Eurasian Journal of Educational Research 73 (2018) 113-130

Eurasian Journal of Educational Research
www.ejer.com.tr

An Evaluation of the Vocabulary in Turkish Textbooks Before and After the Constructivist Approach in the Context of Frequency[1]

Sukran DILIDUZGUN[2]

ARTICLE INFO

Article History:
Received: 10 Mar. 017
Received in revised form: 27.Nov.2017
Accepted: 04 Jan.2018
DOI: 10.14689/ejer.2018.73.7

Keywords
Word frequency, word origin, lemmas, word frequency lists.

ABSTRACT

Purpose: Research shows that students in Turkey are taught to write with much smaller vocabularies than students in other countries. Along with the constructivist approach, the frequency of words used in texts should be high, and unknown words should be at certain levels in order to allow for production, cognizance and creation of information based on existing information.

The aim of this research is to analyze the vocabulary in Turkish textbooks before and after the introduction of the constructivist approach on the basis of frequency and origin.

Method: The research was conducted using a general screening model composed of the words used in informative texts in Turkish textbooks before and after the introduction of the constructivist approach. With the aid of random sampling, informative texts with "Nature and Universe" themes that were used in 2001-2002 and 2013-2014 in secondary schools were chosen as samples. Frequency and percentage proportions were made in a descriptive way in order to compare the numbers of the same and different lemmas. T-test and Mann Whitney U test were used to determine meaningfulness.

Findings: While there is not a significant difference in frequency between TBCA and TACA in the Turkish, Arabic, Persian and western-based words. It has been determined that, among words that do not have synonyms, the same book after the introduction of the constructivist approach shows a preference for words used with less frequency.

Implications for Research and Practice: The word lists for each level should be made in the context of frequency of use in textbooks.

[1] This research, which was supported by Istanbul University Scientific Research Project Department with the number of 44490, was presented in 17th (International Conference on Turkish Linguistics (03 May-05 May, 2015) in Rouen University (France) as a paper.

[2] Istanbul University, Turkey, sdilid@yahoo.com.
ORCID: http://orcid.org/0000-0002-4094-5215

Introduction

One of the primary reasons for not being able to use a language in a competent and fluent way is either not finding the right words or not being able to use the words in an appropriate context. Ozdemir (2012:35) underlines that words are the building blocks of thinking in that they veer away from the reality. "If the name of something is does not exist in our language memory, it does not exist at all (Ozdemir, 2011: 51)." In addition, Vygotsky (1998) states that thinking is like a cloudburst of words, and thinking passes on to words through the aid of meaning.

Two accepted facts about primary school students are that they always use the same words in their essays (Author, 2013), and they are not very successful in using either active or passive vocabulary. Author, 2014, reveals the need for more institutional and application-oriented research into the style and content of the teaching of words. In addition, it has been observed that the number of words and concepts taught to students using course tools and equipment in the mother tongue education is quite low when compared to other countries (Akdogan, 1999, cited as in Ozbay ve Melanlioglu, 2008). Guzel (2006: 323) explains this by saying, "There hasn't been a research focusing on the quantitative analysis of vocabulary of primary school students in our country, and the vocabulary improvement of students is left to coincidence." Karadag's work (2005) titled "A Research on Vocabulary of Students in Primary Education," is a product of this kind of idea, and a non-proportional increase is seen textbooks in terms of total and unique vocabulary in different grades. The same research was applied to secondary education by Kurudayioglu (2005), and it emerged that common words should take their place in textbooks. The report prepared by the Ministry of National Education (MEB) titled "İlk Mektep Kitaplari Tetkik Komisyonu Elifba Kitaplari" also underlines the importance of increasing the vocabulary of students; even though creating committee was proposed to create a list of the words that children use the most, no rating scales or proper vocabulary for a given age and frequency of usage have been created yet (Karadag, 2005).

During the teaching process for both the first tongue and foreign languages for different age clusters, the priority of factors in the textbooks and dictionaries is determined by studies of frequency counts (Aksan, 1982). The "vocabulary control principle," which states that "the most frequent words should be shown first to people who are learning a new language," is shaped by the frequency studies of Thorndike (Aksan, Mersinli ve Yaldir, 2011). In addition, learning the frequently used words may make it easier to learn the words in different subjects and at different levels (Hatch & Brown, 1995). When the frequency of usage of words is taken into consideration, it is clear that the most frequent 1000 to 1500 words correspond to over 90% of that language (Aksan, 1982; Karadag, 2005). According to Nation and Newton (1997), the most frequently used 2000 words in a language form 85% of all the words in a book or newspaper published in that language. The greater the frequency of unknown words in a text, the more the reader will encounter these words and the less understandable the text will become (Ozturk, 2013). Nation (2001: 42) states that "there should be 1 unknown word in every 50 words in order to have pleasure in reading."

In this context, Aksan et al. (2012) compared the vocabulary lists gathered from purpose-made Turkish Textbooks Collection (251,860 words) that were created with Turkish textbooks published between 2005 and 2010 and approved by MEB, and a general collection (260.000 words) which was created using a sample of the Turkish National Corpus (Aksan et al., 2012) in order to see how the vocabulary in Turkish textbooks reflects the general use of language (Aksan, Mersinli & Yaldir, 2011). In addition to creating lists of the words frequently used in Turkish textbooks, the most common 100 words in 3 textbooks were selected to determine how many times they were used in total and in each of the textbooks separately (Arı, 2003). Apaydin (2010), Uludag (2010) and Turhan (2010) analyzed Turkish textbooks from the 6th, 7th and 8th grades, respectively, in terms of the vocabulary in these textbooks.

Aim of the Research

Since 2005, the system of national education in Turkey has been based on a constructivist approach aimed at encouraging an active learning process with innovation, mutual interaction, and a perspective of combining adaptational learning with productive learning (Yurdakul, 2005). During language education, the frequencies of word use in texts should be high, and unknown or unique words should be offered at certain levels in order to have production, cognizance and creation of information based on existing information, in order to place the learning responsibility on the learners.

