

SHUKRU HALUK AKALIN (*Türkiye*)*

RESEARCH ON HEYDAR ALİYEV'S VOCABULARY

Abstract

With the broad vocabulary which he uses in his speeches, writings and books, Azerbaijan's great leader, the wise personality of the Turkic world, the great statesman Heydar Aliyev was an exemplary master of rhetoric, deftly interweaving expressions of precisely distinguished meaning and using effective syntax. As he carried out the important work at the government level of developing and spreading the language of Azerbaijan and gaining it self-esteem, of decision-making and of issuing laws, Heydar Aliyev showed personally in his speeches and writings as much as in his language policies by using language which was an example to everyone how rich a vocabulary the language should have.

Examining the vocabulary of Heydar Aliyev's speeches and writings from this aspect, it is seen that at the base of his capability of expression and effective use of language is a rich vocabulary which covers a very wide range and includes fine distinctions of meaning. In the paper, a corpus is presented, formed with the aim of determining the vocabulary of General National Leader Heydar Aliyev's speeches, declarations and writings entitled *Elm və Təhsil*, and information is given on how this vocabulary is to be scientifically researched.

Keywords: *Heydar Aliyev, rhetoric, vocabulary, semantics, lexicology.*

Introduction

With the broad vocabulary which he used in his speeches, writings and books, Azerbaijan's great leader, the wise personality of the Turkic world, the great statesman Heydar Aliyev was an exemplary master of rhetoric, deftly interweaving expressions of precisely distinguished meaning and using effective syntax.

As he carried out the important work at the government level of developing and spreading the language of Azerbaijan and gaining it self-esteem, of decision-making and of issuing laws, Heydar Aliyev showed personally in his speeches and writings as much as in his language policies by using language which is an example to everyone of how a person from Azerbaijan should speak, and of how rich a vocabulary the language should have. Examining the vocabulary of Heydar Aliyev's speeches and writings from this aspect, it is seen that at the base of his capa-

* Hacettepe University emeritus professor, Azerbaijan National Academy of Sciences Linguistics Institute, and Sumqayıt State University honorary doctor. E-mail: shakalin@shakalin.net, sukruhaluk.akalin@hacettepe.edu.tr, <https://www.shakalin.net>

bility of expression and effective use of language is a rich vocabulary which covers a very wide range and includes fine distinctions of meaning.

What is vocabulary?

Vocabulary is the total of a language's words, expressions, set phrases, proverbs, technical terms and various fixed expressions (TDK, 2011: 2158; Aksan, 2004: 11; Hartmann and James, 1998: 154; Burkhanov, 2010: 261; Karaağaç, 2013: 754). At the same time, vocabulary is also the total of the words which a person knows and uses (Karaağaç, 2013: 745; Hartmann and James, 1998: 154; Burkhanov, 2010: 261). The vocabulary used by the speakers of a language has two dimensions, active and passive (Hartmann and James, 1998: 154; Burkhanov, 2010: 261). The vocabulary known and used by a speaker show the breadth of his or her world. The breadth and richness of a person's vocabulary is reflected in their understanding of the world, their analysis of events, and their conceptualisation of developments, and their expression of their feelings and thoughts will be that much clearer, easier to understand and effective.

In the past, vocabulary research and related studies were conducted by labelling and sorting vocabulary items one by one, assigning meanings, defining and exemplifying, in short by human effort and by hand. Today however, some of this work is performed by computer processing of written or oral texts with the help of various programs. This work, called Natural Language Processing (NLP), allows the process to be carried out by computer in the stages of determining, reviewing and ordering vocabulary. With an increase in the success rate of lemmatisation work by rule-based or deep learning methods, the use of computers in this field has spread progressively. Recently, the use of artificial intelligence in this work has begun, with computer-supported determination, interpretation and finally definition of vocabulary items.

By processing speakers' vocabulary by computer, it will be possible to determine active and passive vocabulary, to show the types of words used, their frequencies and their collocations, and in this way to provide stylistics data.

