and enriched by rich data of corpus, deviant forms will be inescapable. Teachers should show learners how to prepare an effective corpus instead of merely giving them hundreds of examples through a concordancer. However, a concordancer can be used to check whether any used collocation in a classroom setting is written or uttered by native speakers. Learners should be able to revisit and recheck the data that they have extracted and studied. Selective attention of learners may differ from each other in that each learner may attend to different data. Therefore, learners can work together in order to share the data they have chosen during the compilation and selection of lexical collocations (Lewis, 1998). This process will give learners the chance to negotiate the meaning of the data together, which might reinforce learning. By doing so, teachers can give learners the feeling that they are responsible for their own learning, and they learn to be independent while learning a language.

Another implication for ELT is that material writers may have to review their definition of lexical collocations because lexical collocations should also embrace collostructions as well. Material writers should not treat lexis and grammar as separate. Rather, they should show language learners that grammar and lexis can be learnt concurrently (Lewis, 1998; Howarth, 1998). Material writers, in this sense, can help this paradigm change in language learning settings take place. If material preparation contains grammar-lexis activities, then learners will be able to perceive language as holistic and integrative rather than dichotomic.

In terms of testing in ELT, testers should not measure grammar and lexis separately. Rather, they should prepare exams that allow learners to reflect their knowledge of collostructions as well. Since lexical collocations have syntactic functions in language production, it is important to direct learners to focus on these collostructions by developing certain tests containing both collocates and collostructions instead of asking only the meaning of a certain word. It should be borne in mind that each lexeme has its own intrinsic properties that should be perceived by

learners. Therefore, testers should gain an awareness of this new paradigm change in language studies.

As a negative implication of this study, it can be said that language is constantly changing, and the data they have collected may change over the years. In addition, being obsessed with fixed expressions may lead learners not to use their creativity in language. Foreign language learners might be able to use their creativity and make contributions to the target language they learn. Therefore, coming up with creative collocations by foreign or second language learners should not be regarded as something negative. Rather, these creative collocations or collostructions should be perceived as a contribution to the field. Instead of labeling these creative collocations as errors, mistakes or deviances, it is better to treat them as possibly acceptable because each new collocation is a candidate to be a part of language. In this sense, language learners should be encouraged to make use of corpus data and to use their own creativity.

The main issue in corpus studies is that the data collected are increasingly becoming larger and larger. Researchers are incessantly faced with a vast amount of data. The main question as to what researchers can perform through such big data remains quite critical and crucial. As this study dealt with limited amount of data, in the future, by using a larger amount of data with more sophisticated and more easily accessible software tools, some significant questions can be addressed in applied and corpus linguistics. By comparing results of different corpora composed in various research centers, more findings that are robust can be obtained. One of the main aims of corpora studies has been to put intuition aside. However, the amount and comparison of larger amount of data obtained through corpus studies can provide more reliability, validity and objectivity. Verbs alone are rather complicated and intricate. The occurrence of verbs with nouns and nominal phrases is more subtle since verbs are made up of distinct and various layers, and range from single word to idiomatic units. In the future, several steps can be followed:

a. carrying out detailed semantic and syntactic classification and analysis of verbs and nouns;
b. extracting collostructs of verb+noun collocations;
c. conducting more sophisticated statistical analysis to measure the strength of collocates and collostructs;
d. constituting specialized corpora for particular academic genres;
e. integrating the findings in corpus linguistics with the studies in psychology so that some links between conceptual combinations and colllocations can be established, which may help researchers understand human mind at an abstract level;

f. being in cooperation with foreign language studies so that foreign and second language users can make use of findings of obtained data, (vi) performing experimental studies in foreign language settings after introduction of written academic genre corpora is made;

g. performing corpora studies to show that language can be understood on a continuum of lexis and grammar, and last but not least;

h. extending the definition and meaning of collocation and explore new uses and functions of collocations through corpora studies;

i. researching prototypes and lexical priming through the results of these kinds of data-based studies.

References

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

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

3. Bloomfield, L. (1933). Language. New York: Henry Holt.

4. Bolinger, D. (1976). Meaning and memory. Forum Linguisticum I, 1-14.

5. Carter, R. (1998). Vocabulary: Applied linguistics perspectives. London: Routledge.

6. De. Saussure, F. (1916). Nature of the linguistics sign. In C. Bally & A. Sechehaye (Eds.), Cours de linguistique générale, McGraw Hill Education.

7. Evert, S. (2007). Corpora and collocations. Berlin: Mouton de Gruyter.

8. Fillmore, C., J. (1979). On fluency. In C. Fillmore, J. D. Kempler & S.Y. Wang (Eds.), Individual differences in language ability and language behaviour (pp.85-101). New York: Academic Press.