The aim of this research is to analyze the vocabulary in Turkish textbooks before (TBCA) and after constructivist approach (TACA) was introduced on the basis of frequency and origin.

Sub-aims of the research are:

- To determine out the numbers of the same and different lemmas and the words derived from them.
- To determine the percentages of the same and different lemmas in TACA and TBCA in terms of their roots.
- To determine the origins of different lemmas in TACA and TBCA.
- To compare the most repeated 100 words, apart from the same words, with the total percentages of roots.
- To determine whether the difference between the frequencies of different lemmas in TBCA and TACA is meaningful.
- To determine the frequency and roots of synonyms.
- To determine whether there are differences between the conceptual fields analyzed in TACA and TBCA.

Method

Research Design

This research was conducted using a general screening model as a component of the quantitative research method. "Research models are approaches to research that describe a situation as it happened in the past or in the way it is happening currently" (Karasar, 1984: 80). In quantitative research methods, there is a preference for segmenting complicated facts and incidents into analyzable special parts, degrading the data into numerical values and summarizing the conclusions statistically (Buyukozturk, Cakmak, Akgun, Karadeniz & Demirel, 2009).

Research Sample

This research is composed of the words used in informative texts in Turkish textbooks before and after the introduction of the constructivist approach. The research is limited by informative texts, as the literary texts were generally the same in both of periods, and it is thought that there would not be much difference in terms of words used. According to Aksan et al. (2012), as frequency is not a direct linguistic or lexicological property of words, and as they represent a relativist result in a distinct linguistic dataset or a sample, it is important to know the source of text collection and the genres and eras in which the linguistic corpus that the words were collected. Because of that, two corpuses, which have distinct themes and are composed of distinct types of texts, were compared in this research.

Using random sampling, informative texts with "Nature and Universe" themes that were used in the 2013-2014 school year by 6th grade, 7th grade (Ada Press) and 8th grade (MEB Press) students, and informative texts of much the same theme that were used during the 2001-2002 academic year by 6th grade, 7th grade (Altin Kitaplar Press) and 8th grade (MEB Press), were chosen for TACA and TBCA, respectively. A similar number of words was selected for every grade, in order to mitigate any differences in grade levels in terms of number, difficulty and frequency of words. In TBCA, there are five travel essays, two interviews, two articles, and one news article, and in TACA, there are three essays, two articles and one interview. The texts chosen from the textbooks are provided in Table 1 based on their word counts.

Table 1

Texts chosen from TACA and TBCA, Genres and Word Counts

Grade	TBCA	Word count	TACA	Word count
6th grade	Yesil Cigerli Devler (article), Kartalkaya (travel writing) Serin Dere'ye Sicak Yuruyus (travel writing)	881	Bos Arsa (essay) Orman Kustu Bize (essay)	1135

Table 1 Continue

Grade	TBCA	Word count	TACA	Word count
7th grade	Beyaz Karanlik (travel writing) Karinca Yuvalarinin Gizi (article) Van Golu'nun Safagi (reportage), Saroz Korfezi (travel writing)	1339	Kutup Yildizi (article), Ormanda (essay).	1108
8th grade	Abant Cagiriyor (travel writing), Su (interview) Baska Karadeniz Yok (news)	1111	Yanan Ormanlarda Elli Gun-Orman Yanginlarinin Sebepleri (interview) Guney ve Bati Anadolu Ormanlari (article)	1088
Total		3331		3331

Research Instruments and Procedures

The four grammatical categories of nouns, verbs, adjectives and adverbs were included in the scope of the research. L. Tesniere reduces grammatical categories to four by performing functional analysis: Nouns, verbs and their determinants as adjectives and adverbs, respectively (Kiran & Kiran, 2001). This grammatical base, which was built on the contrast and interaction of noun and verb, is suitable for the analysis of Turkish, where words are separated into nouns and verbs at the base level.

Lemmas were used as a base in this research in an aim to develop the expanding vocabulary of students. "Lemma" is defined as the nominative of a word which is purged from the word endings and the morphemes that may be lexical entries in dictionaries (Aksan et al., 2012). For instance, three times "ogrenci-y-di" (He was a student), two times "ogrenci-nin" (student's), and three times "ogrenci-ler" (students), which are all derived from the same word but use different word endings, constitute eight models and three kinds. In this example, there is only one lemma that represents the three kinds of words, which is "ogrenci" (student). The purged versions of words without word endings provide the real frequencies of the words in order to determine which Turkish words will be taught during foreign language learning and reading comprehension education. The research of Ozturk (2013)

reveals that the first 500 kinds correspond to 19% of the corpus, while the first 500 lemmas correspond to 57% of the sub-corpus.

In Turkish, morphemes are divided into the two main categories of "lexical" and "functional" morphemes. "While contextual definitions can be given for lexical morphemes, the definitions for functional morphemes use alikes or functional definitions are made" (Adali, 2004, 26). In this context, in the scope of the research, while the lexical morphemes were used, functional free morphemes such as pronouns (me, self, this, who, etc.), functional verbs (this, else, how many, which, few, etc.), prepositions (for, till, beyond, etc.), rating antecedents (more, most, a lot, etc.), conjunctions (but, however, etc.) and exclamations – Aksan (1982) named as structural words – and words formed with voice suffixes and verbal appendixes (apart from derivational affix functions), demonstratives and numeral adjectives, proper nouns, units of time and length, auxiliary verbs and reflection words were excluded. Research in contemporary era (Ozturk [2013]; Aksan, Mersinli & Yaldir [2011]; Arı [2003]) shows that adjectives (one, this, two, etc.), conjunctions (and, but, etc.), prepositions (like, for, so that, etc.) and pronouns (he, me, etc.) were used the most frequently in texts.

Examination of the words that constitute the compound words and idioms separately has been a limitation of this study. Additionally, words with multiple meanings are provided under the same entry, heading away from the definition of multiple meanings, by Aksan (1997: 58) who states that "It is caused by the need of human beings to explain based on structure, function, aim relevance and closeness of other concepts in order to utter concepts in a more effective, tangible and easy way," during the research, only one of the meanings of the words with multiple meanings was considered. Nevertheless, consistency within the text was observed as well.

