

Understanding the Collocation Issues and Problems in Turkish Language

N. Tayyibe Ateş*

Faculty of Education, Selçuk University, Konya, Turkey

Corresponding author: N. Tayyibe Ateş, E-mail: tayyibeeken@gmail.com

ARTICLE INFO

Article history

Received: June 15, 2022

Accepted: September 08, 2022

Published: October 31, 2022

Volume: 10 Issue: 4

Conflicts of interest: None

Funding: None

Publication Note: This study was created through the doctoral thesis titled "Word Combinations in Narrations: Observations on Turkish Education".

ABSTRACT

Collocations make it possible to use and understand the language in an effective way. In this context, understanding collocation issues is important for both native speakers and language learners. This study aims at revealing certain determinations and problems with regard to the classification of word combinations in Turkish. For this purpose, descriptive analysis, content analysis, and category and frequency analyses have been employed herein. The sample of the study is composed of narrative (tale, story, and novel) texts selected through the cluster sampling method according to the level of 5th-8th grade students. Linguistic input of (+/-) 624,089 words has been entered into the research corpus as a result of the process involving the import of the texts into the electronic environment via a scanner, their digitization through conversion to text format, and their coding to be processed in Turkish language (ISO). The study involves the individual evaluation of word combinations in terms of their syntactic, semantic and pragmatic aspects. Then, the classification of these combinations has been applied as semi-restricted collocations, restricted collocations, figurative idioms, and pure idioms in terms of semantics, according to their structure whether it is syntactic dual or multiple. Generally, the semantic order of word combinations in narrative texts has been found to be as follows: Free combinations > semi-restricted collocations > restricted collocations > figurative idioms > pure idioms. The findings obtained reveal certain problems specific to the Turkish language with regard to the detection of word combinations, especially collocations, via programs. The study contributes to the literature as descriptions of word combinations in Turkish. More importantly it is thought that it will contribute to the development of literacy skills of students both in mother tongue education and foreign language learning.

Key words: Collocation, Word Combinations, Co-occurrence, Turkish Education, Literacy

INTRODUCTION

Collocations¹ appear when words are used in combination and convert into stereotypes without any semantic reason and obtain a nature of a specific indicator in conceptual terms as a result. The common aspects of various definitions created through different approaches to explain the concept of *collocation* in the literature are that these combinations should consist of at least two units and there should be a syntactic relationship between these units².

It is seen that collocation is considered from two different perspectives: First, the *statistically oriented/the frequency-based approach*³; the second one is the *significance oriented/phrasological approach*⁴.

The *statistically oriented approach* was first used by Firth (1951). Firth (1957) suggests that collocations consisting of words used in combination to a certain extent differ from other combinations in terms of frequency of use. Firth (1951, 1957) does not consider collocations as lexemes but believes that each word form is a separate collocation. Halliday (1961, p. 274), on the other hand, suggests that the *lexical item* can

be a morpheme, word, or phrase and includes all variations. Accordingly, collocation, which is defined on a more abstract plane, is independent of the inflectional suffixes (such as plural suffix and possessive suffixes) taken by the units. For example, in Turkish, the combinations of *kapı çal-* [to knock on a door] and *kapı çalın-* [a door to get knocked on] are not considered as two separate nodes, but as different allomorphs of the same unity. Halliday's approach differs from Firth's in that it treats collocations as lexical units rather than as lexemes. In addition, Halliday deals with collocations within the framework of the text. Accordingly, two units in two different sentences or even in different paragraphs can be considered to have collocation. Sinclair (1991, p. 170) defines collocation as two or more words *co-occurring* within a short interval of ± 4 words to the right and left of the *node* in a sentence⁵. For example, if the word *dog* is taken as a node in a sentence, in a sentence such as "s/he had not noticed for a long time that there was a barking dog in the garden of the next-door neighbor", the words *s/he, had, not, noticed, for, a, long, time, that, there, was, a, barking, in, the, garden, of,*

the, next-door, neighbor all have collocations with the word *dog*. In this example, the combination of *a* and *dog* is not determined as meaningful but the words *barking* and *dog* are determined as *meaningful collocations* since they are compatible in terms of semantics⁶. Stating that a single method is not sufficient to explain collocations, Sinclair (1991) proposes two different linguistic interpretation principles: *the open-choice principle* and *the idiom principle*.

Statistically-oriented studies involve the basic criterion as the units to be used together at certain distances. In significance-oriented approaches, on the other hand, for collocation, only being used side by side and frequency is not sufficient, requiring semantic criteria as well. The most important representative of the significance-oriented approach is Cowie (1981, 1988, 1991, 1994, 1998). Cowie (1988, 1994) divides word combinations into *composites* and *formulae* in general and classifies collocations, a sub-type of *composites*, with the criteria of *transparency* and *commutability/substitutability*. According to Cowie (1981), a collocation is by definition a composite unit which permits the substitutability of items for at least one of its constituent elements (the sense of the other element, or elements, remaining constant). Cowie (1981, 1988, 1994) argues that collocations are *transparent* and can be *substituted* in most cases with some restrictions.

Cowie (1994) entitles *free combinations* where substitutability is usual with constraints due to semantic properties. Combinations in which this restriction is somewhat or completely unreasonable are called *collocations*. For example, in the case of *çorba içmek* [*to drink soup*], the verb *iç-* [*to drink*] requires a noun that contains +liquid feature semantically in the noun+verb combination, the unit that can replace *çorba* [*soup*] must have this feature as well. Even though *soup* and *ashura* have a structure that can be consumed with a spoon, in Turkish, *ashura* is used together with the verb *ye-* [*to eat*] and *soup* with the verb *iç-* [*to drink*]. This restriction may be encountered to a certain extent at the semantic level. However, the option for substitution at the usage level is more restricted. Similarly, *boza* and *salep* are used in combination with the verb *to drink*. Even though one of them has a thicker consistency than the other and even though it is not entirely fluid, it is observed that the word that meets both concepts in Turkish is combined with the same verb. This situation, which is a reasonable case for *salep* to some extent, may be considered more unreasonable for *boza*. In this case, it can be said that the combination *boza iç-* [*to drink boza*] forms a collocation in Turkish. This restriction is not the result of the semantic features of the units that make up the combination but the result of the arbitrary consensus in the language. For instance, in case combinations such as *çorba ye-* [*to eat soup*] or *boza ye-* [*to eat boza*] in Turkish were consensually accepted, it could be said that the units are collocations independent of their semantic features.

It is observed that the topic of collocation in Turkish is generally handled as a corpus-based dictionary review (Özkan, 2007; Dedeoğlu and Şen, 2010; Özkan, 2010, 2011, 2014a, 2014b, 2014c; Tüfekçioğlu and Özkan, 2014) or within the scope of text-linguistic studies (Yağcıoğlu, 2002a, 2002b; Taşgüzel, 2004; Adalar, 2005; Çıkrıkçı, 2009; Torun,

2011; Ağca, 2020). In order to make determinations about the processing of word combinations and to obtain generalizations about language, first of all, the appearance of the information in the language must be described. The aim of this study is to address and evaluate the difficulties in determining collocational patterns in Turkish.

Research Questions

The research questions created in line with the subject and aims of the study are as follows:

1. What are the factors affecting the stereotypes of word combinations?
2. What kind of problems arise in Turkish in terms of detecting word combinations?

METHOD

Designed through descriptive and relational models, this study employs descriptive analysis, content analysis, category and frequency analyses⁷. Category analysis is one of the first techniques used in content analysis. Category analysis generally refers to dividing a particular message into units and then grouping these units into specific criteria. Frequency analysis is to reveal the frequency of appearance of units or items in numerical, percentage and proportional manner. This allows to understand the intensity and importance of a particular item. At the end of the frequency analysis, a classification is made based on the importance and frequency of the items. The validity of the study was ensured via *depth-oriented data collection* and in terms of data sources, via *triangulation*, and *expert review* methods⁸. Triangulation using more than one method to collect data on the same topic is a way of assuring the validity of research through the use of a variety of methods to collect data on the same topic, which involves different types of samples as well as methods of data collection. The expert review is an expert-based research method. Multiple experts analyze independently the test product seeking to identify the majority of its usability problems. In this study, two experts on Turkish Education and a linguist analyzed collocation categories independently.

The reliability of the study was ensured through the performance of *consistency and confirmation analyses* with feedback from various experts during the creation of data collection tools, data collection, and analysis stages⁹. Cluster and random sampling methods, which are among the probability sampling methods, were used. The texts were determined randomly from three narrative genres (tale, story, and novel) selected by cluster sampling.

