

Aleksandar Kostić

Faculty of Philosophy
University of Belgrade
Serbia

ELECTRONIC CORPUS OF SERBIAN LANGUAGE FROM 12TH TO 18TH CENTURY

In the mid 1950s at the Institute for Experimental Phonetic and Speech Pathology, Prof. Đorđe Kostić, at that time the director of the Institute, initiated a project of machine translation and automatic speech and text recognition. Professor Kostić, who directed this project, was of the opinion that the problem of automatic text and speech recognition could not be solved algorithmically and that a probabilistic approach is more plausible. However, a probabilistic approach required compilation of an annotated corpus that will provide precise probability estimates of all aspects of language – from phonology to syntax. This was the impetus that brought to the compilation of the *Corpus of Serbian Language*. The Corpus is diachronic, encompassing the Serbian language from the 12th century to the contemporary language, with 11 million words, each word being manually annotated for its grammatical status.

There were two principal aspects of this project. On the one hand, it was necessary to build up the system of annotation, while on the other hand the corpus had to be representative in order to provide reliable probability estimates.

The annotation implied that each word from the corpus should be defined in terms of its grammatical (i.e. morphological status). Available grammars of the Serbian (Serbo-Croatian) language did not seem to satisfy this requirement because at some instances they could not provide a strict specification for each word from the corpus (e.g. constituents of some complex verb tenses). In order to solve this problem a team of linguists expanded the repertoire of standard grammars, building up a system of annotation that distinguished about 2500 grammatical forms.

In order to properly represent the contemporary Serbian language, the Corpus encompassed five functional styles of written language: novels and essays (126 books), poetry (215 books), daily press (*Politika*), scientific texts (136 books) and political texts. Each sample consisted of about one million words: in sum, the sample of contemporary Serbian language consisted of about five million words. This, however, was just one segment of the Corpus. As noted, the Corpus is diachronic, divided into five periods. In addition to the sample of the contemporary language it includes the Serbian language from 12th to 18th century, Serbian language of 18th ct. and the first half of the 19th ct., complete works of Vuk St. Karadžić, and Serbian language of the second part of the 19th century.

The Corpus was manually annotated at the level of inflected morphology with several stages of text preparation and annotation. The final goal was to compile a series of frequency dictionaries that would serve as a probabilistic basis for automatic speech and text recognition

and machine translation. The compilation of frequency dictionaries consisted of 27 distinct operations. Here we outline the most important ones.

- a. A4 sheet of paper was divided into 16 frames and within each frame a word from a book was transcribed. For each given word, its page and line numbers in the original book were specified.
- b. Once the whole text was transcribed and lines and pages annotated, each word was specified for its grammatical status (Figure 1).