Word lists with numerical order were created for the corpus of both time periods examined. Lemmas and model numbers were determined by computer, using the Ctrl+f keys, and they were ordered alphabetically in Microsoft Excel, resulting in two different lists being created for both periods. The lists were controlled using the Turkish Text Frequency Solver created by Kurt (2007) to examine terms of the frequencies of roots and stems of the words. Later, the same lemmas in both of the lists, and the other words which were derived from the roots of these words, were selected in order to determine the words with the same roots in the texts of both periods. Heading away from the idea that "Knowing the meaning of root of the word makes it easier to predict all the words derived from the same root" (Gunes, 2013:12), the words that were derived from the same base were excluded from both of the lists, and by the aid of the Written Turkish Word Frequency Dictionary, which includes 22.693 words and was created by Goz (2003), the vocabulary frequency lists with numbers were created separately for different words and roots by means of the Microsoft Excel program for Turkish, Arabic, and Persian, and especially for the words that come from other languages that were generally affected by French. The work of Goz (2003) was chosen as it includes a textbook category in its corpus, and it is also based on written Turkish and focused on lemmas.

Data Analysis

Frequency and percentage proportions were made in a descriptive way in order to compare the numbers of the same and different lemmas in TACA and TBCA, the percentages of the roots of the same and different lemmas in TACA and TBCA, the numbers and origins of the different lemmas in TACA and TBCA and the total percentages of roots of the 100 most frequent words, excluding the same words.

T-test was used (for total, Turkish, Arabic) in order to determine whether the difference between different words in terms of total and frequency of roots is meaningful, and Mann Whitney U test was used when the number in the vocabulary lists was less than 30 (for Persian and western-oriented words). Nisanyan dictionary, Eren (1999), Turkish Languages Dictionary of Kasgarli Mahmut (TDK, 2003), Tietze (2009), Kanar (2011) and Turkish Dictionary (TDK, 2005) were used to determine the origins of lemmas. In addition, the points of views and conceptual fields in the books for both periods were decided by looking at the most repeated 50 words on the vocabulary frequency lists.

Results

In the research, 841 lemmas that form 1567 words in TACA and 902 lemmas that form 2298 words in TBCA were detected from the 3331 words that were taken in equal numbers within the words that were excluded from the research. There are 347 of the same lemmas used in both of the books. By coincidence, there are 127 words in both of the books that were derived from these words. In summary, there are 474 words which are based on the same roots and which appear in both of the books. When the words from different roots were analyzed, 416 lemmas were observed that were derived from 371 different roots in TBCA and 345 lemmas that were derived from 323 different roots in TACA, as seen in Table 2.

Table 2

The Numbers of the Same and Different Lemmas in TACA and TBCA

Book	Word	lemma	same lemma	same root Lemma	Same root total	Different root	Different lemma
TBCA	2298	902	347	127	474	371	416
TACA	1567	841	347	127	474	323	345
difference	731	61	-	-	-	48	71

Of the lemmas that appear in both books (347), 264 have Turkish, 47 have Arabic, 27 have Persian, 2 have Mongolian, 2 have Armenian, 2 have Greek and 2 have French roots. When the percentages of bases of these words are compared with the percentages of bases of different words in the book, the results in Table 3 are reached.

Table 3

The Percentages of the Same and Different Lemmas in TACA and TBCA in Terms of Their Roots

	Turkish %		Arabic %		Persian %		Other %	
Book	Same Root	Dif. root	Same root	Dif. root	Same root	Dif. root	Same root	Dif. root
TBCA	76	52	13,5	21,6	7,8	7,5	2,3	18,5
TACA		43,6		34,9		8,6		13,6

While it is observed that the percentages of the same lemmas of Turkish origin in both of the books (76%) is higher than the words with other roots, the words of Turkish origin in TACA (43.6%) appear at a lower percentage than TBCA (52%). The percentage of the same Arabic words (%13.5) is lower than the percentage of Arabic words derived from different roots in both of the books; nevertheless, the percentage of different Arabic roots (%34.9) is higher in TACA. In terms of words with Persian origin, there is not a significant difference in the percentages of the same and different roots in the books from both of periods. However, the percentages of the roots of other languages that are the same in both periods (2.3%) is lower than for those from different roots. On the other hand, the percentage in TBCA (18.5%) is greater than the percentage in TACA (13.6%). The origins and the number of roots of different lemmas in TACA and TBCA are shown in Table 4 in detail.

Table 4

Origins and Numbers of Different Lemmas in TACA and TBCA

Book	Tur	Ar.	Pe	Fr	İng.	Gre.	İt.	Arm.	Kur	Ven	Ger.
TBCA	193	81	28	47	7	6	2	1	2	2	1
TACA	141	113	28	19	4	10	5	3	0	0	0
Differ.	52	32	-	28	3	4	3	2	2	2	1

As illustrated in Table 4, the number of Turkish roots (193) in TBCA decreases in TACA (141); the number of Arabic roots increases by 32; and, there is no change in the number of Persian roots. In terms of Western languages, there is a decrease in the number of French roots by 28, English roots by 3, but an increase in Greek roots by 5, Italian roots by 3, and Armenian roots by 2. On the other hand, there are no roots from Kurdish, Venice language or German. The numbers in Table 5 are reached when the total percentage of words in the textbooks, apart from the same lemmas, are compared to the roots of the most repeated 100 words.