Data Collection and Analysis

The database of the study was obtained from the texts of tales, stories, and novels for 5-8 grader primary school students. The texts were selected from among the frequently recommended and read texts at primary education level¹⁰. A database of ± 624,089 words, 204,000 (± 2%) from each text type, was created and the word combinations in the

database were evaluated according to their semantic and usage characteristics. The text types that make up the structure of the study and the number of words they contain are shown in Table 1.

The texts were transferred to the computer via a scanner, converted into TEXT format through optical character recognition software, digitized and coded to be processed in Turkish (ISO). Linguistic input of approximately (+/-) 624,089 words was provided to the corpus of the research. In the study, the token numbers were calculated via the AntConc 3.2.4w 2011 program. In the phase of data analysis, category and frequency analyses related to word combinations were performed. Word combinations are listed through the Microsoft Excel 2011 database and categorized in consideration of the contexts in which they are used.

Category Analysis

TS Turkish Corpus and Turkish National Corpus were used in order for clarifying the unique combinations reflecting the author's style and one-time uses in the database. In case a combination is encountered at least five times in the corpus, it is assumed that the relevant combination is used in the language, otherwise, it is removed from the list. Following this elimination process, the semantic content of the words was clarified via the TDK Contemporary Turkish Dictionary, and first of all, free combinations were separated on the basis of the distinctions determined by Cowie (1994). Units with many substitute options and combinations that can be brought together for logical or semantic reasons were evaluated within the scope of *free combinations* and were excluded from the limitations of the study. Other word combinations were classified into sub-types in line with Cowie's (1994) principles of *transparency* and *commutability/substitutability*. In this study, different from Cowie's (1994) classification, another distinction is made under the title of *semi-restricted collocations* in order to determine less restricted combinations that appear to be *free combinations* according to the *transparency* criterion but are close to *restricted collocations* according to the *substitutability* criterion. The combinations, which can be explained based on the internal structure of the concept in which they are collocated and whose lexical substitution options are more restricted compared to free combinations due to the internal structure of the concept, are classified as semi-restricted collocations. For example, in Turkish, the verb *sokmak* [to sting] in the combination of *arı sok-* ([a bee] to sting) is considered a verb specific to venomous animals such as *insects and snakes*, which means that the name of a venomous creature, *arı* (bee), in this case, is included in the verb's internal structure. Such combinations get combined within the framework of semantic prosody. Yet, the same verb is not used with another venomous animal, the *jellyfish*. In this context, it is classified as collocation due to the restriction in the lexical substitution option. According to the classifications made with semantic restrictions, it was determined that the majority of semi-restricted collocations were brought together on the basis of semantic prosody. For instance, since *water* is a liquid substance, it conceptually exhibits a fluid feature, so it can be logically

Table 1. Dispersion of the texts that make up the database according to their types

TEXT TYPE	Book Title - Author Name	Number of words	Number of texts	
TALES	Allı ile Fırfırı (Oğuz Tansel)	40,382	25	
	Altın Işık (Ziya Gökalp)	9,999	7	
	Az Gittik Uz Gittik (Pertev Naili Boratav)	48,269	48	
	Billur Köşk Masalları (Tahir Alangu)	65,634	14	
	Evvel Zaman İçinde (Eflatun Cem Güney)	21,538	10	
	Keloğlan (Tahir Alangu)	33,307	19	
	Sevdalı Bulut (Nazım Hikmet)	16,926	12	
	TOTAL	236,055	135	
	STORIES	Ay'ı Boyamak (Semih Gümüş)	22,970	20
		Dikkat Kırılacak Eşya (Semih Gümüş)	21,571	20
Eyvah Kitap (Mine Soysal)		19,644	35	
Keskin Naneli Öyküler (Tolga Gümüşay)		17,827	10	
Kırmızı Yoyo (Oktay Akbal)		18,573	16	
Odada Yalnız (Mine Soysal)		18,632	13	
Ötesi Yok (Suzan Geridonmez)		17,936	9	
Öykü Öykü Gezen Kedi (Zeynep Cemali)		32,044	8	
Sessiz Yürek (Necati Güngör)		17,886	7	
TOTAL		187,083	138	
NOVELS	Ankaralı (Zeynep Cemali)	25,183	13	
	Defne'yi Beklerken (Aslı Der)	24,986	22	
	Hazırlıksız (Tolga Gümüşay)	25,484	22	
	İçimdeki Ses (Gaye Boralıoğlu)	57,780	23	
	Sınıfın Yenisi (Behçet Çelik)	23,480	24	
	Yalancı Şahit (Müge İplikçi)	19,978	13	
	Yolun Başındakiler (Cemil Kavukçu)	24,060	17	
TOTAL	200,951	134		
GRAND TOTAL	624,089	407		

explained that it is used in combination with the verb *ak-* [to flow]. This restriction is also stated in the dictionary meaning of the verb *ak-* [to flow]. In the TDK Contemporary Turkish

Dictionary, the semantic content of the verb *ak-* [to flow] is explained as “liquid substances or very fine-grained solids to go from one place to another, liquids to head down, liquid substances to come out of a place, and a container or a place to leak the liquid in or above it”. In the explanations, the displacement of liquid substances, which generally exhibit fluid properties, is emphasized. In other words, there is a logical causality in the semantic internal structures of the two words in question and a requirement relationship arising from the nature of the concept. Some of such combinations may include a more restricted substitution option compared to others. For example, the verb *ak-* [to flow] can be used with many concepts that are liquid and fluid. However, due to its semantic restrictions, the verb *yağ-* [to fall, to rain] has an option of words such as *rain*, *snow*, and *hail* that do not offer many substitution options. The inherent restriction of the verb *yağ-* [to fall, to rain] when used with concepts such as *rain*, *snow*, *hail*, and *avalanche* can be explained within the framework of secondary causation when the same verb is used with the word *money*. Due to the fact that they are transferred, many verbs can gain dual causality as well as causality in different indexes and can be used with metaphors or connotations. In the study, word combinations were taken as the lemma, and quantitative analyses were carried out according to the types and sub-types of the combinations. Word combinations are divided into categories according to their structural (two-word or multiple combinations) and semantic features (*semi-restricted collocation*, *restricted collocation*, *figurative idiom*, *pure idiom*) according to the degree of restriction.

FINDINGS AND REMARKS

When the free combinations (*elma ye-* [to eat apple], *siyah göz* [black eyes], *siyah saç* [black hair]) and co-creations (*mavi zaman* [blue hour], *susuz yağmur* [dry rain]) in the texts were eliminated, 8,690 different word combinations consisting of *semi-restricted collocations*, *restricted collocations*, *figurative idioms* and *pure idioms* were found in the database of the study. This number denotes different word combinations that are considered a lemma, regardless of the frequency of use. Considering the raw frequency values, namely the number of tokens, there are a total of 35,630 word combinations in the texts.

In the study, the number of tokens was calculated via AntWordProfiler 1.4.0w (Anthony, 2013) and AntConc 3.4.3w programs. However, since Turkish is an agglutinative language, the suffixes on the word cause different inflections of the word in question to be counted as different words in such programs. Synonyms also pose a problem in this respect. For example, since the verb *de-* [to say] and the conjunction *de* cannot be labeled separately, these two words, which have two different functions, are detected as a single token by programs, thus resulting in some deviations in the numerical results. Since there is no wordlist for Turkish, the classification of these words into their types cannot be done via such programs, in other words, the analyses for this case do not always give accurate results. Therefore, in order to minimize such mistakes, the words and combinations calculated by the

programs used in the research should be reviewed. In the programs, the number of n-grams depicts the combination value consisting of n number of units in the selected range. Since combinations consisting of two or more words are the subject of this study, a query was made by multiplying the number of n-grams in order to determine combinations containing more than two words, so that the margin of error in the research was attempted to be minimized.

Since the determinations made via the programs take the degree of combination of words according to syntactic combination conditions as criteria, free combinations are frequently encountered in these determinations. Therefore, the elimination of these word combinations according to semantic criteria is possible through review and manual scan of the texts. Accordingly, for example, the frequency of two-word combinations detected by AntConc in this study was 390,315, while the frequency of two-word combinations detected by manual elimination was determined as 31,407.

It is observed that the most frequently used words in the texts are generally function words. This finding is in parallel with the word frequency rankings in Turkish National Corpus (TNC), Aksan et al. (2017) and Göz's (2003) word frequency rankings. However, in this study, since the function words were left out of the limitations, collocation analyses were made on the basis of the combinations formed by the content words in the texts.

