Examination of the Vocabulary Diversity of Basic Level Textbooks for Teaching Turkish as a Foreign Language

Burak Tüfekçioğlu[1]
Eskisehir Osmangazi University, Eskisehir, Türkiye

Abstract

To understand, interpret, shape in our minds, and express in language the objects and concepts we encounter or use in our daily lives, we use words, in other words, vocabulary. Words have semantic and functional roles within sentence structure because they carry referential value. In this regard, vocabulary knowledge is fundamental in language teaching. Therefore, the study has focused on vocabulary. In this context, the Yunus Emre Institute's basic level A1 and A2 textbooks used in the process of teaching Turkish as a foreign language have been compared in terms of vocabulary diversity. For this purpose, the type and token usages of the relevant textbooks were extracted and analyzed by determining the type-token ratios. As a result of the research, 1100 types and 14811 tokens were identified in the Yunus Emre Institute's Yedi İklim Turkish A1 textbook. At the A1 level, the type-token ratio has been determined as 0.0742. At the A2 level, 1251 types and 17634 tokens are used, and the type-token ratio for this textbook has been calculated as 0.0709. Based on these findings, it is understood that the A1 level textbook used in the first stage of the basic level has more lexical diversity compared to the A2 level textbook, which is the next level (0.0742 > 0.0709). In conclusion, the research indicates that the textbooks in question do not have a progressive structure in terms of lexical diversity.

To cite this article:

Tüfekçioğlu, B. (2024). Examination of the vocabulary diversity of basic level textbooks for teaching Turkish as a foreign language. *Innovative Educational Research (INNER), 6,* (2), 213-219

[1] Assoc. Prof. in the Department of Turkish and Social Science Education, Faculty of Education, Eskisehir Osmangazi University, Türkiye. e-mail: burak.tufekcioglu@ogu.edu.tr

Words are indicators that are spread across every area of language and are frequently used in every environment where the language is spoken. Therefore, in language teaching, vocabulary instruction is a process that begins with the first lesson in teaching Turkish as a foreign language and continues in every environment where the foreign language is spoken. In this context, a word is defined as "the smallest meaningful units of language," but when considering words like "for" that do not have meaning on their own but serve a meaningful function when used with other words, it is also defined as "the smallest units of language with grammatical function" (Ergin, 2013, p. 95).

The 'meaning' and 'function' attributes observed in the definitions of words allow for the classification of words within these two attributes. So much so that nouns, adjectives, adverbs, and verbs are members of the class of content words, while prepositions, conjunctions, pronouns, and interjections are members of the class of function words. (Erdem, 2012). Function words are more related to grammar than vocabulary, while content words refer to specific meanings within a sentence and are used less frequently than function words; additionally, vocabulary tests generally focus on content words. (Read, 2000). This classification is particularly useful in the field of education (vocabulary teaching and grammar teaching), especially in the preparation of teaching plans and materials.

When considering the base and derivational properties of words, the term lemma is frequently used. In vocabulary studies, the base and derived forms of a word are referred to as the lemma (Read, 2000). When compared to the term 'word family,' the word family expresses the base form and its inflections, as well as its closely related derivations (manage=manager, management, managerial), while 'lemma' encompasses the base form (manage-) and its inflections (managed, managing, manager) (Read & Yen Dang, 2022). In this context, 'lemma' can be considered as a lexeme. Especially in vocabulary research, the process of lemmatization is frequently carried out. Lemmatization helps distinguish between the types and meanings of words and their derivations that are not clearly related to their base forms (Gardner & Davies, 2014). With the process of lemmatization in the corpus, dominant words are identified. In this regard, the use of lemmas is particularly beneficial in creating word lists for targeted vocabulary teaching.

In vocabulary teaching and vocabulary assessments, vocabulary size, which is important, refers to the quantity of words known by learners (Nation, 2001), while vocabulary depth refers to the quality of word knowledge, including the co-occurrences of known words and collocational knowledge (Read, 1993). Vocabulary size refers to the number of words known, while vocabulary depth or quality refers to how well those words are known. Vocabulary size knowledge refers to the calculation, measurement, and discussions focused on the volume of vocabulary, which are relatively clearly operationalized as knowledge of the structure-meaning connection of known lexical expressions. In second language vocabulary teaching, vocabulary size and depth do not necessarily develop in parallel (Schmitt, 2024).

In language education/teaching, it is necessary to use rich vocabulary elements as much as possible, both in the vocabulary development of students/learners and in the materials produced. At this point, the concept of lexical richness comes to the forefront. Lexical richness is grounded in three dimensions: density, diversity, and sophistication. In instructional studies, the ratio of erroneous words used by second language learners is also

added to these three dimensions (Read, 2000). However, the determination of vocabulary richness directly related to the material is based on the word density, diversity, and comprehensiveness in the texts. Word density refers to the ratio of meaningful words to all words in a text. Due to the density of noun usage, it is a determining factor in the quality of writing in academic and advanced informational texts.

