http://lexikos.journals.ac.za; https://doi.org/10.5788/28-1-1472

Corpus-Based Research on Terminology of Turkish Lexicography (CBRT-TURKLEX)*

Erdoğan Boz, *Center for Lexicography, Turkish Language and Literature Department, Eskişehir Osmangazi University, Eskişehir, Turkey (erdoganboz@ogu.edu.tr)*

Ferdi Bozkurt, *Turkish Language Department, Anadolu University, Eskişehir, Turkey (ferdib@anadolu.edu.tr)*

and

Fatih Doğru, *Turkish Language and Literature Department, Eskişehir Osmangazi University, Eskişehir, Turkey (fdogru@ogu.edu.tr)*

Abstract: In this paper, we introduce an ongoing lexicographic corpus project. The Center for Lexicography, abbreviated as SÖZMER, was established under the aegis of Eskisehir Osmangazi University to support lexicographical projects. SÖZMER decided to initiate a corpus-based Turkish lexicography project. This project will be the first stage of the endeavour aimed at preparing a specialized dictionary for Turkish lexicography. The primary aim of the project is to prepare an electronic corpus for researchers of Turkish lexicography. The secondary aim of the project is to obtain a word list of Turkish lexicographic terms. This paper presents a description of the process of data collection and the methodology employed for building a specialized corpus. The study contains an outline of the project background, needs, problems, and the phases of corpus building.

Keywords: TURKISH LEXICOGRAPHY, TERMINOLOGY, CORPUS LINGUISTICS, DICTIONARY, DATA COLLECTION, DATABASE, TERM EXTRACTION

Opsomming: Korpus-gebaseerde navorsing op terminologie van die Turkse leksikografie (CBRT-TURKLEX). In hierdie artikel word 'n lopende leksikografiese projek bekend gestel. Die Sentrum vir Leksikografie, afgekort tot SÖZMER, is onder die vaandel van die Eskisehir Osmangazi Universiteit tot stand gebring om leksikografiese projekte te ondersteun. SÖZMER het besluit om 'n korpus-gebaseerde Turkse leksikografieprojek te inisieer. Hierdie projek

[*] An earlier version of this article was presented at the 3rd International Conference of Lexicography, which was hosted by SÖZMER (Eskişehir Osmangazi University Center for Lexicography) in Eskişehir Osmangazi University, Turkey, 1–3 November 2016.
This work was supported by Scientific Research Projects Coordination Unit of Eskişehir Osmangazi University. Project number 2016-019056.

http://lexikos.journals.ac.za; https://doi.org/10.5788/28-1-1472

sal die eerste fase vorm van die strewe wat die skep van 'n gespesialiseerde woordeboek vir Turkse leksikografie ten doel het. Die primêre oogmerk van die projek is om 'n elektroniese korpus vir navorsers van die Turkse leksikografie voor te berei. Die sekondêre oogmerk van die projek is om 'n woordelys van Turkse leksikografiese terme te verkry. In hierdie artikel word 'n beskrywing gegee van die proses van dataversameling en die metodologie wat gebruik word vir die bou van 'n gespesialiseerde korpus. 'n Oorsig word gegee van die projekagtergrond, behoeftes, probleme, en die fases van korpusbou.

Sleutelwoorde: TURKSE LEKSIKOGRAFIE, TERMINOLOGIE, KORPUSLINGUISTIEK, WOORDEBOEK, DATAVERSAMELING, DATABASIS, TERMONTTREKKING

1. Introduction

The first dictionary work in Turkish began with Mahmut Kashgar. He started writing his *Divânu Lügati't-Türk* (Dictionary of Turkish Languages) in January 1072 and completed it in February 1074. Turkish lexicography has a long tradition spanning over centuries; however, it is found to be deficient in many aspects, including the realm of theoretical studies which are still not adequate. To date, there is no handbook of lexicography for Turkish lexicographers. Especially considering that for English, there are many handbook studies including Zgusta (1971), Jackson (2002), Van Sterkenburg (2003), Atkins and Rundell (2008), and Svensén (2009). The main reason for the delay in Turkish lexicographical research is the fact that academic institutions that would support field research and researchers have still not reached the desired numbers or the scientific levels. The Turkish Language Institute (*Türk Dil Kurumu*), which was established in 1932, is regarded as a milestone for linguistic research in Turkey. Furthermore, studies in the field of Turkish lexicography began to acquire a scientific character with the establishment of the Turkish Language Institute. Various studies related to the field of Turkish lexicography have been carried out by Turkish researchers (Levend 1957; Parlatır 1995; Aksan 1998 et al.). These studies have made considerable contributions to the development of the Turkish lexicographic literature. Various problems have been discussed in this process, but there are crucial unsolved problems in the field of Turkish lexicography. One of these problems is that a standardized terminology accepted by field experts has not yet been established. The first study about problems in Turkish lexicography was carried out by Tietze in 1976.