Considering the probability of forming genre-specific collocations, two-word and multiple n-gram analysis was performed in order to reach meaningful combination structures. During the performance of these collocation analyses, certain usages specific to the text type were determined. For example, stereotype word combinations specific to the genre draw attention in tale texts. It has been observed that the word combinations in the tales are generally stereotype word combinations such as rhymes, reduplications and formulaic words, certain specific word combinations are preferred at the beginning and of the text (*bir varmış bir yokmuş*. [once upon a time.], *onlar ermiş muradına* [they have come to their will], *gel zaman git zaman* [a long time afterwards], etc.), and the descriptions of the characters specific to the tales (such as *dev anası* [mother giant], *dev baba* [father giant], etc.) form collocations. In addition to these, it is noteworthy that collocations formed with qualifiers like *üç* [three], *kırk* [forty], *bin* [thousand], *yedi* [seven], *ak* [white] such as *ak saç* [white hair], *ak sakal* [white beard], *al yazma* [red turban], *kara yazı* [black destiny] and exaggerated expressions (*kırk gün kırk gece* [forty days and forty nights]) are used, combinations reflecting religious (*namaz kıl-* [to perform salaah], *abdest al-* [to perform ablution]) and cultural elements (*düğün alayı* [wedding procession], *gelin hamamı* [bride bath]) are intense. Collocations usually co-occur in tale texts with intensive and reduplicative structures (such as *kapkara göz* [coal-black eyes], *kara mı kara göz* [eyes black as pitch])¹¹.

As shown in Table 2, according to their syntactic features, certain collocations are seen as two-word combinations, while others may contain three or more units. Most of such collocations may consist of two or three pairs of

Table 2. Examples of collocations in fairy tales

	Examples of collocations in fairy tales
Fairy tale rhymes	bir varmış bir yokmuş... [once upon a time...], onlar ermiş muradına [they have come to their will], gel zaman git zaman [a long time afterwards], etc.
Qualifiers and quantifiers in fairy tales	dev anası [mother giant], dev baba [father giant], ak saç [white hair], ak sakal [white beard], al yazma [red turban], kara yazı [black destiny], kara göz [black eyes], kapkara göz [coal-black eyes], kara mı kara göz [eyes black as pitch], kara baht [black destiny], kırk gün kırk gece [forty days and forty nights], etc.
Religious elements in fairy tales	namaz kıl- [to perform salaah], abdest al- [to perform ablution], etc.
Culturel elements in fairy tales	düğün alayı [wedding procession], gelin hamamı [bride bath], etc.

combinations, or there may be a single collocation in multiple structures. For example, the sequence *to pay close attention* can be decomposed into two collocations: *to pay attention* (a support verb construction) and *close attention* (an intensive, “Magn-like” collocation in ECD words, according to Mel’çuk’s terminology). Though close attention very often collocates with *pay*, this collocation can be encountered in other contexts: a rapid Google search yields for example *close attention is required, to recommend close attention, this issue needs close attention* (For detailed descriptions, see Tutin, 2008)¹². Among the collocations formed with the word *kara* [black], which is one of the most frequently used noun words, there are noun phrases formed through the combination of two words such as *kara baht* [black destiny], *kara göz* [black eye], *kara kız* [dark skinned girl], as well as triple collocation structures formed through the combination of two-word collocation structures such as *kara kara düşün-* [to brood over something], *kara yağız delikanlı* [dark skinned boy/young man]. These are generally characterized as multiple complex collocation structures in which it cannot be determined which part of the combination is the head and which part is the adjunct. Internally multiple collocation structures are also frequently encountered. *Kara kaplı kitap* [black book] is an example of multiple collocation structures that cannot be broken down into smaller parts internally. The collocation of *cehennem azabı çek-* [to suffer hell torment] actually includes two collocation pairs as well: *Azap çek-* [to suffer torment] and *cehennem azabı* [hell torment]. Even though the combination of *cehennem azabı* [hell torment] is often used with the verb *çek-* [to suffer], both units are used with other units in other contexts: *kabir azabı çek-* [to suffer the torment of the grave], *cehennem azabına dön-* [to turn into a hell torment], etc. In certain examples, complex and multiple collocation structures can be reduced to two-word combinations but in others, these multiple-word structures cannot be separated into units. There are some examples of collocations with most frequently used words (nouns and verbs) in the texts below in Table 3 and Table 4.

One of the most frequently used words in the verb type in the texts is the verb *çek-* [to pull, to suffer, to feel, to endure, etc.]. It is observed that this verb is generally used within the framework of the restrictions stated in the dictionary and used in combination with the words “*zorluk, yabancılık, acı*” [difficulty/hardship, alienation, pain] in the meaning of “to endure difficult situations”, with the word “*rutubet*” [moisture] in the meaning of “to take in, to absorb”, with the

word “*sorgu*” [interrogation] in the meaning of “to press someone to understand the background of something”, with the word “*saf*” [line] in the meaning of “to take back someone or something”, with the word “*set*” [barricade/barrier] in the meaning of “to set up any obstacle” and it is stereotyped with the inclination suffix when used with the word *interrogation* [sorguya çek-], and with the possessive and inclination case suffixes when used with the word *line* [safına çek-]. Regardless of the meaning of the word, the verb *çek-* is used in combination with the words *uyku* [sleep], *sıfır* [zero], *rol* [role], and *sine* [bosom]. It is observed that it forms an idiom by being stereotyped with the inclination suffix in its use in combination with the word *sine* [bosom] (*sineye çek-* [to put up] in Turkish), and it forms restricted collocations and figurative idioms with other words.

As another most frequently used verb, it is observed that the verb *al-* is used as a restricted collocation in the form of *satın al-* [to buy] within the framework of the restrictions specified in the dictionary, and as a figurative idiom in the form of *soluk al-* [to breathe in]. In addition to the two-word uses of the verb *al-* such as *ablukaya al-* [to make blockade], *aferin al-* [to win welldone], it is also observed that it is used in triple-word combinations such as *soluğu...DA al-* [to get (somewhere) in no time flat] as well.

In the study, the semantic distinctiveness of suffixes is a significant finding. The semantic difference of “*aklım al-*” [take their mind] and “*akıl al-*” [to consult] makes this issue clearly evident. While distinctiveness is provided by suffixes, in some cases, positive and negative semantic content is presented in the internal structure of the verb. An example to the types of word combinations that have negative semantic connotations in their internal structure is the combination “*akıl almaz*” [inconceivable]. The use of it in the form “*akıl alır gibi değil*” [it isn’t conceivable] can be given as an example to these structures.

In the program analysis of word combination, it can be observed that sometimes, problems occur with syntactic words. For example, in addition to the verb *al-*, other nouns and verbs starting with the noun *al*, which is synonymous with the word *kırmızı* [red], and even the letters “*a*” and “*l*”, cause uncertainty in terms of quantity and quality. The aforementioned types of ambiguity arising from the structure of Turkish are also frequently encountered in the TNC. Ambiguity types and frequencies classified by Aksan et al. (2011) within the scope of TNC Creation Project have been determined and rule-based clarification samples that can eliminate ambiguities have been presented. These

Table 3. Examples of collocations with frequently used verbs in texts

Frequently used verbs	Examples of collocations with frequently used words			
	Semi-restricted collocations	Restricted collocations	Figurative idioms	Pure idioms
çek- [to pull, to suffer, to feel, to endure, etc.]	zorluk çek- [to experience difficulty], yabancılık çek- [to feel alienation], acı çek- [to be in pain]	<i>rutubet çek-</i> [to be in moisture], <i>uyku çek-</i> [to sleep], <i>sıfır çek-</i> [get zero]	<i>set çek-</i> [to build a barrier], <i>sorguya çek-</i> [to interrogate], <i>safına çek-</i> [to take to his/her side]	<i>sineye çek-</i> [to put up]
al- [to get, to take, to have, to gain, etc.]	<i>aferin al-</i> [to win well done]	<i>satın al-</i> [to buy], <i>ablukaya al-</i> [to make blockade], <i>akıl al-</i> [to consult]	<i>soluk al-</i> [to breathe in] <i>dersini al-</i> [to learn one's lesson], <i>makaraya al-</i> [to take the rise out of someone/something], <i>ele al-</i> [to discuss, to go about]	<i>soluğu .DA al-</i> [to get (somewhere) in no time flat], <i>aklını al-</i> [take his/her mind], <i>akıl almaz</i> [inconceivable]
ver- [to give]	<i>tepki ver-</i> [to react], <i>anlam ver-</i> [to give meaning], <i>mola ver-</i> [to take a break], <i>değer ver-</i> [to give value], <i>moral ver-</i> [to give morale, spirits], <i>zarar ver-</i> [to damage, to harm], <i>önem ver-</i> [to give importance]	<i>Yanıt/cevap ver-</i> [to give answer], <i>talimat ver-</i> [to give instructions], <i>karar ver-</i> [to make a decision], <i>izin ver-</i> [to give permit]	<i>pas ver-</i> [to pass the ball/to give somebody the glad eye], <i>renk ver-</i> [to encolour/to enliven], <i>hak ver-</i> [to give right/to justify]	<i>omuz ver-</i> [to support/to give a helping hand], <i>gözdağı ver-</i> [to intimidate]
gel- [to come]		<i>anlamına gel-</i> [to mean], <i>duymazdan gel-</i> [to turn a deaf ear], <i>üstesinden gel-</i> [to overcome], <i>daral gel-</i> [to get fed up with, to get bored], <i>razı gel-</i> [to accept, to consent],	<i>. gibi gel-</i> [to feel like], <i>kendine gel-</i> [to come round/to regain consciousness], <i>başına gel-</i> [to experience]	<i>göze gel-</i> [to be affected by the evil eye], <i>göz göze gel-</i> [to come eye to eye]

language-specific probabilities should be taken into account when listing collocations.

