



Kahraman, M. & Subasi, G. (2022). An Analysis of Verb-Noun Combinations in High Frequency Verbs in Argumentative Essays of Turkish ELT Students: The Case of “Make” And “Do”. *International Online Journal of Education and Teaching (IOJET)*, 9(1). 584-603.

Received : 18.10.2021
Revised version received : 28.10.2021
Accepted : 28.12.2021

AN ANALYSIS OF VERB-NOUN COMBINATIONS IN HIGH FREQUENCY VERBS IN ARGUMENTATIVE ESSAYS OF TURKISH ELT STUDENTS: THE CASE OF “MAKE” and “DO”¹

(Research article)

(corresponding author)

Mehmet Kahraman <https://orcid.org/0000-0002-9789-5642>
mkahraman144@hotmail.com

Gonca Subasi <https://orcid.org/0000-0001-7049-5940>
ELT Department, Anadolu University, Eskişehir, Turkey
gongcas@anadolu.edu.tr

Biodata(s):

Mehmet Kahraman, Ph.D., is an English teacher. His research interests are language teaching, applied linguistics, writing, and data-driven learning.

Gonca Subaşı is an Assistant Professor in the ELT Department at Anadolu University, Turkey. Her research interests include teaching writing skills, vocabulary acquisition, corpora, affective factors in language teaching, language testing and evaluation, and language teacher education.

Copyright © 2014 by International Online Journal of Education and Teaching (IOJET). ISSN: 2148-225X.

Material published and so copyrighted may not be published elsewhere without written permission of IOJET.

¹ This paper is based on the doctoral dissertation titled ‘An Analysis of Verb-Noun Combinations in High Frequency Verbs in Argumentative Essays of Turkish ELT Students: The Case of “Make” and “Do”’ of the first author.

AN ANALYSIS OF VERB-NOUN COMBINATIONS IN HIGH FREQUENCY VERBS IN ARGUMENTATIVE ESSAYS OF TURKISH PRESERVICE TEACHERS: THE CASE OF “MAKE” and “DO”

Mehmet Kahraman
mkahraman144@hotmail.com

Gonca Subasi
goncas@anadolu.edu.tr

Abstract

The present study aims to reveal the use of high-frequency verbs “*make*” and “*do*” when they occur in a verb+noun combination in the argumentative essays of Turkish learners of English. In this context, the present study investigated the grammatical and semantic patterns and erroneous productions in the learner corpus. The investigation made use of a learner corpus and LOCNESS as the reference native corpus for comparison purposes. The findings showed that there were some similarities and dissimilarities among two corpora in terms of grammatical and semantic properties. Based on the findings of the current study and previous studies, the present study shared theoretical and practical implications particularly for language teaching settings.

Keywords: Word combinations, high-frequency words, collocations, corpus

1. Introduction

High frequency verbs have certain characteristics which are common across the languages. They dominate different semantic fields, have equivalent matches in most of the languages, have both universal and language specific meanings and potentially create problems for the learners (Altenberg and Granger, 2001). High-frequency verbs are problematic for learners although they are learnt at the very early stages of language teaching. The problem might be attributed to learners’ negligence of those verbs assuming that they are completely learnt (Hugon, 2008). According to Nation (1990), a learner’s level of vocabulary knowledge might differ when receptive and productive aspects are regarded. Also, knowing a word in a productive sense requires knowing it in full aspects such as spelling, pronunciation, grammatical patterns, frequency, context of use and possible collocations. In this regard, high-frequency verbs are tricky for the learners especially when they are in combination with other words since core meaning of those verbs will not help with productive use of language. Therefore, some learners feel safer with them and ignore the properties of context and collocates, while some others avoid using them when it comes to delexical senses (Altenberg & Granger, 2001; Källkvist, 1998; Nesselhauf, 2005; Sinclair, 1991).

Studies of high-frequency verbs have been of great interest especially in collocational uses; and learner language has been compared to native language in order to reveal the similarities and dissimilarities or developmental stages (Altenberg & Granger, 2001; Juknevičienė, 2008; Nesselhauf, 2003; Wray, 1999). It is possible to see different types of lexical collocations studied in the field. Although there are various word combinations, verb-noun collocations have drawn more attention due to its higher frequency and effect size in communication (Chan & Liou, 2005; Howarth, 1998; Nesselhauf, 2005); and difficulty in learning (Howarth, 1998).

Since collocational knowledge is regarded as an essential indicator of mastery in L2 (Sinclair, 1991; Wray, 1999), contrasting the learner language to native speakers' authentic productions is carried by employing corpora representing two parties. Being one of the most remarkable studies, Altenberg and Granger (2001) examined the uses of "make" by comparing Swedish and French learners of English to a native corpus by using a categorization. The categorization involved both semantic and grammatical factors at the same time. One of the results that the study revealed was that learners and the native corpus differed in their use of "make" in delexical sense and causative pattern. Some other studies (Babanoğlu, 2014; Hugon, 2008; Kim, 2015) as well-made use of the categorization proposed in the study. However, investigating the semantic and grammatical differences between the learner language and native corpus separately might produce more fruitful results. Also, considering the language specific difficulties, investigating more than one high-frequency verbs at once might better illustrate the overall picture. For Turkish learners, for instance, using "do" or "make" is often confusing (Öztuna, 2009). Therefore, investigating the similarities or dissimilarities between the productions of Turkish learners of English and that of native speakers of English is needed.

2. Literature Review

Collocations and Corpus Linguistics

Traditional lexicographic approaches have depended on manual collection and indexing of the language data, which is not an easy way of analysing big bulk of natural data. Fortunately, corpus linguistics has made it easier to study high volume of data with more solid empirical basis. Hence, collocations have been easier to deal with thanks to the concordance lines provided by corpus linguistics tools (Biber, 1993).

Corpus linguistics is a methodological approach which contributes to the language variation and use significantly. What makes it significant are the greater generalizability and validity it offers, which would not be much possible otherwise. Since corpus linguistics deal with the actual language use in texts it is empirical. Also, though they make use of computers extensively for the analysis, corpus linguistic studies can be both quantitative and qualitative as interactive techniques are possible during the analysis. Moreover, corpus-driven approach in corpus linguistics might also yield linguistic parameters that are not yet recognized by linguistic theories (Biber, 2010).

