FREQUENCY DICTIONARY OF 16th CENTURY CYRILLIC
WRITTEN MONUMENTOKSANA NIKA¹ – SVITLANA HRYTSYNA²¹Institute of Philology of Taras Shevchenko National University of Kyiv, Ukraine² Educational and Scientific Center “Institute of Biology and Medicine”
of Taras Shevchenko National University of Kyiv, Ukraine

NIKA, Oksana: Frequency dictionary of 16th century Cyrillic written monument.
Journal of Linguistics, 2019, Vol. 70, No 2, pp. 276 – 288.

Abstract: The article presents the algorithm of the frequency dictionary to an original ancient text, “Otpys” (“Response”) by Kliryk Ostrozkyi (the Cleric of Ostroh) of the late 16th century. Until now, no historical corpus of text of the Ukrainian language has been created; therefore the drafting of metagraphical texts with their subsequent processing in accordance with linguistic tasks can fill this gap. The peculiarity of creating a frequency dictionary based on one written monument is in using the model of frequency dictionaries and describing the specifics of processing the ancient text. These specifics is based on a deep understanding of the state of language in the end of the 16th century and consists in the unification of graphic and spelling variants, as well as in the formation of stems and lemmas. Work results are presented in the form of a Frequency Dictionary of Word Forms of “Otpys” by Kliryk Ostrozkyi according to the frequency decrease and a Frequency Dictionary of “Otpys” by Kliryk Ostrozkyi according to the frequency decrease.

Keywords: frequency dictionary, tokenization, stemming, lemmatization, hapax legomena, written monument of the late 16th century

1 INTRODUCTION

The text of “Otpys” by Kliryk Ostrozkyi, the subject of this analysis, is not common for this type of research, having been written in the end of the 16th century. In 2016, monograph “Dialogism in Historical Dimensions of the Old Ukrainian Time: “Otpys” (“Response”) by Kliryk Ostrozkyi” was published, where along with the photocopies of 1598 old-printed book, metagraphical text and “Index of words and word forms” to this text were provided [5, pp. 1–44]. In particular, the purpose of the publication was the “preparation of a Historical Corpus of the Ukrainian language” [5, p. IV].

In Slavic studies, frequency dictionaries of written monuments often became the basis of the research of historical lexicology [6]. O. Tvorohov developed “Frequency Dictionary of ‘The Tale of Bygone Years’ according to the Lavrentiev List” and showed

the significance of such dictionary for historical and linguistic research as well as the description of the type of dictionary to a literature monument ([7], [8]).

In Ukrainian studies, frequency dictionaries have been elaborated based on texts of the same style or genre, as well as individual texts (primarily literary ones) ([1], [2]). These dictionaries are mainly focused on the texts of the late 20th – early 21st centuries. Some of them are based on texts of the 19th century by Taras Shevchenko, Lesia Ukrainka, Ivan Franko.

In this article, frequency dictionary of an original Cyrillic monument of the late 16th century, “Otpys” by Kliryk Ostrozkyi, is created for the first time.

Taking this into account, it is necessary to clarify the usage of metagraphical text and its variants in the process of creating a frequency dictionary. In the article, the old printings of the late 16th century are called the “authentic text”, the metagraphical text with lowered superscripts, the decoding of some contracted words and sometimes simplified spelling are referred to as the “secondary text”, and the processed text after tokenization is the “working” one.

“Index of words and word forms” of the “Otpys” text was important for the analysis [5, pp. 1–44].

2 CREATION PROCEDURE OF THE WORD FORMS FREQUENCY DICTIONARY OF “OTPYS” BY KLIRIK OSTROZKYI ACCORDING TO THE FREQUENCY DECREASE

Tokenization was applied in the first stage of the text processing. For the purposes of analysis of modern languages, the task of tokenization is “to separate words from syntactic characters, numbers, complexes of letters and numbers, Internet addresses, nicknames, symbols %, +, -, //, etc.” [4, p. 35]. Taking into account historical nature of the analyzed text, tokenization of “Otpys” by Kliryk Ostrozkyi had some specifics and was carried out already with the secondary text.