According to the analysis of the verb *ver-* [to give], which is one of the most frequently used words in the verb type, it is seen that it is most frequently used together with names such as *tepki* [reaction], *anlam* [meaning], *mola* [break], *değer* [value], *cevap* [answer], *moral* [morale, spirits], *gözdağı* [intimidation], *hak* [right], *zarar* [damage, harm], *yanıt* [response], *karar* [decision], *önem* [importance]. It has been determined that the verb *ver-* [to give], which forms predicate relations, is frequently used within restricted collocation structures such as *yanıt ver-* [to give answer], *talimat ver-* [to give instructions], *karar ver-* [to make a decision], *izin ver-* [to give permit], figurative idioms such as *pas ver-* [to pass the ball, to give somebody the glad eye], *renk ver-* [to encolour, to enliven], and pure idioms such as *omuz ver-* [to support, to give a helping hand].

It has been determined that the verb *gel-* [to come], another frequently used verb, forms predicative combinations

such as *anlamına gel-* [to mean], *. gibi gel-* [to feel like], *duymazdan gel-* [to turn a deaf ear], *üstesinden gel-* [to overcome], *göz göze gel-* [to come eye to eye], *başına gel-* [to experience], *daral gel-* [to get fed up with, to get bored], *razı gel-* [to accept, to consent], *kendine gel-* [to come round, to regain consciousness] as well as idioms such as *göze gel-* [to be affected by the evil eye], *göz göze gel-* [to come eye to eye].

One of the most frequently used words in the noun type is the word *göz* [eye]. According to collocation analysis, it can be said that this word is frequently used with verbs such as *at-*, *krp-*, *aç-*, *gezdir-*, *yum-* [throw, crop, open, wander, clench], and nouns such as *hapis* [jail] and *şehla* [asquint]. Accordingly, it is observed that the word *göz* [eye] is frequently seen in idioms (*göz dik-* [to set one's eyes on], *göz kulak ol-* [to look after, to keep eye on]), figurative idioms (*gözleri kan çanağı gibi* [eyes got bloodshot], *göz ucuyla bak-* [to glance at, to cut one's eyes at someone]), restricted collocations (*gözünü dik-* [to gaze upon], *gözden*

Table 4. Examples of collocations with frequently used nouns in texts

Frequently used nouns	Examples of collocations with frequently used words			
	Semi-restricted collocations	Restricted collocations	Figurative idioms	Pure idioms
göz [eye]	şehla göz [skew-eye], göz aç- [to open eyes]	göz hapsi [probation], göz at- [to take a look], göz kırp- [to blink], göz gezdir- [to take a glance], gözünü dik- [to gaze upon], gözden kaçır- [to miss out/to overlook]	göz yum- [to turn a blind eye], gözleri kan çanağı gibi [eyes got bloodshot], göz ucuyla bak- [to glance at/to cut one's eyes at someone] göz açıp kapayınca kadar [in the blink of an eye]	göz dik- [to set one's eyes on], göz kulak ol- [to look after/to keep eye on], göze al- [to take a risk], göze gel- [to be affected by the evil eye], göz göze gel- [to come eye to eye]
yol [way/ road]		yol ağzı [road junction]	dört yol [crossroad], yol al- [to went one's way], yol aç- [to cause/to make way for], yol göster- [to lead the way/to guide], yol ver- [to give way]	
el [hand]	el arabası [handcart, trolley], el çantası [purse, handbag], el hareketi [hand gesture]	el çırp- [to clap]	el salla- [to shake], el koy- [to put] ele al- [to discuss, to go about], elde et- [to achieve, to acquire]	el et- [to wave]
kafa [head]			kafa salla- [to bob], kafasını karıştır- [to confuse], kafaya tak- [to mind, to have a bee in the bonnet], kafasından geçir- [to think off], kafası karış- [to get confused], kafası çalış- [to have a quick mind]	kafa yor- [to puzzle one's brains], kafa patlat- [to rack one's brains], kafayı ye- [to go nuts], kafayı kır- [to trip out], kafayı çek- [to drink heavily]

kaçır- [to miss out, to overlook]), phrasal usages ([göz açıp kapayınca kadar [in the blink of an eye]), and predicate relations (göze al- [to take a risk], göze gel- [to be affected by the evil eye], gözüne iliş- [to catch someone's eye]).

One of the most frequently used nouns is the word *yol* [way/road]. In the analyses made, it has been determined that this word forms restricted collocations and figurative idioms within solid noun phrases such as *yol ağzı* [road junction], *dört yol* [crossroad], as well as verbal combinations such as *yol al-* [to went one's way], *yol aç-* [to cause, to make way for], *yol göster-* [to lead the way, to guide], *yol ver-* [to give way].

Another most frequently used words in the noun type is the word *el* [hand]. It is seen that this word establishes collocations and figurative idioms such as *salla-* [to shake], *çırp-* [to clap], *koy-* [to put], and forms noun phrases such as *el arabası* [handcart, trolley], *el çantası* [purse, handbag], *el hareketi* [hand gesture].

Another frequently used noun is the word *kafa* [head]. It can be said that the noun *kafa* [head] collocates with verbs

such as *yor-*, *patlat-*, *salla-* (*kafa yor-* [to puzzle one's brains], *kafa patlat-* [to rack one's brains], *kafa salla-* [to bob]). It is observed that the word *kafa* [head] also forms verbal combinations such as *kafayı ye-* [to go nuts], *kafayı kır-* [to trip out], *kafasını karıştır-* [to confuse], *kafaya tak-* [to mind, to have a bee in the bonnet], *kafasından geçir-* [to think off], *kafayı çek-* [to drink heavily], *kafası karış-* [to get confused], *kafası çalış-* [to have a quick mind], and the suffixes are stereotyped in these combinations.

Even though collocations are cultural elements, they are shaped according to the phrase structure rules of Turkish in terms of their syntactic features. It is observed that idioms in Turkish are stereotyped to a certain extent. Especially in the case of Turkish, it is seen that suffixes are stereotyped as a part of word combinations and create semantic differences. It has been determined that the possessive and case suffixes on the nouns in the member structure of the verb are stereotyped as a part of the collocation structure and exhibit distinctive features *dersini al-* [to learn one's lesson], *makaraya al-* [to take the rise out of someone/something],