Collocations are one of the mostly discussed language patterns in corpus studies. According to Aisenstadt (1979), all word combinations are either idioms or non-idiomatic phrases. The non-idiomatic phrases include restricted collocations and free phrases. Cowie (2001) argues that word combinations are divided into semantic combinations and pragmatic combinations. The former consists of collocations and idioms, while the latter consists of proverbs and routine formulae. The common ground of these two categorizations is the emphasis on the semantic aspect of the combinations, which is the basic difference between frequency-based approach and phraseological approach. According to phraseological approach, collocations are different from free phrases, due to their restricted word choice. For example, considering the phrase "drink tea", one can easily substitute "tea" with *water*, *beer* ...etc. Nonetheless, considering the phrase "perform a task", substitution of "perform" (e.g.: *make*) is not similarly possible. Collocations are also different from idioms since the elements of the phrase "perform a task" still has semantic relation with "perform" and/or "task" individually. However, "blow the gaff" does not keep semantic cord with the individual elements of the phrase (Nesselhauf, 2005). Apart from the theoretical explications, there are many empirical studies on collocations, as well.

Delexical Verbs

A small group of transitive high-frequency verbs fall into delexical verbs. They take a noun as the object, but this noun has already a verbal sense itself. For example, consider “*make a suggestion*”. This verb-noun combination -more or less- gives the same meaning as “*suggest*” does. “*Make*” is in its delexical sense here, meaning that it loses a significant part of its lexical meaning, and “*suggestion*” is the main focus of the combination. Using a delexical verb preceded by a verbal noun (*make a suggestion*) instead of a simple noun (*suggest*) allows the speakers to give different conceptualizations of the given situation (Allan, 1998).

According to (Altenberg & Granger, 2001), language learners use high-frequency verbs in their delexical senses differently. They evidenced significant amount of underuse and misuse of delexical verbs in their written production.

Empirical Collocation Studies

Grammatical analysis of word combinations has not been much focused independently. In her highly comprehensive study, Nesselhauf (2005) investigated the syntactic patterns in collocations used by native speakers of English and German learners of English. The findings showed that learners’ use of collocations was syntactically more accurate when they are congruent between the target language (English) and native language (German). Hiltunen (1999) examined a huge size of corpus which was piled from several Early Modern English texts. The study showed various grammatical verb-noun phrases in detail, while shedding light on the verbs, verbal phrases, and phrasal verbs. In his study, Hiltunen (1999) defined four grammatical patterns that high frequency verb-noun combinations typically follow, which are also used in the present study. Lareo (2009) concentrated on the nouns used after “*make*” in verb phrases in scientific texts. The researcher compared a mathematic sub-corpus to a fiction text. The analysis revealed that, “*make*” combinations with specific nouns almost doubled the combinations with general ones, implying that academic vocabulary makes use of “*make*” more often than non-academic language. There are also some studies dealing with “*make*” only, due to its various uses in the written and spoken language. Hugon (2008) examined the French learners of English in terms of their use of “*make*” in different semantic and syntactic categories through a comparison of corpora. The results showed that the learner corpus showed a varied degree of accuracy in terms of semantic categorization; however, delexical combinations of “*make*” had many deficiencies.

There are also some other studies trying to illustrate solely semantic differences among collocations and combinations with high-frequency verbs, in particular. Macis and Schmitt (2016), for example, differentiated between literal, figurative and duplex meanings of the collocations. The study contributed to the language teaching by bringing useful insights to the collocations in terms of their semantic nature. As for high-frequency verbs, Liu and Lei (2009) investigated a native corpus and COCA. Then, they highlighted the deep semantic differences among the verb-noun combinations with “*make*”, “*take*”, “*do*” and “*have*”. Lantolf and Tsai (2018) focused only on “*make*” and “*do*” in semantic terms. They used SCBOAs to illustrate the deep semantic difference between “*make*” and *do*, then they asked the participant learners to draw their own SCOBAs for other combinations. The pre-test, post-test and delayed post-test scores revealed a significant improvement in using “*make*” and “*do*” verb noun combinations.

3. The Present Study

The present study focused on the use of high-frequency verbs “make” and “do” when they occurred in a verb+noun word combination. The term “word combination” was deliberately chosen because the present study did not distinguish between word combinations in terms of their restriction levels. In other words, free combinations, collocations, and idioms were all included in the current study. Regarding COCA (Davies, 2008), which consists of more than 600 million words, “do” is the third and “make” is the ninth most frequent verb in English. The present study dealt with “make” and “do” verb-noun combinations only, since they are frequently confused by most of the Turkish EFL learners (Öztuna, 2009), as many other L2 English learners (Altenberg and Granger, 2001).

Basically, a learner corpus was compared to a native corpus in this study. Since the comprehensive and authentic studies of language use cannot rely on small samples or anecdotes, the corpus-based approach was taken as a more feasible alternative to study large amount of natural data (Biber, Conrad and Reppen, 1998, p. 3). A comprehensive and authentic study of language use was necessary for creating a baseline since corpus-based approaches serve a transition to elaboration of better-quality learner input, and thus teachers and researchers are provided with a wider perspective of language as stated by Campoy, Belles and Gea (2010).

The main purpose of this study is to analyze “make” and “do” verb-noun combinations in argumentative essays of Turkish learners of English, who are at B1-B2 CEFR level. The natural uses of “make” and “do” verb-noun combinations in a learner corpus are investigated by taking a native corpus as the reference.

Within this framework, the study attempts to answer the following research questions:

1. What are the grammatical patterns in “make” and “do” verb+noun combinations produced by the learners and native speakers?
2. Which dictionary meaning of the verbs (*make* and *do*) do the learners and native speakers use?

4. Methodology

Considering the scope of the study, the present study, as well, was designed as a corpus-based analysis since it is a feasible way for describing and explaining variations and use in linguistic patterns. Corpus-based research does not aim to discover new linguistic features, rather it aims to discover how pre-recognized linguistic features govern the systematic patterns of use (Biber, 2010). To this end, two different corpora were used for the study. The first one is the learner corpus which contains compilation of essays by Turkish learners of English at intermediate level. The reference corpus is a native corpus containing essays written by native speakers of English. Corpus size and representativeness were taken into consideration before moving to the analysis.