Lexical diversity refers to the number of different single word units in a text or the distribution of words within the text, and it is particularly an indicator of lexical competence and development. Lexical diversity refers to the different words present in a text, avoiding repetitions, and is measured by comparing the number of different words to the total number of words. At this point, the concepts of type and token are important for measuring word diversity, and the study has focused on these concepts. A token refers to the total number of instances of single words used multiple times in a text, while a type refers to the different word forms in a text, where each distinct word is counted as 'one' regardless of its frequency of use. (Read, 2000, s. 18).

When students are asked to write a 500-word composition, this number represents the token count, while the type of count in the texts written by the students will be lower. Type is referred to as "individual words" in various sources, and it is also called an independent word unit that is distinct from the repeating derivatives in the texts, with each representing different lexical units (Coxhead, 2000). Especially in the evaluation of writing skills in terms of word diversity, the type-token ratio is used to determine word diversity. The Type-Token ratio is a value between 0 and 1 obtained by dividing the number of word units in the text by the number of types, and a higher value represents a wider lexical diversity.

Based on this information, the study examined the Yedi İklim Turkish A1 and A2 textbooks used at the basic level of teaching Turkish as a foreign language in terms of vocabulary diversity. The research aims to compare the A1 textbook used in the first stage of the beginner level of Turkish language teaching as a foreign language with the A2 textbook used in the second stage in terms of vocabulary diversity within the scope of the study, the following questions were sought to be answered:

1. How many types are there in the vocabulary of the Yunus Emre Institute Yedi İklim Turkish A1 textbook?
2. How many tokens are there in the vocabulary of the Yunus Emre Institute Yedi İklim Turkish A1 textbook?
3. What is the type-token ratio in the vocabulary of the Yunus Emre Institute Yedi İklim Turkish A1 textbook?
4. How many types are there in the vocabulary of the Yunus Emre Institute Yedi İklim Turkish A2 textbook?
5. How many tokens are there in the vocabulary of the Yunus Emre Institute Yedi İklim Turkish A2 textbook?
6. What is the type-token ratio in the vocabulary of the Yunus Emre Institute Yedi İklim Turkish A2 textbook?

Method

The study is of the descriptive research type. Descriptive research is the process of describing and explaining what currently exists and what has happened (Sönmez and Alacapınar, 2013, p. 48). Within the scope of the study, the 2015 editions of the Yedi İklim Turkish A1 and A2 textbooks published by the Yunus Emre Institute were examined. The word lists of the textbooks were extracted using the kfNgram program, and lemmatization was applied along with their frequencies. In the study, the purposive sampling technique, which is a type of non-probability sampling, was used. In this type of sampling, since the probabilities of the units entering the sample are not the same or unknown, the results cannot be generalized to the universe (Erkuş, 2017, p. 137; Balcı, 2018, p. 99).

Findings

In the study, the type and token ratios of the Yunus Emre Institute's Yedi İklim Turkish A1 and A2 textbooks were extracted, and the findings are presented below.

Table 1

Example of Type and Token List at the A1 Level

Tip	Token	Tip	Token
ünite	870	iletişim	81
ve	209	oku	78
ben	189	yap	76
meslek	188	siz	74
bir	185	ev	70
ne	182	nerede	69
var	167	merhaba	68
bu	150	hanım	63
günlük	148	nasıl	63
tanışma	133	sen	61
çok	130	etmek	60
ulaşım	126	murat	59
hayat	125	gir	58
et-	119	telefon	58
mı	113	bilgi	56
on	112	hava	56
aile	111	en	55
çevremiz	93	kaç	53
ol	92	sonu	52
saat	89	tamamlayalım	51
yaz	89	okul	50
sonra	87	teşekkür	50
hangi	85	ile	49
tatil	83	bayram	48
gel	82	derya	48

Table 1 contains 50 type-token examples at the A1 level. The first column of the table shows the 'types'; the second column shows the 'token' count. The entire table is provided in Appendix 1.

Table 2

Type/Token Ratio at the A1 Level

Tip	1100
Token	14811
Tip/Token	0,0742

Table 2 shows the type-token ratio at the A1 level. Within the scope of the study, it has been determined that 1100 types were used at the A1 level of the Yunus Emre Institute, and

the total number of tokens for these 1100 types was 14811. The type/token ratio at the A1 level is 0.0742.

Table 3

Example of Type and Token List at the A2 Level

Tip	Token	Tip	Token
ünite	758	etkinlik	97
ve	574	ev	92
bir	291	sağlıklı	92
bu	237	duygu	87
ne	195	hanım	87
çok	175	yaz	85
ben	155	Gel-	83
mı	151	zaman	82
güzel	148	Oku-	78
insan	139	iste-	75
için	132	bey	74
de	125	gün	72
ol	125	her	71
tüketim	123	al	70
üretim	123	daha	68
sonra	121	en	68
et-	115	sen	66
git-	111	geç	65
o	111	iyi	64
var	111	dünya	63
sosyal	110	kadar	63
ülke	108	etmek	62
teknoloji	105	gibi	60
yaşam	105	hangi	60
iletişim	104	önce	60

Table 3 contains 50 type-token examples at the A2 level. The first column of the table shows the 'types'; the second column shows the 'token' count. The entire table is provided in Appendix 1.