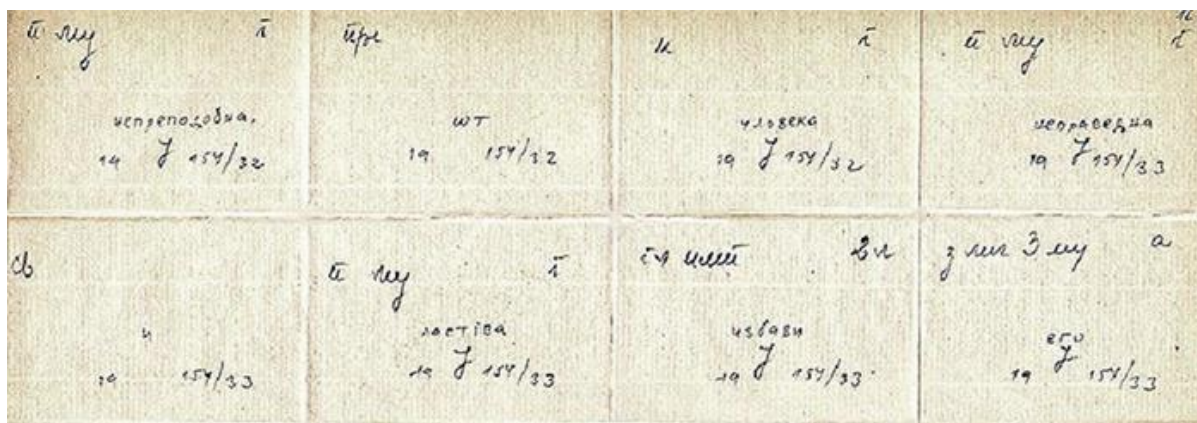


Figure 1: Part of A4 sheet divided into 16 frames with each word within a frame being annotated for its grammatical status, page and line in the book from which it was excerpted.

- c. Grammatical annotation was subsequently monitored by a group of linguistic experts that randomly sampled about 10% of each text. In cases where there was more than a 2% error rate, the grammatical annotation was repeated until the required criterion was reached. Sometimes the procedure had to be executed three or four times to reach the required standards.
- d. Once grammatical annotation was complete, the compilation of series of frequency dictionaries could be initiated. The A4 sheet was cut into 16 frames – one frame for each word. Word frames were then sorted into alphabetical order.
- e. Different grammatical forms (i.e. frames) for each word were sorted according to a specified order (e.g. for the word HOUSE, for example, all nominatives singular were put together, then all genitives singular etc.).
- f. For each word, the frequencies for each grammatical form were counted, as was the total the number of the occurrences of a word entry (e.g. HOUSE appeared 15 times in the nominative singular, 5 times in the genitive singular etc., The word HOUSE, irrespective of its grammatical form (i.e. word entry) appeared 75 times.).
- g. The obtained frequency counts were transcribed into a specialized form and then typed on an A4 sheet. In its final form each frequency dictionary had two distinct versions: one with word entries and grammatical forms for each word being sorted by alphabetical order and the other, with word entries sorted by rank frequency.

From 1957 to 1962 more that 400 collaborators participated in this project. In 1962 the project was suspended for reasons that were not scientific. By that time more than 27,000 pages of various frequency dictionaries had been compiled. In 1996 through mutual efforts of the Institute for Experimental Phonetics and Speech Pathology and the Laboratory for Experimental Psychology, Faculty of Philosophy, University of Belgrade the whole material

was digitalized and thus preserved. A new project, titled *Quantitative Description of Serbian Language Structure*, was initiated with a single aim to finish what had been terminated in the early sixties and to start a systematic statistical and mathematical analysis of the Serbian language based on materials from the project initiated by Prof. Kostić.

In the last fifteen years a seven volume *Frequency Dictionary of Contemporary Serbian Language* has been published by the Institute for Experimental Phonetics and Speech Pathology and the Laboratory for Experimental Psychology and eight volumes of frequency dictionaries of the Serbian language from 12th to 18th century, published by Službeni glasnik. The latter will be described in more detail in the forthcoming text.

Annotated corpus of Serbian Language form 12th to 18th century

Sample: The sample of the Serbian written language from 12th to 18th century consists of about 450.000 words compiled from 18 literary works by seven authors, one manuscript by anonymous authors and vast private and official correspondence of Serbian kings and noblemen. The language and orthography are in their original form, i.e. in Serbian-Slavonic. The whole material has been transferred into an electronic format.

The sample consists of two distinct types of material: a. the lives of Serbian saints (hagiographies), and b. the old Serbian charters and letters. There are also several examples of Serbian medieval church poetry.

Since almost all preserved hagiographies are included in the sample, this means that all significant authors of that period are represented in the Corpus (St. Sava, Teodosije, Domentijan, archbishop Danilo II, Stefan the First-Crowned, patriarch Pajsije and Grigorije Camblak). The lives of saints constitute a distinct genre of the Serbian medieval written literature, and present an interesting (although not always reliable) historical source as well as precious testimony of the Serbian literary language in the period between the 12th and the 18th centuries.