Table 5

Comparison of the Most Repeated 100 words apart from the Same Words with the Total Percentages of Roots

	Turkish%		Arabic%		Persian%		Other%	
Book	First 100	Total	First 100	Total	First 100	Total	First 100	Total
TBCA	64	52	17	21,6	2	7,5	17	18,5
TACA	53	43,6	27	34,9	11	8,6	9	13,6

In TBCA, 64 of the Turkish words in the first 100 appear with great frequency compared to Turkish root percentage (52%), 17 words with Arabic roots are lower than the total Arabic percentage (21.6%), 2 words with Persian roots are lower than the total percentage (7.5%), and the first 100 and the total percentage are quite equal in the "other" category, where mostly western-based words are seen. When TACA is analyzed, the number of Turkish words (53) in the first 100 is observed to be lower compared to TBCA, while the Arabic (27) and Persian (11) are higher. In TACA, the percentage of the first 100 Turkish words (53) is higher than the total percentage (43.6). While the percentages of first 100 Arabic (27) and other (9) words are lower than the total percentage, the percentage of first 100 Persian words (11) is higher than the total percentage (8.6). This result implies that in both of periods, even though the number of words with Turkish roots in first 100 is higher, they are used less frequently in total, and the opposite applies to Arabic words.

Table 6 shows the results of the t-test, which was performed in order to determine whether or not the differences between the frequencies of different lemmas in TACA and TBCA, according to the Written Turkish Word Frequency Dictionary written by Goz (2003), are significant. While applying t-test in a parametric way, Turkish and Arabic words were appropriate in terms of number of lemmas, non-parametric Mann Whitney U test iwass applied for Persian and western languages (French/English) as the number of lemmas was too low.

As a result, no significant difference is observed in total between the lemmas, in terms of word frequency in the text books, between the periods [t(759)=-,617, p>0.05]; and on the other hand, in terms of the frequency of Turkish lemmas [t(759)=-1,190, p>0.05] and the frequency of Arabic lemmas [t(759)=-,442, p>0.05], no significant difference is seen between the words used in TACA and TBCA.

Table 6

The Result of t Test Presenting the Difference between the Frequencies of Different Lemmas in TBCA and TACA.

variables		N	$\bar{x}$	Ss	Sd	t	P
Turkish	TBCA	225	124,26	215,76	374	-1,190	,235
	TACA	151	150,44	198,82			
Arabic	TBCA	87	116,81	149,68	205	-,442	,659
	TACA	120	127,01	173,87			
Total	TBCA	416	123,53	205,51	759	-,617	,538
	TACA	345	132,36	185,22			

Table 7 includes the results of Mann Whitney U test, which was applied in order to observe the significance of the difference between frequencies of western-based and Persian lemmas that are not the same in the books of the two periods.

Tablo 7

The Result of Mann-Whitney U Test for the Various Persian- and Western-based Words in TBCA and TACA

Language	N	S.O.	S.T.	U	Z	P
TBCA						
Persian	29	29,52	856,00			
TACA				420,000	-,008	,994
Persian	29	29,48	855,00			
TBCA						
Western based	57	39,89	2274,00			
TACA				462,000	-,954	,340
Western based	19	34,32	4652,00			

A significant difference is not seen between the word frequencies of Persian-based words (U=420,000, p>0.05) and western-based (French/English) words (U=462,000, p>0.05), which are different in the books of the two periods.

As a result of the evaluation of vocabulary frequency, focusing on the frequency and roots of synonyms in terms of Turkish synonyms that are alternative to loanwords, when the words in the texts in TACA are analyzed according to Çotuksöken (2012), it is observed that either the words with less frequency or more frequency are used, or both of them are used.

Words Containing Both of Synonyms

(Arabic-Turkish): reason [*sebep* (311)] – cause [*neden* (323)] / homeland [*vatan* (58)] – country [*yurt* (105)] / neighbourhood [*etraf* (188)] – environment [*cevre* (708)] / to mention [*bahsetmek* (29)] – to speak of [*soz etmek* (190)] / simple [*basit* (181)] – easy [*kolay* (455)] / answer [*cevap* (381)] – reply [*yanit* (194)] / before [*evvel* (47)] - previous [*once* (1587)] / to suppose [*farz etmek* (6)] – to assume [*varsayma* (46)] / expression [*ifade* (253)] – statement [*anlatim* (933)] / field [*saha* (114)] – area [*alan* (15)] / supply [*tedarik* (1)] – provide [*saglamak* (965)] / fuss [*velvele* (2)] –rumble [*gürültü* (131)]

(Persian-Turkish): Trouble [*dert* (158)] – sadness [*uzuntu* (58)]

Words containing the synonyms with more frequency

(Arabic-Turkish): wreck [*enkaz* (12)] – ruin [*yikinti* (9)] / enthrallment [*esaret* (8)] – tutsaklık [*captivity* (6)] / suspicious [*supheli* (18)] – doubtful [*kuskulu* (15)] / flesh [*vucut* (503)] – body [*beden* (252)]

(*Arabic-Mongolian*): nation [*millet* (211)] – people [*ulus* (50)]

(Arabic-Turkish): possibility [*ihtimal* (87)] – probability [*olasilik* (106)] / generation [*nesil* (64)] – descendants [*kuşak* (100)] / level [*seviye* (103)] – degree [*duzey* (382)] / indigenous [*tabii* (50)] -natural [*dogal* (343)] / exoneration [*beraat* (6)] – be absolved [*aklanmak* (11)] / term [*devre* (138)] – period [*donem* (772)] / needy [*fukara* (14)] – poor [*yoksul* (64)] / longing [*hasret* (37)] –missing [*ozlem* (73)] / invention [*icat* (14)] – discovery [*buluş* (34)] / demonstration [*ispat* (18)] – to prove [*kanitlamak* (53)] / crop [*mahsul* (8)] – product [*urun* (759)] / thriving [*mamur* (3)] – prosperity [*bayindirlik* (12)] / issue [*mesele* (201)] – problem [*sorun* (915)] / destiny [*nasip* (8)] – share [*pay* (152)], income [*kazanc* (65)] / finally [*nihayet* (77)]– at last [*sonunda* (352)] / round [*sefer* (107)] – time [*kez* (642)] / *attester* [sahit (16)] – *witness* [tanik (28)] / *eviler* [ser (14)] – *malignancy* [kotuluk (43)] / task [*vazife* (42)] – duty [*gorev* (522) /] trappings [*ziynet* (4)] – ornaments [*sus* (26)]

