MACHINE-READABLE LINGUISTIC INTERNET RESOURCES AS A BASIS FOR HISTORICAL-PHILOLOGICAL STUDIES

VICTOR A. BARANOV

Abstract

This report presents an experience of development and construction of the information-analytical system “Manuscript” designed for the preparation of the electronic publications of the medieval documents on the Internet (the project portal: http://manuscripts.ru/index_en.html) and also the technique of application of the electronic corpus to historical-linguistic research.
The special system modules interacting with the full-text database help to carry out the entire cycle of works on the preparation of the Internet edition, its annotation and linguistic marking.
The special attention is paid to the system possibilities for the preparation of the search requests and visualization of the retrieval results. The request criteria, various forms of ordering of retrieval units based on the meta-marking of the manuscripts and texts, the annotation of their fragments and the word-by-word parallel analysis of contexts help to the user to get the material for the linguistic and linguistic-textological analysis of the documents.
The electronic collections of Old Russian Service manuscripts of the 11-14th centuries and the collection of M. Lomonosov works are used as examples for demonstration of the system possibilities.

Mathematics Subject Classification 2000: 68N99
General Terms: On-Line Digital Collection
Additional Key Words and Phrases: Complex Technology of Creation of Electronic Collections of Slavonic Manuscripts on the Internet: Multilingual Full-Text Database, Tools for Filling, Access and Publication

1. AIMS AND TASKS

1.1. The last decades of the Russian historical linguistics show a considerable rise in interest in the confessional texts like Gospels, Menaia, collections of sticherons, Triodions and other similar texts and their extant medieval copies. The distinctions of those texts if compared to the original works of Slavonic authors consist in the fact that they can be translated, compiled, and possess relative stability of composition, structure and language. It was considered traditionally and, unfortunately, sometimes it is considered till now that the copies that differ little from one another give little material for studying the history of the Russian literary language, historical lexicology and semantics, and historical grammar and phonetics. One of the causes of such an attitude to the great amount extant medieval Slavonic manuscripts is not only orientation to the absence or insignificance of the discrepancies between the copies but also the difficulty of finding data in the copies of the same works looking really very similar that would give material for revelation of the common and particular regularities that determined the text changes, formation of the literary language and the influence of the informal speech on the written literary language.

1.2. Nowadays the corpus linguistics that enables the investigation of the speech and language facts on the basis of the analysis of great data arrays is in active progress. There are already created continue developing electronic corpora based on the modern documents in the Russian language: “National Corpus of the Russian Language”, the corpora of the Laboratory of General and Computer Lexicology and Lexicography of Moscow State University, Helsinki Annotated Corpus, Uppsala and Tubingen corpora etc. The corpus methods proved to be good in investigation of the language and speech variation, revelation of the facts leading to the origination of innovations and the extinction of archaisms, and in studying the conditions of origination of the differentiation of the linguistic units.

As known, the value of the corpus is in its orientation to the fixation and demonstration of the linguistic units in the amount of documents that would ensure representation of a certain number of them that would be satisfactory for investigation of the textual and linguistic properties, and, first of all, of the variation. As the corpus investigations require a certain level of frequency of unit occurrence for analysis of linguistic phenomena, the corpora of modern texts are created with a volume of one million word usages or more.

At the same time the corpus cannot exist without marking indispensable for its analysis and that enables separation of subcorpora on the basis of various textual parameters, for example, on the basis of genre, function-style, subject, content and other characteristics. The time parameter is extremely important in investigation of the essential characteristics. It is clear that the time differentiation of the documents is very limited within the modern corpus. This limitation is eliminated in the diachronic corpora. The time parameter is one of the most important in them and, as a result, it allows revelation of the common and particular characteristics of a certain speech and/or linguistic phenomenon through their changes on the basis of the observations over the changes of the correlation of the linguistic variants or over the replacement of one unit by another in the subcorpora.

1.3. At the same time there are objective causes due to which the creation of the diachronic corpora comparable with the modern corpora by the volume, completeness and balance of the diachronic corpora that demonstrate the documents of the long time period is doubtful (we distinguish the historical corpus that represents the documents created in a certain time period preceding the XIX century and the diachronic corpus that represents the texts of the pre-national period together with the modern texts). It is clear that the volume and completeness of the historical corpus are limited by the number of extant documents of a certain time period; the same fact limits the possibilities of achieving the corpus balance.

There are some more causes that nowadays considerably brake the creation and use of diachronic corpora: a very insignificant number of electronic machine-readable copies of the documents of the previous periods, big labour expenses for their preparation, the absence of the historical corpora comparable with the modern corpora by the most important corpus parameters, in particular, by the volume.

A special type of corpus is the author’s corpus which main characteristics are also determined by the volume of extant works of the author, their theme and genre-style characteristics.

Despite the fact that during the creation of the historical and author corpora it is often impossible to obtain such a number of units (for example, words or syntactic structures) that is required for getting statistically important values, the availability of all works of a specific author or all texts of a certain time period in such corpora gives grounds to consider them not less important and considerable for studying the language than the modern corpora of great volume: the author corpus represents the author sublanguage that is part of the language of the specific epoch; the historical corpus of a certain time period gives facts for description of the linguistic system that is comparable with the systems of the previous and following periods.

The consideration of the properties of entrance and comparability makes it possible to take off the limitations due to the lack of representation of linguistic phenomena in the texts and considering even single and statistically unrepresentative cases of variation as reliable enough against the wider time and/or author linguistic background.

1.4. The essential differences of the author and historical corpora from the modern corpora dictating the use of special methods for material analysis require the use of not only data demonstration techniques already firmly established in the investigation of corpora, but also specialized tools directed to the analysis of the historical and author speech and linguistic phenomena that ensure data visualization according to the aims and tasks of the researcher work.

Returning to the question of "immobility" of the confessional texts with traditional content that was risen in the first paragraph of this work, it is possible to state based on the above that: if a corpus or a collection represents not only different works, but also the copies of the same text made in different time periods then the discrepancies found between the copies on the background of the mass structure and speech coincidences become considerably more simple, their systematization more demonstrative, interpretation more convincing, and the conclusions on the factors, causes and mechanisms of origination and replication more true.

2. INFORMATION-ANALITICAL SYSTEM "MANUSCRIPT"

2.1. We repeatedly told in [2; 3; 4; 5; 7; 8; 9] and other publications about the creation of the electronic collections of medieval Slavonic manuscripts in the framework of project "Manuscript". This is why here I will only mention that by the efforts of the linguists and programmers of Izhevsk State Technical and Izhevsk State Universities 1) created was the information-analytical system intended for the preparation, demonstration on the Internet and investigation of medieval Slavonic manuscripts that are complex by structure, composition and encoding, 2) publicated were and continue being supplemented several collections including the transcriptions of the manuscripts of the XI-XVI centuries and also the works of the XVIII century (the portal link "Manuscript: Slavonic Written Heritage": http://manuscripts.ru/index_en.html). The preparation of the material is supported by the modules of input, editing, publication and representation on the Internet of the electronic transcriptions of the documents and reference materials created on their basis.

The undiminishing interest in the medieval manuscripts is understandable. The ancient church books are a valuable, and in a range of cases, an irreplaceable source not only for studying the medieval literary-writing norms, but also for studying the processes of formation of the norms of the Russian language on the whole. Until recently the texts of the Slavonic books of the XI-XIV cent. were considered only as the secondary source of studying of the history of the Russian language by its importance (after the original medieval Russian works): their proper textual and linguistic traits have not been studied enough till now.

The transcriptions of the documents and software created by the research group for work with them help introducing the manuscripts the access to which is limited in research, giving a possibility of getting acquainted with them not only to a wide circles of amateurs of medieval Slavonic books, but also a possibility of using the collections for searching and selecting materials for research.

2.2. Main System Modules and Their Functions.

The basis of the system is a database which model makes it possible to save any document objects, their values and relationships in the hierarchic form.