Thus, tilde “~” (a superscripted diacritical mark used to reduce frequently repetitive words) was removed from the text. With the help of a computer program (developed by O. Malin, engineer, mechanic, researcher) according to the section “Words with tildes” [5, pp. 63–65], an algorithm of the correlation of words with the tilde to the corresponding words without the tilde was developed. As a result, they were replaced by those with a similar meaning, but complete in their form in the index, and then, for the accuracy of calculation, the respective words were replaced also in the “Otpys” text.

Parentheses in the authentic text serving as punctuation marks were also removed. Remarkably, parentheses “()”, that were introduced in the process of working on the secondary text, remained (they contained superscript letters, and some of ligatures were decoded with their help (^)), as they did not prevent the recognition of individual words that a sequence of letters between spaces was considered to be. In addition, the text was checked for the presence of square brackets that neither were a part of the

authentic text, but contained research information. Among the signs removed there were also dash, hyphen, comma, full stop, semicolon. For the accuracy of calculation, explanatory words in modern Ukrainian language were also removed from the text (for example, “gloss”). Moreover, foliations in the authentic and in the secondary text were removed. In addition, letters with a numerical value were not taken into account, as their sequence did not form words.

The removal of homonymy was no less important. A graphic symbol was added to either word from a homonym pair, thus helping the program to consider these two words as different ones.

With the help of a specially computer program developed by engineer, mechanic, and researcher O. Malin, a working text was prepared and automatic calculation of absolute frequency of each word form was carried out. As a result, the frequency dictionary of word forms of the “Otpys” was developed according to the frequency decrease.

2.1 The lexicographic database of the GDC

The lexicographic database of the GDC is a separate section of the corpus. This platform has functions such as collecting, processing and converting lexicographic data into dictionaries.

The lexicographic database includes:

- Digitized online dictionaries from earlier printed dictionaries
- Various lexicographic data collected from lexicographic fieldwork
- Lexicographic data published by various authors
- Lexicographic data extracted from the existing linguistic and/or ethnographic studies.

The lexicographic database grows continually, with new texts being added over time. The database covers over 10 dialects and lists about 60 000 entries. This database has been developed based on the traditional lexicographic principles and methods. The research team will follow this methods in compiling comprehensive online dictionary of other Georgian dialects. Overall, four online dictionaries has been published so far. The published dictionaries are: Fereydanian, Ingiloan, “Chveneburebi”, and Laz dictionaries. New dictionaries with corresponding lexicographic data will be added to the corpus interface.

3 CREATION PROCEDURE OF THE FREQUENCY DICTIONARY OF “OTPYs” BY KLIRIK OSTROZKYI ACCORDING TO THE FREQUENCY DECREASE

Lemmaization is a traditional stage of creating a frequency dictionary. V. Jongejan and H. Dalianis define it as “the process of reducing a word to its base form, normally the dictionary look-up form (lemma) of the word” [3, p. 45].

In this study, it comprised several stages. The first step was to determine parts of speech, and then to apply the algorithm of stemming to stem words of different parts of speech. According to B. Jongejan and H. Dalian, “Stemming conflates a word to its stem. A stem does not have to be the lemma of the word, but can be any trait that is shared between a group of words” [3, p. 45].

On the basis of the “Index of words and word forms”, words were reduced to their stems by removing endings, and sometimes suffixes. For example, the word forms of the noun **голова**/holova ‘head’ (**головахъ**/holovah, **головою**/holovoju, **головъ**/holov, **головы**/holovy, **головѣ**/holovi) were combined in the stem **голов** (holov). With the old Ukrainian language having variations of the graphic writing of the same word, the stemming of adjectives **дѹховной**/duchovnoj, **дѹховномѣ**/duchovnom, **дѹховномъ**/duchovnom, **дѹхъ(в)ными**/duchovnymi ‘spiritual’ resulted in two separate stems **дѹховн**, **дѹхъ(в)н**/duchovn. In the example above, the writing of **ов** (o), **в(в)** (v) is parallel. In addition, different writing of letters identical in sound can be referred to as graphic unifications: **зз** (z) (**закон**/zakon – **закон**/zakon ‘law’), **уѣ** (u) (**мудр**/mudr – **мѣдр**/mudr ‘wise’) and others.

