SOME DISCUSSION ABOUT COMPILING TEXT DICTIONARY OF EPOS "ARZYGUL"

Khadjimusaeva Nilufar Nematovna

Teacher of Djizak State Pedagogical Institute. City Djizak, Uzbekistan.

Abstract

This article is devoted to the field of lexicography which is compiled on computer programs developed by S.Karimov and others, Islam poet Nazar speaks about the principles of compiling alphabetical, frequency and inverted dictionaries of the poem "Arzygul". The article cites scientific facts in the article "Arzygul" that the Islamic poet used 2,384 key words in 6679 words and used 22,660 times. The principle of the formation of frequency dictionaries is similar to the alphabetical dictionaries. Even in this type of dictionary, the total number of usage words and statistics on the page is clearly indicated. Only the words are given in alphabetical order, not in alphabetical order. The three types of dictionaries, based on the text of the epigraph "Arzygul" epos, and a brief description of them, in a sense, can demonstrate the importance of lexicographical and statistical research in textual analysis, and give rise to certainty that these analyzes are more valid. In addition, the frequency and percentage of top-down declines in the phrase "Arzygul" are presented in the table. The table shows that the frequency of use of the vocabulary and the activity are inversely proportional.

In modern Uzbek linguistics, in particular in the Uzbek lexicography, the development and implementation of computer-based word processing programs for various linguistic research remains relevant.

The computer programs developed by Prof. S.Karimov and others, we created an alphabetical, frequency and inverse dictionary of the "Arzygul" epos. The compiled dictionaries allowed to identify linguistic and grammatical units characteristic of the creative poem of the poet Nazar Nazar, who sings "Arzygul" and compare it with the language and style of Fazil shair Yuldash ogly, representative of the second largest school of bakhshi – Bulungur school. The basis of real material has been the separation of units of the common language or dialectal lexicon of Islamic poetry, and the contribution of Bakhshi to the development of the Uzbek language.

When the electronic version of the poem was introduced to the computer, the computer showed that it had 23036 dictionary units. In the compiled dictionary this number is 22,660. The difference is 376. This is 0.016%. Hence, the relative error in the interval is d = 0.016. In other words, the difference is only one in every 100 words. The reason for this relative error is that the computer registers the number of texts as a single unit. S.Rizaev said in this regard: "The more the sample size (N) in linguistic research, the more reliable and effective the expected result" [2, p. 21]. In this sense, the relative error in the dictionary's expression in a dictionary of just 114 pages is insignificant.

Here is an example of an alphabetical dictionary based on the text of "Arzygul" epos:

abzalini 1 294(1)
avaylab 1 346(1)
avval 16 282(1) 291(1) 310(4) 319(2) 320(2) 321(1) 329(1) 348(1) 352(1) 361(1) 364(1)

avval-oxir 3 291(1) 297(1) 343(1)
avvalgi 3 276(1) 300(1) 325(1)
avvaldan 2 313(1) 355(1)
avvaliga 1 320(1)
avlodir 1 277(1)
agar 29 278(1) 280(1) 291(2) 295(1) 306(1) 309(2) 312(1) 314(1) 318(1) 321(2) 322(1) 329(1) 336(1) 338(2) 339(1) 347(1) 349(1) 352(1) 355(2) 361(1) 367(1) 375(1) 380(2) and other.

If the dictionary is taken care of, the dictionary unit is presented as it is written in the text: adolatda, adolatimni, adolatli, adolatni. The second thing to notice is the number after the word. This number represents the total number of words used in the text. For example, the word agar is used 29 times in the poem. Then there is another number, next to it, and another number in brackets: 29 278 (1) 280 (1) 291 (2) 295 (1). The number after the number 29 that indicates how many times it has been used, for example 278 (1), shows the page where it is coming from. The number in parentheses refers to the number of times a word appears on a page. For example, if the word occurs once on page 280, page 291 appears twice, and so on. This order continued from the beginning to the end.

If you look up the word agar in the dictionary, we will see it and you will have an idea of how many times, where and how it appears in the text "Arzygul" epos. The next step is to search for the poem the word agar there are any similar items. At present, this task is also performed by the computer itself. The word is presented in the dictionary in the same way as in the text, without any changes. For example: adashgan, adashganlar, adashdi, adashib, adashmayik, adolatda, adolatimni, adolatli, adolatni and etc.