In the second part of the 19th century Đura Daničić, reputed Serbian philologist and expert on the old Serbian literature, edited and published manuscripts of the most prominent Serbian medieval authors (Domentijan, Teodosije and archbishop Danilo II). Daničić not only published those manuscripts, but through comparing the available medieval transcripts, had standardized the Serbian-Slavonic orthography. The material that constitutes the sample is in the most part based on Daničić's editions as well as on the editions of Janko Šafarik, Stojan Novaković and the extensive editorial endeavor of historian Vladimir Ćorović (St. Sava and Stefan the First-Crowned).

The old Serbian charters and letters encompass written documents that span five centuries of Serbian history. They include correspondence and charters of Serbian kings and noblemen, as well as contracts and other administrative documents and could be treated as valuable and reliable historical source. Unlike the lives of Serbian saints, the language in which they were written (also Serbian-Slavonic) is much closer to the Serbian vernacular of that period. The old charters and letters were collected and edited at the beginning of the 20th century in their original form by Ljubomir Stojanović, one of the most prominent Serbian philologists. The material in the Corpus is based on those editions. The list of manuscripts that constitute the sample of Serbian medieval written language is given in the Appendix.

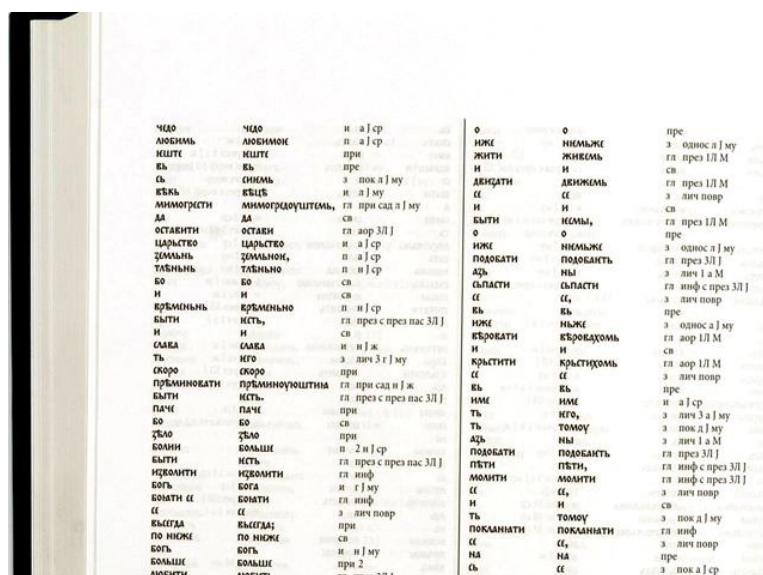
Transcript and grammatical tagging of the material was carried out in the mid 1950's by Prof. Đorđe Sp. Radojičić, at the time the utmost expert in the medieval Serbian literature and the Serbian-Slavonic language, and by a group of experts in Serbian-Slavonic from the University of Novi Sad.

The grammar (i.e. the inflected morphology) of Serbian-Slavonic is somewhat different from the grammar of the contemporary Serbian language, although those differences are not as great as might be expected given a time interval of almost eight centuries. The most conspicuous and common differences are in grammatical number (e.g. dual, which does not exist in contemporary Serbian language), use of participle and some pronouns. The constraint in the grammatical tagging of Serbian-Slavonic was that no radical change of coding system is permissible in order to enable comparisons with the contemporary language. Therefore, the system of tagging for the contemporary language has been projected on Serbian-Slavonic, with few minor changes that do not disturb the overall system.