Other researchers such as Aksan (1990), Boz (2006), Boz (2011), Bozkurt (2017) have published various studies on Turkish lexicography, however, standardization of the specialized terminology of Turkish lexicography — both practical and theoretical — have not been provided by these studies.

Language for specific purposes (LSP) dictionaries such as those by Hartmann and James (1998), Burkhanov (1998) and the glossaries appended at the end of research studies such as those by Robinson (1983), Van Sterkenburg (2003) and Jackson (2013) have been very useful in standardization of lexicographical terminology.

To date, no significant research has been published covering all terms related to Turkish lexicography. The absence of a comprehensive list of terminology or a dictionary of Turkish lexicography has given rise to standardization problems among researchers.

Despite the increase in the number of research and educational centers such as universities and research institutes, especially in the period of the Turkish Republic, terms in the field of Turkish lexicography could not be gathered together, and the usage of lexicographical terms was not presented scientifically and systematically.

Instead of studies utilizing intuitive approach; studies that will allow the use of corpus linguistics, statistics, and computer-aided linguistics operation modes will generate more objective and more scientific results. Hence, in recent years, it has given rise to the so-called "corpus revolution" (Rundell and Stock 1992; Bergenholtz and Tarp 1995; Krishnamurthy 2002, 2008; Hanks 2012). A systematic, principled, scientific terminology study needs to be carried out by researchers for the development of the quality of the texts in the field of Turkish lexicography.

Term preference in cases of multiple terms for a single concept in Turkish lexicography is based on subjective approaches, or small discussions in academic communities of several people. Hence, extensive studies in the field of lexicography will increase the quality of terminology usage. Furthermore, there is no Turkish lexicography platform where researchers can agree on the usage of lexicographical terms by analyzing the tendencies in the corpus. Bowker and Pearson (2002: 12) state that "A special purpose corpus is one that focuses on a particular aspect of a language. It could be restricted to the LSP of a particular subject field, to a specific text type, to a particular language variety or to the language used by members of a certain demographic group (e.g. teenagers)." Therefore, an LSP corpus for Turkish lexicography is important with regard to providing term unity in the field of Turkish lexicography.

2. Aim of CBRT-TURKLEX

The main aim of the CBRT-TURKLEX is to build a lexicographical corpus for researchers that consists of master dissertations, doctoral theses, published presentations, news, books, articles, and reviews about the field of Turkish lexicography.

The secondary aim of the project is to obtain a word list of Turkish lexicographic terms, and to determine polysemy, synonymy, and term preferences among authors.

3. Method of CBRT-TURKLEX

The CBRT-TURKLEX project consists of five main phases.

3.1 Determination of corpus content and scope

There is no academic journal which relates only to Turkish lexicography in Turkey. However, there are many academic journals addressing grammar and linguistics research studies. Topics related to Turkish lexicography are generally published in the linguistics and grammar journals.

The articles, published presentations, books, master dissertations, doctoral theses, news and reviews were considered for CBRT-TURKLEX by the project researchers. Texts produced between 1932 (the year of the establishment of the Turkish Language Institution) and 2016 (the year of the project initiation) were collected for the corpus.

The texts containing the keywords "sözlük" (dictionary), "lügat" (dictionary, an old usage), "sözlükbilim" (lexicography), "sözlük bilim" (lexicography), "sözlükbilimi" (lexicography), "sözlük bilimi" (lexicography), "sözlükçülük" (synonym with lexicography), "leksikografi" (lexicography) were included in the corpus. A total of 1003 texts were identified as a result of this search. The types and the number of the texts included in the corpus are presented in Table 1.