Advantages of preserving originality, the ease of finding a specific element in the text, and comparing forming units are its advantages. But these five different forms of speech are actually derived from the same root – adashgan, and should be regarded as a lexeme. Leaving phrases in the dictionary, taking into account both sides, is convenient for research. Arranging keywords based on vocabulary is not a problem.

Each letter example can be displayed as follows:

Letters	А	Б	В	Г	Д	Е	Ё	Ж	З	И	Й
Words	232	685	48	85	253	61	155	140	73	155	95
Letters	К	Л	М	Н	О	П	Р	С	Т	У	Ф
Words	559	35	248	122	387	109	52	491	599	144	56
Letters	Х	Ц	Ч	Ш	Э	Ю	Я	Ў	Қ	Ғ	Ҳ
Words	153	-	165	113	184	123	70	273	628	46	140
Wordforms	*6679*										
Total frequency	*22660*										

Keywords (roots) statistics are as follows:

Letters	А	Б	В	Г	Д	Е	Ё	Ж	З	И	Й
Words	103	195	21	31	84	18	54	56	34	61	27
Letters	К	Л	М	Н	О	П	Р	С	Т	У	Ф
Words	129	17	107	68	122	50	25	168	232	48	23
Letters	Х	Ц	Ч	Ш	Э	Ю	Я	Ў	Қ	Ғ	Ҳ
Words	73	-	78	49	65	28	39	81	210	20	68
Total root words = 2384											

During the creation of the dictionaries it was learned that in the "Arzygul" the Islam shair used 2,384 main words in 6679 words and used 22,660 times.

The principle of the formation of frequency dictionaries is similar to the alphabetical dictionaries. Even in this type of dictionary, the total number of usage words and statistics on the page is clearly indicated. Only the words are given in alphabetical order, not in alphabetical order. For example:

bir 403 267(5) 268(3) 269(4) 270(5) 271(3) 272(2) 273(3) 274(4) 275(3) 276(3) 277(3) 278(3) 279(6) 280(7) 281(6) 282(3) 283(4) 284(7) 285(2) 286(4) 287(3) 288(7) 289(5) 290(3) 291(1) 292(3) 293(5) 294(5) 295(2) 296(2) 297(3) 298(3) 299(2) 300(4) 301(9) 302(5) 303(3) 304(6) 305(2) 306(2) 307(5) 308(3) 309(2) 310(1) 311(4) 312(5) 313(3) 314(2) 315(3) 316(6) 317(4) 318(2) 319(3) 320(3) 321(3) 322(3) 324(1) 325(5) 327(2) 329(5) 330(2) 331(3) 332(4) 333(1) 334(4) 335(3) 336(7) 337(6) 338(6) 339(4) 340(7) 341(4) 342(5) 343(3) 344(2) 345(1) 346(3) 347(3) 349(1) 350(3) 351(4) 352(3) 353(2) 354(2) 355(11) 357(4) 358(10) 359(2) 360(7) 361(6) 362(2) 363(7) 364(9) 365(4) 366(4) 367(1) 368(2) 369(3) 370(2) 371(2) 372(2) 373(2) 374(2) 375(3) 376(5) 377(2) 378(2) 379(3) 380(3);

dedi 266 268(2) 269(1) 270(2) 271(1) 272(2) 273(3) 274(2) 275(1) 277(1) 278(2) 279(3) 280(3) 281(2) 282(3) 283(1) 284(1) 285(6) 286(1) 287(1) 288(5) 289(2) 290(2) 291(1) 292(1) 293(1) 294(2) 295(1) 296(2) 297(2) 300(1) 301(2) 302(1) 303(1) 304(3) 305(2) 306(1) 307(3) 308(4) 309(7) 310(4) 311(2) 312(3) 313(3) 314(1) 315(2) 316(2) 317(3) 318(1) 320(2) 321(7) 322(1) 324(1) 325(2) 329(4) 330(6) 332(3) 333(1) 334(2) 336(4) 337(4) 338(4) 339(7) 340(3) 341(3) 342(4) 343(3) 344(2) 345(4) 346(1) 347(1) 350(3) 351(6) 352(2) 353(2) 354(1) 355(8) 357(1) 358(3) 359(5) 360(4) 361(7) 362(2) 363(1) 364(3) 365(3) 367(4) 368(5) 369(5) 370(2) 371(4) 372(2) 373(4) 374(1) 375(4) 376(2) 377(1) 378(2) 379(4) 380(3) and etc.