The learner corpus examined was a compilation of argumentative essays written by the first-year university students studying in ELT department at a state university in Turkey between the years 2009 and 2019. The students were those successfully completing preparatory English year prior to their first year at their departments. Based on their end-term exam scores, the students were of B1-B2 CEFR proficiency level in English. A total of 166 essays, composed of 150,404 words, were included in the study. Although the clue words of the essays were various, the themes were education, technology and social life, health, media, and art. They

were extracted from an electronic assignment submission system following the ethical procedure required.

For comparison purpose, the study made use of Louvain Corpus of Native English Essays (LOCNESS). The corpus was compiled by Granger (1998). Only the argumentative essays written by native American university students were used in the present study so that the essay types and the number of words were compatible with the learner corpus. A total of 176 essays, composed of 149,574 words, were included in the study. Although the clue words show a wide range of variety, majority of the essay topics were education, technology and social life, health, media, art, sports, environment, politics, and monetary issues. The table below illustrates the writing topics included in the essays.

Table 1. *The corpora used in the study*

	Learner Corpus	LOCNESS (native)
Contributors	Turkish university students at intermediate level of proficiency	American university students who are native speakers of English
Essay Genre	Argumentative	Argumentative
Number of Essays	166	176
Total Corpus Size	150,404 words	149,574 words
Average Essay Length	906 words	850 words
Topics	Education Technology and social life Health Media Art	Education Technology and social life Health Media Art Sports Environment Politics and monetary issues

For RQ-1, learner corpus was analysed syntactically. For this purpose, Hiltunen's (1999) grammatical patterns for high frequency verb-noun combinations were used. The patterning instructed by Hiltunen (1999) was followed in the present study, due to its sole concentration on high-frequency verbs. This patterning was also used by Lareo (2009). In accordance with the framework, the extracted verb-noun combinations were categorized into the corresponding patterns (P) in Table 2.

Table 2. *The grammatical categorization of the verb+noun combinations*

Pattern Numbers	Patterns	Examples
P1	Verb + a / an + (Modifier / s) + Noun	<i>... make a dramatic change ...</i>
P2	Verb + (Modifier / s) + Noun	<i>... make money ...</i>
P3	Verb + the + (Modifier / s) + Noun	<i>... make the biggest mistake ...</i>
P4	Verb + (the) + (Modifier / s) + Nounplural	<i>... make the most delicious cookies ...</i>

Hiltunen's (1999) grammatical typology of the verb+ noun structures has been used to make comparisons among various periods in the history of the English language or to analyse and compare different written genre in English (Hiltunen, 1999; Koskenniemi, 1977; Lareo, 2009; Nickel, 1968; Visser, 1963). However, the current study takes this analysis as another way to express the similarity/ dissimilarity between the native and learner corpora, keeping the above-mentioned variables (historical period and genre) stable. This analysis aimed to reveal the structural differences among the native and non-native data. The erroneous productions of the learners were not excluded in the figures since their erroneous productions might still give an idea about the trend in grammatical patterning in their interlanguage.

For RQ-2, the online version of Oxford Learner's Dictionary, Cambridge Dictionary and Macmillan Dictionary were used. The definitions given in the dictionaries were synthesized by the researcher and one other language expert, who is an experienced English Language instructor. Thus, one single meaning categorization was created. The final categorization is presented in the table below (see Table 3).

Table 3. *Definitions of "do" and "make"*

No	Definitions of "do"	Definitions of "make"
1	to perform an action, activity, or job	to create/produce something
2	to clean something, or to make a place tidy	to cause to be formed by breaking, cutting, or tearing an object or by pushing one object into or through another
3	to have a good or harmful effect	someone performs the action referred to by the noun usually in fixed phrases
4	to study a subject	to arrange something
5	to spend an amount of time doing something	to earn/get money
6	to provide a service or product for customers to buy	to give the result of a mathematical calculation
7	to make something	to cause something to be successful
8	to move a particular distance or at a particular speed	to have right qualities for something/ to achieve something by reaching a necessary standard
9	to copy someone's voice, manner, or way of moving, in order to entertain people	to reach a place- to be able to be present at a particular event
10	to cheat someone	
11	to use illegal drugs	
12	to apprehend, arrest	
13	to visit a famous place as a tourist	
14	with some adjectives	

Various definitions of "make" and "do" were analysed and both native corpus and learner data were matched with the corresponding definitions given in the categories. This analysis is supposed to reveal whether more common, in the sense of simpler and more frequent, uses of "make" and "do" are used by the learners comparing to native speakers as argued by Hugon (2008) and Lennon (1996).

5. Results

The first research question aimed to find out whether the type of “make”/”do” collocations differs across the native corpus and learner data. In order to find out the distribution of aforementioned combinations, each “make”+noun and “do”+noun combination in both corpora was tagged according to the patterns suggested by Hiltunen (1999).

The distribution of the “make”+noun and do+noun combinations according to the patterns (Ps) is tabulated for each corpus (see Table 4 and Table 5).

Table 4. *Grammatical patterns in LOCNESS*

	P1	%	P2	%	P3	%	P4	%	TOTAL	%
make	79	28.6	80	28.9	32	11.5	85	30.7	276	100
do	18	18.1	48	48.4	13	13.1	20	20.2	99	100

Table 5. *Grammatical patterns in the learner corpus*

	P1	%	P2	%	P3	%	P4	%	TOTAL	%
make	31	18.2	67	39.4	6	3.5	66	38.8	170	100
do	12	11.7	42	42.1	7	6.8	40	39.2	102	100

Consequently, the results revealed that learners followed P2 and P4 more frequently in both do+noun and “make”+noun combinations. On the other hand, they used P1 and P3 less frequently again in both “do”+noun and “make”+noun combinations. As for natives, comparing their within group do+noun and “make”+noun combinations, they seemed to have followed divergent patterns. In do+noun combinations, they used P2 at the highest and P3 at the lowest frequency. In “make”+noun combinations, native students followed a fairer distribution among the patterns. They used P4, P2, and P1 at a similar rate. Yet, they still underused P3. The overall results showed that P2 was the pattern observed at the highest frequency in all four cases (both do+noun and “make”+noun combinations in both corpora), which is a commonality between two corpora. However, learners did not make a discrimination in their use of do+noun and “make”+noun combinations regarding the grammatical patterns they followed. They followed a similar grammatical patterning in their productions regardless of do+noun or “make”+noun combination. However, the findings in the LOCNESS corpus revealed that the native students had a tendency of adjusting the grammatical patterning considering the high-frequency verb in the combination.

