The GORAZD Project: An Old Church Slavonic Digital Hub

English edition

The Czech Academy of Sciences

SLOVANSKÝ ÚSTAV

Science 76
Around us
What is it...

The **Institute of Slavonic Studies of the Academy of Sciences of the Czech Republic, v. v. i.**, was founded at the initiative of President T. G. Masaryk through a law passed by the Parliament of the Czechoslovak Republic on January 25th, 1922, although due to various bureaucratic barriers, it could not begin its activities until 1928. It focused on two fields: social sciences and the national economy. The renowned archaeologist and ethnographer Lubor Niederle was elected its first chairman, and Lobkowicz Palace in Prague's Malá Strana district became its headquarters (this building today houses the German Embassy). The Institute's aim was to cultivate all the fields of Slavic research, from archaeology and history to philology and linguistics. The Institute quickly achieved international renown; however, the German occupation was at first crippling for its activities, and then brought its complete closure in 1943. After liberation, the Institute of Slavonic Studies was reinstated as a scholarly institution, and in 1953, it was integrated into the newly founded Czechoslovak Academy of Sciences, including the Brno branch founded one year earlier. The Institute of Slavonic Studies took over – besides other tasks – the already ongoing work on composing the *Old Church Slavonic dictionary*. Yet despite the Institute's unquestionable research successes, in 1963 it was once again shut down, with the excuse of "lingering Masarykism," and its staff was scattered across other academic institutions. Shortly after the Velvet Revolution, in 1992, it was renewed thanks to the enthusiastic efforts of certain of its former employees. After this renewal, it was initially a joint site of the Czech Academy of Sciences (as an autonomous unit of the archive) and the Faculty of Arts of Charles University. However, this arrangement did not prove itself fruitful, and so as of January 1st, 1998, the Institute of Slavonic Studies was definitively established as one of the independent institutes of the Czech Academy of Sciences.

At present, the activities of the Institute of Slavonic Studies of the Czech Academy of Sciences, v. v. i., center upon research into Slavic languages and literature in the European cultural context. The Institute works from a traditional comparative conception, comparing individual languages and literatures among each other, and it enriches this view with intercultural and interdisciplinary topics stretching out beyond the Slavic language region towards other European cultures. It publishes three peer--reviewed international scholarly periodicals: *Slavia* (since 1921), *Byzantinoslavica* (since 1929), and *Germanoslavica* (since 1931). It has three divisions, which reflect its three basic areas of Slavic studies: diachronous linguistics, synchronous linguistics, and literary studies. The Institute is specific for certain subdisciplines that are cultivated here alone. Its research infrastructure, meanwhile, is supplemented by the publicly accessible Library of the Institute of Slavonic Studies, counting nearly 80,000 pieces of the field's literature, and a number of domestic and foreign scholarly magazines.

The **Department of Old Slavonic and Byzantine Studies** is dedicated to the study of the Old Church Slavonic and Church Slavonic languages and literature; the Byzantine section, unique nationwide, is mainly oriented towards research into the history of the Byzantine Empire and Great Moravia, and into Byzantine-Slavic relations. Past work by this department includes the unique four-volume *Old Church Slavonic Dictionary* (*Slovník jazyka staroslověnského*), for which the *Supplements to Volume I of the Dictionary of Old Church Slavonic* (*Dodatky k I. dílu Slovníku jazyka staroslověnského*) were completed in 2016, and the *Old Church Slavonic Dictionary According to the Manuscripts of the 10th–11th Centuries* (*Staroslavjanskij slovar' po rukopisjam X–XI vekov*), whose second edition is currently in preparation. At present, work on the *Greek-Old Church Slavonic Index* is also ongoing, and since 2016, the department's activities linked to the *GORAZD: An Old Church Slavonic Digital Hub* project as well. This project will lay the foundations for digitally processing and publishing material from this field. We also cannot omit the department's publishing activities; let us mention for example the two-volume *Forty Gospel Homilies of Gregory the Great* (*Čtyřicet homilií Řehoře Velikého na evangelia*; 2005, 2006), the largest known Church Slavonic relic of Czech origin. Out of its activities within Byzantine studies, we