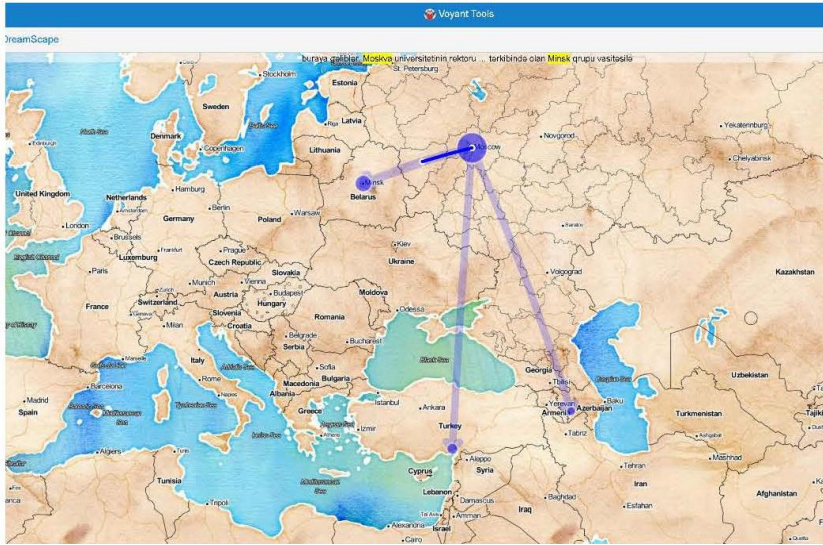
Heydar Aliyev's vocabulary

In this year in which we celebrate the centenary of his birth, we aim with this study which we have conducted to determine the vocabulary of the speeches, statements, writings and other works of Azerbaijan's great leader Heydar Aliyev to show his mastery of language, his skill at the use of language, his frequency of use of words, and his preferred collocations, and in this way to provide data for research on Heydar Aliyev's stylistics. As the first step in this study, 80 writings, speech texts and statements in Heydar Aliyev's work entitled *Elm və Təhsil in Heydər Əliyev İrsi Beynəlxalq Elektron Kitabxana*, on the website (<https://lib.aliyev-heritage.org/az/742736.html>), which contain his talks, statements and writing on science and education, have been converted into text files. Topic headings and insertions by others in the text files were extracted, and only talks,

statements and writings by Heydar Aliyev were considered. Depending on both oral and written material enabled the use of realistic data on vocabulary. Written or oral texts published at various times in the 20-year period from September 1993 to September 2003 can provide data for chronological vocabulary research. This text file was transferred to the program Voyant Tools and converted into database form for vocabulary research. As a result of the analysis conducted, it was seen that the text consisted of 104 954 words, with 14 782 different vocabulary items. This number shows the value obtained before lemmatisation.

Following a review of the talks, statements and writings entitled *Elm və Təhsil* by Heydar Aliyev in the Research on Heydar Aliyev's Vocabulary which I have created in the section where I am conducting corpus studies on my personal web page using the program Voyant Tools at <https://www.shakalin.net>, a review of talks, writings and statements on other topics will be produced. This review can be accessed at <https://aliyev-sozvar.shakalin.net>.

The figure below contains the data frames according to the preferred vocabulary analyses of this program. Frame 1 shows the word cloud formed according to the frequency of use of words in the texts of Heydar Aliyev included in the review. The first hundred most frequently used words are shown in sizes corresponding to their proportion of use, and each word is shown in a different colour. When the button below this frame is pulled to the right, the number of words in the sequence can be increased. In this way, word clouds of, for example, Heydar Aliyev's 200 or 300 most used words can be formed. In Frame 2, the word *Azərbaycan* was searched for in the vocabulary data table, and the words most found along with this word were listed. Frame 3 shows the rate of use of the word searched for in places where it occurs in the text. Frame 4 contains information on the coverage and date of creation of the prepared text. When new files are added to the database, the numerical data here will increase. As mentioned above, the material converted into text consisted of a total of 104 954 words, with 14 782 vocabulary items before lemmatisation, and this is shown in Frame 4. Frame 5 of the program shows the collocations of the word searched for. For example, five n-gram collocations can be seen on the left and right of the searched-for word *Azərbaycan*. The n-gram can be increased by pulling the button below to the right. Clicking on a searched-for word on each line opens a blue field in the frame, and it can be seen how that word appeared in the paragraph. In this way, the collocations of each word in Heydar Aliyev's vocabulary can be seen.