It is worth noting that the same lemma could comprise words with the interchange of consonants, for example: **гѣ** (g/z) (**дорог**/dorohu – **дорозѣ**/dorozi ‘way’), **гз** (g/z) (**ноги**/nohy – **нозѣ**/nozi ‘legs’); different degrees of adjectives comparison (**менша**/menša ‘less’ – **наменшого**/namenšoho ‘the least’); different forms of pronouns: **я**/ja ‘I’, **ѣ**/ja ‘I’, **мене**/mene ‘me’, **мѣ**/mia ‘me’, **мнѣ**/mni ‘to me’, **мною**/mnoju ‘by me’, **ми**/mi ‘me’. All of them formed different stems, for instance: **дорог**/doroh, **дороз**/doroz; **менш**/menš, **наменш**/namenš etc.

Different parts of speech were analyzed separately. Word forms of nouns and adjectives with the same meaning, but different in terms of gender, number, and case were referred to nouns and adjective lemmas. Pronoun lemmas also included word forms, divided in two groups according to the type of declension: those comparable with nouns and others comparable with adjectives. The first group had case forms, and in some instances, indications of gender.

Verbal lemmas with the same lexical meaning united word forms different in tense, mood, aspect. Verbs with prepositional **сам** ‘self’ were analyzed in one paradigm (**приводила сам**/pryvodylasia – **сам приводила**/sia pryvodyla ‘lead’). Verbal forms ending with **-но**/no, **-то**/to were considered as separate lemmas. Participle lemmas combined word forms different in the categories of tense, aspect, voice, as well as gender, number and case, and adverbial participle lemmas – in the categories of aspect and tense. Adverbs and conjunctions were also defined by separate lemmas.

Thus, sets of stems corresponding to one lemma (with the same lexical meaning) were combined in groups. This resulted in a sequence of stem lists, each of them corresponding to one particular lemma.

On the basis of the frequency dictionary of word forms, the program merged word forms (taking into account their frequency) by stems corresponding to

a particular lemma. Thus, the total frequency of word forms which correlate with a certain lemma determined the frequency of the lemma itself. As the result of this stage, frequency dictionary of “Otpys” by Kliryk Ostrozkyi according to the frequency decrease was formed.

4 STATISTICAL REVIEW

1. In the text “Otpys” by Kliryk Ostrozkyi, containing 11 222 words (N), 2 640 lemmas (V) were recorded, which is the total amount of words reduced to the original form, and 4 677 word forms (Vf), the variations of inflexion within the same lemma.

$$N=11\ 222$$

$$V=2\ 640$$

$$Vf=4\ 677$$

2. The diversity of the dictionary (B) of “Otpys” by Kliryk Ostrozky, namely the index of diversity (the ratio of the lemma dictionary volume (V) to the text volume (N)) is 0,235.

$$B=\frac{V}{N} ; B=\frac{2640}{11222}=0,235$$

3. Average repetition of a word in the text (A), that is, the ratio of the text volume (N) to the lemmas dictionary volume (V) for “Otpys” is 4,25. This means that on average, each word is repeated in the text 4 times.

$$A=\frac{N}{V} ; A=\frac{11\ 222}{2\ 640}=4,25$$

4. Calculation of the quantity of words, called hapax legomena (Greek hapax legomena – ‘name once’), is one of the most common stages. It is generally accepted that these are the words used by the author in the text only once: “in the broad sense, hapax legomena can refer to lexical items used only in one written monument, from among all texts representing a significant period in the history of a language” [10, p.12]. Given the absence of a historical corpus of the Ukrainian language until now, it is not possible to compare the usage of hapax legomena in other ancient texts. Therefore, this article deals only with the usage of unique words in the “Otpys” by Kliryk Ostrozkyi. Consequently, the number of words with the frequency 1 (V1) in the analyzed text amounts to 1 632.

$$V1=1\ 632$$

5. The index of uniqueness for the text (Et), that is, the ratio of the quantity of words with the frequency 1 (V1) to the volume of the text (N) of “Otpys” is 0,1454.

$$Em=\frac{V1}{N} ; Em=\frac{1\ 632}{11\ 222}=0,1454$$

6. The index of uniqueness for the dictionary (Ed), that is, the ratio of the quantity of words with the frequency 1 (V1) to the dictionary volume (V)) equals 0,6181.

$$Ed=\frac{V1}{V} \; ; \; Ed=\frac{1 \; 632}{2 \; 640}=0,6181$$

7. The index of concentration is opposite to the previous results. The index of concentration in the text (Ect) is the ratio of the number of the most frequent words in the text with frequency of 10 and more (V10_t) to the total text volume (N). V10_t = 137, N = 11 222, respectively, the index of concentration in “Otpys” by Kliryk Ostrozkyi is 0,012.