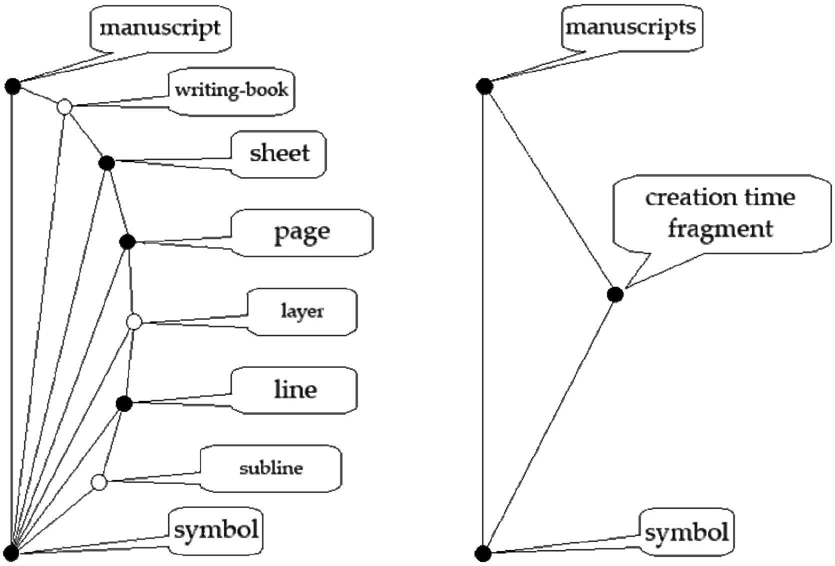


Fig. 1. An example of database structure: geometric and analytic hierarchy.

The database has foreseen special types of objects – reference books and dictionaries – that are invariants of the document objects.

Input and editing of the objects is performed with the use of the special-purpose editor OldEd interacting immediately with the database. The editor functions also enable assigning values to the objects, editing them, establishing relationships between the objects, and splitting the documents into fragments [1].

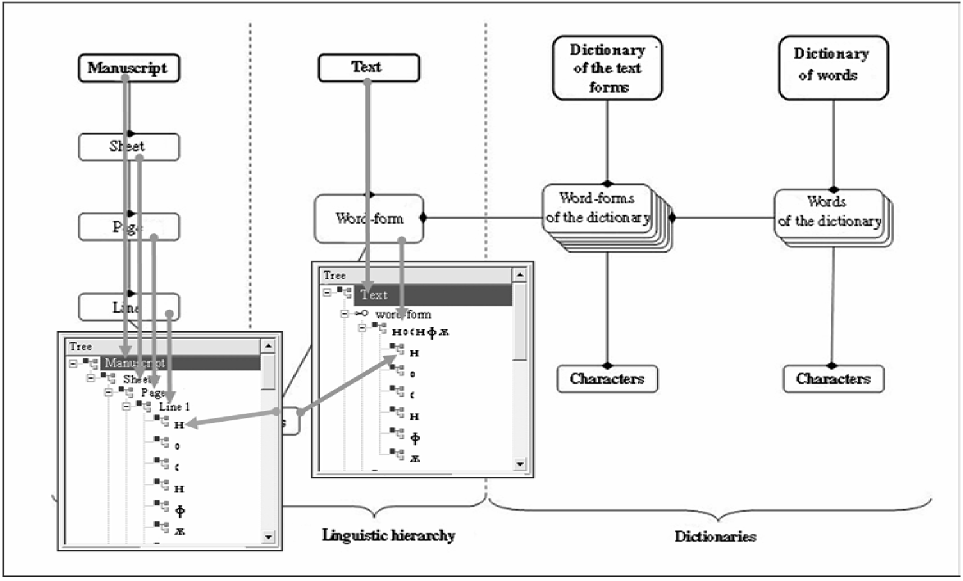


Fig. 2. An example of the relationship between objects and their invariants and the implementation of hierarchies in a specialized editor Olded.

The web-module of access to separate texts are the so-called single-text query forms that help to prepare a retrieval in the form of text fragment or indexes of word forms – direct, inversed or quantitative on the basis of simple request criteria (range of sheets, word form mask and some other). See, for example, the collection of the 11th century Slavonic Manuscrtips that is organized as several single-text query forms (URL: http://manuscripts.ru/mns/portal.main?p1=8&p_lid=2).

To get materials on the basis of several texts, the so-called multi-textual query form is used that makes it possible to obtain comparative reference materials over several manuscripts selected by the user at a time. The two-step interface of the internet-module enables to find and choose the manuscripts or texts and their fragments, to set the search parameters of the linguistic units and visualize the retrieval in the form of contexts, lists of tables (URL: http://manuscripts.ru/mns/srch.simple?p_lang=EN). This module makes it possible, in particular, to reveal the correlation of the alternative forms of the same grammar category in the large corpus of manuscripts and to analyze this correlation depending on the fragment type both in the same manuscript and various manuscript.

The preparation of complicated retrievals required for analysis of great data arrays is carried out with the help of the retrieval module (URL: http://manuscripts.ru/forms90/f90servlet?form=men2.fmx&config=MNS_SEC, authorization is required). The module is designed for the implementation of the complex queries and retrievals on the basis of the textological, linguistic, dictionary and other values of the units of the database “Manuscript” (for details refer, for example, to [4]).

There is created in the framework of the project the database of the morphological analyzer of the Old Russian that is intended for the automatic lemmatization (bringing of word forms to the initial form) of the ancient Slavonic texts. Four versions of the web-modules are available on the Internet (URL: http://manuscripts.ru/mns/portal.main?p1=16&p_lid=2&p_sid=1).

3. ELECTRONIC COLLECTION OF WORKS BY M. V. LOMONOSOV

3.1. The project includes the creation of a corpus of works by M. V. Lomonosov. The work goals are 1) demonstration of the literary heritage of M. V. Lomonosov, 2) offer to the users of the specialized web-services for work with the corpus – forms for the creation of queries, interfaces for visualization of retrievals and ordering of their units, means for viewing contexts etc.

The complete works of M. V. Lomonosov in 11 volumes [10] served as the material for the creation of the electronic collection.

3.2. A model of meta- and analytical marking of documents was developed specially for that collection and the activities on the creation of the web-pages of queries and output of findings, as well as the activities on the creation of the morphological analyzer for lemmatization of the corpus of the XVIII cent were conducted.

3.3. The meta- and analytical information on the corpus documents is a set of properties and their values that ensure the description of the document (the unit of the printed edition or work) and its fragments from the point of view of the main characteristics. The minimum unit of the marking is the symbol, but actually a text fragment or a fragment of the volume of the complete works is set as the minimal object. The set of fragment characteristics is determined by the fragment type according to its structural-functional properties.

A model of data of the corpus documents was developed on the basis of the analysis of works and their fragments, the meta- and analytical parameters and values required and satisfactory for material search and retrieval were revealed:

- the metadata of the volume of complete works: name, continuation of the name, parallel name, conventional name, place of publication, year of publication, publishing house;

- the metadata of the document: author, name, continuation of the name, parallel name, conventional name, language, capacity for being translated, time of creation beginning, time of creation end, place of creation, genre, text theme, text subjects, ordinal number;

- the analytical attributes of the work: title-paragraph structure, composition-theme structure; genre-style composition, non-textual fragments.

The input and editing of the meta- and analytical information were carried out with the help of the special-purpose editor OldEd that interacts directly with the System database.

3.4.1. The requirements for the query web-forms and retrieval representations were formulated as follows:

– selection of search attributes with the use of the logic conditions,

– selection of formal and grammar characteristics of the linguistic units;

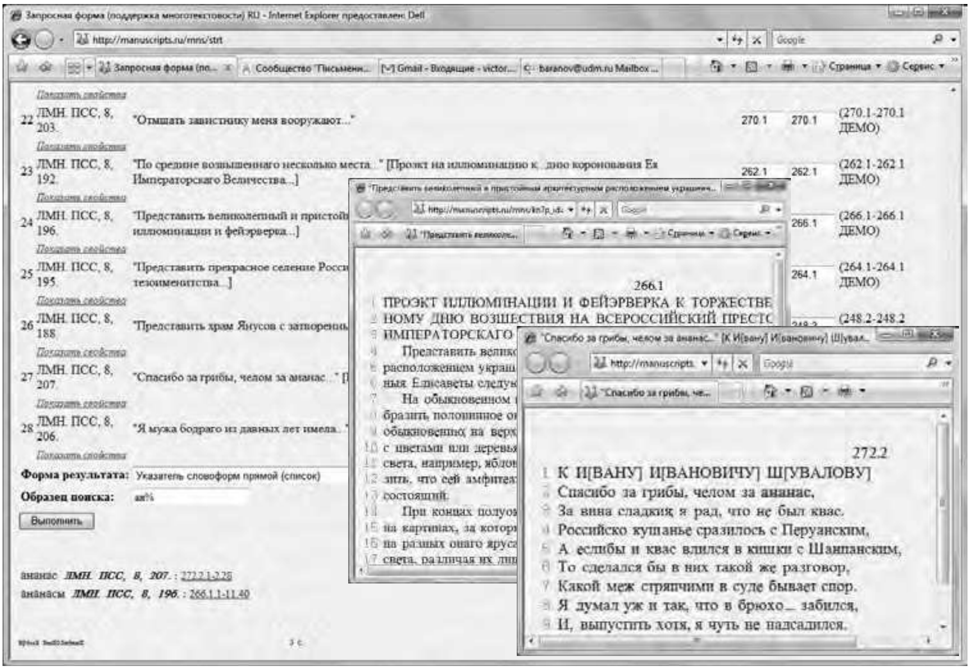
– representation of indexes on the basis of the query parameters,

– representation of indexes on the basis of several works,

– representation of direct, reverse, and quantitative indexes,

– jump from indexes to contexts.