(Persian-Turkish): merrymaking [cumbus (4)] – entertainment [eglence (115)] / remedy [care (113)] – solution [cozum (248)] / to coincide [rastlamak (95)] / to come accross [karsilasmak (214)]

(*Arabic-Persian):* spleen [garaz (1)] – resentment [kin (33)] / hostile [hasim (6)] – enemy [dusman (137)] / while [vakit (195)] – time [zaman (23)]

(Arabic- Arabic): to consume [sarf etmek (2)] – to spend [harcamak (149)] / round [sefer (107)] – time [kere (211)]

Words Containing the Synonym with Less Frequency

Differences were discovered between the conceptual fields analyzed in TACA and TBCA. Content words are in the upper range of the frequency lists in the corpus prepared by narrowing down subjects (Kennedy, 1998, cited in Aksan, Mersinli and Yaldır, 2011). When the most frequent 50 words in the numerically-ordered lists that are created for the textbooks of both periods are analyzed, it is seen that forest (88), tree (20), fire (17), to burn (22), cause (10), and reason (10) occupy the first rows in the new books, and "forest fires" and "the harm that human beings to nature" are underlined; conversely, the older books focus on water (32), sea (29), lake (29) and otherwise desert (10) and "warm". In addition to this, while the older books pay equal attention to mountain (28), forest (22) and hill (19), they also focus on different geographical terms and concepts such as sun (13), hillside (13), snow (12), storm (11), shore (10), stream (10), soil (10), dream (9), spike (8), rock (8), countryside (8) and tree (8). In the newer books, the frequency of words such as city (16), field (16), soil (13), house (12), land (8), man (8), villager (8), and to stay (17) draws attention to the dominance of human beings over nature. New books refer to plant and animal species such as leaf (11), bird (9), pine (8), and goat (8) repeatedly, while not paying much attention to mountain (7). Newer books repeatedly use sky (10), sphere (8), star (16) and planet (6), but these concepts are not seen in the old books, which show human beings who are finished with earth and seek their future in the sky. The older books underline "seeing (21)" nature and the universe, while the newer books are more didactic, referring to "saving" from burning (22), as can be understood from the frequencies with which "to burn" (22) and "to fire" (7) appear.

Discussion and Conclusion

Regardless of their origins, words with a high frequency of use should be included in the textbooks during the primary educational years. There is not a significant difference between TBCA and TACA and the frequency of use of Turkish, Arabic, Persian and western-based words. The use of identical texts from the same writers, such as Yasar Kema,l in both of the periods may be the reason for the lack of significant difference. There is a great difference between the synonyms in TACA where both of the synonyms are included [supply (1) – provide (965) / fuss (2) – rumble (131)]. Among the words that do not have synonyms in the same book in TACA, the words used with less frequency are preferred; however, in the case of synonyms, the use of words with higher frequency will increase the rate of understanding and decrease the number of unknown words. The same problem is seen in Turkish coursebooks for foreigners. The results of Ozdemirel's (2017) research evaluating words in Turkish and English coursebooks for foreigners in terms of frequency of occurance showed that English coursebooks present more frequently used words than their Turkish counterparts.

In TACA, concepts related to nature are seen less frequently, and instead of offering the students an understanding of the joy of life, curiosity; the books direct students to more negative acts and adopts a didactic and accusatory manner. Loving

nature is a prerequisite for preserving nature. In textbooks, including different concepts in a theme will generate ideas and eventually expand the range of ideas.

Unknown words should be chosen carefully by taking the frequencies of words in textbooks into consideration in the specified rates and deciding on the words that students should know based on their grade levels. Word frequency for Turkish can be obtained from TUD created with written texts, and Ts Corpus created with written correspondence such as newspapers, forms and conversations in virtual environments (Karaoglu, 2014). In addition, more vocabulary enhancement activities should take place in textbooks. The results of experimental research by Topkaraoğlu and Dilman (2013) showed a significant difference between a control group that followed the regular curriculum, which included learning the second one thousand most frequently used words in English, and the experimental group, which had a fourteen-week schedule of vocabulary enhancement activities including integrating the same second one thousand words into the regular curriculum. Institutions should also create criteria and data surrounding this topic. At this point, while deciding on the words to be taught based on grade levels, frequency studies may play a crucial role (Nation & Newton 1997).

References

Adali, O. (2004). *Turkiye Turkcesinde bicimbirimler* [Morphemes in Turkish](2. Basım). Istanbul: Papatya Yayıncılık Egitim.

Aksan, D. (1982). *Her yonuyle dil-Ana cizgileriyle dilbilim* [Language in all Perspective-Linguistics within Basic Lines]. Ankara: Turk Dil Kurumu.

Aksan, D. (1997). *Anlambilim: Anlambilim konulari ve Turkcenin anlambilimi* [Semantics: Topics in semantics and Turkish semantics]. Ankara: Engin Yayınevi.

Aksan, Y., Mersinli, U. & Yaldir, Y. (2011). İlkogretim Turkce ders kitaplari derlemi ve Turkce ulusal dil derlemi orneklemindeki sozcuk sikliklari [The corpus of Turkish textbooks and the word frequency in the sample of Turkish National Corpus]. In: Günay, D. et al (Eds). (2011). *Turkce ogretimi uzerine calismalar*, 397- 408. İzmir: Dokuz Eylul Universitesi.

Aksan, Y. et al. (2012). Construction of the Turkish National Corpus (TNC). *Proceedings Of The Eight İnternational Conference On Language Resources And Evaluation* (LREC 2012). Istanbul. Turkiye. http://www.lrec-conf.org/proceedings/lrec2012/papers.html

Apaydin, N. (2010). *6. sinif Turkce ders kitaplarinin soz varligi acisindan incelenmesi [The evaluation of Turkish textbooks in the context of vocabulary].* [M.A. Thesis, Marmara Universitesi, Egitim Bilimleri Enstitusu].