Based on dictionary articles we can say that the phrase "Orzigul" contains more than one hundred wordforms used, it is 15: bir (403), dedi (266), deb (262), bu (218), men (206), Orzigul (194), bilan (167), soʻz (153), shunda (132), Suvonxon (126), shu (118), qarab (115), olib (114), ham (110), endi (104) and etc.

In this way, the frequency with which words are reduced is determined. For example:
- from 90 to 100 – 1: Suvon (97);
- from 80 to 90 – 2: boʻldi (86), boʻlib (86);
- from 70 to 80 – 5: kelib (77), ekan (74), u (73), edi (71), uch (70);

- from 60 to 70 – 10: kun (69), qilib (69), mening (68), bo'lsa (67), podsho (67), keldi (66), menga (66), Oqtosh (66), borib (63), ikki (61);
- from 50 to 60 – 10: bor (57), ot (57), qirq (57), Olloyor (55), bo'lsin (54), va (54), har (54), bolam (53), beklar (52), ko'p (52);
- from 40 to 50 – 17: Orzigulning (49), sen (48), Suvonning (48), qildi (48), hech (48), qoshiga (47), bobo (46), dushman (46), Dilmurod (45), chiqib (45), oldi (44), ketdi (43), meni (43), berdi (42), nima (41), Ernazar (41), o'z (41);
- from 30 to 40 – 31: Sultonxon (39), bog'bon (38), kelgan (38), senga (38), gul (37), o'n (37), qo'liga (37), alla (36), sening (36), turib (36), yo'l (35), yo'q (35), Sarimiroxo'r (35), tashlab (35), so'rayman (33), yaxshi (33), Qoraxon (33), der (32), tushdi (32), cho'pon (32), ko'rib (31), Suvonxonning (31), turgan (31), ushlab (31), quloq (31), bordi (30), dono (30), kirib (30), kishi (30), oldiga (30), uchun (30).

In short, the frequency and percentage of top-down phrases in the "Arzygul" epos are as follows:

No.	*frequency of change*	*number*	*percent*
1	More 100	15	0,002246
2	from 90 to 100	1	0,00015
3	from 80 to 90	2	0,000299
4	from 70 to 80	5	0,000749
5	from 60 to 70	10	0,001497
6	from 50 to 60	10	0,001497
7	from 40 to 50	17	0,002545
8	from 30 to 40	31	0,004641
9	from 20 to 30	53	0,007935 (1)
10	from 10 to 20	249	0,037281 (4)
11	from 1 to 10 – 6286. from this:		0
12	words that appear 9 times	58	0,008684 (1)
13	words that appear 8 times	74	0,01108 (1)
14	words that appear 7 times	91	0,013625 (1)
15	words that appear 6 times	117	0,017518 (2)
16	words that appear 5 times	167	0,025004 (3)
17	words that appear 4 times	255	0,038179 (4)
18	words that appear 3 times	467	0,069921 (7)
19	words that appear 2 times	1041	0,155862 (16)
20	words that appear 1 times	4016	0,601288 (60)
	Total	*6679*	*100*

From the table it can be seen that the frequency of use of the wordforms and their activity are inversely proportional. That is, words, that appear one time 60%, 2-time 16%, 3-time 7%, 4-time and 10-to-20 words 4%, 5-time 3%, 6-time 2%, 7-of-9 and 20% to 30 are 1%.

Although not only the number of active and passive words in the story, but also the units that they are, it would be a mistake to consider them as a sign of skill. What are the benefits of this analysis? The activeness of the words makes it possible to identify their peculiarities to the language of the poems, and the peculiarities of the genre of folklore.

The three types of dictionaries, based on the text of the epigraph "Arzygul" epos, and a brief description of them, in a sense, can demonstrate the importance of lexicographical and statistical research in textual analysis, and give rise to certainty that these analyzes are more valid. But there is something else. What is the purpose of the material analysis? For example, in this work, we also have a specific goal - to analyze a particular poem in a linguistic and linguistic way – to draw conclusions about the style of Islam shair and its use of the native language.