3.4.1. The web-interface for search of texts and their fragments over the corpus by their meta- and analytical descriptions was developed and created in two modes – simple and advanced. The first mode uses the mechanism of search improvement that enables search detailing up to the required level (URL: http://manuscripts.ru/mns/srch.simple?p_lang=EN&p_ed_id=50584966). The filter creation in the second mode is possible simultaneously on the basis of several main objects of the collection – edition, work and the fragments of the texts combined by the logic conditions and on the basis of the parameters of the objects (URL: http://manuscripts.ru/mns/srch.complex?p_lang=EN&p_ed_id=50584966) (see [4; 8; 9] for details.


Запросная форма (поддержка многотекстовости) RU - Internet Explorer предоставлен: Dell
http://manuscripts.ru/mns/strt
Google
Запросная форма (по...
Сообщество "Письменн...
Gmail - Входящие - victor...
baranov@udm.ru Mailbox ...
Страница
Сервис
Показать свойства
Выполнить
"Представить великолепный и пристойный архитектурным расположением украшен...
http://manuscripts.ru/mns/kn?p_id=
"Представить великоле...
"Спасибо за грибы, челом за ананас..." [К И[вану] И[вановичу] Ш]увал...
http://manuscripts.
"Спасибо за грибы, че...

22	ЛМН. ПСС, 8, 203.	"Отмщать завистнику меня вооружают..."	270.1	270.1	(270.1-270.1 ДЕМО)
23	ЛМН. ПСС, 8, 192.	"По средине возвышеннаго несколько места..." [Проэкт на иллюминацию к дню коронования Ея Императорскаго Величества...]	262.1	262.1	(262.1-262.1 ДЕМО)
24	ЛМН. ПСС, 8, 196.	"Представить великолепный и пристойн... иллюминации и фейэрверка...]		266.1	(266.1-266.1 ДЕМО)
25	ЛМН. ПСС, 8, 195.	"Представить прекрасное селение Росси... тезоименитства...]		264.1	(264.1-264.1 ДЕМО)
26	ЛМН. ПСС, 8, 188.	"Представить храм Янусов с затворенн...		248.2	(248.2-248.2
27	ЛМН. ПСС, 8, 207.	"Спасибо за грибы, челом за ананас..." [			
28	ЛМН. ПСС, 8, 206.	"Я мужа бодраго из давных лет имела..."			

Форма результата: Указатель словоформ прямой (список)

Образец поиска: ан%

ананас *ЛМН. ПСС, 8, 207.* : 272.2.1-2.28
ананасы *ЛМН. ПСС, 8, 196.* : 266.1.1-11.40

266.1

ПРОЭКТ ИЛЛЮМИНАЦИИ И ФЕЙЭРВЕРКА К ТОРЖЕСТВЕ
НОМУ ДНЮ ВОЗШЕСТВИЯ НА ВСЕРОССИЙСКИЙ ПРЕСТО
ИМПЕРАТОРСКАГО
Представить велик
расположением украш
ныя Елисаветы следу
На обыкновенном
бразить полошнное о
обыкновенных на верх
с цветами или деревь
света, например, яблон
зить, что сей амфитеа
состоящий.
При концах полуо
на картинах, за котор
на разных онаго ярус
света, различая их лиц

272.2

К И[ВАНУ] И[ВАНОВИЧУ] Ш[УВАЛОВУ]
Спасибо за грибы, челом за ананас,
За вина сладкия, я рад, что не был квас.
Российско купанье сразилось с Перуанским,
А еслибы и квас влился в кишки с Шампанским,
То сделался бы в них такой же разговор,
Какой меж стряпчими в суде бывает спор.
Я думал уж и так, что в брюхо... забился,
И, выпустить хотя, я чуть не надсадился.

Fig. 3. Inquire Form Lomonosov collection, sampling and contexts

3.5. An automated lemmatizer was created on the basis of the grammar dictionary of the modern Russian language of the System (URL: http://manuscripts.ru/apex/f?p=104:1, authorization is required) for the corpus lemmatization.

The lemmatizer has a web-interface that enables to lemmatize Lomonosov's texts, control the lemmatization process and edit its data in the geographically distributed mode.

The lemmatization results are stored in the database of the grammar dictionary of Lomonosov language.

3.5.1. Functions of Lemmatization Module:

- performs lemmatization of arbitrary fragments of the texts of the System;
- gives possibilities of editing of the relationships of the text precedents and dictionary units manually;
- saves the obtained results in the System database with the possibility of further viewing through the web-interface;
- has an interface for input of new lexemes and word forms in the System grammar dictionary for lemmatization of unrecognized text precedents that have no form in the grammar dictionary;
- the lemmatization of the texts is done in succession;
- the module controls the queue of tasks (viewing, cancellation, deletion) through the web-interface etc.

3.5.2. The automated morphological analyzer (lemmatizer) has a three-level architecture and comprises the following components: the client application, the server application and the server of the database.

3.5.2.1. The client application is an interface component designed for work of the end user. It has no direct connections with the database and not loaded with the main logic of the system. This level ensures:

- user authorization;
- control over the process of automated lemmatization;
- viewing of the lemmatization results.

The client application of the lemmatizer has a web-interface and is supported in full by all up-to-date web-browsers. The free environment for software development based on Oracle - Oracle Application Express DBMS was used for its creation.

3.5.2.2. The second level of the architecture has located the server of applications. The main part of the system logic is concentrated here. The server functions are the following:

- creation and support of the dictionary and texts cashes in the actual state;
- control over the user queue of tasks for lemmatization;
- preliminary processing of data, preparation of them for lemmatization;
- proper process of automated analysis of word forms;
- post-processing of analysis data.

3.5.2.3. The third level of the architecture has arranged the database server. The server ensured storing of all system data: normalized dictionaries used in lemmatization, texts and also lemmatization results.

3.6. The word- and form index has a special composition and structure and is created by the user with the help of the query form: the index comprises the grammar marks of lemmas and word forms, the results of automated lemmatization and elimination of homonymy are demonstrated within one dictionary entry. The word forms with the eliminated and not eliminated homonymy are placed in different parts of the dictionary entry: first, all forms with eliminated homonymy are given, then with not eliminated homonymy. The index structure has envisaged demonstration of compound word forms and linguistic units (for example, prepositional-case form).

3.6.1. Components of Word- and Form Index:

- text abbreviation,
- text fragment abbreviation,
- address in the manuscript,
- lemma,
- linguistic unit,
- component of the compound linguistic unit,
- number of linguistic units,
- grammar marks (stem signs, paradigm signs, word form signs),
- priority signs,
- homonym number and some others.

3.6.2. Sequence Order of Index Components:

initial_form homonym_number "word lexical mark" (total_number_of word_forms) – *lemma grammar marks* (paradigm_abbreviation) *word_form_grammar_marks* (number_of_linguistic_units) priority_sign **linguistic_unit** / **component_of_compound_linguistic_unit**: **compound_linguistic_unit** TEXT_ABBREVIATION [fragment_abbreviation] address_in_manuscript & **linguistic_unit_with_not_eliminated_homonymy** / **component_of_compound_linguistic_unit_with_not_eliminated_homonymy**: **compound_linguistic_unit_with_components_with_not_eliminated_homonymy** *grammar_signs_of_word_form_with_not_eliminated_homonymy* /

grammar_signs_of_compound_linguistic_unit_with_not_eliminated_homonymy TEXT_ABBREVIATION [fragment_abbreviatio] address_in_manuscript.