Work in progress 2002–2012: In 2002 at the Laboratory for Experimental Psychology a group of experts for Serbian-Slavonic, with professor Đorđe Trifunović as a chief consultant, started the project of transferring hand written grammatically annotated Serbian-Slavonic texts into electronic format, with subsequent lemmatization that was not finished in the early sixties. This was the basis for compilation of frequency dictionaries and construction of application that enables various kinds of text analyses to be described in the forthcoming paragraphs. Up to now eight volumes of frequency dictionaries have been published by Službeni glasnik:

- Domentijan: *The life of St. Simeon and the Life of St. Sava* (three volumes – 3100 pages)
- Teodosije: *The Life of St. Sava* (one volume – 1250 pages)
- Archbishop Danilo II: *The Lives of Serbian Kings and Archbishops* (two volumes – 1733 pages)
- Vitae Canon Chronicles (writings of St. Sava, Stefan the First-Crowned, Grigoriје Camblak, Patriarch Pajsije, count Đorđe Branković and Founding charter of Hilandar) /two volumes - 2921 pages/

Dictionary for each author includes the original Serbian-Slavonic text, grammatically annotated and lemmatized text (Figure 2), frequency dictionary with lemma frequency and frequencies of all forms for a given lemma. The frequencies are given as raw frequencies, proportions and \log_2 transform of raw frequencies (Figure 3).



ЧЕЛО	ЧЕЛО	н а ср	О	О	пре
ЛЮБИМЬ	ЛЮБИМОУ	н а ср	НИЖЕ	НИЖЕ	э однос н му
КИШЕ	КИШЕ	при	ЖИТИ	ЖИТЕМЬ	гл през 1 Л М
КЪ	КЪ	пре	И	И	св
СЪ	СНЕМЬ	э пок а му	ДЕНГАТИ	ДЕНИЖЕМЬ	гл през 1 Л М
КЕКА	КЕКА	н а му	СЪ	СЪ	э лич повр
МИМОГРЕШЕ	МИМОГРЕШЕ	гл при сад н му	И	И	св
ДА	ДА	св	БЫТИ	ИМЫ	гл през 1 Л М
ОСТАВТИ	ОСТАВТИ	гл аор 3 Л	О	О	пре
ЦАРСТВО	ЦАРСТВО	н а ср	НИЖЕ	НИЖЕ	э однос н му
ДРАМА	ДРАМА	н а ср	ПОДОВАТИ	ПОДОВАТИ	гл през 3 Л
ТАБЕЛЬ	ТАБЕЛЬ	н а ср	НА	НА	э лич 1 а М
БО	БО	св	СЫПАТИ	СЫПАТИ	гл инф с през 3 Л
И	И	св	СЪ	СЪ	э лич повр
КРЕМЕНИ	КРЕМЕНИ	н а ср	КЪ	КЪ	пре
БЫТИ	ИСТЬ	гл през с през пас 3 Л	НИЖЕ	НИЖЕ	э однос а му
И	И	св	КРЕОВАТИ	КРЕОВАТИ	гл аор 1 Л М
САКА	САКА	н а ж	И	И	св
ТЪ	ИГО	э лич 3 г му	КРСТИТИ	КРСТИТИ	гл аор 1 Л М
ОКО	ОКО	при	СЪ	СЪ	э лич повр
ПОВИНОВАТИ	ПОВИНОВАТИ	гл при сад н ж	КЪ	КЪ	пре
БЫТИ	ИСТЬ	гл през с през пас 3 Л	НИЖЕ	НИЖЕ	э лич 3 а му
ПЛАЧЕ	ПЛАЧЕ	при	ТЪ	ТЪ	э пок д му
БО	БО	св	АДЪ	АДЪ	э лич 1 а М
УЛА	УЛА	при	ПОДОВАТИ	ПОДОВАТИ	гл през 3 Л
БОМЫ	БОМЫ	н а ср	ПВТИ	ПВТИ	гл инф с през 3 Л
БЫТИ	ИСТЬ	гл през с през пас 3 Л	МОМТИ	МОМТИ	э лич повр
ИДЕАТИ	ИДЕАТИ	гл инф	СЪ	СЪ	св
БОГЪ	БОГЪ	н а му	И	И	св
БОГАТИ	БОГАТИ	гл инф	ТЪ	ТОМОУ	э пок д му
СЪ	СЪ	э лич повр	ПОКЛНАТИ	ПОКЛНАТИ	гл инф
ВЪГДА	ВЪГДА	при	СЪ	СЪ	э лич повр
ПО НИЖЕ	ПО НИЖЕ	св	НА	НА	пре
БОГЪ	БОГЪ	н а му	СЪ	СЪ	э пок а ср
БОМЫ	БОМЫ	при 2			
МОМТИ	МОМТИ	гл пок 3 Л			