Ari, G. (2003). *İlkogretim altinci sinif ders kitaplarindaki kelimeler ve bu kelimelerin kullanilabilirligi [The vocabulary in the 6th Grade in primary education and the usage of these].* [M. A. Thesis, Marmara Universitesi, Istanbul]

Buyukozturk, S, Cakmak, K. E., Akgun, O.E., Karadeniz, Ş. & Demirel, F. (2009). *Bilimsel arastirma yontemleri [Scientific research methods]*. Ankara: Pegem Akademi.

Cotuksoken, Y. (2012). *Turkce anlamdaslar ve karsitanlamlilar sozlugu [The dictinary of Turkish synonyms and antonyms]*. Istanbul: Papatya.

Author. (2013). The efficiency of secondary school Turkish students in using cohesive devices. *13. Uluslararası Dil, Yazin ve Deyişbilim Sempozyumu: Basit Uslup*, 1169-1182.

Author. (2014). Turkçe ogretiminde sozcuk ogretim yontemlerinin yeterliligi [The efficiency of vocabulary teaching methods in Turkish language teaching]. *Adiyaman Universitesi Sosyal Bilimler Enstitusu Dergisi*, 17, 233-258.

Eren, H. (1999). *Turk dilinin etimolojik sozlugu [The etymological dictionary of Turkish]*. Ankara: Bizim Buro Basim Evi.

Goz, I. (2003). *Yazili Turkcenin kelime sikligi sozlugu [The frequency dictionary of written Turkish].* Ankara: TDK

Gunes, F. (2013). *Turkçe ogretimi. Yaklaşimlar ve modeller [Turkish language teaching. Approaches and models.]*Ankara: Pegem Akademi.

Guzel, A. (2006). İlkogretim Birinci Kademe Ogrencilerine Ortak Kelime Hazinesi Kazandirmanin Onemi [The Importance of Improving Common Vocabulary of Students in Primary Education]. In: C. Yıldız & L. Beyreli (Eds). *Dilbilim, Dil Ogretimi ve Ceviribilim Yazilari*, 323-331. Ankara: Pegema Yayinlari.

Hatch, B. & Brown, C. (1995). *Vocabulary, semantics and language education.* New York, ABD: Cambridge University Press.

Kanar, M. (2011). *Osmanli Turkcesi sozlugu [The dictionary of Ottoman Turkish].* Istanbul: SAY Yayinevi.

Karadag, O. (2005). *İlkogretim 1. kademe ogrencilerinin kelime hazinesi uzerine bir araştirma [A research on the vocabulary of students at the 1st level in primary education]*. [Ph. D. dissertation, Gazi Universitesi, Egitim Bilimleri Enstitusu].

Karaoglu, S. (2014). Turkçe Cevrimici Derlemler Uzerine. *KMU Sosyal ve Ekonomik Araştirmalar Dergısi, 16 (Ozel Sayi I)*, 181-188.

Karasar, N. (1984). *Bilimsel araştirma yontemi: Kavramlar, ilkeler, teknikler [Scientific research method: Notions, principles and techniques.* (3.Baskı). Ankara: Bilim

Kasgarli M. (1071). *Divan-i Luqat-i it-Turk [The dictionary of Turkish].* TDK (Ed.). Divan-i Luqat-i it-Turk Dizini, 2003.

Kiran, Z. & Kiran A. (2001). *Dilbilime giris [Introduction of linquistics].* Ankara: Seckin Yayincilik San. ve Tic. A.S.

Kurt, A. (Agustos 2007). *Metin analizi [Text analysis].* http//ce.istanbul.edu.tr/atakan/metinanalizi/ce.istanbul.edu.tr.]. 2014, May 15.

Kurudayioglu, (2005). *İlkogretim II. Kademe ogrencilerinin Kelime Hazinesi Uzerine Bir Arastirma* [*A research on the vocabulary of Students in Secondary Education*]. Ankara: Gazi Universitesi Egitim Bilimleri Enstitusu (Yayimlanmamiş Doktora Tezi).

Nation, P. & Newton, J. (1997). Teaching vocabulary . In: J. Coady & T. Huckin (Eds). *Second Language Vocabulary Acquisition*, 238-255. USA: Cambridge University Press.

Nation, P. (2001). Using small corpora to investigate learner needs. In: Ghadessy, M., Henry, A. & Robert L. Roseberry, R. L. (Eds.) *Small Corpus Studies and ELT,* 31-47. Amsterdam: John Benjamins.

Nisanyan, S. *Nisanyan Sozluk [Nişanyan's dictionary].* https://www.nisanyansozluk.com. 2015, October 10.

Ozbay, M. ve Meylanlioglu, D. (2008). Turkce egitiminde kelime hazinesinin onemi [The importance of vocabulary in Turkish education]. *Yuzuncu Yil Universitesi Egitim Fakultesi Dergisi, V*(1), 30-45.

Ozdemir, E. (2011). *Yuzler ve sozcukler [Faces and words].* Ankara: Bilgi Yayinevi.

Ozdemir, E. (2012). *Anlatim Sanati [The art of telling]*(17.basım). Ankara: Bilgi Yayinevi.

Ozdemirel, A. Y. (2017). *Yabancı Dil Olarak Turkçe ve İngilizce Ders Kitaplarindaki Sozcuklerin Kullanim Sikligi Baglaminda Degerlendirilmesi* [*The Evaluation of The Words in The Textbooks for Teaching Turkish and English as A Foreign Languages in Terms of Frequency of Occurance].*Yayimlanmamis Yuksek Lisans Tezi. Istanbul Universitesi. Sosyal Bilimler Enstitusu.

Ozturk, H. (2013). Cagdas Turk romaninda bir sozvarligi incelemesi [An evaluation of vocabulary in contemporary Turkish novel]. *E-Dil Dergisi*, 157-173.

Tietze, A. (2009). *Tarihi ve etimolojik Turkiye Turkcesi lugati* (1914-2003)*[The historical and etymological Turkish dictinary]*. Wien: Verlag der Osterreichischen Academic der Wissenschaften.