Text Type	*Number of Texts*
Master dissertations	39
Doctoral theses	12
Published presentations	310
News	21
Books	3
Articles	468
Reviews	150
Total	**1003**

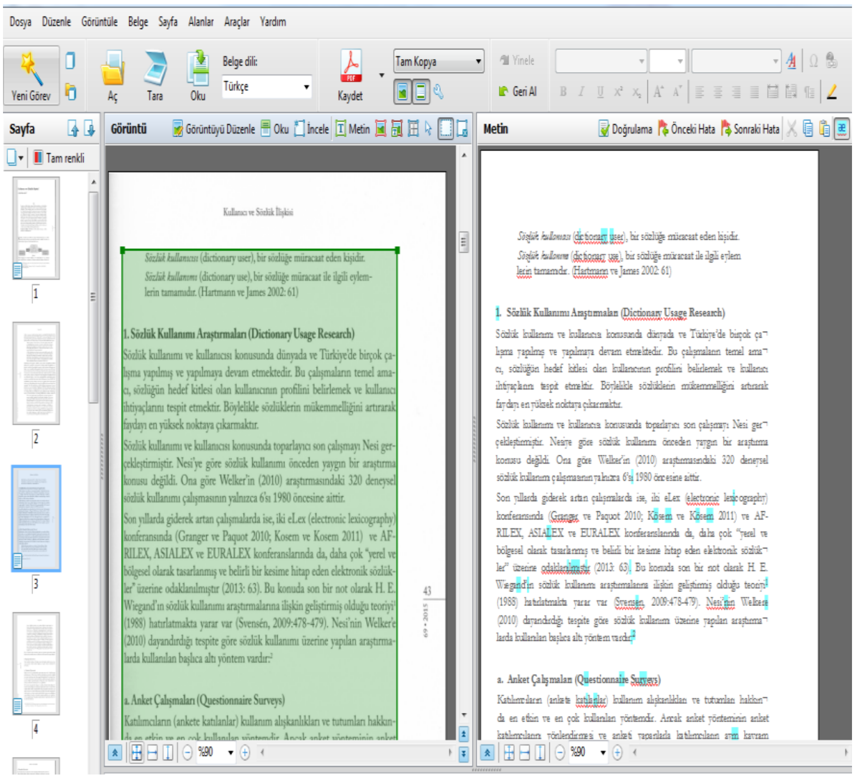
Table 1: Text types included in the corpus database

3.2 Digitization of printed texts

Some of the specified texts were in print format and others were in portable document format (PDF). Printed texts were transferred to the digital medium by means of optical character recognition (OCR) scanning. Texts in PDF were converted to OCR format by Abbyy Finereader 11© software.

In the process of conversion to OCR format, information such as bibliography, name of the journal, and page number in each text were deleted. An article page which was imported into Abbyy Finereader 11© software is presented in Figure 1.

http://lexikos.journals.ac.za; https://doi.org/10.5788/28-1-1472



Kullanıcı ve Sözlük İlişkisi

Sözlük kullanıcısı (dictionary user), bir sözlüğe müracaat eden kişidir.

Sözlük kullanımı (dictionary use), bir sözlüğe müracaat ile ilgili eylemlerin tamamıdır. (Hartmann ve James 2002: 61)

1. Sözlük Kullanımı Araştırmaları (Dictionary Usage Research)

Sözlük kullanımı ve kullanıcısı konusunda dünyada ve Türkiye'de birçok çalışma yapılmış ve yapılmaya devam etmektedir. Bu çalışmaların temel amacı, sözlüğün hedef kitlesi olan kullanıcının profilini belirlemek ve kullanıcı ihtiyaçlarını tespit etmektir. Böylelikle sözlüklerin mükemmelliğini artırarak faydayı en yüksek noktaya çıkarmaktır.

Sözlük kullanımı ve kullanıcısı konusunda toparlayıcı son çalışmayı Nesi gerçekleştirmiştir. Nesi'ye göre sözlük kullanımı önceden yaygın bir araştırma konusu değildi. Ona göre Welker'in (2010) araştırmasındaki 320 deneysel sözlük kullanımı çalışmasının yalnızca 6'sı 1980 öncesine aittir.

Son yıllarda giderek artan çalışmalarda ise, iki eLex (electronic lexicography) konferansında (Granger ve Paquot 2010; Kosem ve Kosem 2011) ve AFRILEX, ASIALEX ve EURALEX konferanslarında da, daha çok "yerel ve bölgesel olarak tasarlanmış ve belirli bir kesime hitap eden elektronik sözlükler" üzerine odaklanılmıştır (2013: 63). Bu konuda son bir not olarak H. E. Wiegand'ın sözlük kullanımı araştırmalarına ilişkin geliştirmiş olduğu teoriyi[1] (1988) hatırlatmakta yarar var (Svensén, 2009:478-479). Nesi'nin Welker'e (2010) dayandırdığı tespite göre sözlük kullanımı üzerine yapılan araştırmalarda kullanılan başlıca altı yöntem vardır:[2]

43

69 • 2015

a. Anket Çalışmaları (Questionnaire Surveys)