Figure 2: Page of dictionary with grammatically annotated and lemmatized text

Frequency dictionaries are divided into several segments: a. lemmas and forms of lemmas given in alphabetical order with their respective frequencies, b. lemmas given by rank frequency, c. lemmas given by rank frequency within part of speech (e.g. rank frequency for nouns, verbs, adjectives etc.), d. frequencies of onomastic terms, and e. frequency of grammatical forms.

прѣспѣти — прѣходити			
одредница/облик	граматички облик	F	p - log ₂ (p)
прѣспѣти	гл	2	,000026 15,2311
прѣспѣхъ	гл аор 1Л J	2	,000026 15,2311
прѣспавити се	гл	14	,000180 12,4397
прѣспавити се	гл инф	1	,000013 16,2311
прѣспави се	гл аор 3Л J	7	,000090 13,4397
прѣспави се	гл рп плп 3Л J му	1	,000013 16,2311
прѣспави се	гл при про г J му	1	,000013 16,2311
прѣспави се	гл при про г M му	1	,000013 16,2311
прѣспави се	гл при про д J му	3	,000039 14,6462
прѣспавеник	и	51	,000657 10,5718
прѣспавеник	и н J ср	3	,000039 14,6462
прѣспавеница	и г J ср	1	,000013 16,2311
прѣспавеница	и г J ср	16	,000206 12,2451
прѣспавеник	и д J ср	21	,000271 11,8494
прѣспавеник	и д J ср	2	,000026 15,2311
прѣспавеник	и а J ср	7	,000090 13,4397
прѣспавеник	и л J ср	1	,000013 16,2311
прѣспавити	гл	2	,000026 15,2311
прѣспави	гл при сад н J му	2	,000026 15,2311
прѣспави	и	109	,001404 9,4762
прѣспави	и н J му	3	,000039 14,6462
прѣспави	и н M му	1	,000013 16,2311
прѣспави	и г J му	12	,000155 12,6554
прѣспави	и д J му	9	,000116 13,0736
прѣспави	и а J му	65	,000837 10,2225
прѣспави	и инс J му	1	,000013 16,2311
прѣспави	и л J му	18	,000232 12,0736
прѣспавити	гл	3	,000039 14,6462
прѣспавити	гл инф	2	,000026 15,2311
прѣспави	гл при про н J му	1	,000013 16,2311

Figure 3: Page of frequency dictionary with raw lemma frequencies, raw frequencies of all forms for a given lemma, frequencies given as proportions, and log₂ transforms of raw frequencies

Д. Велектронски корпус архиепископа Даниила II - Животи краљева и архиепископа српских