Logo of the *GORAZD Project:
An Old Church Slavonic
Digital Hub*

Project's profile on Facebook:
www.facebook.com/projekt.gorazd

The GORAZD Project: An Old Church Slavonic Digital Hub

In the last several decades, providing translation dictionaries in digital form has become standard practice. The advantages of this approach are clear – a digital dictionary (no matter how high its headword count) is available to its users in every time and place, and without the need for the often arduous manipulation of bulky, heavy printed dictionaries. Moreover it lets one nearly instantly find a desired headword – and often other dictionary material as well. Such a dictionary is also better for authors: its contents are easy to update, and so the authors do not have to waste time, money, and nerves on preparing new printed editions. This digitization trend has also begun to affect dictionaries of historical languages. However, it comes with numerous pitfalls. For example, many such languages used unusual alphabets or characters that were only fully incorporated into internationally accepted digital standards recently, if at all. Not long ago, this problem still also existed for Old Church Slavonic. A number of mutually incompatible fonts were used for digitally expressing it, and generally one could not read Old Church Slavonic texts at all without installing these fonts. Without a universal standard, it was also significantly harder to develop advanced software capable of working with Old Church Slavonic texts, as software was also dependent on the character distribution within a given font. The situation was not made any simpler by the fact that Old Church Slavonic has two different alphabets: Glagolitic and Cyrillic. The Unicode digital character encoding standard did not contain most of the Cyrillic and Glagolitic characters needed for recording Old Church Slavonic texts until Unicode 8.0, released in June of 2015. Thanks to this standard, Old Church Slavonic can now be recorded in a font-independent way. This has opened the door to systematic, internationally standardized digital processing of Old Church Slavonic materials.

The Department of Old Church Slavonic and Byzantine Studies at the Institute of Slavonic Studies of the Czech Academy of Sciences has thus been able to commence work on its long-intended project for digitizing the Old Church Slavonic dictionaries that it has produced or is producing. The result is the project named **GORAZD: An Old Church Slavonic Digital Hub**, which has been financially

Josef Kurz (1901–1972), the first main editor of the *OCSD*

supported for 2016–2020 by the Ministry of Culture of the Czech Republic's NAKI II program (project number DG16P02H024). Within this project, the *Old Church Slavonic Dictionary* (*Slovník jazyka staroslověnského*, volume I–IV, 1966–1997, *OCSD* below) will be digitized, and it will directly incorporate material from *Supplements and Corrections to Volume I of the Old Church Slavonic Dictionary* (*Slovník jazyka staroslověnského V: Dodatky a opravy k I. dílu*, 2016, *Supplements to Volume I* below). Also to be digitized is the *Dictionary of the Oldest Old Church Slavonic Manuscripts* (*Slovník nejstarších staroslověnských památek*; *DOOCSM* below), on the basis of the prepared second expanded and supplemented edition (the first edition was published in 1994 in Moscow under the name *Staroslavjanskij slovar' po rukopisjam X–XI vekov*, reprint 1999). By the time of printing of the second edition, a sample portion will be publicly available, with a scope covering the letter ⰰ. For the *Greek-Old Church Slavonic Index* (*Řecko-staroslověnský index*, *GOCSI* below), which is still in preparation, initially its Volume I (2014) will be made publicly available. Its other volumes will be added after their completion. *GORAZD* also includes the scanning and publishing of a unique Old Church Slavonic sheet-based card index with roughly one million sheets. This index formed the foundation for

both of the above-mentioned Old Church Slavonic dictionaries, and its digital archival will serve to guarantee its preservation for future generations of scholars.

This project's software solution

The texts of the *Dictionary of the Oldest Old Church Slavonic Manuscripts,* the *Greek-Old Church Slavonic Index Volume I*, and *Supplements to Volume I of the Old Church Slavonic Dictionary* were already accessible to the *GORAZD* project team in digital form, but not in a Unicode-compatible form, and thus they had to be converted. Moreover, the text for all four volumes of the *Old Church Slavonic Dictionary* existed in printed form only. It was not feasible to manually rewrite this extensive text in full, as that would have required an enormous investment of labor and time. Therefore, the implementation team chose to proceed based on optical character recognition (OCR), which was a pioneering step to take in relation to a multilingual dictionary utilizing four alphabets (of which two are historical) and three languages (of which three are historical). No specifications existed for the automatic recognition of Old Church Slavonic and Old Greek, and thus the implementation team had to newly create such specifications. Thus one output of the project is the *Tools for Old Church Slavonic Optical Character Recognition*, designed for use within *ABBYY FineReader 12*. These tools include a definition for Old Church Slavonic as the user language and a database-style Old Church Slavonic dictionary, based on canonical and non-canonical texts, for better recognition. These tools have enabled us to achieve satisfactory automatic recognition, allowing us to forge ahead with the OCR-based approach. The tools will be made freely available to the scholarly community after the project ends.