Example:

честь (2) - *сущ. жен.* (парадигма № 33978791) *ед. вин.* **честь** ЛМН. ПСС, 8, 193 263.1.1-4.3 & ***чести*** *ед. род.* | *ед. дат.* | *ед. предложн.* ЛМН. ПСС, 8, 269 392.2.1-15.28

3.7.3. The hypertext jumps from the entry components available in the indexes enable to get additional information on the retrieval units. Thus, the jump from the lemma paradigm index enables to get the list of words with the paradigm that is similar to the lemma; the jump from the lemma gives the lemma paradigm with visualization of the word forms represented in Lomonosov texts by the precedents, from the paradigm word form – to the text precedents etc.

3.8. The research group has ahead a lot of work on the elimination of homonymy in the corpus. Now this labour-consuming part of the preparation of the corpus is carried out manually. The analysis of the results of the automated lemmatization only on the basis of the grammar dictionary of the modern Russian shows that the number of textual precedents for which there are found several homonymic forms varies from 43% to 55% and more in the texts that are different by genre and theme, and the number of not lemmatized textual precedents varies from 4% to 10%. For example, 46% from 91654 of textual precedents of the eighth volume are brought to the lemma identically, 45% have several variants of morphologic analysis, 9% have no lemma; 46% from 16847 textual word forms of the tenth volume have one variant of analysis, 47% have several variants, 6% have no variant.

At the same time, the use of the grammar dictionary for lemmatization which paradigms have added the morphologic indexes of the variable parts of speech found in Lomonosov texts more frequently decreases the percentage of not lemmatized textual precedents already on the first stage.

4. ELECTRONIC CRITICAL EDITION OF SERVICE MENAION FOR MAY

4.1. One year ago the activities on the creation of the first version of the critical edition were begun within the framework of the project. Nowadays some Slavonic manuscripts of the XI-XIV cent. of the Service Menaion for May and its Greek text are the materials for the development and creation of the procedures of data processing and the edition web-interfaces.

4.2. As known, the traditional printed critical edition of the medieval manuscript is a scientific publication where the text features are represented by its versions realized in the extant copies. The edition aimed at the comparison of the variants is structured so that to enable to see the discrepancies between the manuscripts that were introduced during correction, editing, insertions, losses, during copying, and, finally, to investigate the text history, textological, linguiostic and other signs of each specific copy. If the text is available in various languages, it is possible to find the correlation between the translation and the original, which is important for the historian, linguist, and culture expert.

The existing critical editions, despite the considerable variation in the ways of presenting the material, are stable in their main traits: the presence of the main text, bringing of the variant readings by the copies chosen for that, inclusion of appendixes in the form of reference books, indexes, commentaries, and also, if the text can be translated, the original in a different language.

4.3. The requirements for the electronic critical edition formulated in [6] are based both on the traditions of the printed edition and the opportunities provided by the information technologies.

Thus, when preparing the printed critical edition, the author of the publication chooses one of the copies as the main and shows the others in the form of discrepancies (variant readings) both on the level of the structure, composition, sequence order of fragments and on the level of linguistic units: word forms and their combinations. It is clear that in the electronic critical edition the user may select any copy as the main relative to which the discrepancies are shown.

The discrepancies of the structure and composition of the copies and of the sequence order of the fragments in the printed edition can be described only in the form of a commentary or represented in the form of a summarized table of fragments which volume may be big. In the electronic edition the presence or absence of the fragments corresponding to one another and their sequence in the manuscript may be also represented with the help of graphic means.

The search and selection of a specific material in the printed edition is carried out during reading or on the basis of the reference apparatus prepared by the publication author. The user of the electronic edition may have a possibility of selecting independently the material of interest, ordering and grouping it.

The list of differences can be continued.

4.4. The goal of the work is the creation of the electronic publication of the Service Menaion for May that contains materials for investigation of:

- the discrepancies in the composition and structure of the Slavonic copies of the text;

- the correlations and variant readings on the level of the vocabulary, morphology, syntax, semantics, graphics and orthography;

- the linguistic relationships between the Greek original and the Slavonic translation.

Several specific problems were and have been solved during the creation of the critical edition:

- the database with the invariants of fragments and linguistic units and their relationships with the textual precedents was designed and is being created;

- the textologic and linguistic analysis of the copies is being carried out with the aim of finding correlations on the level of the textual fragments and linguistic units;

- there are developed and created the web-interfaces of access to data that enable search, ordering and visualization of retrieval.

The electronic critical edition of the project “Manuscript” is being created on the basis of the full-text database containing transcriptions of several copies of the Menaion for May and the reference books necessary for work – the dictionary of fragments and the list of word forms that can be replenished – that are, in essence, the invariants of the fragments and word forms in the manuscripts.

4.5. The creation of the database of the critical edition is being carried out with the help of the special-purpose editor. The main task of the editor in this case is the support of the tools for establishing the relationships between the units of the documents and the reference books of the invariants.

To indicate the correspondences between the manuscripts on the level of fragments, the dictionary of fragments is used to tie the textual fragments corresponding to one another to their units. The units of the dictionary of fragments may have values that are inherited by the units of the manuscripts.

To establish the correspondences between the linguistic units of the different manuscripts, a special unit is used that plays the role of the reconstructed text which we call a ‘prototext’. The prototext is an auxiliary unit of the critical edition that ensures the relationships between the corresponding linguistic units of the copies. The visible representation of the prototext units is established by the author of the edition taking into account a certain system of graphic-orthographic normalization that can be used for visualization of the deviation from the system.

4.6. The prototext has a structure that is similar to the documents: it is split into the sheets, pages, and lines. The prototext comprises word forms and/or syntactic fragments and symbols that are subordinate units of the line. The volume (length) of the line is determined by the author of the critical edition. The sequence order of the prototext word forms is determined by the prototext author on the basis of the analysis of the manuscripts. The visible representation of the prototext unit in creation can be formed automatically and then edited. It is possible to view and change the properties both for the units of the reconstructed text and for the units of the manuscripts to be compared. There is foreseen indication of the necessary comments for any units and their relationships.

4.7. The scheme of preparation of the critical edition with the use of the editor looks as follows:

- the creation of the necessary prototext and filling it with the required number of sheets in the automated mode;
- opening of two or several texts to be compared;
- the creation of the unit of the required type in the reconstructed text or by indicating the required properties manually, or with the use of the automated creation on the basis of a unit of one of the texts;
- indication of the relationships (variant readings) between the units of the texts to be compared relative to the prototext units.

The result of the work of the author of the critical edition is the database containing the information on the connections between the dictionary of fragments and the fragments of the manuscripts and the prototext units and the linguistic units of the texts.

4.8. The web-interface of the critical edition of the Service Menaion for May is one of the module of the information-analytical system “Manuscript” and is available on the portal “Manuscript: Slavonic Written Heritage” (URL: http://manuscripts.ru/mns/cred.cred).

4.8.1. The manuscripts available in the edition are not represented evidently or by a list, and, this is why, to get the list of manuscripts, the user should introduce the search mask in the query field. After getting the list of manuscripts, the user gets a possibility of choosing the parameters of visualization of the edition page.

The edition has a Two-window interface. The upper window is intended for the determination of the query parameters (what? where?) and the form of data visualization (how?).

What?

- fragments or linguistic units,
- correspondences or variant readings.

Where?

- number of manuscripts,
- sheets or fragments.

How?

- form of visualization,
- method of showing the correspondences and variant readings.

The lower window is intended for the data visualization. There are foreseen several types and ways of demonstration of correspondences and variant readings:

- visualization of the composition and structure of the manuscripts on the level of fragments – songs,
- visualization of the composition and structure of the manuscripts on the level of linguistic units,
- visualization of discrepancies in the sequence order of the fragments and linguistic units,
- visualization of discrepancies in the unit values.