The second research question aimed to find out which meanings of “make” and “do” are considered among various dictionary definitions. The analysis here only focuses on the “make” and “do” in combinations with a noun. In order to use as the reference, Oxford Learner’s Dictionary, Cambridge Dictionary and Macmillan Dictionary were examined due to their wide use among language learners. Considering the criteria of selecting “make”/”do”+noun combinations at the beginning of the analysis, irrelevant meanings such as causatives or phrasal verb structures were eliminated. Consulting with a language expert, who is an experienced English teacher and an academic at a university, given definitions of “make” and “do” were put into categories. Since it seemed more practical to assign numbers to each individual definition category (D) and explanations were clearer, the style of Macmillan Dictionary was

followed by taking the other dictionaries into consideration, as well. The final meaning categorization for “make” and “do” is presented below (see Table 6 and Table 7).

Table 6. *Meaning categories for “make”*

No	Definition	Example
D1	to create/produce something	<i>Let's make coffee.</i>
D2	to cause to be formed by breaking, cutting, or tearing an object or by pushing one object into or through another	<i>The rain made a hole on the ground.</i>
D3	someone performs the action referred to by the noun usually in fixed phrases	<i>We couldn't make a progress yesterday.</i>
D4	to arrange something	<i>I want to make an appointment</i>
D5	to earn/get money	<i>She makes 75 dollars a day.</i>
D6	to give the result of a mathematical calculation	<i>Five and two makes seven</i>
D7	to cause something to be successful	<i>His songs made the show.</i>
D8	to have right qualities for something/ to achieve something by reaching a necessary standard	<i>A good story makes a good film</i>
D9	to reach a place- to be able to be present at a particular event	<i>We cannot make the conference hall on time.</i>

Table 7. *Meaning categories for “do”*

No	Definition	Example
D1	to perform an action, activity, or job	<i>I do karate at the weekends.</i>
D2	to clean something, or to make a place tidy	<i>I will do the bedroom after lunch.</i>
D3	to have a good or harmful effect	<i>Sunlight will do good for your body.</i>
D4	to study a subject	<i>She is doing chemistry and biology.</i>
D5	to spend an amount of time doing something	<i>I did three years in New York</i>
D6	to provide a service or product for customers to buy	<i>We do sandwiches for parties.</i>
D7	to make something	<i>The paintings were done by him.</i>
D8	to move a particular distance or at a particular speed	<i>They did 500 km last night</i>
D9	to copy someone's voice, manner, or way of moving, in order to entertain people	<i>He did Michael Jackson at the party.</i>
D10	to cheat someone	<i>You paid £50 for this? You have been done!</i>
D11	to use illegal drugs	<i>She doesn't do drugs like other guys.</i>
D12	to apprehend, arrest	<i>He was done for shooting a guy at the bar.</i>
D13	to visit a famous place as a tourist	<i>We can go back to hotel after we do the museum.</i>
D14	with some adjectives	<i>I always advised her, but she did the opposite.</i>

Taking given definitions in Table 6, the distribution of the meaning categories within the extracted “make”+noun combinations are given in numbers and percentages below (see Table 8).

Table 8. *The distribution of the meaning categories across corpora for “make”*

No	Definition	LOCNESS		LEARNER	
		N	%	N	%
D1	to create/produce something	32	11.5	26	15,3
D2	to cause to be formed by breaking, cutting, or tearing an object or by pushing one object into or through another	2	0,7	0	0,0
D3	someone performs the action referred to by the noun usually in fixed phrases	182	65.9	129	75,9
D4	to arrange something	3	1.0	1	0,6
D5	to earn/get money	47	17.0	14	8,2
D7	to cause something to be successful	2	0.7	0	0,0
D8	to have right qualities for something/ to achieve something by reaching a necessary standard	8	2.8	0	0,0
D9	to reach a place- to be able to be present at a particular event	0	0,0	0	0,0
TOTAL		276	100	170	100

Ultimately, it can be concluded that both the natives and learners refer to the same meaning category (D3) as the primary meaning of “make”. Nevertheless, they still differ in the frequency count. Also, the secondary meanings referred to “make” are divergent. Learners tend to use “make” in D1 as the secondary meaning attributed to it, while the secondary meaning attributed to “make” is the one defined in D5 for the native students. Also, the learners do not present variety of meaning categories in their essays as much as native students do.

Taking given definitions in Table 7, the distribution of the meaning categories within the extracted do+noun combinations are given in numbers and percentages below (see Table 9).

Table 9. *The distribution of the meaning categories across corpora for “do”*

No	Definition	LOCNESS		LEARNER	
		N	%	N	%
D1	perform an action, activity, or job	65	65.6	96	94.1
D2	to clean something, or to make a place tidy	0	0.0	0	0.0
D3	to have a good or harmful effect	9	9.0	1	0.9
D4	to study a subject	0	0.0	0	0.0
D5	to spend an amount of time doing something	0	0.0	0	0.0
D6	to provide a service or product for customers to buy	1	1.0	0	0.0
D7	make something	4	4.0	2	1.9
D8	move a particular distance or at a particular speed	0	0.0	0	0.0
D9	to copy someone’s voice, manner, or way of moving, in order to entertain people	0	0.0	0	0.0
D10	to cheat someone	1	1.0	0	0.0
D11	to use illegal drugs	0	0.0	0	0.0
D12	apprehend, arrest	0	0.0	0	0.0
D13	to visit a famous place as a tourist	0	0.0	0	0.0
D14	with some adjectives	19	19.1	3	2.9
TOTAL		99	100	102	100

Overall, it can be concluded that both native students and learners take D1 as the primary meaning of “do” in their do+noun combinations. In fact, this category dominates over all other possible meanings in both corpora. On the other hand, D3 and D14 reveal a difference between two corpora. Native students consider those categories as the secondary meanings of “do”, while the same categories are hardly ever referred in learners’ essays.