The open-source digital-library solution INVENIO, developed by the European Society for Nuclear Research (CERN), was selected for managing the database system in which the individual dictionary databases and card indexes are stored. This system is being adapted for the needs of *GORAZD* by AiP Beroun, s. r. o., a company with rich INVENIO experience. The Gulliver system, developed by that same company, will be used to provide web access to the databases' contents and to run their search tools. The implementation team is developing three specialized programs for converting the dictionary texts into a structured, specialized database and for working with that database effectively. The first of them is the *Headword Generator*. It can use defined rules to convert unstructured text output from OCR or a text file into individual headwords with various depths of structuration. The outputs are XML files containing individual headwords, which are then imported into INVENIO. From there, they can be processed using another of the team's software outputs – the *Headword Editor*, which works as an extension to INVENIO. Within this editor, users can both display image documents and edit XML files' content using specialized tools. This makes the Headword Editor an ideal environment for proofreading dictionary text obtained from OCR or editing and updating already-processed headwords. As the Headword Editor is a web application, it is accessible to users anywhere and on any platform, with no need for them to install any software. The implementers' third application is the *Export Module*, through which users can export selected records from the INVENIO environment as continuous,

formatted HTML text that can be opened and edited in ordinary text editors (e.g. Microsoft Word). This kind of output is very useful for preparing printed editions of digitized dictionaries. Once the development of these three applications is complete, their source code will be made available to interested experts as building blocks that can be used in solutions for future similar projects. Thus after some adaptation, they will be a boon for the digitization of a historical or encyclopedic dictionary or monolingual explanatory dictionary of any kind.

The history of the *Old Church Slavonic Dictionary*

The first efforts to capture the vocabulary of Old Church Slavonic texts in the form of a dictionary appeared quite early: at the beginning of scholarly research into Old Church Slavonic, in the first half of the 19th century. Among the more serious of such efforts we can above all cite the dictionaries of A. K. Vostokov (1781–1864), F. Miklošič (1813–1891), and I. I. Sreznevsky (1812–1880), which – despite their individual imperfections – represent an important information source that is frequently used even today. Meanwhile the beginnings of the specifically Czech lexicography of Old Church Slavonic are linked to the names of important national personalities in Slavic

Zoe Hauptová (1929–2012), *OCSD's* main editor from 1973 and until her death, the main editor of the *Supplements to Volume I of the Old Church Slavonic Dictionary*

studies. One of these was professor Václav Vondrák (1859–1925), who was preparing an Old Church Slavonic dictionary focused on evangelical texts, but whose work was cut short by his death; another was Miloš Weingart (1890–1939). Weingart and his students prepared a selection of excerpts from the Old Church Slavonic texts at Charles University, working from 1926 up until his death at the outbreak of World War II.

The year 1943 was a milestone year for Czech and international Old Church Slavonic lexicography, as in this year, the Commission for the *Old Church Slavonic Dictionary* (*OCSD*) was founded under the chairmanship of Bohuslav Havránek (1893–1978). Its purpose was to create a design and methodology for that which it would name the *Old Church Slavonic Dictionary*, and to oversee its preparation. This shared work and responsibility were intended to ensure that work on the dictionary would continue even if its authors were to shrink in number.

The commission's first step was to define the range of manuscripts that would be included in the dictionary. Soon afterwards, the actual excerpting of lexical material began; that material then formed the foundation for a sheet-based Old Church Slavonic card index. This large index, unique worldwide, contains roughly 1.6 million sheets and represents the total vocabulary provably contained in the excerpted texts, sorted both by Old Church Slavonic words and by their Greek and Latin equivalents.

Josef Kurz (1901–1972) was named as the dictionary's main editor. In 1953, editorship was transferred to the Institute of Slavonic Studies of the Czechoslovak Academy of Sciences, and simultaneously the processing of headwords was begun. Three years later these efforts bore fruit: the publishing of the *OCSD* sample book, which was very well received among researchers. The *OCSD* was then published in parts (a total of 52 during more than 40 years) that were, over the course of the work, finalized into today's four volumes (а–ħ 1966, к–ѻ 1973, п–ρ 1982, с–ѵ 1997). After the death of J. Kurz, Zoe Hauptová (1929–2012) became the main editor of the *OCSD*. Under her leadership, the dictionary's third and fourth volumes were completed, successfully concluding the work of two generations of Czech scholars of Old Church Slavonic.