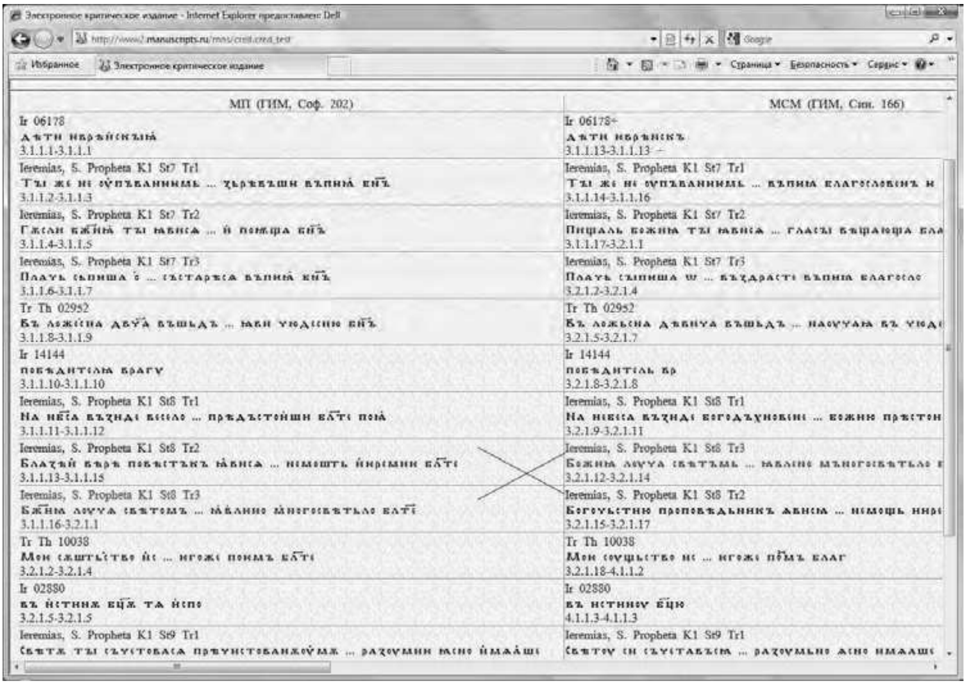
The means of unit identification and navigation between the units corresponding to one another are used for data visualization.

The identification means:

- manuscript abbreviation,
- fragment unique name,
- unit adress.

The navigation means:

- arrangement of units,
- connecting lines,
- highlighting in color and fixation of the highlighting.


Электронное критическое издание - Internet Explorer предоставлен Dell
http://www2.manuscripts.ru/mns/cred.cred_test
Google
Избранное
Электронное критическое издание
Страница
Безопасность
Сервис

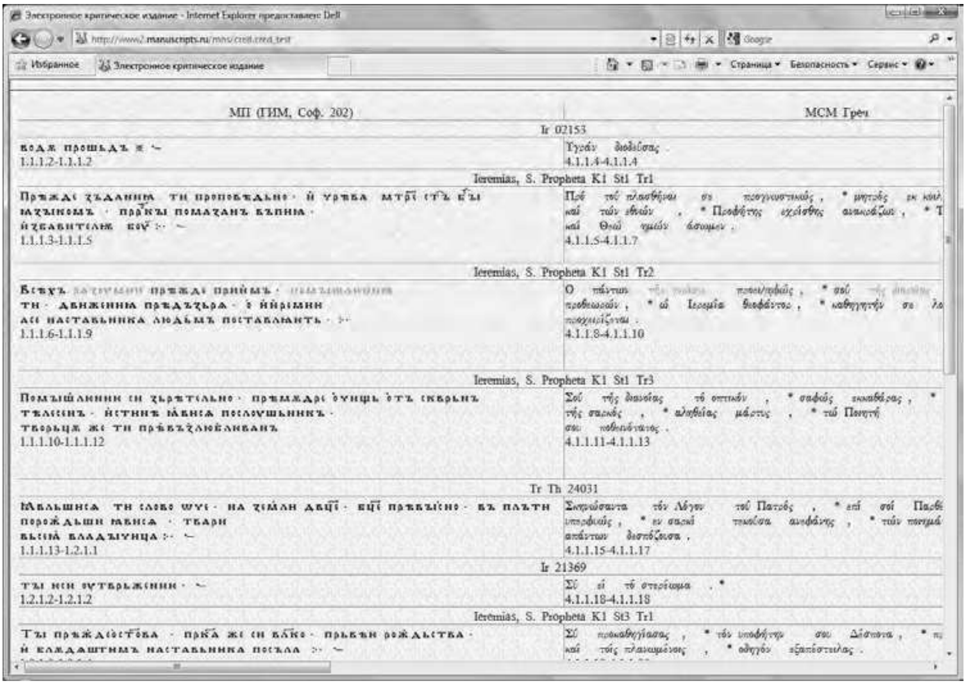
МП (ГИМ, Соф. 202)	МСМ (ГИМ, Син. 166)
Ir 06178 дѣти нєврѣйскыя 3.1.1.1-3.1.1.1	Ir 06178+ дѣти нєврѣнскъ 3.1.1.13-3.1.1.13 –
Ieremias, S. Propheta K1 St7 Tr1 Ты жє нє ѹпъванниимь ... ҳърѣвъши въпия бн҃ъ 3.1.1.2-3.1.1.3	Ieremias, S. Propheta K1 St7 Tr1 Ты жє нє оупъванниимь ... въпиꙗ благословєнъ и 3.1.1.14-3.1.1.16
Ieremias, S. Propheta K1 St7 Tr2 Гѫсли бж҃ия ты явися ... и помоща бн҃ъ 3.1.1.4-3.1.1.5	Ieremias, S. Propheta K1 St7 Tr2 Пищаль божия ты явися ... гласы вѣщающа бла 3.1.1.17-3.2.1.1
Ieremias, S. Propheta K1 St7 Tr3 Плачь сьпиша ѿ ... състаръся въпия бн҃ъ 3.1.1.6-3.1.1.7	Ieremias, S. Propheta K1 St7 Tr3 Плачь съпиша ѡ ... въздрасте въпия благосло 3.2.1.2-3.2.1.4
Tr Th 02952 Въ ложєсна дв҃а въшьдъ ... яви чюдєсно бн҃ъ 3.1.1.8-3.1.1.9	Tr Th 02952 Въ ложьсна дѣвича въшьдъ ... научая въ чюдє 3.2.1.5-3.2.1.7
Ir 14144 побѣдитєля врагу 3.1.1.10-3.1.1.10	Ir 14144 побѣдитєль вр 3.2.1.8-3.2.1.8
Ieremias, S. Propheta K1 St8 Tr1 На нб҃са възидє вєсєло ... прѣдъстоиши бл҃тє поя 3.1.1.11-3.1.1.12	Ieremias, S. Propheta K1 St8 Tr1 На нєбєса възидє богодъхновєнє ... божию прѣстои 3.2.1.9-3.2.1.11
Ieremias, S. Propheta K1 St8 Tr2 Блаꙁый вѣрѣ повѣстъкъ явися ... нємощь ѥирємии бл҃тє 3.1.1.13-3.1.1.15	Ieremias, S. Propheta K1 St8 Tr3 Божия лоуча свѣтъмь ... явлєно мъногосвѣтьло б 3.2.1.12-3.2.1.14
Ieremias, S. Propheta K1 St8 Tr3 Бж҃ия лоуча свѣтомъ ... явлєно многосвѣтьло бл҃тє 3.1.1.16-3.2.1.1	Ieremias, S. Propheta K1 St8 Tr2 Богочьстию проповѣдьникъ явися ... нємощь ниє 3.2.1.15-3.2.1.17
Tr Th 10038 Моє сѫщьство ис ... ѥгожє поємъ бл҃тє 3.2.1.2-3.2.1.4	Tr Th 10038 Моє соущьство нє ... ѥгожє пѣмъ благ 3.2.1.18-4.1.1.2
Ir 02880 въ истинѫ бцѫ та испо 3.2.1.5-3.2.1.5	Ir 02880 въ истиноу бцю 4.1.1.3-4.1.1.3
Ieremias, S. Propheta K1 St9 Tr1 Свѣтѫ ты съчєтоваса прѣчистованѫоумѫ ... разоумни ясно имааши	Ieremias, S. Propheta K1 St9 Tr1 Свѣтоу си съчєтавъся ... разоумьно ясно имааши

Fig 4. Incipitarium of Manuscripts

4.8.2. Display (visualization) of the correspondences on the level of linguistic units (word forms and syntactic fragments) is done in various forms:

- vertical text,
- parallel fragments,
- text with brought variant readings.