6. Discussion

6.1. Grammatical Patterns of “Make” and “Do” Verb+Noun Combinations Produced by the Learners and Native Speakers

The current study used the grammatical patterning summarized by Hiltunen (1999) in order to show the similarity or dissimilarity among the native and learner corpora in terms of “make”+noun and “do”+noun combinations.

As detailed in the “Results” section, the most frequent pattern observed in both corpora (native and learner) and both combination type (“do”+noun and “make”+noun) was P2 (see Table 2 for details). On the other hand, it was also evidenced that the native students followed a relatively different patterning regarding the combination types whereas learners were stick to similar patterns (P2 and P4) in their combinations regardless of the combination type.

It should be noted that many other studies on the grammatical or semantic aspect of high-frequency verbs (Altenberg and Granger, 2001; Babanoğlu, 2014; Kim, 2015; Laporte, 2012; Lin, 2019) mostly used the categorization suggested by Altenberg and Granger (2001), so that they could present a single picture depicting the grammatical and semantic properties at once. The current study, within the framework of the first research question, focused on the grammatical aspect individually in order to find a more robust answer. Therefore, Hiltunen's (1999) categorization of the grammatical patterns was used unlike many other studies. In essence, the categorization used in the current study has already been used to identify the English language across genres and historical periods (Hiltunen, 1999; Koskenniemi, 1977; Lareo, 2009; Nickel, 1968; Visser, 1963). Learner language was not addressed in those studies. Thus, the current literature does not provide any previous study using Hiltunen's (1999) categorization comparing the native English and learner language in terms of collocations or high-frequency verbs. Only that of Lareo (2009) might be compared to the current one regarding its concentration on the verb+noun collocations. Comparing a Maths corpus to a fiction corpus, Lareo (2009) reported that P2 was the most or one of the most frequently used patterns in both corpora, which is in line with the results of the current study. When the learner corpus is taken into consideration in isolation, one can infer that the learners produced “do” and “make” combinations which were very similar to the native speakers in terms of grammatical patterning.

Interestingly, Hiltunen (1999) argues that P1 (see Table 2 for details) is the most common pattern today, but both the current study findings and that of Lareo (2009) yielded conflicting results with this argument. One possible explanation for this conflict might be that the genres analysed in the above-mentioned studies were different from each other. The current study made use of the argumentative essays written by university students in various topics and Lareo (2009) used fiction writings and science (Mathematics) articles. This diversity of the genres might explain the differing grammatical patterns in the corpora. At the same time, it might also imply that grammatical patterning is not governed by genre of the writing.

Taking only the learner data into consideration, the discussion might be further elaborated by recalling the specifications of P2. As stated above, P2 is a pattern in which one can use a noun without any definite or indefinite article. In fact, it is only possible for generic, abstract or non-countable nouns in English. Emphasising on the abstract nouns, Hiltunen (1999) explains this as an effect of French language on English in the course of time. Considering that the learners' L1 was Turkish, this might bring forward the issue of L1 effect on learners' verb+noun combinations. In Turkish, a verb+noun combination is possible with zero article as well as the accusative case (definite article "the") and indefinite article case. However, it is not acceptable in English.

E.g.: *Ali bir kek yaptı* *Ali made a cake*
 Ali keki yaptı *Ali made the cake*
 Ali kek yaptı **Ali made cake*

As can be seen in the above examples, Turkish learners of English might have produced some erroneous "make"/"do" combinations without any definite (a/an) or indefinite article (the) since it is acceptable in their L1. Thus, the frequency of P2 -Verb + (Modifier / s) + Noun-might have increased (Üstüenalp, 2013). The mismatch between the languages might account for the use of P2 pattern more than other patterns.

6.2. Semantic Distribution of "Make" and "Do" Verb+Noun Combinations Produced by the Learners and Native Speakers

The study findings revealed that both natives and learners assigned one common primary definition category (D) for each of "do" and "make" in their essays. For do+noun combinations, this common ground for both corpora were D1 - perform an action, activity, or job. This meaning category can be regarded as the core meaning of the verb "do". As for "make"+noun combinations, the highest frequency was observed in D3 - someone performs the action referred to by the noun usually in fixed phrases- in both native and learner corpora. Although native and learner corpora yielded seemingly similar results, the results are more remarkable in terms of differences. Firstly, it was observed that these common primary Ds were used relatively at higher percentages in the learner corpora and D1 and D3 (for "do" and "make" respectively) were so frequent in the learner corpus that the other meaning categories were hardly considered, which decreased the diversity in their productions. Native corpora, on the other hand, showed more diversity in terms of definitions attributed to "do" and "make" in their combinational uses. Secondly, while D1 for do+noun combinations can be regarded as the core meaning of the verb do, D3 for "make"+noun combinations are not the core meaning of the verb "make" but is a delexical (e.g., *make a judgement*) definition of it. Comparing the percentages of "make" productions within groups, one might, deceptively, argue that the learners showed more examples of delexical "make" in their writings. However, when the frequency counts are reconsidered, the figures show that native speaker used higher number of "make"+noun combinations either in total or D3 in particular. It means that native students showed more examples of delexical "make" in their writings while presenting diversity at the same time. Finally, some senses of the high-frequency verbs were used at a relatively lower frequency in the learner corpus. For example, D3 (i.e.: *do the biggest wrong*) and D14 (i.e.: *do the opposite*) for do, and D5 (i.e.: *make money*) and D8 (none) for "make" were rare in the learner corpus in comparison with the native corpus.