The *OCSD*'s importance for international research is attested to by – among other things – the reprinting of the entire work in 2006 in St. Petersburg. However, just as with the original edition, this print run is by now sold out entirely. The *OCSD* has also served as the methodological foundation for Church Slavonic dictionaries produced abroad: *The Church Slavonic Dictionary of Croatian Redaction* (*Rječnik crkvenoslavenskoga jezika hrvatske redakcije*, Zagreb 1991–) and *The Church Slavonic Dictionary of Macedonian Redaction* (*Rečnik na crkovnoslovenskiot jazik od makedonska redakcija*, Skopje 2000–), as well as one work produced through international cooperation, the *Comparative Index for the Dictionaries Written in the Framework of the International Committee of Slavonic Scholars' Commission for Church Slavonic Dictionaries* (*Sporedben indeks kon rečnicite obrabotuvani vo ramkite na komisijata za crkvoslovenski rečnici pri MKS,* ed. Z. Ribarova, Skopje 2015–). The *OCSD* also served as the material foundation for the *Etymological Dictionary of Old Church Slavonic Language* (*Etymologický slovník jazyka staroslověnského*), whose still-ongoing publishing by the Etymological Department of the Institute of the Czech Language of the Czech Academy of Sciences in Brno began in 1989.

The Old Church Slavonic Dictionary
in four volumes
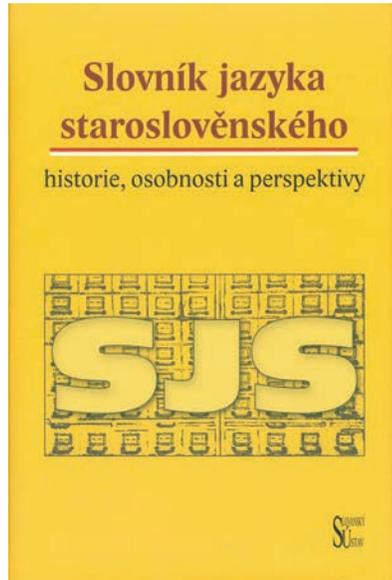
The Old Church Slavonic Dictionary:
History, People, and Perspectives
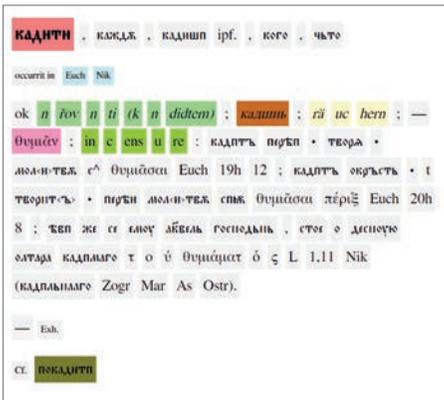
◄ Example of the headword **кадити**
from *OCSD* after the processing by the OCR
technology and the Headword Generator

▼ Example of the same headword after
the corrections made in the Headword
Editor
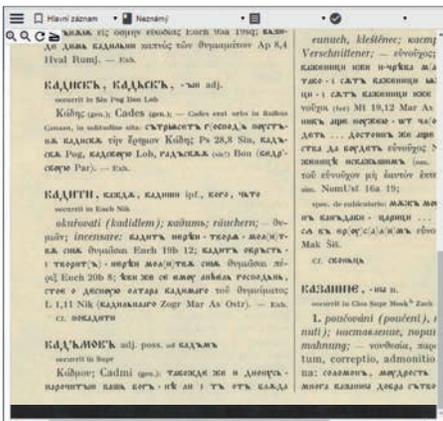
We must note that throughout its creation this dictionary struggled with the adversities of the Communist era, and especially during the post-1968 crackdown, its team had to stand up to pressures and ill will from many corners. A detailed description of this troubled history of Czech Old Church Slavonic lexicography, along with brief biographies of significant persons connected with the *OCSD*, was incorporated into the large publication *The Old Church Slavonic Dictionary. History, People, and Perspectives* (*Slovník jazyka staroslověnského: historie, osobnosti a perspektivy*), published by the Institute of Slavonic Studies in 2016. This publication also offers more detailed information on the other lexicographic works that are connected with the *OCSD*.