The visualization forms differ in the arrangement of units, way of visualization of the correspondences, and the number and composition of additional information. It is possible to highlight the units in the main manuscript in any form and to display the correspondences in the others.


Электронное критическое издание - Internet Explorer предоставлен Dell
http://www2.manuscripts.ru/mns/cred.cred_test
Google
Избранное
Электронное критическое издание
Страница ▾ Безопасность ▾ Сервис ▾

МП (ГИМ, Соф. 202)	МСМ Греч
Ir 02153	
водѫ прошьдъ ж ∼ 1.1.1.2-1.1.1.2	Ὑγράν διοδεύσας . 4.1.1.4-4.1.1.4
Ieremias, S. Propheta K1 St1 Tr1	
Прѣжде зъданиꙗ ти проповѣдьно · и утробѣ мтрѣ стъ бы ꙗзыкомъ · пркы помазанъ бъпиꙗ · избавителю бу ⁘ ∼ 1.1.1.3-1.1.1.5	Πρό τοῦ πλασθῆναι σε προγνωστικῶς , * μητρὸς ἐκ κοιλ καὶ τῶν ἐθνῶν , * Προφήτης ἐχρίσθης ἀνακράζων , * Τ καὶ Θεῷ ἡμῶν ᾄσωμεν . 4.1.1.5-4.1.1.7
Ieremias, S. Propheta K1 St1 Tr2	
Вьсѣхъ [illegible] прѣжде приимъ · [illegible] ти · движениꙗ прѣдъзьра · о иереміи лѣ наставьника людьмъ поставлꙗють · ⁘ 1.1.1.6-1.1.1.9	Ο πάντων [illegible] προσεκτικῶς , * σοῦ [illegible] προθεωρῶν , * ὦ Ἱερεμία θεοφάντορ , * καθηγητήν σε λα προχειρίζονται . 4.1.1.8-4.1.1.10
Ieremias, S. Propheta K1 St1 Tr3	
Помышлениꙗ си зьрѣтельно · прѣмѫдре очищь отъ скврьнъ тѣлесънъ · истинѣ ꙗвиса послоушьникъ · творьцѫ же ти прѣвъзлюбленаванъ 1.1.1.10-1.1.1.12	Σοῦ τῆς διανοίας τὸ ὀπτικόν , * σαφῶς ἐκκαθάρας , * τῆς σαρκός , * ἀληθείας μάρτυς , * τῷ Ποιητῇ σου ποθεινότατος . 4.1.1.11-4.1.1.13
Tr Th 24031	
Ꙗвльшиса ти слово оца · на земли двце · еже прѣвыше · въ плъти порождьши ꙗвиса · твари вьсеꙗ владычица ⁘ ∼ 1.1.1.13-1.2.1.1	Σαρκώσαντα τὸν Λόγον τοῦ Πατρός , * ἐπὶ σοὶ Παρθ ὑπερφυῶς , * ἐν σαρκὶ τεκοῦσα ἀνεφάνης , * τῶν ποιημά ἁπάντων δεσπόζουσα . 4.1.1.15-4.1.1.17
Ir 21369	
ты еси оутвьрждение · ∼ 1.2.1.2-1.2.1.2	Σὺ εἶ τὸ στερέωμα . * 4.1.1.18-4.1.1.18
Ieremias, S. Propheta K1 St3 Tr1	
Ты прѣждесътвова · прка же си бже · прьвѣи рождьства · и блѫдащимъ наставьника посъла ⁘ ∼	Σὺ προκαθηγήσας , * τὸν ὑποφήτην σου Δέσποτα , * π καὶ τοῖς πλανωμένοις , * ὁδηγὸν ἐξαπέστειλας .

Fig. 5. Parallel Fragments

4.8.3. The structure and composition of the manuscripts can be seen in two somewhat different forms:

- in the form of parallel fragments where one of the identifying component are the incipit and the explicit;
- in the form of parallel fragments which identifying components are their unique names and addresses.

4.8.4. The user has a possibility of choosing the sheets of the main manuscript for viewing and changing the sequence order of the manuscripts. The sequence order of the manuscripts is set by changing the priority in the window on the left of each manuscript.

The user also has a possibility of choosing one of two ways of representation of the relationships between the corresponding units. The first way of displaying the various positions of the units uses the lines connecting the fragments of the manuscripts arranged next. With the second way the line connects the fragments of each manuscript with the relevant unit of the main manuscript.

The material selection can be done not only by choosing the range of sheets of the main manuscript, but also by choosing the text fragments.

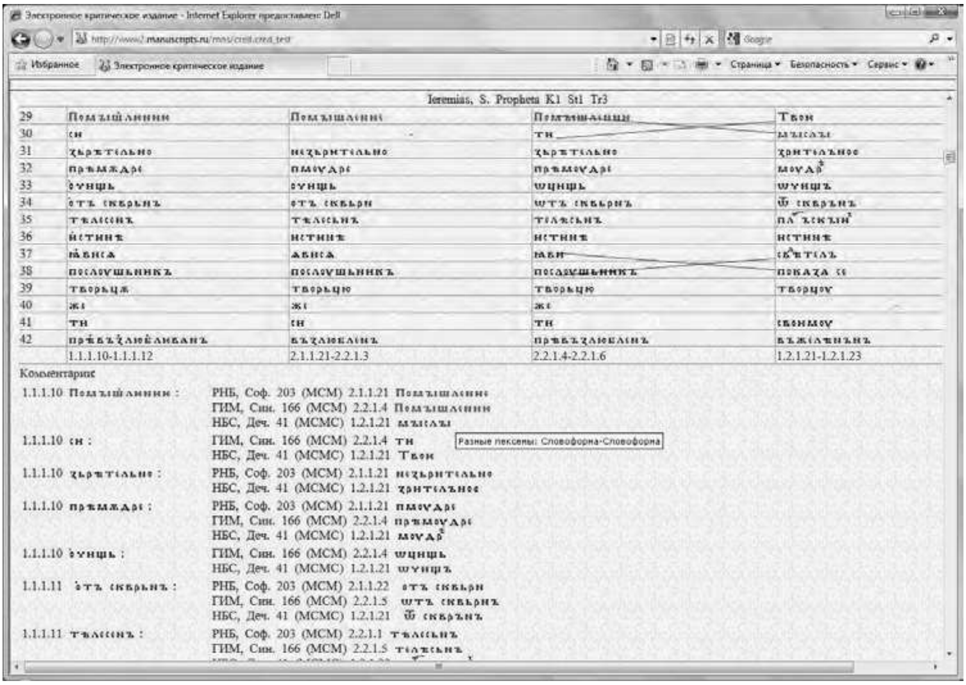
4.8.5. Nowadays the edition has foreseen four forms of demonstration of the correspondences and variant readings of the linguistic units:

- demonstration of the arrangement of correspondences,

- demonstration of the form and value of the correspondences,
- demonstration of the arrangement of the variant readings,
- demonstration of the form and value of the variant readings.

The discrepancies in the forms consist in:

- the type of the demonstrated relationships – visualized are all corresponding units or only those that have discrepancies;
- the arrangement of the units – the textual precedents are arranged vertically or by lines;
- in the presence of additional information – the form can include information on the type of variant readings;
- in the presence of the summarized list of variant readings of the retrieval.