ele al- [to discuss, to go about], *el et-* [to wave], *elde et-* [to achieve, to acquire], etc.). This stereotype involves certain semantic and structural restrictions. It is observed that the suffixes in the stereotypes cannot be substituted in the collocations and idioms, they are structurally stereotyped and they have a distinctive feature, as in the examples of (*birinden*) *taraf ol-* [to stand up for someone/something]/ (*birinin*) *tarafında ol-* [to be on somebody's side]. Halliday (1961) suggested that collocations must be evaluated independently of the suffixes they take. Accordingly, the examples of “*gönlünü aldı* [made up to him/her]”, “*gönlünüzü alamam* [I cannot make it up to you]” should be described as different variants of the collocation *gönül al-* [to make it up to]. All of these examples are considered different forms of the same collocation, and since they do not carry any semantic or conceptual distinction, they are considered as different lexical forms depending on the lexeme *gönül al-* [to make it up to]. On the other hand, it is observed that the combinations of *göz dik-* [to set one's eye on] and *gözünü dik-* [to bore into] or the word combinations of *ad koy-* [to name, to call] and *adını koy-* [to make clear, to give a name] are formed with the same words, but due to the suffixes they take, they form different collocation structures. All these examples clearly show that suffixes have distinctive semantic and usage characteristics and that in Turkish, they are stereotyped as a part of collocation. It can be said that all word combinations that exhibit a stereotyped structure in Turkish also contain suffixes in this stereotype structure and become stereotypes through suffixes. For example, in the combination of nouns and adjectives, formal changes occur according to the position of the words. In cases where the adjective comes before the noun, there is no adjunct but if the adjective comes after the noun, it is observed that the possessive suffix is stereotyped and is a part of the collocation. For example, when the units are substituted in the combination *kara göz* [black eye], a possessive suffix must come on the noun, resulting in the word *gözü kara* [reckless, fearless]. In such cases, such a lexical substitution in collocation units also carries a distinctive feature. It is observed that the combinations of *gözü kara* [reckless, fearless] and *kara göz* [black eye] form different collocations. Among these combinations, it can be clearly observed that the possessive and case suffixes are stereotyped, especially in fixed combinations such as idioms. When such combinations are evaluated according to both semantic criteria and substitution criteria, they are characterized as different structures. Just as is the case in the possessive suffixes and the noun case suffixes, the negation suffix is also observed as a part of collocations in certain cases [such as *burnundan kıl aldırılmamak* [to have tickets on oneself], *yüzüne bakmamak* [to not look at one's face]). However, it has been observed that these structures do not always give the meaning containing the negative suffix with suffixes, and that there must be an intuition with the meaning of negativity. For example, instead of “*geldiğimizden beri yüzümüze bakmadı* [s/he hasn't looked at us since we came here], “*geldiğimizden beri yüzümüze bakan olmadı* [there is nobody looking at us since we came

here]” is frequently used in text contexts. Again, word combinations such as *adını ağzına al-* [to mention someone's name], *adını söylemeye/demeye dili var-* [to have the urge to mention someone's name] bear negativities in their internal structures and it is observed that they are used in contexts with negative semantic content, with the negative suffix such as “*adını ağzına almam* [I won't mention his/her name], *adını söylemeye dili varmaz* [I don't have the urge to mention his/her name]” or “*adını ağzına alırsam namerdim!* [I'll be damned if I mention his/her name once again!]”.

While the conditional suffix mostly creates free combinations, it is observed that the so-called indefinite objects often form collocations. However, this may sometimes differ depending on the context. This situation can be embodied by examples (1) and (2). The use of accusative case suffix in example (1) causes the structure to be classified as a free combination. In example (2), however, it is observed that the same suffix and the same combination structure form restricted collocation this time. This situation shows the importance of context in the clarification of collocational meaning.

- (1) Çarşı bedesteninde bir halı dükkânımız, bir de fırınımız vardı. Babam halıcılık eder, ilâ amcam da fırını çalıştırırlardı. Ama hepsi de **hesabı** dedeme verirdi. [We had a carpet shop and a bakery in the covered bazaar. My father was a carpet maker, and my uncle used to run the bakery. But they all **gave the account** of their business to my grandfather.] (Sessiz Yürek)
- (2) Bugün burada olmamın **hesabını** size verecek değilim. [I will not **give the account** of my being here today to you.] (Sessiz Yürek)

The importance of context in the determination of word combinations can also be supported by the following example. In example (3), the combination of *burnundan solu-* [to be steamed up] can be considered as an idiom when taken into account independent of the context. However, when evaluated in the context in which it is used, it is observed that it is a free combination.

- (3) Aslan, durum değerlendirmesi için birazcık soluklanıyor. Yavrunun aşağılarda kaldığını, kendisinin bu kavalık alana çekildiğini anlıyor. Geyik hâlâ birkaç adım ötesinde. **Burnundan** soluyor kesik kesik. Aslan, avının soluğunu yüzünde hissediyor gibi. [The lion gets a second wind a little to assess the situation. It gets that the cub is left below and that it is drawn to this rocky area. The deer is still a few feet away. Its lungs work nonstop like a fast bellow. It is **gasping through its nose** intermittently. The lion seems to feel the breath of its prey on its face.] (Sessiz Yürek)

Such structures were considered free combinations in the collocation classification made in this study, therefore they did not affect the numerical results of the study. However, in collocation determinations made via programs, such structures, especially encountered side by side, can be accepted as collocations in case they are classified only according to the statistical data, regardless of their meaning. One of the problems arising in the determination

of collocations in all text types covered via computer programs is the probability of determining all adjacent units as collocations. This is clearly seen in Example (4). The adjacentness of the nouns *tatlı* [sweet] and *uyku* [sleep] causes these units to be perceived as collocations, thus misleading when determining via programs. However, here, “*tatlı*” [sweet] is not an adjective describing the noun *uyku* [sleep], but rather a part of reduplication in the descriptive function of the verb *sız-* [to fall into sleep, to leak, to slip in].

- (4) **Tatlı tatlı uykuma sızıyordu**, Kendimi çocukluk yıllarımın sıcak pazar sabahlarından birinde zannettim. [It was creeping into my sleep sweetly, I thought I was in one of the warm Sunday mornings of my childhood years.] (Hazırlıksız)

In some combinations, it is not possible to determine whether words are transferred or not independently of the context of the text. Examples (5) and (6) can be given that the use of the same word combination in different text contexts leads to collocational errors. It is observed that the word combination *yelkenleri indir-* [to lower sails, to come round], which is encountered in the context of two separate texts, is used as a free combination in the context of the first text and as a figurative idiom in the context of the second text.

- (5) “Geminin kaptanı korkar, **yelkenleri indirir**, demir atar. Tayfaları ve yolcuları toplar; sorup soruşturur” [The captain of the ship gets afraid, he **lowers the sails** and anchors. He gathers crew and passengers and asks about the situation] (Billur Köşk Masalları)
- (6) İşte, onun bütün afrası tafrası buraya kadardı. Biri çıksın, ona hak versin, hemen **yelkenleri indiriverir**. [You see, all his bluster was thus far. Just as someone deems him right, he **comes around** immediately.] (Ankaralı)

Different variants of the same collocation can sometimes appear in the same text, and sometimes in different texts. There are two different uses of the collocation *hayal kırıklığı* [disappointment] in example (7).

- (7) Yine de sabah ve akşam etütlerini, gündüz ders gördüğüm sınıfta yapacağımı öğrendiğimde **hayal kırıklığına** uğramadım. Daha doğrusu, öylesine paramparçaydım ki, böyle bir konu için **kırılacak hayalim kalmamıştı**. [However, I was not **disappointed** when I learned that I would be doing the morning and evening studies in my daytime class. Or rather, I was so shattered that I couldn’t get more **disappointed** for such a thing] (Hazırlıksız)

In the context of intertextuality, collocations related to other texts can be found within a text type. For example, it is known that the phrase “*kırk katır mı istersin kırk satır mı?*” [do you choose forty mules or forty cleavers?] in the example below chosen from a novel type text is a frequently encountered usage in tale texts. This situation is also emphasized in the text.

- (8) Masallardaki kurbanlara tanınan seçme hakkı gibi yani: **Kırk katır mı istersin, kırk satır mı?** [It’s like the right of choice given to the victims in tales: **Do you choose forty mules or forty cleavers?**] (Hazırlıksız)

It is seen that certain collocation examples are used together with dialect features as in Example (9) in narrative text, especially in tales.

- (9) **Gusura galma** amma, merak etmişimdir. [I’m sorry but I was just wondering.] (Ankaralı)

Depending on the structural features of Turkish, elliptical structures are frequently encountered in narrative texts. In Example (10), it is observed that the object that is a part of the collocation *fotoğraf çek-* [to take a photograph] is reduced and a reference is made to the first sentence.

- (10) Çenesinden aşağı inen yaş, kucığına damlarken yine gülümsedi, “Babanın albümünde olmayan **fotoğrafları** görmek ister misin?” Kurumla kasıldı. “Hepsini ben **çektim**.” [He smiled again as the tears trickled down his chin, “Would you like to see **photos** that aren’t in your dad’s album?” He got into a tizz. “I **took** all of them.”] (Ankaralı)

Such situations, which can be encountered due to the elliptical feature of Turkish, may cause problems for studies that take into account the single-sentence determination range of collocations between two points. Quantitatively, the use of two collocations can also be overlooked by such elimination.

It is observed that the verbal combination *ad koy-* [to name] in the example (11) consists of the verbal combination of the noun in the previous sentence and the verb in the next sentence. In other words, the combination *ad koy-* [to name] has undergone object omission in the second sentence, again due to the structure of Turkish. It is observed that the name in the member structure of the verb, which is a part of the collocation, is used by omission. However, native language speakers can complete and make sense of these elliptical structures in collocations. There is a similar situation in Example (12). Since the word combinations *söz ver-* [to give a promise] and *söz tut-* [to fulfill one’s promise] are collocations with the same object, they are used in a sentence with a single noun by forming an elliptical structure.