Katılımcıların (ankete katılanlar) kullanım alışkanlıkları ve tutumları hakkında en etkin ve en çok kullanılan yöntemdir. Ancak anket yönteminin anket

Figure 1: Converting PDF files to OCR format

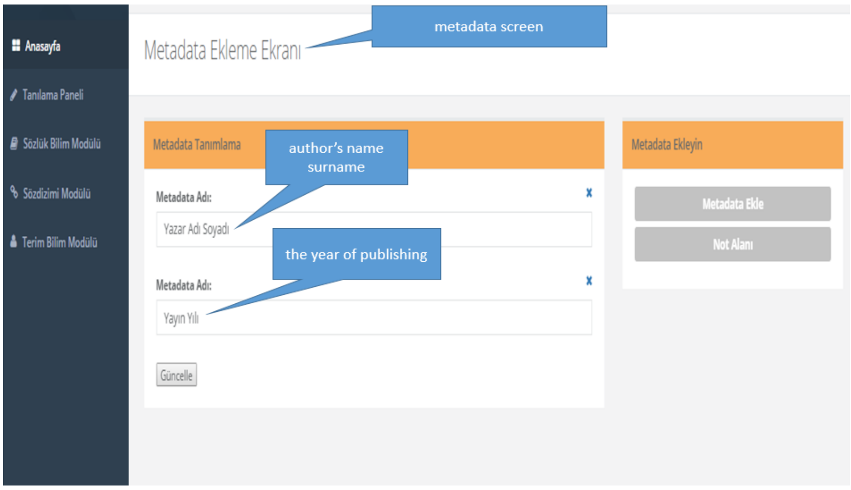
The text page contains details such as the name of the article "Kullanıcı Sözlük İlişkisi", the number of the page, the year of the publication and the volume of the journal in which the article was published. These details were not included in the text corpus due to the software considering this information as junk.

3.3 Uploading of texts into the corpus

Once the conversion phase was complete, the machine-readable texts were uploaded into the corpus in the following stages.

3.3.1 Determination of metadata for the corpus

Information about the texts in the corpus means metadata, in other words metadata is data about data. This information may include the title, author, publisher and date of a written text, or details of the speakers in a spoken text (Baker et al. 2006: 115). Authors' names/last names and the publication year of the text were identified as the metadata in the corpus for Turkish lexicography. The metadata screen is shown in Figure 2.



Metadata Ekleme Ekranı

metadata screen

Metadata Tanımlama

Metadata Adı: Yazar Adı Soyadı

author's name surname

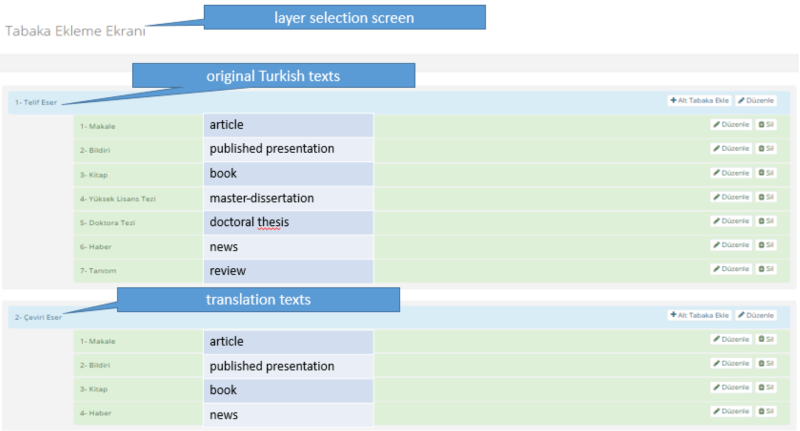
Metadata Adı: Yayın Yılı

the year of publishing

Figure 2: Metadata of the texts

3.3.2 Determination of layers for the corpus

In corpora, it is necessary to decide at the beginning on correct clustering of the texts for reporting corpus findings (Kupietz 2016: 68-70). The texts related to the field of lexicography are classified into seven different types in the database as shown in Figure 3.