Table 4

A2 Düzeyi Tip/Token oranı

Tip	1251
Token	17634
Tip/Token	0,0709

Table 4 shows the type-token ratio at the A2 level. Within the scope of the study, it has been determined that 1251 types were used at the A2 level of the Yunus Emre Institute, and the total number of tokens for these 1251 types was 17634. The type/token ratio at the A2 level is 0.0709.

Conclusion

The concept of a word is functionally used in language and mind, and it continues to develop and change in the mind in terms of thought and expression capacity, related to vocabulary development in language, in connection with different sub-disciplines of linguistics such as phonetics, morphology, syntax, and semantics. Therefore, the ways in which words exist in language, the processes and styles through which individuals acquire/use these words, and the teaching/learning of words are at the intersection of language teaching and linguistics. Linguistics conceptually describes the 'word' according to its related sub-disciplines, and language teaching also utilizes vocabulary in the stages of material production,

lesson planning, measuring vocabulary level, and assessing language proficiency, particularly by using terms produced based on definitions made within the scope of vocabulary teaching. Therefore, knowledge of vocabulary concepts and usage is essential for language teachers.

In language education/teaching, it is necessary to use as many rich vocabulary elements as possible in both the vocabulary development of students/learners and the materials produced. Therefore, the question of how many words is taught in language education courses, along with the question of how many different types of words are taught, is also important.

In terms of teaching Turkish as a foreign language, A1 and A2 levels are considered basic levels, with A1 level being introductory. Therefore, with the increase in the number of words (types) taught in the A2 level textbooks, it is expected that the diversity of vocabulary taught for the development of a qualitative lexicon will also increase. For this reason, because of the type/token ratio calculation conducted in the study, it is understood that despite the increase in the number of words taught between the A1 and A2 levels of the Yunus Emre Institute's Yedi İklim Turkish textbooks, there is a decrease in word diversity. It is evaluated that this situation will pose a problem in terms of the quality of vocabulary teaching.

Suggestions

When preparing textbooks and materials, it is observed that the increase in word counts in the selection of texts in textbooks is taken into consideration. However, it is understood that the selected texts should also be evaluated in terms of their composition of different words and their diversity of vocabulary.

References

Barın, E., Çobanoğlu, Ş., Ateş, Ş., Balcı, M. and Özdemir, C. [Ed.]. (2015). *Yunus Emre Enstitüsü Yedi İklim Türkçe A1*. Türkiye Diyanet Vakfı Yayıncılık.

Barın, E., Çobanoğlu, Ş., Ateş, Ş., Balcı, M. ve Özdemir, C. [Ed.]. (2015). *Yunus Emre Enstitüsü Yedi İklim Türkçe A2*. Türkiye Diyanet Vakfı Yayıncılık.

Coxhead, A. (2000). A New academic word list. *TESOL Quarterly, 34*(2), 213-238

Erdem, M. (2012). Biçim Bilgisi. In (Ed. Nurettin Demir-Emine Yılmaz) *Türk Dili Yazılı ve Sözlü Anlatım,* (pp. 91-108). Nobel Yayınları.

Ergin, M. (2013). Ede*biyat ve Eğitim Fakültelerinin Türk Dili ve Edebiyatı Bölümleri İçin Türk Dil Bilgisi*. Bayrak Yayınları.

Gardner, D. & Davies, M. (2014). A New academic vocabulary list. *Applied Linguistics, 35*(3), 305–327.

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press.

Read, J. & Yen Dang, T. N. (2022). Measuring depth of academic vocabulary knowledge. *Language Teaching Research, 11* July. Pp: 1–27. https://doi.org/10.1177/ 136216882211059

Schmitt, N. (May, 2024). *Size and Depth of Vocabulary Knowledge: What the Research Shows*. https://core.ac.uk/download/pdf/162656631.pdf

Author Information

Burak TÜFEKÇİOĞLU: Assoc. Prof. Tüfekçioglu graduated from the Hacettepe University PhD program in Teaching Turkish as a Foreign Language in 2018. His areas of expertise are teaching Turkish as a foreign language and vocabulary. He is currently serving as a faculty member in the Department of Turkish Language and Social Sciences Education at Eskişehir Osmangazi University.

Conflict of Interest

There is no conflict of interest.

Funding

Ethical Standards

The study is a type of document review conducted on textbook review and vocabulary analysis, and it does not involve any participants. Therefore, ethical committee approval is not required.