Frequency of use of words and phrases in the text of the poem is known. We have a clear idea of the most or least used vocabulary and grammatical units, and we can show them exactly as they are. Now, let's talk about how the level of activity or inactivity of language units can play a role in the way Islam dictates poets or poets.

For example, the "Arzygul" poem contains 4329 phrases containing 33 words at the base of 4,529 times, accounting for 20% of the total words used in the poem.

The word with the highest frequency used the word bir is used 403 times. Its occurrence is rarely seen in the text of the epos: Barnogul, kanizlaridan yashirib, bir lagan tillani va qizini ham momaning qo'liga berdi (268) and etc. For the most part, he participates in the connotative meanings. That is why in the text of the poem it is possible to observe all the subtext, such as hypothesis, amplification, conjecture, accuracy, emphasis, humility, counting on the semantics and stylistics of word bir. We are limited to just one example: Bu dunyoning o'tarini bilgandir, Bu joylarga bir ko'p mehmon kelgandir (364). In some of them there is a certain resilience of traditional folklore: Bu dunyoda bir armonim shu edi, O'ynasam der edim senday yor bilan (301), Yov boshiga balolar yog'dirayik, Birin ko'zlab, o'nin boshin olayik (330).

This unit has been shown elsewhere as a necessary element of the storytelling method. If it is removed, this method is broken: Qorong'ilik Qushqanot oldida bir ajdarho bor edi, Orzigulni damiga tortar bo'ldi (286). U yerda bulbuligo'yo nomli bir qush bor edi (286) and etc.

An example of word bir's traditionalism in the text of the poems is its inclusion in the phrases of bir so'z aytmoq, bir so'z demoq. This is how it came to 122nd place in the text. About 30% of words are used in these sentences gul, nargis, dilbar, qalamqosh. The combination of one with flowery, narcissistic, charming, pen-like words is one of the most characteristic features that provide its services in providing expressiveness: Bog'dan bir ajoyib gulning isi podshoning dimog'iga urdi (270). Bog' ichida bir gul yuzli dilbarni ko'rdi (270.) Boqib ko'rdim, bog' ichida bir dilbar, Yor ishqiga xoning bo'ldi giriftor (271). Sahar vaqti noz uyquda yotuvdim, Bir qalamqosh kelib kirdi qoshima (289). Ko'rar ko'zim, bir nargisim sen eding, Bu yo'lingda sarg'aytirding betimni (312) and etc. These examples illustrate the peculiarities of the language of the poem in the example of the word bir in "Arzygul" epos.

The following elements dedi (266), claiming to be the next to follow deb (266), are not stylistic. But only in the form of a bakhshi comment: Ernazar bogʻbonning xotini eriga bir soʻz dedi (270), Shunda Haydarbek Ernazar bogʻbonga qarab, sovchiligini ma'lum qilib, bir soʻz dedi (272) and etc. The use of this sentence is also due to the high frequency of words bir, soʻz, dedi, phrase in the epos. As well as, the verb деб only one place – Qoʻzichoq deb, yuzlarimga toʻymasding, Uchunar deb, yolgʻiz uyga qoʻymasding (348), it is observed that the syntactic anaphor is expressive in its structure.

There is no need to look for the expressiveness of the active use of the names of the heroes – Arzygul (194) and Suvanxon / Suvon (126/97) and Olloyor (55). The same can be said for the onomastic names in general. Except for names like Qingir, which are in the talking category.

Statistical data shows that the auxiliary bilan uses 167 seats. It is doubtful that the word is so high. We associate it with the writer's actions.

In fact, the style of the poems is typical of auxiliary minan and man versions: Qilolmadim sizminan kengashim, Qulogʻingizga borarmikan nolishim (349); Doʻstlariman jam boʻlgancha, Qorongʻi boʻlib kun botdi (299), Ushlab olib jonivorni men soʻyib, Qirq yigitman oʻz vaqtimni xushlayman (300), Qaytib kelib, senman oʻynab-kulayin, Javob ber, dilbarim, shunga borayin (320) and etc. Well, these analyzes, in turn, continue to clarify aspects of the poems or the language and style of the performer.

References

1. Orzigul (Dostonlar). – Toshkent: Gʻ.Gʻulom nomidagi Adabiyot va san'at nashriyoti, 1975.
2. Rizaev S.A. Oʻzbek tilshunosligida lingvostatistika muammolari. – Toshkent: Fan, 2006.