Tabaka Ekleme Ekranı

layer selection screen

original Turkish texts

1- Telif Eser

1- Makale	article
2- Bildiri	published presentation
3- Kitap	book
4- Yüksek Lisans Tezi	master-dissertation
5- Doktora Tezi	doctoral thesis
6- Haber	news
7- Tanıtım	review

translation texts

2- Çeviri Eser

1- Makale	article
2- Bildiri	published presentation
3- Kitap	book
4- Haber	news

Figure 3: Layer selection screen

http://lexikos.journals.ac.za; https://doi.org/10.5788/28-1-1472

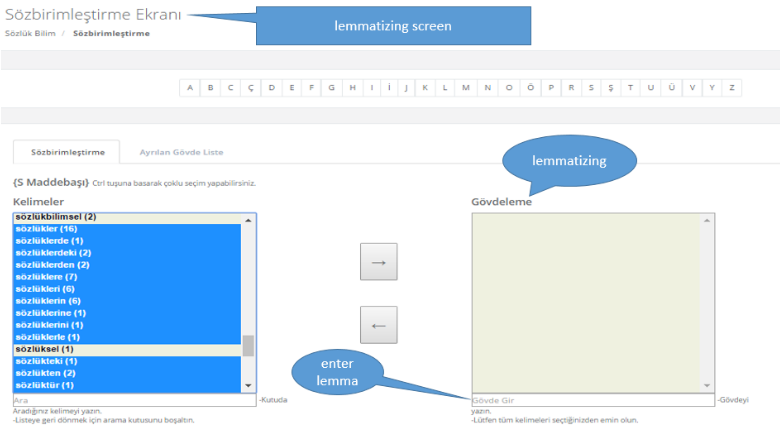
Layers of the corpus are articles, published presentations, books, masters-dissertations, doctoral theses, news and reviews. It is possible to report the frequency and dispersion of the terms according to the text types through these layers.

3.4 Lemmatizing of words

Francis and Kučera (1982: 1) define a lemma as a 'set of lexical forms having the same stem and belonging to the same major word class, differing only in inflection and/or spelling. Inflected forms of WALK as a lemma are given by Francis and Kučera. These are *walk*, *walked*, *walking* and *walks*.

A lemmatization process was necessary for CBRT-TURKLEX due to the fact that Turkish is an agglutinative language. There are two kinds of suffixes in this language. Some of the suffixes are inflectional suffixes and the others are derivational suffixes. Derivated words are accepted as separate lemmas, but inflected ones are not considered as separate lemmas.

Various inflected forms for the lemma SÖZLÜK (dictionary) lemma are given in Figure 4.



Sözbirimleştirme Ekranı

Sözlük Bilim / Sözbirimleştirme

{S Maddebaşı} Ctrl tuşuna basarak çoklu seçim yapabilirsiniz.

Kelimeler

sözlükbilimsel (2)
sözlükler (16)
sözlüklerde (1)
sözlüklerdeki (2)
sözlüklerden (2)
sözlüklere (7)
sözlükleri (6)
sözlüklerin (6)
sözlüklerine (1)
sözlüklerini (1)
sözlüklerle (1)
sözlüksel (1)
sözlükteki (1)
sözlükten (2)
sözlüktür (1)

Ara -Kutuda Aradığınız kelimeyi yazın.
-Listeye geri dönmek için arama kutusunu boşaltın.

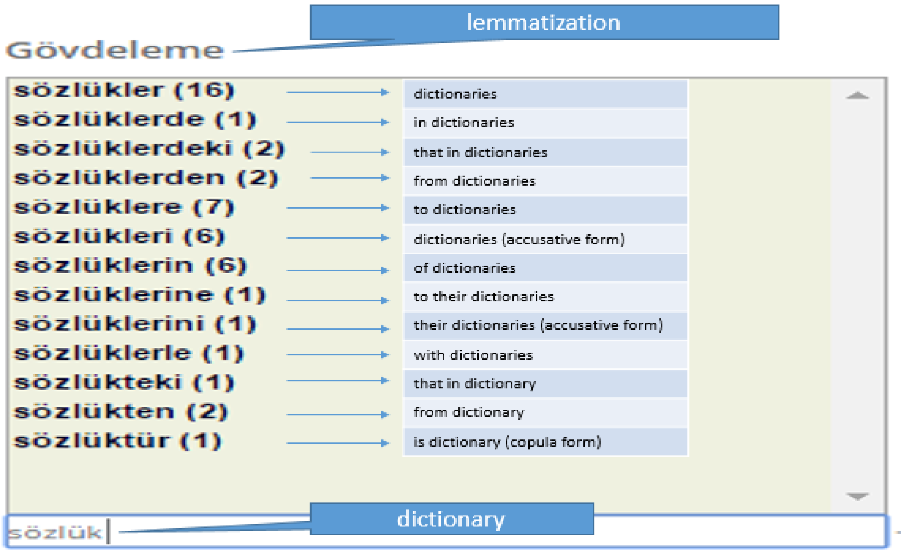
Gövdeleme

Gövde Gir -Gövdeyi yazın.
-Lütfen tüm kelimeleri seçtiğinizden emin olun.