Related words can be seen on the Correlations tab:

Document Terms		Correlations			
Term 1	←	→	Term 2	Correlation ...	Significanc ...
saxlamaq			öz	0.8941795	0.00048196 ...
fikir			öz	0.8531632	0.00169694 ...
isə			öz	0.83348584	0.00273686 ...
müstəqilliyi...			öz	0.80742675	0.004734845
nasilər			öz	0.8049837	0.00496404 ...
bu			öz	0.79923123	0.005534863
yalnız			öz	0.7992047	0.00553780 ...
hadisədir			öz	0.790868	0.006448074
müharibə			öz	0.7849621	0.00715053 ...
müstəqil			öz	0.78472635	0.00717949 ...
isteyir			öz	0.77260906	0.00879775 ...
15			öz	0.7652864	0.009891613
sahədə			öz	0.76318485	0.01022244 ...
müqəddəs			öz	0.76124924	0.01053396
yanışı			öz	0.7608201	0.010603922
vezfəsədir			öz	0.7597508	0.01077966 ...
müharibədən			öz	0.7565779	0.01131319
nədənsə			öz	0.7565779	0.01131319
elimizi			öz	0.7565779	0.011313189
qrupunun			öz	0.7565779	0.011313189
asli			öz	0.7546868	0.011639654
odur			öz	0.75412536	0.011738094
aydın			öz	0.75352484	0.011843817
olması			öz	0.75292933	0.011949316
sosial			öz	0.75002223	0.012473806

6,525 minimum coverage (%100)

As the program works with the UTF-8 character set, each letter of the Azerbaijani alphabet can be seen appropriately in its correct form. Also, when performing searches, it is necessary to use the letters of the Azerbaijani alphabet and to write words in the search boxes according to the rules of writing the Azerbaijani.

After removing the stop list words¹ of Azerbaijani, the 20 words used with the highest frequency in the talks, statements and writings of Heydar Aliyev on the topic of science and education were listed as follows:

1. <i>Azərbaycan</i>	6. <i>yüksək</i>	11. <i>müəllim</i>	16. <i>sahə</i>
2. <i>təhsil</i>	7. <i>elm</i>	12. <i>həyat</i>	17. <i>məktəb</i>
3. <i>böyük</i>	8. <i>müstəqil</i>	13. <i>xalq</i>	18. <i>universitet</i>
4. <i>ali</i>	9. <i>inkışaf</i>	14. <i>tarix</i>	19. <i>ziya</i>
5. <i>dövlət</i>	10. <i>milli</i>	15. <i>istifadə</i>	20. <i>respublika</i>

These words clearly show the topics to which the great leader Heydar Aliyev attached importance, the matters to which he paid attention, and the areas to which he was sensitive. Collocations and correlations of these words on the database which we have created will be evaluated, and results which are of interest will be presented.

Conclusion

Working on the vocabulary of all of the legacy works of the great leader of Azerbaijan and wise personality of the Turkic world Heydar Aliyev and displaying in a concrete way what were the words of his world, what words he used most frequently over the years, which words he preferred to use together, and the place names and other names occurring in his speeches will be possible thanks to computer-supported language research and natural language processing work. Our work and research on this topic continues. The results which we have obtained and the material which has emerged will be presented in articles to be published in the coming months for the appreciation of the world of science and culture.