Электронное критическое издание - Internet Explorer предоставлен: Dell
http://www2.manuscripts.ru/mns/crit.crit_test
Google
Избранное
Электронное критическое издание
Страница
Безопасность
Сервис

Ieremias, S. Propheta K1 St1 Tr3

29	Помꙑшлѧнини	Помꙑшлєниѥ	Помꙑшлєнии	Твои
30	си	-	ти	мꙑслꙑ
31	ꙁьрѣтєльно	нєꙁьритєльно	ꙁьрѣтєльно	ꙁритєлъноє
32	прѣмѫдрє	пмоудрє	прѣмоудрє	моудрⷭ҇
33	оуньшь	оуньшь	ѡуньшь	ѡуньшъ
34	отъ скврьнъ	отъ скврьни	ѡтъ скврьнъ	ѿ скврънъ
35	тѣлєснъ	тѣлєсьнъ	тѣлѣснъ	пл҃ъскъıи
36	истинѣ	истинѣ	истинѣ	истинѣ
37	ꙗвиса	ꙗвиса	ꙗви	свѣтєлъ
38	послоушьникъ	послоушьникъ	послоушьникъ	показа сє
39	творьцѧ	творьцю	творьцю	творцоу
40	жє	жє	жє	
41	ти	си	ти	своимоу
42	прѣвъꙁлюблєнамъ	въꙁлюблєнъ	прѣвъꙁлюблєнъ	въжєлѣнънъ
	1.1.1.10-1.1.1.12	2.1.1.21-2.2.1.3	2.2.1.4-2.2.1.6	1.2.1.21-1.2.1.23

Комментарии

1.1.1.10 Помꙑшлѧнини :	РНБ, Соф. 203 (МСМ) 2.1.1.21 Помꙑшлєниѥ ГИМ, Син. 166 (МСМ) 2.2.1.4 Помꙑшлєнии НБС, Деч. 41 (МСМС) 1.2.1.21 мꙑслꙑ
1.1.1.10 си :	ГИМ, Син. 166 (МСМ) 2.2.1.4 ти Разные лексемы: Словоформа-Словоформа НБС, Деч. 41 (МСМС) 1.2.1.21 Твои
1.1.1.10 ꙁьрѣтєльно :	РНБ, Соф. 203 (МСМ) 2.1.1.21 нєꙁьритєльно НБС, Деч. 41 (МСМС) 1.2.1.21 ꙁритєлъноє
1.1.1.10 прѣмѫдрє :	РНБ, Соф. 203 (МСМ) 2.1.1.21 пмоудрє ГИМ, Син. 166 (МСМ) 2.2.1.4 прѣмоудрє НБС, Деч. 41 (МСМС) 1.2.1.21 моудрⷭ҇
1.1.1.10 оуньшь :	ГИМ, Син. 166 (МСМ) 2.2.1.4 ѡуньшь НБС, Деч. 41 (МСМС) 1.2.1.21 ѡуньшъ
1.1.1.11 отъ скврьнъ :	РНБ, Соф. 203 (МСМ) 2.1.1.22 отъ скврьни ГИМ, Син. 166 (МСМ) 2.2.1.5 ѡтъ скврьнъ НБС, Деч. 41 (МСМС) 1.2.1.21 ѿ скврънъ
1.1.1.11 тѣлєснъ :	РНБ, Соф. 203 (МСМ) 2.2.1.1 тѣлєсьнъ ГИМ, Син. 166 (МСМ) 2.2.1.5 тѣлѣснъ

Fig. 6. Vertical Text and Variant Readings

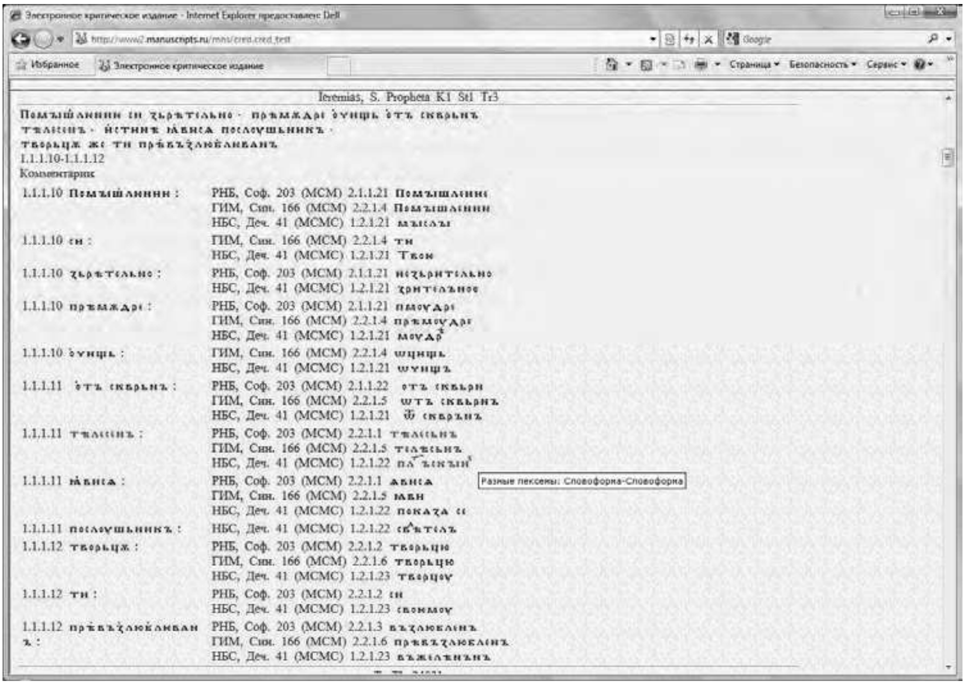
4.8.6. Let us give an example of the page structure for the mode “On the level of the linguistic units (the form and value of the variant readings)”:

Abbreviation_of manuscript1 (M1)
Abbreviation_of fragment1
Text_of fragment1
address_of fragment1
address_of unit2_M1 unit2_M1: M2 address_of unit_M2 unit_M2
M3 address_of unit_M3 unit_M3 ...
address_of unit4_M1 unit4_M1: M2 address_of unit_M2 unit_M2
M3 address_of unit_M3 unit_M3 ...
...
Abbreviation_of fragment2
Text_of fragment_2

address_of fragment_2
address_of unit2_M1 unit_M2:   M2 address_of unit_M2 unit_M2
                               M3 address_of unit_M3 unit_M3 ...
address_of unit4_M1 unit_M4:   M2 address_of unit_M2 unit_M2
                               M3 address_of unit_M3 unit_M3 ...
...
M1 - M2, M3 (all fragments of the retrieval)
unit_M2 address_of unit2_M1:   M2 address_of unit_M2 unit_M2
                               M3 address_of unit_M3 unit_M3 ...
unit_M4 address_of unit_M1:   M2 address_of unit_M2 unit_M2
                               M3 address_of unit_M3 unit_M3 ...

This demonstration form shows only the discrepancies; the complete coincidences are not shown. The form enables highlighting and fixation of the corresponding elements of the retrieval – the correlated units in the manuscripts, and the correlated units in the comment. All the word forms of the retrieval of the last chapter are arranged alphabetically

The settings enable also to get the materials required for the analysis of the correlation between the Slavonic translation and Greek original.


Электронное критическое издание - Internet Explorer предоставлен Dell
http://www2.manuscripts.ru/mns/cred.cred_test
Google
Избранное
Электронное критическое издание
Страница
Безопасность
Сервис

Ieremias, S. Propheta K1 Stl Tr3

Помꙑшлании си зьрѣтельно · прѣмѫдрѣ ѹчищь ѿтъ скврьнъ тѣлесинъ · истинѫ ꙗвиса послоушьникъ · творьцѫ же ти прѣвъзлюбленаанъ
1.1.1.10-1.1.1.12
Комментарии