- (11) “**Adı** da var demek,” dedi büyükanne, kinayeli kinayeli. “Kırçıl! O mu söyledi?” “Yoo,” dedi çocuk, saf saf, “ben **koydum**.” Annesine döndü: “Ekmek versem mi?” [So it has a **name** as well, huh,” said the grandmother allusively. “Kırçıl! Did it say that?” “Nah,” said the kid, and continued, “I **named** it.” He turned to his mother and said, “Should I give it some bread?”] (Kırçıl: Dikkat Kırılacak Eşya).

- (12) Dayım unutmazdı, **söz verirse tutardı**. [My uncle didn’t forget, when he **gave a promise**, he **fulfilled** it.] (Tiyatrolar, Aynalar: Dikkat Kırılacak Eşya).

One of the conditions of collocation is that a limited number of units can be inserted between the words that make up the combination. However, in the texts covered, it is observed that many words can enter especially between collocation units. According to Sinclair’s (1987, p. 325) views that collocations are not adjacent structures, this makes it possible to make certain determinations regarding the degree of restriction of collocations. The collocations

ilgi çek- [to arouse interest], *aklına gel-* [to come to mind], *içinden geçir-* [to think to oneself] in the contexts of the texts in (13), (14), (15), and (16) function as examples to this situation.

(13) İsmetin **ilgisini** en çok, içinde çizimler ve formüller olan fizik ile tek sözcüğünü bile anlamadığı Fransızca kitapları **çekmişti**. [The thing that **aroused the interest** of İsmet the most was physics, which contained drawings and formulars, and French books, which he did not understand a word of.] (Yolun Başındakiler)

(14) **Aklıma** en değerli eşyamız, borcu henüz biten renkli televizyonumuz **geldi**. [Our most valuable item, our color television whose debt has just been paid off **came to my mind**.] (Hazırlıksız)

(15) Vahit, yuvarlak masanın başına geçip otururken **içinden**, ‘Oğlum, Jale’yi düşün. Acaba Bir gün sabah kahvaltısı hazırlamış mı Cavit’e?’ diye **geçirdi**. [As Vahit sat at the round table, he **thought to himself**, ‘Think of Jale. I wonder if she prepared a single breakfast for Cavit.’] (Ankaralı)

(16) Galip Pehlivanoglu Gündüz Bakımevi ve Kız Yurdu’nu gezerken **içinden**, onunla nasıl yalnız kalırım, yüz yüze nasıl konuşabilirim diye **geçirdim** durdum. [As I visited the Galip Pehlivanoglu Daytime Nursing Home and Girls’ Dormitory, I **thought to myself** about how I can get alone with her, how I can talk to her in person.] (Ankaralı)

It is also observed that sometimes another collocation pair can be inserted in such a range between a collocation pair in the same sentence. Collocations in the sentence in (17) can be given as an example.

(17) Tam o sırada **karşımıza, geride bıraktığımız** hayalet yapıdan daha ufak, arka cephesi İslî bir bina **çıktı**. [Just then, we **came across** a building with a sooty back façade, smaller than the ghostly structure we left behind.] (Hazırlıksız)

The findings reveal that the classification of word combinations by reviewing them one by one according to their use in the context of the text will give more precise results. The analyses made according to the frequency of use of words and word combinations in the texts show that word combinations cannot be determined based on frequency alone.

DISCUSSION AND CONCLUSION

Word combinations in the database were classified as free combinations, semi-restricted collocations, restricted collocations, figurative idioms and pure idioms according to their semantic features on the basis of Cowie (1994), and free combinations were excluded from the restrictions of the study. Considering word combinations on the basis of their semantic features, it can be said that there is a hierarchy that we can observe between the examples of *çiçek ver-* [to give flowers] - free combination - < *acı ver-* [to inflict pain, to grieve] - semi-restricted collocation - < *selam ver-* [to greet] - restricted collocation - < *ders ver-* [to give a lesson] - figurative idiom - < *akıl ver-* [to give advice to someone] - pure idiom in terms of the degree of

stereotyping among their types. It is observed that all of these word combinations have the same verb (*ver-*) but in line with the semantic and substitutional limits, the degree stereotyping and accordingly the type of each combination differs.

Semantic and syntactic restrictions can be counted among the factors affecting the degree of stereotyping of collocations. It is observed that the restricted exhibits a gradual structuring according to the substitutability of the words in the combination, the comprehension of the meaning of the word combinations from the meaning of the words that make up the combination, and the ability to add another word between the collocation.

In addition to this, it is noteworthy that the negation suffix in Turkish, as seen in the examples *adını ağzına alma-* [to not mention one’s name], *aklının ucundan geçme-* [to never occur to], *aman verme-* [aman vermez/*aman verir] [to give no quarter/*to give quarter]¹³ and the negative, possessive, and case suffixes exhibit semantic and usage-based distinctive features, as seen in the examples *kılmı kıpırdat-/*kıl kıpırdat-* [to move a muscle], *aklı havada* [to be head-in-the-clouds], *başta gel-* [to experience], *başı şiş-* [to get confused], *başından at-* [to brush off] and therefore that it is stereotyped as a part of the combination. This stereotyping structure can also be embodied through the semantic distinction in the examples of *el et-* [to wave a hand] and *elde et-* [to achieve, to acquire], *el ver-* [to lend a hand] and *ele ver-* [to blow the whistle on].

As it will be noted in example such as *u dönüşü* [u-turn], there is a change in the type and difficulty levels of the combinations in direct proportion to the transfer of meaning of the units that make up the word combination. Accordingly, if a hierarchy including semantic transference is made, it can be said that free combinations (*süt iç-* [to drink milk], *çiçek al-* [to take flowers]) are the least restricted, easiest and therefore most frequently used combinations, collocations as more restricted combinations (*sigara iç-* [to smoke cigarette], *abdest al-* [to perform ablution]), and figurative idioms (*ders al-* [to take a lesson], *u dönüşü* [u-turn]) and pure idioms (*aklına al-* [to fascinate], *nalları dik-* [kick the bucket]) as the least used and most difficult units. As the substitution option in free combinations decreases, semi-restricted collocations emerge, and as the semantic restrictions in semi-restricted collocations increase, restricted collocations emerge. It is known that as one moves towards idioms, the effects of socio-cultural structure on the language are reflected more. Figurative idioms are mostly metaphor-based word combinations and restricted collocations are described as word combinations that contain completely different restrictions from idioms and free combinations in terms of clarity and modifiability. The combination of units in semi-restricted combinations can generally be explained logically and semantically to a certain extent, as in free combinations but semi-restricted collocations include a more restricted substitution option than free combinations. This stereotype structure can be embodied through the collocation *at bin-* [to ride a horse] and the free combination

ata bin- [to mount a horse]. In terms substitutability, the first one creates a more restricted collocational combination, and the second one forms a relatively less restricted combination due to the presence of the inclination suffix. In this context, the fact that examples such as *arabaya bin-* [to ride a car], *eşeğe bin-* [to ride a donkey], *deveye bin-* [to ride a camel], etc. can be reproduced is an indication that there is no restriction on the substitutability of the noun as in the example of *at bin-* [to ride a horse]. Since the substitution option is relatively higher in the combination *ata bin-* [to mount a horse], the restriction works inversely. It is seen that substitutions such as **katır bin-* [to ride mule], **eşek bin-* [to ride donkey] are not possible in the combination *at bin-* [to ride a horse]. It can be said that the degree of restriction between these two examples differs depending on the orientational suffix.

Semi-restricted collocations include collocations that do not allow phrasal or stereotyped uses, have few substitution options, and have many restrictions but have fewer restrictions compared to restricted collocations. Idioms, on the other hand, correspond to the structures that contain the most phrasal usage and exhibit stereotypes with the opposite grade. Most word combinations accepted as idioms in the Turkish dictionary were not evaluated as idioms in this context, but were handled in restricted collocations within the framework of semantic and substitution-based criteria (such as *ad tak-* [to nickname], *ad ver-* [to name], *ad koy-* [to give a name], *güven ver-* [to reassure], *haber ver-* [to let know, to inform], *hakkını ver-* [to do justice to], *fotoğraf çek-* [to take a photograph] etc.).