Изворни текст Конкорданције Колокације Фреквенцијски речник Упутство за употребу

дело Животи краљева и архиепископа српских страна 1

жѣтѣ и жѣтѣ и похѣти боѣоудѣнѣхъ дѣлѣни скѣтѣхъ
 благѣбѣнѣхъ и христомѣнѣхъ **краљ** сѣбѣскѣ и поморѣскѣ
 дѣлѣ, кѣ нѣхѣхъ прѣбѣлаго начѣлѣника нѣхъ похѣлѣмъ свѣтѣлаго оуѣла
 нашѣго Симѣна Немѣно некалаго мирѣоуѣца сѣбѣскалаго, и сѣна нѣмоу
 прѣбѣскалаго архѣпископа кѣрѣ Сѣвоу, и бѣра сѣмоу
 прѣбѣскалаго прѣбѣскалаго краља Стефѣна, и сѣна нѣмоу краља
 Рѣдѣсѣла и Кѣдѣсѣла, и бѣра нѣхъ благѣскалаго и великаго краља
 Оуѣроша, и похѣлѣнѣ нѣхъ благѣскалаго краљѣ Оуѣроша,
 нарѣчѣноу кѣ нѣхѣсѣлѣмъ оуѣрѣ мѣнахѣ Немѣна, и сѣна нѣхъ
 прѣбѣскалаго и кѣрѣскалаго краља Стефѣна, такожде нарѣчѣнаго кѣ
 нѣхѣсѣлѣмъ оуѣрѣ Тѣбѣсѣлѣ мѣнахѣ, и бѣра нѣмоу благѣскалаго
 и христомѣнѣлаго и самѣоуѣскалаго сѣ бѣгомъ кѣсѣ сѣбѣскѣ и
 поморѣскѣ дѣлѣмъ господѣна краља Стефѣна Оуѣроша трѣтнаго, и
 сѣна нѣмоу кѣрѣскалаго и самѣоуѣскалаго краља Стефѣна, бѣноуѣла
 свѣтѣлаго краља Стефѣна Оуѣроша. похѣла жѣтѣна нѣхъ сѣстаѣмѣна
 прѣбѣскалѣмъ архѣпископомъ кѣрѣ Дѣнѣиѣмъ, пѣмѣтѣ
 прѣбѣскалѣмъ нѣхъ, кѣдѣмъ скѣтѣмъ благѣскалѣмъ.

жѣтѣмъ благѣскалаго краља Оуѣроша нарѣчѣнаго великаго.

Бѣдѣмѣмъ христомѣнѣмъ, сѣ оуѣ нѣмѣмъ кѣбѣрѣкѣмъ оуѣмъ
 малѣмъшѣмъ сѣмѣмъ хѣдѣмъ и мѣрѣмъ пѣмъ мѣмъ, хѣтѣ похѣлѣмъ
 дѣстѣмънѣмъ принѣсѣти Хѣрѣста бѣдѣмънѣмъ нѣмъ, нѣдѣмъ оуѣмъ кѣмъ
 нѣмъ гдѣмъ сѣмъмъ нѣмѣмъ. нѣмъ прѣбѣмъ оуѣмъ похѣмъ оуѣмъ, и
 кѣдѣмъ сѣмъ, и пѣмъ кѣ сѣмъ припѣдѣмъ и мѣмъ сѣмъ и гѣмъмъ: бѣмъмъ
 и прѣмъмъмъмъ и кѣсѣмъмъмъ и нѣдѣмъмъмъмъмъмъ, нѣ хѣтѣ
 сѣмъмъмъмъмъмъ, нѣ хѣтѣ кѣсѣмъмъмъмъмъ кѣмъмъмъмъмъ, рѣмъмъ кѣ
 сѣмъмъмъмъмъмъмъмъмъ: прѣсѣмъ и дѣсѣмъ сѣмъмъ, и пѣмъмъ: нѣмъ
 ашѣмъ кѣсѣмъмъмъмъмъмъмъмъ, кѣсѣмъ сѣмъ дѣсѣмъ сѣмъмъ: тѣмъмъмъ

одредница F 426

облик F 5

вѣста рѣчи граматички облик

имѣница и г M му

избор граматичког облика

бѣмъ похѣмъмъмъмъмъмъмъмъ 426

облики	граматички облик	F
краљ	и н J му	160
краља	и н D му	1
краља	и г J му	94
краља	и г M му	5
краљѣ	и г M му	1
краљѣ	и д J му	109
краља	и а J му	30
краљѣ	и а M му	1
краљѣ	и в J му	8
краљѣ	и инс J му	9
крали	и л J му	8

Figure 4: Panel with the original Serbian-Slavonic text and windows with lemma frequency, form frequency, grammatical status of a form and all forms of a given lemma in given text.