Topkaraoglu, M. & Dilman, H. (2013). Sozcuk gelistirme aktivitelerinin ogrencilerin sahip oldukları genel dil seviyeleri uzerindeki etkisi [The effect of the word developing activities on language level of students]. *Bati Anadolu Egitim Bilimleri Dergisi,* 4 (8), 15-29.

Turhan, H. (2010). *8. sinif Turkce ders kitaplarindaki metinlerin soz varligi acisindan incelenmesi [Evaluation of the texts in 8th grade Turkish textbooks in the context of vocabulary].* [M.A. Thesis, Marmara Universitesi Egitim Bilimleri Enstitusu].

Turk Dil Kurumu (TDK, 2005). *Turkce Sozluk* [Turkish Dictionary]. Ankara: Turk Dil Kurumu.

Uludag, C. (2010). *7. sinif Turkce ders kitaplarinin soz varligi acisindan incelenmesi [Evaluation of the texts in 8th grade Turkish textbooks in the context of vocabulary]*. [M.A. Thesis, Marmara Universitesi Egitim Bilimleri Enstitusu].

Vygosky, L. S. (1998). *Dusunce ve dil* [Mind and language]. Istanbul: Toplumsal Donusum Yayinlari.

Yurdakul, B. (2005). *Yapilandirmacilik [Constructivism]*. In: O. Demirel (Ed). Egitimde Yeni Yonelimler. Ankara: Pegem Yayincilik.

Yapılandırmacı Yaklaşım Öncesi ve Sonrası Türkçe Kitaplarındaki Sözcüklerin Sıklık Bağlamında Değerlendirilmesi

Atıf:

Diliduzgun, S. (2018). An evaluation of the vocabulary in Turkish textbooks before and after the constructivist approach in the context of frequency. *Eurasian Journal of Educational Research, 73,* 113-130, DOI: 10.14689/ejer.2018.73.7

Özet

Problem Durumu: Türkiye'de ana dili eğitimi bağlamında yapılan araştırmalarda ilköğretim öğrencilerinin yazılı metinlerinde hep aynı sözcükleri kullanmasının, ne etken ne de edilgen sözvarlıkları açısından başarılı olmalarının, ders kitaplarında kullanılan sözcük ve kavram sayısının diğer ülkelere oranla oldukça düşük olmasının ve ders kitaplarında sözcüklerin toplam ve farklı sözcük bakımından sınıf düzeyleri arasında oranlı olmayan bir artış göstermesinin bulgulanması sözcük öğretiminin içeriği ve biçimi üzerine daha çok kuramsal ve uygulamaya dönük araştırmalar yapılmasının gerekliliğini ortaya koymaktadır. Türk Eğitim Sisteminde 2005 yılından beri benimsenen yapılandırmacı yaklaşım doğrultusunda dil öğretiminde de, bilginin idraki, üretimi ve eski bilgilere dayanılarak oluşturulması için metinlerde geçen sözcüklerin kullanım sıklıklarının yüksek olması gerekmektedir

Araştırmanın Amacı: Yapılandırmacı yaklaşım öncesi (YÖTDK) ve sonrası (YSTDK) Türkçe ders kitaplarında yer alan metinlerdeki sözcükleri sıklık ve köken değişkenleri bağlamında karşılaştırmaktır.

Araştırmanın Yöntemi: Genel tarama modelinin temel alındığı bu nicel araştırmanın evrenini yapılandırmacı yaklaşım öncesi ve sonrası Türkçe ders kitaplarında yer alan

bilgi iletici metinlerdeki sözcükler oluşturmaktadır. Örneklem olarak; seçkisiz, tesadüfi örnekleme yöntemi ile 2013-2014 öğretim yılında "Doğa ve Evren" teması altında yer alan bilgi iletici metinlerle ile 2001-2002 öğretim yılında okutulan ders kitaplarında, üniteler temaya göre ayrılmadığı için, bu tema altına alınabilecek eşit sayıda sözcüklü bilgi iletici metinler seçilmiştir. Tesniere'in dörde indirgediği dilbilgisi ulamlarından ad, eylem, sıfat ve belirteçlerden hareketle, özgür biçimbirimler kapsamında ve başsözcük temelli düzenlenen derlemlerde sayısal sıralı sözcük listeleri oluşturulmuştur. Yapısal yaklaşım öncesi ve sonrası derlemler kendi aralarında ve Türkçe Ulusal Derlemi'ndeki kullanım sıklıkları bağlamında karşılaştırıldıktan sonra sıklık sayısı yüksek sözcükler köken ve sözcük türü yüzdeleri açısından da incelenmiştir. Ayrıca iki dönemin kitaplarında sözcük çalışmalarında hedeflenen sözcükler de sıklık, köken ve tür olarak karşılaştırılmıştır.

Araştırmanın Bulguları: YÖTDK'de 2298 sözcük çeşidinin oluşturulduğu 902 başsözcük, YSTDK'de ise 1567 sözcük çeşidini oluşturan 841 başsözcük saptanmıştır. Her iki kitapta da ortak olarak kullanılan 347 başsözcük vardır. YÖTDK'de 371 farklı kökten oluşmuş 416, YSTDK'de ise 323 farklı kökten oluşmuş 345 baş sözcük olduğu görülmüştür. YÖTDK'de Türkçe kök sayısının (193) YSTDK'de (141) düştüğü, Arapça kök sayısında 32 artış olduğu, Farsça kök sayısında ise bir değişiklik olmadığı görülmüştür. Batı dilleri bağlamında ise YSTDK'de Fransızca köklerin 28 ve İngilizce köklerin 3 farkla azaldığı, Yunanca köklerin 4, İtalyanca köklerin 3, Ermeniceden gelen kök sayısının da 2 farkla arttığı görülürken Kürtçeden, Venedik dilinden, Soğdçadan ve Almancadan gelen kök yoktur. YÖTDK'de ilk yüzde yer alan 64 Türkçe kökenli sözcüğün toplam Türkçe köken yüzdesinden (%52) fazla olduğu, ilk yüzde yer alan 17 Arapça kökenli sözcüğün toplam Arapça yüzdesinden (21,6) az olduğu, ilk yüzde yer alan Farsça kökenli 2 sözcüğün toplam yüzdeden (%7,5) az olduğu ve özellikle batı kökenli sözcüklerin yer aldığı "diğer" kategorisinde ise ilk yüz ile toplam yüzdenin hemen hemen eşit olduğu görülmektedir. YSTDK'de de ilk yüz Türkçe kökenli sözcüğün oranı (53), toplam orandan (43,6) çoktur, ancak Arapça (27) ve diğer kökenli (9) ilk yüz sözcüğün oranı, toplam oranlardan azken Farsça ilk yüz sözcüğün oranı (11) toplam orandan (8,6) çoktur. Bağımsız t testi ve Mann Whitney U testlerinin sonucunda da iki döneme ait olan kitaplarda ortak olmayan baş sözcüklerin kelime sıklığı açısından toplamda anlamlı bir fark göstermediği gibi [t(759)=-,617, p>0.05], Türkçe baş sözcüklerin kelime sıklığı [t(759)=-1,190, p>0.05], Arapça kökenli baş sözcüklerin kelime sıklığı [t(759)=-,442, p>0.05], Farsça kökenli sözcüklerin kelime sıklıkları (U=420,000, p>0,05) ve batı kökenli (Fransızca/İngilizce) sözcüklerin kelime sıklıkları (U=462,000, p>0,05) açısından YÖTDK ile YSTDK arasında anlamlı bir fark görülmemiştir.