The fact that Turkish has a flexible syntactic structure provides a basis for the conclusion that word combinations can be determined and their types can be clarified only by questioning at the sentence and context level. The type of a word combination can only be determined depending on the context. For example, combinations such as *burnundan solu-* [to be steamed up], *yelkenleri indir-* [to lower the sails] may exhibit idiom, collocation, and free combination features depending on the context. On the other hand, contrary to the assumption adopted in statistically-oriented approaches, it draws attention in the process of determining word combinations that collocations in Turkish can include words according to the degree of restriction and the number of these words cannot be specified with definite limits due to the flexibility in the syntax of Turkish. Word combinations in Turkish can be clarified at the sentence level and context-dependent.

In general, the headword is expected to be a verb in noun + verb combinations. However, for Turkish, especially in the case of word combinations formed with borrowed words, the opposite is the case. It is observed that the headword in combinations formed with pure actions is the noun, not the verb. This situation is explained on the basis that the verb is a light verb when associated with its power in combination. This result is supported by the features and functions stated by Akşehirli (2013) regarding light verbs.

It is observed that predicative combinations formed with borrowed words are mostly formed with light verbs (*yap-*, *et-*, *ol-* [do, make, be], etc.). Words coined from other languages, regardless of their lexical functions/tasks in the source language, are mostly used by nomenclature when transitioning to Turkish as a target language, therefore they need a light verb or an auxiliary verb in order to be used in the verb function. Combinations formed by auxiliary verbs or light verbs constitute another problem. It is observed that auxiliary verbs are sometimes handled within the category of idioms and sometimes under compound words formed with auxiliary verbs, and there is no consistent explanation on this issue either. Collocations consisting of borrowed words should be queried in larger corpus studies regarding the degree of restriction, determinations should be made and the results should be supported by etymological findings.

In Turkish dictionaries, word combinations such as *paha biç-* [to estimate a price, to value], *acemilik çek-* [to suffer from inexperience], *çile çek-* [to suffer a lot], *yanıt ver-* [to respond], *ilişki kur-* [to establish a relationship, to make contact], *fotoğraf çek-* [to take a photograph] are sometimes described and explained as compound words and sometimes as idioms. For this reason, certain problems may arise in terms of language users' access to collocations and their use of these structures. For instance, *güneş çarpması* [sunstroke] is given as a compound noun in the Turkish dictionary under the compound words section, while *güneş çarpmak* [to have sunstroke], which is a different variant of the same lexical unit, is given in the idioms section. The combination *kura çekmek* [to draw lots] is under the section of compound words but the combinations *acı*, *azap*, *ağrı çekmek* [to feel pain, to suffer torment, to feel discomfort, etc.] are handled under the idioms section.

Synonyms also pose a problem in some cases. For example, it is striking that there is no consistency in the dictionaries between the combinations in which *ad* [name] and *isim* [name, noun] words take place. In the dictionary of idioms, even though there are combinations established with the word *ad* [name] such as *adı duyulmak* [to become prominent], *adı geçmek* [to be mentioned], *adı kalmak* [to be remembered after one is dead], *adı karışmak* [to be involved], *ad takmak* [to nickname], *ad yapmak* [to make a name], it is seen that there are no combinations such as *ismi duyulmak*, *ismi kalmak*, *ismi karışmak*, *isim vermek* [be heard, stay in the name, mix the name, submit name]. While *satın almak* [to buy] combination is at the item head of the *al-* verb, it is observed that combinations *ad almak*, *isim almak* [to be given a name] are not. In addition to this, it is seen that while the word combination *soluk almadan* [in one breath] is included as the headword of the verb *al-* among the compound words, *soluk al-* [to take a breath] combination is not included there. In fact, the contexts in which the collocation *soluk al-* [to take a breath] and the collocation *soluk almadan* [in one breath] are used and their meanings are also different from each other. It is

noteworthy that some word combinations are not included in Turkish dictionary. For example, the combination *zengin kalkışı* [(to take) a french leave] is neither found in the dictionary of idioms nor in the current Turkish dictionary. In addition, there is no consistent explanation in Turkish dictionaries of the subject of which word combinations should be searched based on which part of the combination in the dictionary. For instance, the combination *ekmek parası* [bread and butter] is not found as the headword of the word *para* [money] but only as the headword of the word *ekmek* [bread], while the combination *kan parası* [blood money] is found both as the headword of *kan* [blood] and *para* [money]. In dictionaries, it is a problem that should be clarified that such word combinations should be evaluated within the scope of which word combination type and at which article they should be included. It should be clarified that the collocations should be reached on the basis of which word of combination.

Since Turkish is an agglutinative language, the agglutinations on the word cause different conjugations of the word in question to be counted as different words in such programs. Synonyms also mostly pose a problem in this respect. For example, since the verb *de-* [to say] and the conjunction *de* cannot be labeled separately, these two words, which have two different functions, are detected as a single token by programs, thus resulting in some deviations in the numerical results¹⁴. Since Turkish does not have a comprehensive wordlist¹⁵ as in English, the analyses made do not always give accurate results, since these words cannot be separated into their types by computer programs¹⁶. Therefore, in order to minimize such mistakes, the words and combinations calculated by the programs should be reviewed.

Collocations make it possible to use and understand the language in an effective way. In this context, specific-purpose collocation dictionaries should be created by making use of the outputs of such descriptive studies. This way, the problems that students/native speakers may experience while internalizing collocations can be minimized. As with all the rules of the language, lexical-based rules are also intuitively acquired and used by native speakers. However, teaching stereotyped word combinations such as collocations and idioms will help native speakers to classify phrases and patterns they use in daily life in their memory. For this reason, such combinations must first be defined fully and accurately. Teaching under what conditions word combinations are collocations, under what conditions idioms, and under what conditions are compound words will increase students' awareness of the words they use. It will also contribute to the development of literacy skills of students both in mother tongue education and foreign language learning.

NOTES

1. It is considered appropriate to use the term collocation in this study for this concept, which is found in different terms in the literature such as lexical chunks (Lewis, 1993), formula/formulae (Ellis, 1994), multiword

- items/units/lexical phenomena, routine, set phrases, stereotypes, stereotyped phrases, stock utterances, unanalyzed multiword chunks/units (Wray, 2002).
2. For detailed information on the definition and classification of word combinations and collocations, see Eken (2016).
3. It has been adopted by researchers such as Firth, Halliday, and Sinclair.
4. It has been adopted by researchers such as Cowie, Howart, Mel'çuk.
5. While Sinclair (1966) considers three units on each side (at the beginning and the end) of the node, in 1970s, four units on each side of the node were considered and collocation determinations were made accordingly (Hori, 2004, p. 5).
6. Brought together by nature, such word combinations are considered within the framework of semantic prosody and to a certain extent as causal combinations in this study. For example, *dog* is a concept that is used together with the verb *to bark* by its nature, in other words, there is the presupposition that the word *to bark* is an action performed by a *dog* by nature. Therefore, such words are used in combination within the framework of *semantic prosody*.
7. Analyses were performed based on Bilgin (2006) and Büyüköztürk et al. (2013).
8. Validity and reliability checks were performed based on Yıldırım and Şimşek (2011).
9. Validity and reliability checks were performed based on Yıldırım and Şimşek (2011).
10. See Okur & Arı (2014).
11. The translation of collocational expressions is sometimes made by combining a verb and a noun, sometimes by finding the counterpart words used in the target culture, and sometimes by giving new meaning to words. See Sarıkaş (2006), Seymen&Kalkan (2019) and Soyer (2017) for more information about translation of collocations.
12. According to Hausmann (2007), these are called "*collocational chains*" According to Spohr (2005), they are called "*collocational clusters*". For detailed descriptions, see Tutin (2008)
13. "**aman verir*" [to give quarter] is unacceptable to use in Turkish. It is only acceptable with negative suffix as "*aman vermez*".
14. For studies and detailed information on the subject, see Mersinli and Aksan (2011), Aksan et al. (2011).
15. Such as BNC (British National Corpus)
16. For detailed information on morpheme and word type marking in Turkish, see Mersinli and Aksan (2011). For the classification and distribution of uncertainties in morpheme studies, see Aksan et al. (2011).

REFERENCES

- Adalar, D. (2005). Arapça ve Türkçe Ders Kitaplarındaki "Yardımlaşma" ve "Arkadaşlık" Konulu Metinlerin Karşılaştırılması: Bir Eşdizimsel Örüntüleme Çözümü Örneği. *Dil Dergisi*, 129, 63-84.