$$Ect=\frac{V10t}{N} \; ; \; Ect=\frac{137}{11 \; 222}=0,012$$

8. The index of concentration in the dictionary (Ecd) is the ratio of the number of the most frequent words in the dictionary (with the frequency of 10 and more) (V10 = 175) to the total dictionary volume (V). In the frequency dictionary of “Otpys” it amounts to 0,066.

$$Ecd=\frac{V10}{V} \; ; \; Ecd=\frac{175}{2 \; 640}=0,066$$

Results of the frequency distribution of lemmas, calculated on the basis of the data contained in the frequency dictionary of “Otpys” according to the frequency decrease, and *word forms*, calculated on the basis of the data of frequency dictionary of the “Otpys” word forms according to the frequency decrease, are represented in Table 1.

Interval of frequencies	Number of word forms	Number of lemmas
500 – 999	1	1
400 – 499	0	0
300 – 399	0	0
200 – 299	1	2
100 – 199	4	10
90 – 99	2	4
80 – 89	1	4
70 – 79	1	3
60 – 69	4	6
50 – 59	2	8
40 – 49	14	12
30 – 39	10	9
20 – 29	23	28
10 – 19	74	88
9	16	14

Interval of frequencies	Number of word forms	Number of lemmas
8	18	13
7	26	40
6	2	4
5	66	69
4	94	83
3	168	165
2	531	414
1	3575	1586

Tab. 1. Results of the frequency distribution of lemmas and word forms

The table shows that the text of “Отпыс” is characterized by 1 word form and 1 lemma with the frequency in the range of 500–999; 1 word form and 2 lemmas with the frequency of 200–299; 4 word forms and 10 lemmas with the frequency of occurrence 100–199, etc.

It is worth noting that there are no word forms or lemmas with the frequencies of 300–399 and 400–499. The largest number of lemmas and word forms has the frequency of usage 1, 2 and 3 times (see Figure 2).

To illustrate this, examples of the lemmas and word forms according to the frequency of their usage are provided below. The most frequent word forms are *и/i* ‘and’ (573), *не/ne* ‘not’ (280), *в(ъ)/v* ‘in’ (169), *на/na* ‘on’ (151), *а/a* ‘and’ (147), *до/do* ‘to’ (108), *в(т)/ot* ‘from’ (93), *ко/jako* ‘as, like’ (93), *але/ale* ‘but’ (85), *же/že* ‘that’ (70), *то/to* ‘that’ (69), *бы/by* ‘would’ (67), *што/što* ‘what, that’ (64), *его/jeho* ‘him’ (62), *ижъ/iž* ‘that’ (59), *або/abo* ‘or’ (51), *такъ/tak* ‘so’ (49), *и/i* ‘and’ (48), *по/по* ‘by’ (48), *за/za* ‘for, with, in, according to’ (46), *в/v* ‘in’ (46), *хто/čto* ‘who’ (45), *сам/ia* ‘self’ (43), *ихъ/ich* ‘their, them’ (42), *църкѣн/cerkvi* ‘church’ (41), *з(ъ)/z* ‘with’ (41), *с/s* ‘with, from’ (40), *о/o* ‘oh’ (40), *есть/jest* ‘to be’ (40), *если/jesli* ‘if’ (40) and others.

The most frequent lemmas are *и/i* ‘and’ (621), *не/ne* ‘not’ (280), *въ/v* ‘in’ (207), *той/toj* ‘that’ (190), *на/na* ‘on’ (151), *а/a* ‘and’ (147), *который/kotoryj* ‘who, which’ (142), *был/byl* ‘was’ (138), *свій/svij* ‘own’ (121), *они/ony* ‘they’ (111), *до/do* ‘to’ (108), *что/čto* ‘that, what’ (107), *онъ/on* ‘he’ (101), *оний/onyj* ‘that’ (97), *ко/jako* ‘as, like’ (97), *в(т)/ot* ‘from’ (93), *есть/jest* ‘to be’ (91), *згода/zhoda* ‘agreement’ (85), *але/ale* ‘but’ (85), *ижъ/iž* ‘that’ (84), *църкѣн/cerkov* ‘church’ (82), *о/o* ‘oh’ (79), *з(ъ)/z* ‘with’ (71), *же/že* ‘that’ (70), *то/to* ‘that’ (69), *бы/by* ‘would’ (67), *богъ/boh* ‘God’ (65), *такъ/tak* ‘so’ (64), *всеъ/vsi* ‘all’ (64), *божій/božij* ‘God’s’ (62), *хто/čto* ‘who’ (56), *свѣтъ/sviatyj* ‘holy’ (56), *за/za* ‘for, with, in, according to’ (56) etc.