Frequency dictionaries are accompanied by CD with application that allows for number of different types of search and analyses. The application has four panels, each of them being devoted to a particular type of analysis.

The first panel offers the original Serbian-Slavonic text that can be scrolled and searched for lemma, grammatical form of lemma or grammatical form per se (Figure 4).

In the second panel the user can get concordances for the selected lemma, form of lemma or grammatical form, i.e., all sentences where the selected lemma, form of lemma or grammatical form appears. For each selected item from the list of concordances, the full sentence is available, with page number of the book from which the selected item was excerpted. Lemma frequency, form frequency, form's grammatical status and all the forms of the selected lemma in a given text are also available (Figure 5).

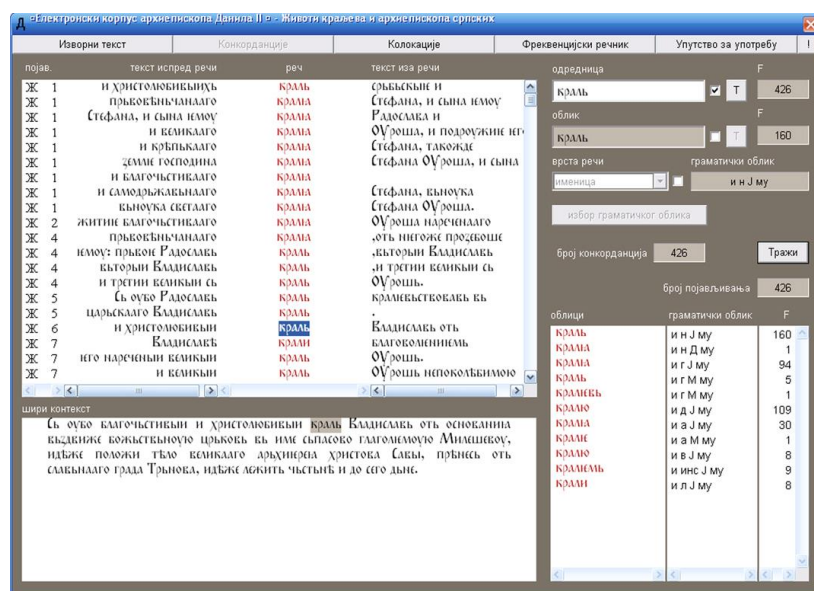


Figure 5: Panel with the list of concordances, the full sentence for a chosen item, lemma frequency, form frequency, grammatical status of a form and all forms of the selected lemma in a given text

The third panel offers collocations (i.e. the word that precedes or follows the target word) for each lemma, grammatical form of a lemma or grammatical form per se (Figure 6).

Collocations are ordered by rank frequency and for each word-word combination. It is possible to get the whole sentence as well as a grammatical specification for each word pair. Collocations could be selected by several criteria: a. particular lemma that precedes or follows the target word, b. part of speech (e.g. nouns, prepositions) preceding or following the target word, and c. grammatical form preceding or following the target word.

The fourth panel offers frequency dictionary with number of possible ways of sorting lemmas and grammatical forms of lemmas (Figure 7). Lemmas and grammatical forms of lemmas could be sorted by alphabetical order, with their respective frequencies. Likewise, it is possible to sort only lemmas by rank frequencies or to get rank frequencies of lemmas within part of speech (e.g. rank frequencies for nouns, verbs, adjectives etc).

The application offers specification of frequency range, for example, it is possible to extract 5% of most frequent words, or 10% of most frequent nouns etc. The application is still in developing stage. The final version will offer graphs and some elementary statistics.