9. Fillmore, C. (1992). Corpus linguistics or computer-aided armchair linguistics. In J. Svartvik (Ed.), Directions in Corpus Linguistic, Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991 . ACM, Berlin, New York: Mouton de Gruyter.

10. Firth, J. (1951). Modes of meaning. Papers in Linguistics, 5, 190–215. Firth, J. (1957). Papers in Linguistics. Oxford: Oxford University Press.

11. Firth, J. (1957). Papers in Linguistics. Oxford: Oxford University Press.

12. Golberg, A. (2006). Constructions at work: The nature of generalization in language. Oxford: Oxford University Press.

13. Gries, S.Th. (2011). Corpus data in usage-based linguistics: What's the right degree of granularity for the analysis of argument structure constructions? In M. Brdar, M. Žic Fuchs, & S. Th. Gries (Eds.), Cognitive linguistics: convergence and expansion, (pp. 237-256). Amsterdam & Philadelphia: John Benjamins.

14. Gries, S.Th. (2013). Data in construction grammar. In G. Trousdale & T. Hoffmann (Eds.), The Oxford handbook of construction grammar (pp.93-108). Oxford: Oxford University Press.

15. Gries, S.Th. & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus based perspective on alternations. International Journal of Corpus Linguistics, 9, 97-129.

16. Gries, S.Th. (2012). Collostructions. In P. Robinson (Ed.), The Routledge Encyclopedia of Second Language Acquisition (pp.92-95). London & New York: Routledge, Taylor & Francis Group.

17. Hausmann F. J. (1989). "Le dictionnaire de collocations". In F.J. Hausmann, O Reichmann, H.E.Wiegand., L. Zgusta (Eds.), Wörterbücher : Ein internationales Handbuch zur Lexicographie, Dictionaries, Dictionnaires (pp.1010-1019). Berlin/New-York, De Gruyter.

18. Hoey, M. (2001). Textual interaction. London, Routledge.

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

20. Hoey, M. (2005). Lexical Priming: A new theory of words and language. London: Routledge, Taylor and Francis Group.

21. Howarth, P. (1998). Phraseology and second language proficiency. Applied Linguistics, 19(1), 24 – 44.

22. Hunston, S. & Gill F. (2000). Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English. Amsterdam: John Benjamins.

23. Hyland, K. (1999). Disciplinary discourses: writer stance in research articles. In C. Candlin, & K. Hyland (Eds.), Writing: Texts, processes and practices (pp. 99-121). London: Longman.

24. Hyland, K. (2000). Disciplinary discourses: Social interactions in academic writing. London: Longman.

25. Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. Journal of Second Language Writing, 16 (3), 148-164.

26. Hyland, K. (2008). Genre and academic writing in the disciplines. Language Teaching, 41 (4), 543-562.

27. Hyland, K. & Tse, P. (2005). Hooking the reader: A corpus study of evaluative that in abstracts. English for Specific Purposes, 24 (2), 123-139.

28. Hyland, K. & Tse, P. (2007). Is there an "academic vocabulary? TESOL Quarterly, 41 (2), 235-254.

29. Hymes, D. H. (1962). The ethnography of speaking. The Hague: Mouton.

30. Leech, G. (1992). Corpora and theories of linguistic performance. In J. Svartvik (Ed.) Directions in Corpus Linguistic, Proceedings of Nobel Symposium 82, Stockholm, (pp.4-8). ACM, Berlin, New York: Mouton de Gruyter.

31. Lewis, M. (1998). Implementing the lexical approach: Putting theory into practice. Hove: Language Teaching Publications.

32. Murphy, G. L. (2004). The big book of concepts. Cambridge, MA: MIT Press.

33. Scott, M. (2010). Word Smith tools version 5.0. Oxford: Oxford University Press.

34. Sinclair, J. (1991). Corpus, concordance, collocation. Oxford: Oxford University Press.

35. Sinclair, J. (1996). The search for units of meaning. Textus, 9(1), 75-106.

36. Sinclair, J. (2004). Trust the text: Language, corpus and discourse. London: Routledge.

37. Sinclair, J. (2005). Corpus and text: Basic principles. In M.Wynne (Ed.), Developing linguistic corpora: A guide to practice (pp.1-16). Oxford: Oxbow Books.

38. Thompson, G. & Ye, Y. (1991). Evaluation in the reporting verbs used in academic papers. Applied Linguistics, 12, 365-382.

39. Watt, R.J.C. (2012). Retrieved from http://www.concordancesoftware.co.uk/index.html on 17th June, 2012.

40. Wray, A. (2002). Formulaic language and the lexicon. Cambridge: Cambridge University Press.

41. Wray, A. (2008). Formulaic language: Pushing the boundaries. Oxford: Oxford University Press.