Figure 4: Lemma selection

As shown in Figure 5, "sözlükler" (dictionaries), "sözlüklerde" (in dictionaries), "sözlüklerdeki" (that in dictionaries), "sözlüklerden" (from dictionaries), "sözlüklere" (to dictionaries), "sözlükleri" (dictionaries, accusative form), "sözlüklerin" (of dictionaries), "sözlüklerine" (to their dictionaries), "sözlüklerini" (their dictionaries, accusative form), "sözlüklerle" (with dictionaries), "sözlükteki" (that in dictionary), "sözlükten" (from dictionary), "sözlüktür" (is dictionary). As shown in Figure 5, "SÖZLÜK" (dictionary) is the lemma of these inflected forms.

http://lexikos.journals.ac.za; https://doi.org/10.5788/28-1-1472



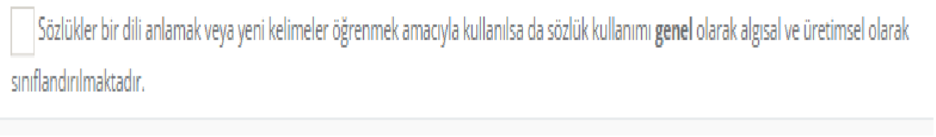
sözlükler (16)	dictionaries
sözlüklerde (1)	in dictionaries
sözlüklerdeki (2)	that in dictionaries
sözlüklerden (2)	from dictionaries
sözlüklere (7)	to dictionaries
sözlükleri (6)	dictionaries (accusative form)
sözlüklerin (6)	of dictionaries
sözlüklerine (1)	to their dictionaries
sözlüklerini (1)	their dictionaries (accusative form)
sözlüklerle (1)	with dictionaries
sözlükteki (1)	that in dictionary
sözlükten (2)	from dictionary
sözlüktür (1)	is dictionary (copula form)

Figure 5: Screen for lemmatization

3.5 Tagging terms (identification and extraction of terms)

Some of the words in the corpus such as "bu" (this) and "güzel" (beautiful) cannot be lexicographic term candidates. At this stage, term candidates related to the field of lexicography will be selected from the sample sentences by means of "term extraction tab". For instance, the word "genel" (general) can be a lexicographic term or not, according to context.

A sample sentence which includes the word "genel" is illustrated in Figure 6. The sentence is not marked since the "genel" word is regarded as a non-lexicographic term.

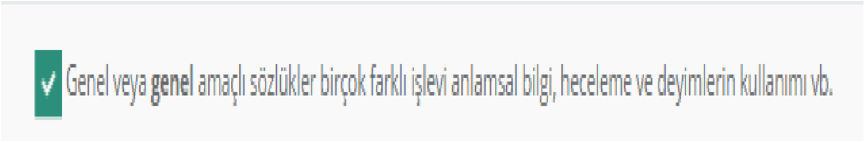


□ Sözlükler bir dili anlamak veya yeni kelimeler öğrenmek amacıyla kullanılsa da sözlük kullanımı **genel** olarak algısal ve üretimsel olarak sınıflandırılmaktadır.

Figure 6: Term extraction tab (□ is not a term)

A sample sentence which includes the word "genel" is illustrated in Figure 7. The sentence is marked since the word "genel" is regarded as a lexicographic term.

http://lexikos.journals.ac.za; https://doi.org/10.5788/28-1-1472



☑ Genel veya **genel** amaçlı sözlükler birçok farklı işlevi anlamsal bilgi, heceleme ve deyimlerin kullanımı vb.

Figure 7: Term extraction tab (☑ is a term)

This decision procedure was followed for all of the term candidates in the corpus.

Not only single-word terms but also multi-word terms appear in the field of Turkish lexicography. Collocations, in which two or more words constitute or enter into a syntactic unit, also had to be marked in the corpus (Bergenholtz and Tarp 1995: 118).

Collocations were determined with collocation screen as can be seen in Figure 8. Word collocations could be listed on the screen. Collocational relations could be provided for the left and right of the center word.