Lemma	Lemma's frequency	Lemma's word forms	Word forms frequency
БО 'because'	39	БО	39
ГДЕ 'where'	38	ГДЕ ГДѢ	34 4
СЛОВО 'word'	37	СЛОВО СЛОВА СЛОВѢ СЛОВО(М) СЛОВЫ СЛОВЪ СЛОВАХЪ СЛОВА(Х)	8 18 1 2 3 3 1 1
МИЛОСТЬ 'grace'	36	МИЛОСТЬ МИЛОСТИ МИЛОСТЬЮ	15 20 1
Петръ 'Peter'	34	Петръ Пе(т)ръ Петра Петроу Петромъ Петрови Петровъ	22 2 4 1 1 2 2

Tab. 2. Fragment of the frequency dictionary

The results of the analysis showed that among word forms with the frequency of 40 and higher, auxiliary parts of speech, pronouns (*хто/hto* 'who', *ихъ/ich* 'their, them'), linking verb (*естъ/jest* 'to be') and nouns (*църкѣви/cerkvi* 'church') prevail.

Among the most frequent lemmas, there are various parts of speech: in addition to the auxiliary parts of speech, there are also pronouns (*той/toj* 'that', *онъ/on* 'he', *мы/my* 'we'), adjectives (*божій/božij* 'God's', *святый/sviatyj* 'holy'), nouns (*църкѣвъ/cerkov* 'church'), numerals (*один/odyn* 'one'), copular verb (*естъ/jest* 'to be').

Fig. 1. Graphic representation of the distribution of lemmas and word forms according to the frequency decrease ranged from 999 to 4 based on Table 1

Fig. 2. Graphic representation of the distribution of lemmas and word forms according to the frequency decrease ranged from 3 to 1, according to Table 1

The next step was the calculation of ratio between word ranking and the volume of text which they cover (C). Further analysis requires to define the “ranking” of

a lemma. A lemma with the first ranking has the highest frequency, a lemma with the second ranking is the one following, based on the frequency decrease, and so on.

$$C=\frac{F}{N} \times 100\%;$$

$$F=F_1+ F_2+...+ F_n = \sum_{k=1}^n F_k$$

where **n** is the ranking of a lemma, **F** is the sum of frequencies from the 1st to the n-ranking.

For example, if n = 1, F is equal to the frequency of lemma with the first ranking, that is, the lemma with the highest frequency; if n = 2, then F equals the sum of frequencies of the first and second rankings; if n = 3, then F is equal to the sum of frequencies of the first, second and third rankings, and so on.

Number of first rankings (n) in calculation of the sum of frequencies (F)	% of coverage, (C)
1	6
5	13
10	18
25	30
50	41
75	48
100	52
200	63
300	68
400	72
500	75
1000	85
1500	89
2000	94
2640	100

Tab. 3. Percentage of text coverage by lemmas by sequences of rankings

Fig. 3. Graphic representation of percentage of text coverage by lemmas by sequence of rankings according to Table 2

The table and the figure show that *и* 'and' as a lemma with the first ranking, covers 6% of the text. The calculations were carried out as follows: a lemma with the first rankings *и* 'and' has the frequency of usage 621, that is:

$$F = F_1 = 621.$$

the percentage of text coverage (C) with the first ranking lemma is:

$$C = \frac{F}{N} \times 100\% = \frac{621}{11\ 222} \times 100\% = 5,53\ \% \approx 6\%$$

Calculation of the percentage of text coverage with the lemmas ranking from 1 to 5, namely: first ranking lemma *и* 'and' with the frequency of 621 (F_1), second ranking lemma *не/ne* 'not' with the frequency of 280 (F_2), third ranking lemma of *въ/v* 'in', with the frequency of 207 (F_3), fourth ranking lemma *той/toj* 'that' with the frequency of 190 (F_4), fifth ranking lemma *на/na* 'on' with the frequency of 151 (F_5). The sum of their frequencies (F) is equal to:

$$F = F_1 + F_2 + F_3 + F_4 + F_5 = 621 + 280 + 207 + 190 + 151 = 1449$$

So, the text coverage with the first 5 ranking lemmas is:

$$C = \frac{F}{N} \times 100\% = \frac{1\ 449}{11\ 222} \times 100\% = 12,9\ \% \approx 13\%$$

Therefore, the lemmas *и* 'and', *не/ne* 'not', *въ/v* 'in', *той/toj* 'that', *на/na* 'on' cover 13% of the text.