Ödünç sözcüklerin yerine geçebilecek yerleşmiş eş anlamlı Türkçe kökenli sözcükler bağlamında kitaplarda ya ikisine de yer verilmiş ya kullanım sıklığı çok olanlar ya da kullanım sıklığı az olanlar tercih edilmiştir. Her iki döneme ait kitaplar için oluşturulan sayısal sıralı listelerde en çok yinelenen ilk 50 sözcüğe bakıldığında yeni kitaplarda orman (88), ağaç (20), yangın (17), yak-(22), neden (10), sebep (10) ilk sıralarda yer alırken "orman yangınları" ve "insanların doğaya verdiği zarar vurgulanmış", eski kitaplarda ise daha çok su (32), deniz (29), göl (29) ve aksi

durumda çöl (17) ve sıcak (10) olma durumu üzerinde durulmuştur. Bunun yanı sıra eski kitaplarda doğa ile ilgili dağ (28), orman (22), tepe (19) kavramlarına da eşit oranda yer verilirken güneş (13), yamaç (13), kar (12), rüzgâr (11), kıyı (10), dere (10), toprak (10), serap (9), diken (8), kaya (8), köy (8) ve ağaç (8) gibi farklı coğrafi terimler ve farklı kavramlar üzerinde de durulmuştur. Yeni kitaplarda ise kent (16), tarla (16), toprak (13), ev (12), arsa (8), adam (8), köylü (8), kal- (17) sözcüklerinin sıklığı insanların doğa üzerindeki hâkimiyetine dikkat çekmektedir. Yeni kitaplarda gökyüzü (10), gök (8), yıldız (16) ve gezegen (6) sözcükleri de bulunmaktadır. Eski kitaplar doğa ve evreni "gör- (21)" meyi vurgularken, yeni kitaplarda yak-(22) ve yan- (7) eylemleri sıklık göstermektedir.

Araştırmanın Sonuçları ve Önerileri: YÖTDK'de 48 daha fazla kökten 71 daha fazla başsözcüğe rastlanmıştır, yani daha fazla kavram yer almaktadır. YÖTDK'de daha fazla Türkçe kökenli ve Batı kökenli sözcük varken YSTDK'de daha fazla Arapça sözcük bulunmaktadır. YSTDK'ye bakıldığında ise metinlerde en çok kullanılan ilk yüzde Türkçe kökenli sözcük sayısı YÖTDK'den daha azdır, Arapça ve Farsça kökenli sözcük sayısı ise daha çoktur. Her iki dönemde de Türkçe kökenli sözcüklerin ilk yüzde sayılarının çok olmasına rağmen toplamda daha az kullanıldığı, Arapça sözcüklerin ise ilk yüzde oranları daha azken toplamda daha çok kullanıldığı görülmektedir. Her ne kadar her iki dönemde de köken olarak bazılarının oranları yükselip bazılarınınki azalıyorsa da YÖTDK ile YSTDK arasında Türkçe, Arapça, Farsça ve batı kökenli sözcüklerin kullanım sıklıkları açısından anlamlı bir fark bulunamamıştır. Yapılandırmacı yaklaşım bağlamında sözcüklerin kökenlerinden çok sıklık oranlarına göre ders kitaplarında yer almaları önemlidir.

YSTDK'de eşanlamlı sözcüklerin her ikisini de bulundurduklarında sözcüklerin sıklıkları arasında büyük fark vardır, eşanlamı kitapta bulunmayan sözcüklerden ise kullanım sıklığı az olanların daha çok tercih edildiği görülmüştür. Eşanlamlı sözcüklerden kullanım sıklığı yüksek olanların tercih edilmesi anlamı bilinmeyen sözcük sayısının oranını düşürecek ve anlama oranını yükseltecektir.

YSTDK'de doğa ile ilgili daha az kavram yer almakla birlikte YÖTDK'de olduğu gibi öğrencilere yaşam sevinci katmak, merak arttırmak, güzellikleri hissederek korumayı sezdirmek yerine daha didaktik bir tavır sergilendiği görülmektedir.

Düzeylerine göre öğrencilerin anlamını bildikleri sözcükler belirlenerek saptanan oranlarda anlamı bilinmeyen sözcükler metinlerde kullanım sıklığı dikkate alınarak seçilmelidir. Bu konuda ilgili kurumlar tarafından bir ölçüt ve veritabanı oluşturulmalıdır. Bu noktada düzeylere göre öğretilecek sözcüklere karar verilirken sıklık çalışmalarının önemli yardımları olabilmektedir.

Anahtar sözcükler: Sözcük sıklığı, sözcük kökeni, başsözcük, sıklık listeleri.