The Old Church Slavonic Dictionary and its entry into the digital era

The *Old Church Slavonic Dictionary* is currently the largest dictionaries of the oldest written Slavic language. Its four published volumes contain all the vocabulary of the Old Church Slavonic texts that are reflected within it. These texts can be divided by origin into three main groups:
- the "canonical" Old Church Slavonic texts preserved in the oldest manuscripts from the 10th and 11th centuries,
- the texts that, while preserved via newer transcriptions, also have their actual origins dated to the Cyril and Methodius period of Old Church Slavonic literature,
- texts of Czech Old Church Slavonic origin that in all likelihood were written on Czech territory before the expulsion of Slavic monks from the Sázava monastery in 1096, and that are mostly documented via newer copies – Russian or occasionally South Slavic.

In terms of genre, these are primarily Biblical texts (the entire New Testament, a psalter from the Old Testament, and other extracts contained primarily in what is called a "prophetologion"), liturgical texts (based on the Byzantine and Roman rites), sermons, theological treatises, legends (the lives of the saints), and collections of church law. We will also note that most of the texts reflected are translations of Greek or Latin originals, although a small portion of them also comes from Old High German originals. Meanwhile original texts (e.g. the poem Proglas, or the legends of Cyril and Methodius) also comprise a quite small, yet historically significant, portion of the texts.

The dictionary is thus exceptional in its source-material selection. Yet it is also – and above all – exceptional in its composition. Each of its headwords is introduced with a heading, whose form is normalized based on Old Church Slavonic practice. A complete list of the texts in which the given word is found is provided under the heading by way of established abbreviations. The reader can thus quickly evaluate how widely it is used and roughly place it in time. One major plus of the *OCSD* is that it is multilingual: this is exceptional even within historical lexicography. Besides the corresponding phrase from the language of the source text, the reader

Václav Konzal (*1931), a long-termed scientific secretary of the *OCSD's* editorial board, later one of the main authors of the *Supplements to Volume I of the OCSD*

finds here a translation of the word and its individual sub-meanings into Czech, Russian, and German. If no Latin precursor or parallel is known for the given word, a translation into Latin is given as well. One important and often invaluable aid is the listing of synonyms at the end of each headword. Representative listings of examples from among the source texts are presented including sufficient context and, for those texts that are translations, the texts of their originals as well. Headwords citing an exhaustive list of the extant occurrences of a given word are marked with the abbreviation exh. One important aspect of the *OCSD* is that it is encyclopedic and explains headwords' meanings; this is primarily beneficial for descriptions of liturgical terminology and for proper nouns – most often including the names of persons, cities, regions, and ethnicities. The *OCSD* can thus to some extent also support the study of biblical or liturgical real-world contexts.

The digitization of the *OCSD* is beneficial mainly in that it will be accessible to researchers worldwide and that revising and correcting it and supplementing its content will become easier. And indeed, the headwords from the *OCSD's* first volume will see significant reconstruction already during its digitization; this volume will be updated based on the *Supplements to Volume I*. The dictionary will be made more accessible for the international public by changing the language of its explanations from the original Latin to English; moreover, these English equivalents will newly be supplemented with translations of individual headwords' meanings into other modern languages.

The advanced search functions will bring an important change for users vis-à-vis printed dictionaries. Besides ordinary searches by heading, searches by equivalents

in modern languages (Czech, English, German, and Russian) and three historical languages (Greek, Latin, and Old High German) will also be enabled. Users will thus also be able to use the digital *OCSD* for translating between Old Church Slavonic and Czech, English, etc., which in itself opens up a number of possibilities for its use by not only researchers, but also the public. Users will likewise be able to filter search results by other criteria – more specifically, by selected grammatical properties (e.g. part of speech, declension type, or conjugation type) and individual source texts or groups of them. Full-text search will naturally be included; among other things, this enables searching by a word's specifics forms within works. The scope and combinability of the search criteria will make the *OCSD* much more convenient than printed dictionaries, and thus it will offer new perspectives on Old Church Slavonic vocabulary.

The provision of the digitized *OCSD* over a web interface and the other adjustments made to simplify and greatly broaden its use make this dictionary one the most important global feats in electronic historical lexicography. Thus it offers not only researchers, but also the public at large, easy access to outputs from several generations of Czech scholars of Old Church Slavonic and other languages. Yet while the digital *OCSD* will *be* a strengthening of this globally respected tradition, it will *become* much more: a living platform for the further development, expansion, and supplementation of the data it stores.