5 CONCLUSIONS

Thus, the article presents the algorithm of creating the frequency dictionary of the Cyrillic written monument of the 16th century, the analysis of the text “Otpys” by Kliryk Ostrozkyi according to the frequency of its elements. The research resulted in creation of the Frequency Dictionary of Word Forms of “Otpys” by Kliryk Ostrozkyi according to the frequency decrease and the Frequency Dictionary of “Otpys” by Kliryk Ostrozkyi according to the frequency decrease.

Research perspectives lie in the ambit of developing frequency dictionaries on the basis of other written monuments of that time and comparison of the obtained results according to the frequency of the usage of words, analysis of the obtained results in lexicology, morphology, stylistics, corpus linguistics and computer lexicography.

References

- [1] Buk S., and Rovenčák A. (2007). Častotnyi slovnyk romanu Ivana Franka “Perechresni stežky”. [Frequency dictionary of Ivan Franko’s novel “Cross Paths”]. In Stežkamy Frankovoho tekstu (komunikatyvni, stylistyčni ta leksyčni vymiry romanu “Perechresni stežky”), pages 138–369, Lviv.
- [2] Častotni slovnyky. [Frequency dictionaries.] Accessible at: <http://www.mova.info/Page.aspx?l1=57>.
- [3] Jongejan, B., and H. Dalianis. (2009). Automatic training of lemmatization rules that handle morphological changes in pre-, in- and suffixes alike. In the Proceeding of the ACL-2009, Joint conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, August 2–7, pages 145–153, Singapore.
- [4] Matvejeva S. A. (2018). Tokenizacija jak sposib obrobky korpusnogo tekstu [Tokenization as a method of text corpus processing] I Mižnarodna nauko-vo-prykładna konferencija “Prykladna i korpusna linhvistyka: rozrobleńna tehnolohij novoho pokolińńa”, pages 35–36, Kyiv.
- [5] Nika O. I. (2016). Dialohizm v istoryčnych vymirach staroukrajinskoho času: “Otpys” Kliryka Ostrozkoho [Dialogism in Historical Dimensions of the Old Ukrainian Time: “Otpys” by Kliryk Ostrozkyi (the Cleric of Ostroh).] Kyiv, Education of Ukraine Publishers.
- [6] Tvorogov O. V. (1984). Leksičeskij sostav “Povesti vremennyh let” (slovoukazateli i častotnyj slovnyk) [The lexical structure of the “The Tale of Bygone Years” (the word pointers and frequency dictionary).], pages 211–217, Kyiv. Accessible at: <http://litopys.org.ua/lavrlet/lavrdod20.htm>.
- [7] Tvorogov O. V. (1967). O primenenii častotnyh slovarej v istoričeskij leksikologii russkogo jazyka [About the usage of frequency dictionaries in the historical lexicology of the Russian language.] In Voprosy jazykoznanija, 2, pages 109–117.
- [8] Tvorogov O. V. (1961). O tipe slovaria k literaturnomu pamiatniku [The type of dictionary to the literary monument.] In Vestnik Leningradskogo universiteta, 14(3). Serija istorii, jazyka i literatury, pages 119–129.

- [9] Tvorogov O. V. (1962). Slovar'-kommentarij k "Povesti vremennyh let" [Dictionary-commentary on the "Tale of Bygone Years"] Author's abstract. dis. candidate of Philological Sciences, 20 p.
- [10] Tvorogov O. V. (1995). Gapaksy "Slova" [Hapaxes of "Word"] In Enciklopedija "Slova o polku Igoreve", In 5 Vol., Vol. 2. G – I, p. 12–15, St. Petersburg. Accessible at: <http://feb-web.ru/feb/slovinc/es/es2/es2-0121.htm>.