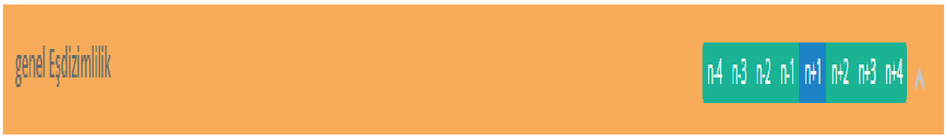
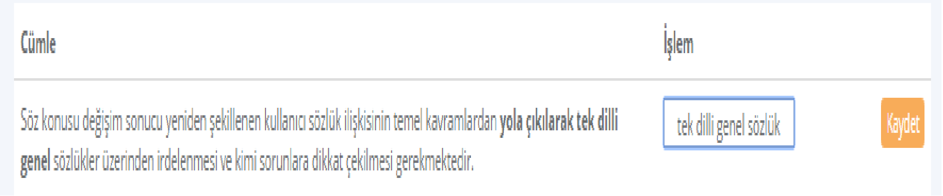


Figure 8: Collocation screen

As can be seen in Figure 9, the query for the word "genel" was input as n-4. The four words to the left of "genel", "yola çıkılarak tek dilli" turned as results and were shown in bold in the query screen. As a result of this query the word "sözlük" to the right of "genel" was deduced to be related by the researcher. The lexicographical term in this context was determined as "tek dilli genel sözlük" meaning "monolingual general dictionary".

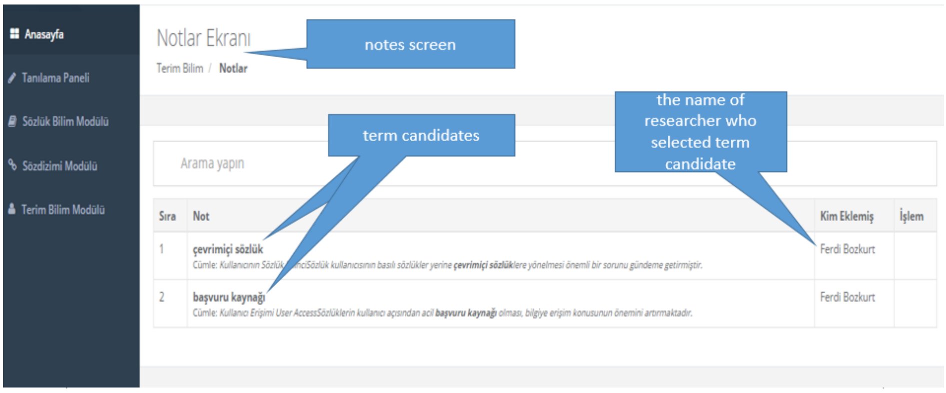


Cümle	İşlem
Söz konusu değişim sonucu yeniden şekillenen kullanıcı sözlük ilişkisinin temel kavramlardan **yola çıkılarak tek dilli genel** sözlükler üzerinden irdelenmesi ve kimi sorunlara dikkat çekilmesi gerekmektedir.	tek dilli genel sözlük

Figure 9: Collocational words query screen

Tagging terms in the corpus is conducted by multiple project researchers to eliminate individual mistakes and decisions based on intuition. Figure 10 shows the screen for notes. The project researcher's decision, whether a word is a term or not, can be followed in the notes screen.

http://lexikos.journals.ac.za; https://doi.org/10.5788/28-1-1472



Sıra	Not	Kim Eklemiş	İşlem
1	**çevrimiçi sözlük** Cümle: *Kullanıcının Sözlük [illegible]nciSözlük kullanıcısının basılı sözlükler yerine **çevrimiçi sözlük**lere yönelmesi önemli bir sorunu gündeme getirmiştir.*	Ferdi Bozkurt	
2	**başvuru kaynağı** Cümle: *Kullanıcı Erişimi User AccessSözlüklerin kullanıcı açısından acil **başvuru kaynağı** olması, bilgiye erişim konusunun önemini artırmaktadır.*	Ferdi Bozkurt	

Figure 10: Notes screen

4. Conclusion

In this article a research project, namely Corpus-Based Research on Terminology of Turkish Lexicography, has been presented. The project is conducted by the Center for Lexicography at Eskişehir Osmangazi University.

The processes for the determination of the terms within the scope of the study are presented in this article. Totally 1003 texts were determined on the field of Turkish lexicography. 329 texts were in printed form. These were scanned to PDF. 674 texts were already in PDF. All of the PDF texts were converted to OCR format.

The corpus was built on October 10th, 2017. The website of the corpus is available at www.tsd.ogu.edu.tr for lexicographers. The corpus contains 1003 texts. It comprises 42.831 sentences, 703.986 orthographic words, and 86.368 types.