The results are partially or completely in line with some of the significant studies comparing the learner language and the native language in this regard (Allerton, 1984; Altenberg and Granger, 2001; Cobb, 2003; Hugon, 2008; Lennon, 1996; Sinclair, 1991). Although detecting

the pitfalls of the learner language and generalizing over the differences from the native English is not focused majorly, the current study still provided some evidence for the differences between the learner language system and native language within overlapping issues. In an effort to discover the semantic diversity produced by the learners in high-frequency verbs, the current study revealed that although “make” and “do” are typically learned at the beginning of the EFL instruction, the learners (mixture of pre-intermediate and intermediate proficiency) seemed to fail in presenting various dictionary meanings attributed to these verbs in their academic writing. In a more global perspective, Cobb (2003) stated that even advanced learners of English have difficulty in discovering the full phrasicon in English, and they tend to repeat the same phrases whereas the native speakers implicitly know it, and thus they show more diversity in their language use. The current study presented evidence for this statement in terms of semantic diversity. Also, Sinclair (1991; 79) argued that learners avoid using common words and instead “they rely on larger, rarer, and clumsier words which make their language sound stilted and awkward”. This tendency was particularly observable in two instances in the current study. In the case of “do”, the learners used D3 -to have a good or harmful effect- only once, however it was possible to find D3 9 times in the native corpus. Moreover, the learners unacceptably used the lemma forms of “give harm” instead of “do” harm, cause harm or harm”. Doing this several times, the learners decreased their D3 frequency while making their productions awkward. As for “make”, learners seemed to have preferred “earn” as an alternative to “make” in the sense of D5 -to earn/get money-. Although the same interchange was observed in the native corpus (7 occurrences), the learners used that alternative more often (11 occurrences). Therefore, it can be said that Sinclair’s argument was confirmed in the current study.

A similar claim was made by Lennon (1996). It was claimed that although learners have a broad idea of verb meaning, they have a more limited knowledge of some other important aspects, such as polysemy, semantic boundaries or collocational restrictions. Thus, their productions are mostly based on the core meaning of the verbs, though the verbs “make” and “do” have broader boundaries and collocational possibilities. As exemplified above, the current study shows that although learners are well familiar with the words “make” and “money” separately, they still prefer “earn money” at a higher frequency but “make money” at a lower frequency in comparison with the native corpus. A similar underuse of “make” in the sense of make money was also found in some other studies (Altenberg and Granger, 2001; Babanoğlu, 2014; Hugon, 2008). Accordingly, it can be concluded that learners do not prefer “make” as an alternative to earn in the sense of make money. As stated by Lennon (1996), the problem is not just a verb-choice error, in fact “earn money” in the current study is not erroneous at all. The essence of the problem is that learners have a great tendency of sticking to the core verb meanings and they are unable to extend their knowledge to the delexical usage of a verb. Although high-frequency verbs allow for various uses, learners still feel restricted to the core meanings of them. Nevertheless, it should be noted that Shaw (2001) reported contrary findings in this regard. Comparing the writings of Chinese learners of English and a combination of two native sub-corpora, the researcher evidenced that the learner corpus contained higher number of “make” in the sense of make money than the native corpora. In the current study and above-mentioned concordant studies dealt with Turkish, Japanese, Swedish and French learners of English, while Shaw (2001) examined Chinese learners in his research. Although, there is not an explanation for this contrast in Shaw (2001), the difference among the findings might imply an L1 effect or the language instruction.

As cited in Altenberg and Granger (2001), Allerton (1984) argues that although there are grammatical, syntactic and semantic restrictions which have already been defined, it is still worth considering some sort of arbitrariness in the selection of some words such as high-

frequency verbs. According to Altenberg and Granger (2001), learners might be aware of this arbitrariness, and thus they avoid using semantically unmotivated high-frequency verbs particularly when a high-frequency verb does not match with its L1 equivalent. The case can be exemplified with D8- to have right qualities for something/ to achieve something by reaching a necessary standard - of the verb “make”. This meaning category was not observed at all in the learner corpus while there were 8 occurrences of it in the native corpus. One plausible explanation of this divergence might be the fact that this sense of “make” in English does not make a similar sense in the L1 of the learners, Turkish. A similar motivation can account for some other differences across the corpora, such as underuse of “make” in the sense of D5 – make money- or “do” in the sense of D3 – do good/bad-. Therefore, it can be concluded that the current study confirmed the argument (Allerton, 1984) and explanation (Altenberg and Granger, 2001) by previous studies.

7. Conclusion

The study showed that when the grammatical patterns are regarded, although P2 and P4 were the most frequent patterns in both corpora, learners seemed to use almost the same portions of the patterns regardless of their being “make” or “do” combinations. However, the native corpus revealed somewhat varying degrees of frequency considering the difference between “make” and “do” combinations. It implies that learners do not distinguish between “do” and “make” in this regard and they apply the same grammatical patterning not considering these items individually. Also, the high frequency of P2 in the learner corpus could be attributed to use of erroneous zero article nouns under the influence of L1 (Turkish). When the semantic properties are concerned, the study revealed that although both “do” and “make” have various dictionary meanings, the learner corpus did not show variety in using them in combination with nouns. They tended to stick to core meanings of them. These findings imply that learners are not much aware of the collocational possibilities or extended meanings of “make” and “do”, which make their writing clumsier as argued by Sinclair (1991, p.79).

8. Implications

Considering the current study findings and the relevant previous studies in the field, one major conclusion to be drawn is that knowing a word involves collocational uses of the word as well its core meaning and grammatical properties attached to it in usage (Thornbury, 2002, p.16). This brings the implication that the teachers, learners, and material designers should have the awareness of collocations in the target language (Babanoğlu, 2014; Bıçkı, 2012; Cobb, 2003; Gilquin, 2007; Nesselhauf, 2004; Wray, 1999). Secondly, as stated by Lennon (1996), not only quantitative but also qualitative vocabulary gain should be addressed in language teaching. It means instead of continuously teaching new but undigested vocabulary, imperfectly acquired vocabulary items should be better consolidated so that the learners are able to use their pile of vocabulary productively (Cobb, 2003).

Considering the above-mentioned implications, the methodologies in ESL/EFL settings should put more emphasis on the literal and figurative meanings of the target vocabulary, not the core meanings only. Especially, extended meanings and collocational uses of the words should be explicitly taught since mere exposure has little or no effect in this respect (Nesselhauf, 2003). Since the learners failed to show productive variation both in grammatical and semantic aspects, the explicit teaching should keep a good balance of form-focused and meaning-focused activities. Also, in collocation teaching, congruency among the native language and the target language should be considered, as well. Some collocations the in the

target language can be directly translated into the native language keeping the same semantic property. However, it is not possible for incongruent ones. Thus, they create a big problem for the learners especially in productive sense. Therefore, the teachers should give priority to the incongruent collocations by explicitly contrasting the literal, figurative or register specific meanings in language teaching (Bahns, 1993). The current study evidenced that the learner corpus involved many occurrences of L1 translation strategies which did not work.