1.1.1.10 Помꙑшлании :	РНБ, Соф. 203 (МСМ) 2.1.1.21 Помꙑшлѣние ГИМ, Син. 166 (МСМ) 2.2.1.4 Помꙑшлении НБС, Деч. 41 (МСМС) 1.2.1.21 мꙑслꙑ
1.1.1.10 си :	ГИМ, Син. 166 (МСМ) 2.2.1.4 ти НБС, Деч. 41 (МСМС) 1.2.1.21 Твои
1.1.1.10 зьрѣтельно :	РНБ, Соф. 203 (МСМ) 2.1.1.21 незьрительно НБС, Деч. 41 (МСМС) 1.2.1.21 зрителное
1.1.1.10 прѣмѫдрѣ :	РНБ, Соф. 203 (МСМ) 2.1.1.21 пмоудрѣ ГИМ, Син. 166 (МСМ) 2.2.1.4 прѣмоудрѣ НБС, Деч. 41 (МСМС) 1.2.1.21 моудр
1.1.1.10 ѹчищь :	ГИМ, Син. 166 (МСМ) 2.2.1.4 ѡчищь НБС, Деч. 41 (МСМС) 1.2.1.21 ѡучищъ
1.1.1.11 ѿтъ скврьнъ :	РНБ, Соф. 203 (МСМ) 2.1.1.22 отъ скврьни ГИМ, Син. 166 (МСМ) 2.2.1.5 ѡтъ скврьнъ НБС, Деч. 41 (МСМС) 1.2.1.21 ѿ скврънъ
1.1.1.11 тѣлесинъ :	РНБ, Соф. 203 (МСМ) 2.2.1.1 тѣлесьнъ ГИМ, Син. 166 (МСМ) 2.2.1.5 тѣлѣсьнъ НБС, Деч. 41 (МСМС) 1.2.1.22 плътьскꙑи
1.1.1.11 ꙗвиса :	РНБ, Соф. 203 (МСМ) 2.2.1.1 ависа Разные лексемы: Словоформа-Словоформа ГИМ, Син. 166 (МСМ) 2.2.1.5 ꙗви НБС, Деч. 41 (МСМС) 1.2.1.22 показа се
1.1.1.11 послоушьникъ :	НБС, Деч. 41 (МСМС) 1.2.1.22 свѣтелъ
1.1.1.12 творьцѫ :	РНБ, Соф. 203 (МСМ) 2.2.1.2 творьцю ГИМ, Син. 166 (МСМ) 2.2.1.6 творьцю НБС, Деч. 41 (МСМС) 1.2.1.23 творцоу
1.1.1.12 ти :	РНБ, Соф. 203 (МСМ) 2.2.1.2 си НБС, Деч. 41 (МСМС) 1.2.1.23 своимоу
1.1.1.12 прѣвъзлюбленаанъ :	РНБ, Соф. 203 (МСМ) 2.2.1.3 възлюбленъ ГИМ, Син. 166 (МСМ) 2.2.1.6 прѣвъзлюбленъ НБС, Деч. 41 (МСМС) 1.2.1.23 въжделѣнънъ

Fig. 7. Fragment and Variant Readings

4.9. Thus, the critical edition prepared in the framework of the project "Manuscript" gives to the reader the opportunities of:

- getting acquainted with the copies of the Slavonic Service Menaion for May and the Greek text of the Menaion,
- getting acquainted with their structure and composition,

- obtaining information on the correspondences between the structures and linguistic units of the manuscripts,
- obtaining information on the variant readings between the manuscripts,
- getting material for analysis of the correlations between the Greek and Slavonic texts, i.e. about everything that is represented in the printed edition.

At the same time the electronic critical edition makes it possible for the user to form a page of the electronic edition independently by selecting the required fragments of the texts or the pages of the manuscripts, arranging the fragments or word forms in the required order and choosing the required form of data visualization.

Finally, the user has a possibility of adjusting representation of the materials chosen for comparison so that to be able to solve a wide circle of problems associated with the text history and linguistic peculiarities of its copies.

ACKNOWLEDGEMENTS

The creation of the medieval collections of the system “Manuscript” is supported by the Russian Foundation for Basic Research (project # 09-06-00298a); the corpus of works by M. Lomonosov was created with the support of the Russian Humanities Foundation (project # 07-04-12147в); research on the historical-linguistic corpus is supported by the Russian Humanities Foundation (project # 07-04-00369a).

REFERENCES

[1] BARANOV, V., AND GNUTIKOV, R. 2006. Redaktor OldEd kak specializirovannyj instrument dlya redaktirovaniya dokumentov v baze dannych «Manuscript» (Editor OldEd as a specialized tool for editing documents in the database "Manuscript"). In *Sovremennyje informatsionnyje tekhnologii i pis'mennoje nasledije*: Ot drevnikh rukopisej k elektronnym tekstam. 2006. (Modern Informational Technologies and Written Heritage: From Ancient Manuscripts to Electronic Texts). Proceedings of the International Conference. Izhevsk, Russia, June 2006, 43-46.

[2] VICTOR A. BARANOV. 2007. The ideology and technology of creating online full-text digital collections of ancient and medieval Slavonic manuscripts. In *International Conference on Applied Natural Sciences*, Trnava, November 2007, 199-207. ISBN 978-80-89220-91-5. Trnava, UCM, 2007.

[3] BARANOV, V., AND GNUTIKOV, R. 2007. Up-to-date means of access to full-text databases. In *Proceedings of the* 19th Joint International Conference of the Association for Computers and the Humanities, and the Association for Literary and Linguistic Computing, Urbana-Champaign, USA, June 2007, 74–76. (Electronic resource: http://www.digitalhumanities.org/dh2007/abstracts/xhtml.xq?id=199).

[4] VICTOR A. BARANOV. 2008. Polnotekstovye bazy dannyx kak osnova dlya elektronnych izdanij srednevekovych rukopisej v Internete: trebovaniya, realizaciya, perspektivy (Full-text database as the basis for electronic editions of medieval manuscripts in the Internet: Requirements, Implementation, Prospects). In *Scripta & e-Scripta.* The Journal of Interdisciplinary Mediaeval Studies. Vol. 6. Sofia, 2008, 47-64, 422. ISSN 1312-238X.

[5] BARANOV, V. 2008. Proekt «Manuscript»: predvaritel'nye itogi (Project "Manuscript": preliminary results). In *Sovremennyje informatsionnyje tekhnologii i pis'mennoje nasledije*: Ot drevnikh tekstov k elektronnym bibliotekam. 2008. (Modern Informational Technologies and Written Heritage: From Ancient Texts to Electronic Libraries). Proceedings of the International Conference. Kazan, Russia, August 2008, 32-36.

[6] BARANOV, V., AND GNUTIKOV, R. 2008. Elektronnoe kriticheskoe izdanie srednevekovogo teksta: postanovka zadachi, osnovnye trebovaniya i instrumental'naya podgotovka (Electronic critical edition of a medieval text: statement of the problem, the basic requirements and tools). In *Sovremennyje informatsionnyje tekhnologii i pis'mennoje nasledije*: Ot drevnikh tekstov k elektronnym bibliotekam. 2008. (Modern Informational Technologies and Written Heritage: From Ancient Texts to Electronic Libraries). Proceedings of the International Conference. Kazan, Russia, August 2008, 36-44.

[7] VICTOR A. BARANOV, ALEKSEY N. MIRONOV, ALEKSEY N. LAPIN, IRINA S. MELNIKOVA, ANASTASIYA A. SOKOLOVA. 2008. Development of the Processing and Visualization Technologies for the Linguistic Information in the Manuscript System: Lemmatization. In *JADT 2008*: actes des 9es Journées internationales d’Analyse statistique des Données Textuelles: Proceedings of 9th International Conference on Textual Data statistical Analysis, Lyon, March 2008, 137-145. ISBN: 978-2-7297-0810-8 (pb.).

[8] BARANOV, V.A., ANIKINA R.A., KOKORINA T.V., OSHCHEPKOV, S.V., AND SOKOLOVA, A.A. 2008. Metainformaciya v kollekcii M. V. Lomonosova na portale «Manuscript: Slavyanskoe pis'mennoe

nasledie» (Meta-information in the collection of Lomonosov on the portal "Manuscript: Slavonic Written Heritage). In *Sovremennyje informatsionnyje tekhnologii i pis'mennoje nasledije*: Ot drevnikh tekstov k elektronnym bibliotekam. 2008. (Modern Informational Technologies and Written Heritage: From Ancient Texts to Electronic Libraries). Proceedings of the International Conference. Kazan, Russia, August 2008, 23-27.

[9] BARANOV, V.A., VOTINTSEV, A.A., VOTINTSEV, P.A., AND SOLOMENNIKOV, I.S. 2008. Internet-sredstva poiska i vizualizacii dannyx dlya lingvisticheskogo analiza informacionno-analiticheskoj sistemy "Manuscript" (Internet search tools and visualization of data for linguistic analysis of Information-analytical system "Manuscript"). In *Sovremennyje informatsionnyje tekhnologii i pis'mennoje nasledije*: Ot drevnikh tekstoc k elektronnym bibliotekam. 2008. (Modern Informational Technologies and Written Heritage: From Ancient Texts to Electronic Libraries). Proceedings of the International Conference. Kazan, Russia, August 2008, 64-68.

[10] LOMONOSOV, M. V. 1950-1083. Polnoe sobranie sochinenij (Complete Works). Moscow, Leningrad, 1950–1983. Vol. 1-11.

Victor Baranov
Department of Linguistics,
Izhevsk State Technical University,
Izhevsk, Studencheskaya Str. 6, Izhevsk, 426069, Russian Federation
victor.a.baranov@gmail.com

Received September 2009