The frequency, dispersion, and the author's word preferences of term candidates were examined in the corpus. 1.616 lexicographic terms were determined in the corpus by the project researchers.

Future Work

A Dictionary of Turkish Lexicography will be compiled through the corpus.

Acknowledgement

1. The software of the corpus was utilized by Assoc. Prof. Dr. Bülent ÖZKAN's "TÜBİTAK-SOBAG - 114E791 Do It Yourself Corpora" project.
2. The project was completed on October 10th, 2017. The website of the corpus is available at www.tsd.ogu.edu.tr.

http://lexikos.journals.ac.za; https://doi.org/10.5788/28-1-1472

References

Aksan, D. 1990. *Her Yönüyle Dil*. Ankara: Türk Dil Kurumu Yayınları.

Aksan, D. 1998. Türklerde Sözlükçülük, Bugün Türkiye`de Sözlük. *Kebikeç Dergisi* 6: 115-118.

Atkins, B.T.S. and M. Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford/New York: Oxford University Press.

Baker, P., A. Hardie and T. McEnery. 2006. *A Glossary of Corpus Linguistics*. Edinburg: Edinburgh University Press.

Bergenholtz, H. and S. Tarp. 1995. *Manual of Specialised Lexicography: The Preparation of Specialised Dictionaries.* Vol. 12. Amsterdam/Philadelphia: John Benjamins.

Bowker, L. and J. Pearson. 2002. *Working with Specialized Language: A Practical Guide to Using Corpora.* London: Routledge.

Boz, E. 2006. Sözlük ve Sözlükçülük Sorunu. *Türkçenin Çağdaş Sorunları*: 9-46. İstanbul: Divan Yayınları.

Boz, E. 2011. Leksikografi Teriminin Tanımı ve Türkçe Karşılığı Üzerine. *Dil ve Edebiyat Araştırmaları Dergisi* 4: 9-14.

Bozkurt, F. 2017. *Sözlükselleşme: Genel Sözlükler için Sözlük Birim Seçimi.* İstanbul: Kesit Yayınları.

Burkhanov, I. 1998. *Lexicography: A Dictionary of Basic Terminology*. Wydawn: Wyższej Szkoły Pedagogicznej w Rzeszowie.

Francis, W.N. and H. Kučera. 1982. *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston: Houghton Mifflin.

Hanks, P. 2012. The Corpus Revolution in Lexicography. *International Journal of Lexicography* 25(4): 398-436.

Hartmann, R.R.K. and G. James. 1998. *Dictionary of Lexicography*. London/New York: Routledge.

Jackson, H. (Ed.). 2013. *The Bloomsbury Companion to Lexicography*. London: Bloomsbury.

Jackson, H. 2002. *Lexicography. An Introduction*. London/New York: Routledge.

Krishnamurthy, R. 2002. The Corpus Revolution in EFL Dictionaries. *Kernerman Dictionary News* 10: 23-27.

Krishnamurthy, R. 2008. Corpus-driven Lexicography. *International Journal of Lexicography* 21(3): 231-242.

Kupietz, M. 2016. Constructing a Corpus. Durkin, Philip (Ed.). 2016. *The Oxford Handbook of Lexicography*: 62-75. Oxford: Oxford University Press.

Levend, A.S. 1957. Türkçe Sözlük Üzerine. *Türk Dili* VI(67): 365-367.

Parlatır, İ. 1995. Türkçe Sözlük Çalışmaları ve Sorunlarımız. *Türk Dili. Dil ve Edebiyat Dergisi* I(517): 3-19.

Robinson, J. 1983. A Glossary of Contemporary English Lexicographic Terminology. *Dictionaries* 5: 76-114.

Rundell, M. and P. Stock. 1992. The Corpus Revolution 3. A Consideration of the Prospects and Potential of Corpus-and-concordance Lexicography (third article of three). *English Today, The International Review of the English Language* 8(4): 45-51.

Svensén, B. 2009. *A Handbook of Lexicography: The Theory and Practice of Dictionary-making.* Cambridge/New York: Cambridge University Press.

Tietze, A. 1976. Problems of Turkish Lexicography. Householder, Fred W. and Sol Saporta (Eds.). 1976. *Problems in Lexicography*: 263-272. Third edition. Bloomington: Indiana University.

http://lexikos.journals.ac.za; https://doi.org/10.5788/28-1-1472

Van Sterkenburg, P. 2003. *A Practical Guide to Lexicography*. Amsterdam/Philadelphia: John Benjamins.

Zgusta, L. 1971. *Manual of Lexicography*. Berlin/New York: Walter de Gruyter.