As for the classroom procedures, there are some practical implications with some caveats. Firstly, it should be remembered that classroom instruction might lead learners to produce grammatically correct but unidiomatic utterances due to the lack of sensitivity of collocational associates (Wray, 1999). Although there are some basic formulaic expressions covered in many course books, collocations with high-frequency verbs have a wider coverage in the language. It takes years for learners to learn these seemingly easy verbs, especially in delexical sense. As evidenced in the current study and many other studies, even upper-intermediate and advanced level learners cannot deal with high-frequency verbs when it comes to collocational uses. Therefore, the teachers should go beyond the course books and try to expose students to more real-life examples of the language. This is highly possible thanks to the concordance software nowadays. Especially, web-based concordance programs make it possible to observe a particular word with its numerous collocates in various authentic texts and evidently increase the vocabulary development (Akıncı and Yıldız, 2017; Conrad, 1999; Daskalovska, 2015).

Testing students' collocational knowledge is also important since it is closely linked to processing, comprehension and use of language (Almacioğlu, 2018). It should be noted that collocation knowledge was evidenced to be in correlation with both vocabulary score (Mutlu, 2015) and writing score (Hsu, 2007). Hence, testing collocation knowledge might give a novel and better understanding of learners' vocabulary and writing development, which is sometimes not much observable by repeating the same traditional tests in the classroom. Therefore, allocating time and energy for collocation cannot be regarded as an extra burden for teachers or students.

Ideally, textbooks and dictionaries should consider learners' native language while selecting the target collocations. Mainstream textbooks are, in nature, unable to consider numerous native languages in the world. Since one size-fits-all approach is not very helpful in collocation teaching, teachers should take more responsibility so that learners are exposed to the exclusively selected collocations rather than random ones. Alternating textbooks with corpus-based collocation teaching seems a very effective technique as stated above. As for dictionaries, Hugon (2008) suggested that they should provide the learners with more contextual information (formal/informal, frequency... etc.) about high-frequency verbs. Although it is labour-extensive for dictionary authors, finding such information in a dictionary, of course, would have benefits for the learners.

9. Limitations and Future Directions

One major limitation of the current study was on the proficiency levels of the students who are the contributors of the non-native learner corpus used in the present study. As mentioned above, the students were at B1-B2 level regarding their base passing scores at the end-term exams of English preparation class in previous year. However, it should be noted that since the learner data were compiled from year 2009 to 2019, the students contributing to the learner corpus in the present study might not be very homogenous in terms of proficiency level because the exam format changed a few times in this period and thus the student profile, too, might have changed in the course of time. Therefore, the proficiency level of B1-B2 might be mixed

with some higher achievers such as upper-intermediates or advanced students, though they are very limited in number. Secondly, the corpora used in the study are composed of argumentative essays only. Considering that written language or even academic writing is not limited to argumentative essays, further studies can include other genres of writing. Hence, broader picture of the use of word combinations can be observed.

This study focused on one single proficiency level. A further study can focus on more than one proficiency level such as A1-A2 and C1-C2 levels. Comparing two learner corpora among each other and to one native reference corpus might yield important results on the developmental factors in collocations with high-frequency verbs. Alternatively, the current learner group might be asked to write argumentative essays parallel to the current ones in terms of representativeness again in their fourth year at the department and their development in terms of collocations with high-frequency verbs can be observed in a longitudinal way.

The present study considered “make” and “do”, only due to the potential problem they create for Turkish learners. A further study can focus on other high-frequency verbs such as take, have, get... etc. The high-frequency verbs can be studied all together, individually, or selectively based on certain criteria. Also, the current study considered only verb+noun combinations regardless of their restriction level. A further study can focus on other types of combinations such as adjective+noun combinations. Even, a certain restriction level can be focused such as idioms, collocations, and free combinations.

References

- Aisenstadt, E. (1979). Collocability restrictions in dictionaries. *International Journal of Applied Linguistics*, 45(1), 71–74.
- Akıncı, M., & Yıldız, S. (2017). Effectiveness of corpus consultation in teaching verb + noun collocations to advanced ELT students. *Eurasian Journal of Applied Linguistics*, 3(1), 91–109.
- Allan, Q. (1998). Delexical verbs and degrees of desemanticization. *WORD*, 49(1), 1–17. <https://doi.org/10.1080/00437956.1998.11673876>
- Allerton, D. J. (1984). Three (or four) levels of word cooccurrence restriction. *Lingua*, 63(1), 17–40. [https://doi.org/10.1016/0024-3841\(84\)90030-5](https://doi.org/10.1016/0024-3841(84)90030-5)
- Almacioğlu, G. (2018). Examining Turkish ESL learners' receptive collocational knowledge. *Gaziantep University Journal of Social Sciences*, 17(3), 796–812.
- Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of make in native and non-native student writing. *Applied Linguistics*, 22(2), 173–194.
- Babanoğlu, M. P. (2014). A Corpus-based study on the use of make by Turkish EFL learners. *International Journal of Education and Literacy Studies*, 2(2), 43–47. <https://doi.org/10.7575/aiac.ijels.v.2n.2p.43>
- Bahns, J. (1993). Lexical collocations: A contrastive view. *ELT Journal*, 47(1), 56–63.
- Biber, D. (1993). Co-occurrence patterns among collocations: A tool for corpus-based lexical knowledge acquisition. *Computational Linguistics*, 19(3), 531-538.
- Biber, D. (2010). Corpus-based and corpus-driven analyses of language variation and use. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 159–192). Oxford: Oxford University Press.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. New York: Cambridge University Press.
- Bıçkıcı, A. (2012). Acquisition of English collocations by adult Turkish L2 learners. Unpublished Doctoral Thesis. Adana: Çukurova University, Institute of Social Sciences.
- Campoy, M. C., Belles, B., & Gea, M. L. (2010). Introduction to corpus linguistics and ELT. In M. C. Campoy, B. Belles, & M. L. Gea (Eds.), *Corpus-based approaches to English language teaching*. Continuum International Publishing Group.