- Ağca, M. (2020). Eski Türkçe Metinlerde Sıfatların Eşdizimliliği [Unpublished Doctoral Thesis]. Pamukkale University Social Sciences Institute, Denizli.
- Aksan, Y., Mersinli, Ü., Yaldir, Y., & Demirhan, U. U. (2011). Türkçe Ulusal Dil Derlemi Projesi Biçimbirim Çalışmalarında Belirsizliklerin Sınıflandırılması ve Dağılımı. In 25th National Linguistic Conference (pp. 219-226).
- Aksan, Y., Aksan, M., Mersinli, Ü. & Demirhan, U. U. (2017). A frequency dictionary of Turkish-Core vocabulary for learners. New York: Routledge.
- Akşehirli, S. (2013). Türkçede Et- Katkısız Eyleminin Sözlüksel İşlevleri: *Turkish Studies - International Periodical For The Languages, Literature and History of Turkish or Turkic Volume 8/9 Summer 2013*, (pp. 481-494).
- Anthony, L. (2014). AntConc (Version 3.4.3w) [Computer Software]. Tokyo, Japan: Waseda University. Available from: http://www.laurenceanthony.net/software/antconc343/AntConc_readme.pdf
- Bilgin, N. (2006). *Sosyal Bilimlerde İçerik Analizi -Teknikler ve Örnek Çalışmalar*. Siyasal Press.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. & Demirel, F. (2013). *Bilimsel Araştırma Yöntemleri* (14th ed.). Pegem Akademi Publishing.
- Cowie, A. P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics*, 2(3). (pp. 223-235).
- Cowie, A. P. (1988). Stable and creative aspects of vocabulary use. In R. Carter and M. McCarthy (Eds.), *Vocabulary and Language Teaching* (pp. 126-139). Longman.
- Cowie, A.P. (1991). Multiword Units in Newspaper Language. In Granger, S (Ed.), *Perspectives on the English Lexicon: A tribute to Jaques van Roey* (pp. 101-116). CILL 17, 1-3.
- Cowie, A. P. (1994). Phraseology. In R. E. Asher (Ed.), *The encyclopaedia of language and linguistics* (pp. 3168-3171). Pergamon.
- Cowie, A. P. (Ed.). (1998). *Phraseology: theory, analysis, and applications*. Oxford University Press.
- Çıkrıkçı, S. (2009). Türkçede Boyut Sıfatlarının Sözlükte Temsil Edilen Eşdizimsel Görünümleri. In S Ay, Ö Aydın, İ Ergenç, S Gökmen, S İşsever & D Peçenek (Eds.), *Essays on Turkish Linguistics* (pp. 167-176). Wiebaden.
- Dedeoğlu, E., & Şen, G. (2010). *İngilizce-Arapça-Türkçe Eşdizim*. Fecr Publishing.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford University Press.
- Firth, J. R. (1951). Modes of meaning. In Palmer, F.R. (Ed.), *Papers in linguistics* (pp. 1934-1951). Oxford University Press.
- Firth, J. R. (1957). Modes of meaning. In Palmer, F.R. (Ed.), *Papers in Linguistics* (pp. 190-215). Oxford University Press.
- Halliday, M. A. K. (1961). Categories of the theory of grammar. *Word*, 17, 241-292.
- Hausmann, F. J. (2007). Lexicographie française et phraséologie. *Collocations, phraséologie, lexicographie*. Aachen, Shaker Verlag, 121-153.
- Hori, M. (2004). Investigating Dickens' Style. A collocational Analysis. Basingstoke: Palgrave Macmillan.
- Lewis, M. (1993). *The Lexical Approach: The State of ELT and a Way Forward*. Language Teaching Publications.
- Mersinli, Ü. & Aksan, M. (2011). Türkçenin Biçimbirim ve Sözcük Türü İşaretleme. Ç. Hatipoğlu & Ç. Sağın-Şimşek (Ed.), In 24th National Linguistic Conference, 367-376. Ankara: ODTÜ Press.
- Okur, A. & Arı, G. (2014). Öğrencilerin İlköğretim 100 Temel Eseri Okuma Durumu. *Türkiye Sosyal Araştırmalar Dergisi*, 173, 307-328. Retrieved from <https://dergipark.org.tr/en/pub/tsadergisi/issue/21486/230317>.
- Özkan, B. (2007). *Türkiye Türkçesinde Belirteçlerin Fiillerle Birliktelik Kullanımları ve Eşdizimliliği* [Unpublished Doctoral Thesis]. Çukurova University Social Sciences Institute. Department of Turkish Language And Literature. Adana.
- Özkan, B. (2010). Türkçenin Öğretiminde Sıfatların Eşdizim Sözlüğü –Yöntem ve Uygulama-: *e-İnternational Journal of Educational Research*, 1(2), 51-65.
- Özkan, B. (2011). *Türkiye Türkçesinde Belirteçlerin Fiillerle Birliktelik Kullanılması ve Eş Dizimlilikleri -Derlem Tabanlı Bir Uygulama-* Ankara: Turkish Language Society Press.
- Özkan, B. (2014a). Türkiye Türkçesi Eşdizim Sözlüğünün Sayısallaştırılması. In B. Özkan, B. T. Tahiroğlu & A. E. Özkan (Eds.), *Türkçe Üzerine Derlem Dilbilim Uygulamaları* (pp. 23-38). Karahan Publishing.
- Özkan, B. (2014b). Türk Yazın Dilinde Argo-Derlem Tabanlı Bir Uygulama. In B. Özkan, B. T. Tahiroğlu & A. E. Özkan (Eds.), *Türkçe Üzerine Derlem Dilbilim Uygulamaları* (pp. 123-174). Karahan Publishing.
- Özkan, B. (2014c). Türkiye Türkçesi Söz Varlığında Sıfatların Eşdizimliliği- Derlem Tabanlı Bir Uygulama. In B. Özkan, B. T. Tahiroğlu & A. E. Özkan (Eds.), *Türkçe Üzerine Derlem Dilbilim Uygulamaları* (pp. 175-224). Karahan Publishing.
- Sarıkaş, Ö. F. (2006). Problems In Translating Collocations. *Elektronik Sosyal Bilimler Dergisi*, 5(17), 33-40.
- Seymen, A., & Kalkan, H. K. (2019). Kollokation Als Evidentes Übersetzungsproblemkollokationen In Zweisprachigen Wörterbücherndeutsch-Türkisch/Türkisch-Deutsch. *Zeitschrift für die Welt der Türken*, 11, 53-78.
- Sinclair, J. (1987). Collocation: a progress report. In Steele, R. and Threadgold, T. (eds.), *Language topics: Essays in honour of Michael Halliday*, 319-331.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Soyer, S. (2017). Translation of cultural expressions and its importance in foreign language teaching. *Turkish Studies International Periodical for the Languages, Literature and History of Turkish or Turkic Volume*, 12(15), 567-578.
- Taşgüzel, S. (2004). İlköğretim Türkçe Ders Kitaplarında Öğretici Nitelikli Metinlerdeki Eşdizimsel Örüntülerin Görünümü. *Dil Dergisi*. 125, 72-87.

- Torun, Y. (2011). Dede Korkut Hikâyelerinde Barınma ile İlgili Sözcükler ve Bu Sözcüklerin Birlikte Kullanımları Üzerine: Turkish Studies. *International Periodical For The Languages, Literature and History of Turkish or Turkic*, 6(3), 1251-1263.
- Tutin, A. (2008). *For an extended definition of lexical collocations*. Eurolex. 27 March 2009 Retrieved May 05, 2009 Manuscript Available at hal.archives-ouvertes.fr/docs/00/37/14/18/PDF/Euralex_2008_tutin.pdf
- Tüfekçioğlu, B., & Özkan, B. (2014). Derlem Tabanlı Çevrim İçi Türkçe Öğrenici Sözlüğü-Önadlar A Maddebaşı. In B. Özkan, B. T. Tahiroğlu & A. E. Özkan (Eds.), *Türkçe Üzerine Derlem Dilbilim Uygulamaları* (pp. 225-262). Karahan Publishing.
- Türkçe Sözlük. (2005). TDK Press: 549.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge University Press.
- Yağcıoğlu, S. (2002a). Metinsel Bağdaşıklığın Sağlanması da Eşdizimin İşlevi. *Türkçede Bilgi Yapısı ve Bilimsel Metinler*, German (Essen), pp.107-114.
- Yağcıoğlu, S. (2002b). Sosyoloji Metinlerinde Konu Sürekliliğinin Eşdizimsel Örüntü Açısından İncelenmesi: Edimbilimsel Bir Yaklaşım. *Türkçede Bilgi Yapısı ve Bilimsel Metinler*, German (Essen), pp.187-202.

Web References

- <http://www.tdk.gov.tr/tdksozluk>.
<http://www.tudweb.org/index.php?dil=1>
<http://tscorpus.com/tr>