прво појав	прева реч	друга реч	број појав
Ж 1	Богочастиниѣ	Краљ	107
Ж 45	Православниѣ	Краљ	96
Ж 1	Христомовниѣ	Краљ	47
Ж 58	Богочастиниѣ	Краљ	13
Ж 1	Богочастиниѣ	Краљ	8
Ж 1	Самодержавниѣ	Краљ	6
Ж 183	Младъ	Краљ	6
Ж 1	Православниѣ	Краљ	4
Ж 1	Крѣпкѣ	Краљ	4
Ж 118	Славыѣ	Краљ	2
Ж 179	Богочастиниѣ	Краљ	2
Ж 143	Блаженъ	Краљ	2
Ж 227	Домовиѣ	Краљ	1
Ж 1	Свѣтъ	Краљ	1
Ж 142	Свѣтоудѣлѣ	Краљ	1
Ж 229	Свѣтоудѣлѣ	Краљ	1
Ж 58	Срѣдѣ	Краљ	1
Ж 228	Окрѣпѣ	Краљ	1
Ж 131	Младѣ	Краљ	1

Figure 6: Panel with collocations ordered by rank frequency

одредница/облик	граматички облик	F
Господъ	и	513
Богъ	и	453
Краљ	и	426
Смѣна	и	258
Господиѣ	и	213
Сила	и	211
Житиѣ	и	210
Царѣ	и	208
Дѣнь	и	208
Глаголъ	и	203
Мѣсто	и	192
Аобы	и	191
Црѣкѣ	и	177
Дрѣвѣ	и	173
Врѣмѣ	и	171
Сынѣ	и	168
Срѣдѣ	и	156
Христосъ	и	153
Име	и	149
Слава	и	145
Тѣло	и	144
Доуша	и	142
Отѣцѣ	и	140
Отѣцѣ	и	133
Дѣло	и	128
Слово	и	128
Богѣ	и	123
Градѣ	и	120
Православниѣ	и	117
Рочѣ	и	115
Блаженъ	и	114
Родитѣ	и	112
Православъ	и	109

Figure 7: Panel with frequency dictionary

At this point each frequency dictionary has its own application restricted to a single author or group of authors. Once the whole sample of the Serbian language from 12th to 18th century is lemmatized and frequency dictionaries being published, the whole period will be available within a single application, allowing cross-referencing and direct comparison of different authors at the level of vocabulary and morphology (grammar). This, on the other hand, is a good basis for compiling a Serbian-Slavonic dictionary since concordances for each lemma are available at hand, thus offering an insight to all possible meanings of a given lemma. This, however, is beyond the scope of the current project.

APPENDIX

List of manuscripts that constitute the sample
of the Serbian written language from 12th to 18th century

The Old Serbian Charters and Letters, Collected by Ljubomir Stojanović, Belgrade, 1929.

Teodosije: *The Life of St. Sava*, Edited by Đura Daničić, Belgrade 1860.

Domentijan: *The Lives of St. Simeun and St. Sava*, Edited by Đura Daničić, Belgrade 1865.

Archbishop Danilo II: *The Lives of Serbian Kings and Archbishops*, Edited by Đura Daničić, Zagreb 1866.

Stefan Prvovenčani: *The Life of Simeon Nemanja*, Vladimir Ćorović: Svetosavski Zbornik 2, pp. 1–76.

Grigorije Camblak: *The Life of Stefan Uroš*, Janko Šafarik: Glasnik Društva srbske slovesnosti 11, 1859, pp. 43–94.

The Writings of St. Sava, Vladimir Ćorović: Belgrade – Sremski Karlovci, 1928.

Patriarch Pajsije: *The Life of Emperor Uroš*, Ilarion Ruvarac: Glasnik Srpskog učenog Društva 5, Vol. 22, 1867, pp. 209–239.

Anonymous: *The Life of Joachim from Sarandapor*, Stojan Novaković: Glasnik Srpskog učenog Društva 5, Vol. 22, 1867.

Excerpts from *The Chronicles of Despotes Đorđe Branković*, Stojan Novaković: Glas Srpskog učenog društva 33, 1872 pp. 242–264.

Encomium to St. Simeon and St. Sava, Đorđe Sp. Radojičić: Južnoslovenski filolog 21, Vol. 1–4, 1955-56, 137–151.

akostic@f.bg.ac.rs