- Chan, T. P., & Liou, H. C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, 18(3), 231–250. <https://doi.org/10.1080/09588220500185769>
- Cobb, T. (2003). Analyzing late interlanguage with learner corpora: Québec replications of three European studies. *Canadian Modern Language Review*, 59(3), 393–423. <https://doi.org/10.3138/cmlr.59.3.393>
- Conrad, S. M. (1999). The importance of corpus-based research for language teachers. *System*, 27, 1–18.
- Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, 28(2), 130–144. <https://doi.org/10.1080/09588221.2013.803982>
- Davies, M. (2008). The Corpus of Contemporary American English (COCA): 600 million words, 1990-present. <https://www.english-corpora.org/coca/>
- Gilquin, G. (2007). To err is not all: what corpus and elicitation can reveal about the use of collocations by learners. *Zeitschrift Für Anglistik Und Amerikanistik*, 55(3), 273–291.
- Granger, S. (1998). The computer learner corpus: A versatile new source of data for SLA research. In S. Granger (Ed.), *Learner English on computer* (pp. 3–18). New York: Routledge.
- Hiltunen, R. (1999). Verbal phrases and phrasal verbs in Early Modern English. In L. J. Brinton & M. Akimoto (Eds.), *Collocational and idiomatic aspects of composite predicates in the history of English* (pp. 133–165). Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Howarth, P. (1998). The phraseology of learners' academic writing. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 161–186). Oxford: Oxford University Press.
- Hsu, J. (2007). Lexical collocations and their impact on the online writing of Taiwanese college English majors and non-English majors. Online Submission.
- Hugon, C. (2008). High-frequency verbs : Starting block or stumbling block for advanced L2 communication? Insights from native and learner corpora. In G. Rawoens (Ed.), *Taal aan den lijve. Het gebruik van corpora in taalkundig onderzoek en taalonderwijs* (pp. 69–98). Gent: Academia Press.

- Juknevičienė, R. (2008). Collocations with high-frequency verbs in learner English: Lithuanian learners vs native speakers. *Kalbotyra*, 59(3), 119–127.
- Källkvist, M. (1998). Lexical infelicity in English: the case of nouns and verbs. In *Perspectives on lexical acquisition in a second language*. Lund University Press.
- Kim, S. J. (2015). The effect of learner proficiency and L1 transfer on the use of make by Korean EFL learners of English. *Language Research*, 51(1), 139–166.
- Koskenniemi, I. (1977). On the use of verbal phrases of the type ‘to take revenge’ in English Renaissance drama. *Poetica*, 7, 80–90.
- Lantolf, J. P., & Tsai, M.-H. (2018). L2 developmental education and systemic theoretical instruction: The case of English verb+noun collocations. In A. E. Tyler, O. Lourdes, M. Uno, & H. I. Park (Eds.), *Usage-inspired L2 Instruction. Researched pedagogy* (49th ed., pp. 29–54). Amsterdam / Philadelphia: John Benjamins Publishing Company. <https://doi.org/doi.10.1075/llt.49>
- Laporte, S. (2012). Mind the gap! Bridge between world Englishes and learner Englishes in the making. *English Text Construction*, 5(2), 265–292.
- Lareo, I. (2009). Make- collocations in nineteenth-century scientific English. *Studia Neophilologica*, 81(1), 1–16. <https://doi.org/10.1080/00393270802083067>
- Lennon, P. (1996). Getting easy verbs wrong at the advanced level. *IRAL - International Review of Applied Linguistics in Language Teaching*, 34(1), 23–36. <https://doi.org/doi:10.1515/iral.1996.34.1.23>
- Lin, C. (2019). Grammatical and lexical patterning of make in Asian learner writing: A corpus-based study of ICNALE. *The Southeast Asian Journal of English Language Studies*, 25(3), 1–15.
- Liu, D., & Lei, L. E. I. (2009). Teaching idiomatic expressions of make, do, have, and other common verbs. *The TESOL Encyclopedia of English Language Teaching*. <https://doi.org/10.1002/9781118784235.eelt0786>
- Macis, M., & Schmitt, N. (2016). The figurative and polysemous nature of collocations and their place in ELT. *ELT Journal*, 71(1), 1–10. <https://doi.org/10.1093/elt/ccw044>
- Mutlu, G. (2015). Knowledge and perceptions of collocations: The cases of Turkish EFL students and teachers. Unpublished MA Thesis. İstanbul: Marmara University, Institute of Educational Sciences.

- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston: Heinle Cengage Learning.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223–242.
- Nesselhauf, N. (2004). What are collocations? In D. Allerton, N. Nesselhauf, & P. Skandera (Eds.), *Phraseological units: Basic concepts and their application* (pp. 1– 21). Schwabe.
- Nesselhauf, N. (2005). *Collocations in a learner corpus* (N. Nesselhauf (ed.)). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Nickel, G. (1968). Complex verbal structures in English. *IRAL: International Review of Applied Linguistics in Language Teaching*, 6(1), 1.
- Öztuna, S. (2009). Effects of input flood and negative evidence on learning of make/do collocations: A study with seventh grade Turkish EFL students. Unpublished MA Thesis. Eskişehir: Anadolu University, Institute of Educational Sciences.
- Shaw, P. (2001). Investigating learner vocabulary : A possible approach to looking at EFL / ESL learners qualitative knowledge of the word. *IRAL - International Review of Applied Linguistics in Language Teaching*, 39(January 2001), 171–194. <https://doi.org/10.1515/iral.2001.001>
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Thornbury, S. (2002). *How to Teach Vocabulary*. London: Pearson Longman.
- Üstüenalp, İ. (2013). An analysis of verb-noun collocational error types and error sources in written narrative texts of students majoring in English. Unpublished MA Thesis. Mersin: Mersin Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Visser, F. T. (1963). *An historical syntax of the English language: Part one, syntactical units with one verb*. Leiden: EJ Brill.
- Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching*, 32(4), 213–231. <https://doi.org/10.1017/S0261444800014154>