RESOURCES

1. Adalı, E. (2020). *Türkçe Doğal Dil İşleme*. Ankara: Akçağ Yayınevi.
2. Aksan, D. (2004). *Türkçenin Söz Varlığı*. Ankara: Engin Yayınevi.
3. Burkhanov, Igor. (2010) *Lexicography A Dictionary of Basic Terminology*. Wydawnictwo Uniwersytetu Rzeszowskiego.
4. Hartmann, R. R. K. (1998). *Dictionary of Lexicography*. London- New York: Routledge.
5. Heydər Əliyev İrsi Beynəlxalq Elektron Kitabxana. <https://lib.aliyev-heritage.org/az/742736.html>
6. Karaağaç, G. (2013). *Dil Bilimi Terimleri Sözlüğü*. Ankara: Türk Dil Kurumu.
7. TDK. (2011). *Türkçe Sözlük*. Ankara: Türk Dil Kurumu

¹ In NLP studies, vocabulary items which are accepted as functional and meaningless words, which have no direct effect on meaning such as conjunctions, prepositions or pronouns like *and, or, a, this, other, on, at, etc.*, are known in English as stop list words. Dr. Eşref Adalı, who has conducted important work in the field of NLP in Türkiye, has suggested *dilsiz sözcük* as a Turkish equivalent (Adalı, 2020: 313). Stop list words in Azerbaijani were determined taking into account the lists in Adalı's *The Stop List Words of Turkish* (Adalı, 2020: 739) and *The Stop List Words of English* (Adalı, 2020: 742).

Şükrü Halûk Akalın (Türkiyə)

Heydər Əliyevin söz varlığıyla bağlı araşdırmalar

Xülasə

Azərbaycanın ulu öndəri, türk dünyasının müdrik şəxsiyyəti, böyük dövlət xadimi Heydər Əliyev nitqlərində, yazılarında və kitablarında işlətdiyi geniş söz ehtiyatı, incə məna detalları ilə toxunmuş ifadələri, təsirli sintaksisi ilə nümunəvi nətiqlik ustasıdır. Azərbaycan dilinin inkişafı və yayılması, özünəməxsusluğunun qazanılması üçün dövlət səviyyəsində çox mühüm işlər aparın, qərarların və qanunların qəbul edilməsini təmin edən Heydər Əliyev bu dil siyasəti ilə bərabər, həm də çıxışlarında və əsərlərində hər kəsə örnək olacaq bir dildən istifadə edərək, zəngin söz ehtiyatının necə olması lazım olduğunu özündə göstərmişdir.

Heydər Əliyevin nitq və əsərlərinin lüğət tərkibinə bu aspektdən nəzər saldıqda onun ifadə qabiliyyətinə, təsirli dildən istifadəsinə görə son dərəcə əhatəli və incə semantik təfərrüatları özündə ehtiva edən zəngin lüğətə malik olduğu görünür. Məqələdə ümummillə Lider Heydər Əliyevin elm və təhsillə bağlı nitqlərinin, açıqlamalarının və yazılarının lüğət tərkibini müəyyən etmək üçün yaradılmış korpus təqdim edilmiş, bu söz ehtiyatının elmi aspektdən tədqiqi yolları haqqında məlumat verilmişdir.

Açar sözlər: *Heydər Əliyev, nitq mədəniyyəti, lüğət tərkibi, semantika, leksikologiya.*

Шукру Халук Акалын (Турция)

Исследования по лексике Гейдара Алиева

Резюме

Великий лидер, государственный деятель Азербайджана, мудрец тюркского мира Гейдар Алиев в своих устных и письменных регах, книгах, используя богатую лексику, выражения с различными смысловыми оттенками, действенный синтаксис, выступил мастером ораторского искусства. Наряду с проведением очень важных работ на государственном уровне с целью развития и распространения азербайджанского языка, приобретения им своего статуса, принятием соответствующих постановлений и законов Гейдар Алиев на примере образцового языка своих вступлений и работ продемонстрировал, каким должен быть богатый словарный запас.

Анализ выступлений и работ Гейдара Алиева в данном аспекте показал богатство его всеобъемлющего словаря, охватывающего тонкие семантические нюансы. В статье представлен корпус, созданный с целью выявления словарного состава устных и письменных речей общенационального лидера Гейдара Алиева, связанных с наукой и образованием, даны сведения о путях научного исследования данного словарного объема.

Ключевые слова: *Гейдар Алиев, ораторское искусство, словарный состав, семантика, лексикология.*