Supplements to Volume I of the Old Church Slavonic Dictionary

Over forty years elapsed between the publication of the first and last volume of the *OCSD*, and two generations of researchers took part in its preparation. It is thus entirely natural that the methodology for its drafting changed (e.g. in the shaping of its headwords and in the material presented) during the course of work upon it. The selection of Old Church Slavonic texts upon whose basis it was prepared changed as well – new manuscripts were discovered over the years, and certain other texts were newly published – and we also must mention the new discoveries of Greek and Latin originals whose materials could not be utilized at the *OCSD*'s time of publishing. Thus over time the need for an overall revision of the dictionary and its method of preparation began to manifest; this especially applies for its Volume I, within which the methodological heterogeneity is the most visible. Soon after the *OCSD*'s completion in 1997, a possibility appeared for publishing the entire dictionary digitally; it would have taken the form of cooperation between the Institute of Slavonic Studies and the LEDA publishing house. This proposal anticipated the supplementation of further materials, as well as further adjustments to the dictionary; e.g. English translations for the meanings were to be filled in. However, this was not yet the right moment for a new, revised edition of the *OCSD*: there was a lack of time, finances, and people to do the job. Thus the overall revision of Volume I of the *OCSD* had to wait until 2008–2016. It was revised in these years in the framework of the *Supplements to Volume I of the Old Church Slavonic Dictionary* project (*Dodatky k I. dílu Slovníku jazyka staroslověnského*).

Example of the headword **бєспрѣстани** from the *OCSD's* original edition

Example of the innovations, brought to the same headword by the *Supplements*

Example of the same headword in the *GORAZD* interface after the processing of the data from the *Supplements* and the addition of the English language

The main goal of the *Supplements to Volume I* was thus to supplement the work's first volume based on newly obtained material and unify its drafting methodology with that of later volumes. Therefore the *Supplements to Volume I* provide supplementations and corrections to the existing headwords in the *OCSD*'s Volume I; at the same time, new Old Church Slavonic headwords based on previously untapped sources in the Old Church Slavonic card index that had not been included in the *OCSD*'s Volume I for a variety of reasons (new discoveries of Old Church Slavonic manuscripts, revised excerpts from original source documents, etc.) are also processed here. The authorial team published eight parts totaling 504 pages, whose progressive numbering (parts 53–60) ties into the original dictionary, and thus the *Supplements to Volume I* overall form an integral part of the *OCSD* as its Volume V: *Old Church Slavonic Dictionary V: Supplements and Corrections to Volume I* (*Slovník jazyka staroslověnského V: Dodatky a opravy k I. dílu*), eds. Z. Hauptová, V. Konzal, Š. Pilát, Prague 2016.

Within the *GORAZD: An Old Church Slavonic Digital Hub* project, the material from the *Supplements to Volume I* is interspersed directly within the structure of the electronically processed *OCSD* headwords, thus enriching the original headwords through the newly processed manuscripts – they are supplemented with material they previously did not reflect – while simultaneously unifying the methodology of their processing. In selected cases other, smaller formal adjustments have also been made to the headwords (to make the explanations more logical or the presented material more complete). English translations for meanings and explanations have been added to the headwords as well. Thus the incorporation of the *Supplements to Volume I* is directly fulfilling the project's goal of keeping the contents of the *Old Church Slavonic Dictionary* up to date with the current scholarly delineation of Old Church Slavonic vocabulary.

The Dictionary of the Oldest Old Church Slavonic Manuscripts

The idea of producing a single-volume dictionary focused exclusively on the vocabulary of the oldest, so-called "canonical" Old Church Slavonic texts has been on the collective minds of Old Church Slavonic scholars in Prague since this community finished producing excerpts for the *OCSD*'s card index in the mid-20th century. However, because all of its available manpower at that time was devoted to preparing the *OCSD*, it was not until the Institute of Slavonic Studies of the Russian Academy of Sciences offered its cooperation at the end of the 1950s that the impulse existed for actually preparing such a dictionary. This was followed by the difficult work of defining its conception, and then the broader work needed to bridge past the historical events of 1968; the dictionary's preparation thus only truly began in 1973. Emilie Bláhová and Radoslav Večerka were named as the editors for the Czech side, while Ralya Mikhailovna Tseitlin was named for the Russian side. The dictionary's manuscript, with a total of 10,000 headwords, was completed and sent to press in 1988, but it was not published until 1994, in Moscow, under the name *Staroslavjanskij slovar' (po rukopisjam X–XI vekov) (Old Church Slavonic*

Ralya Mikhailovna Tseitlin (1920–2001), a Russian linguist and co-editor of the *Old Church Slavonic Dictionary According to the Manuscripts of the 10th–11th Centuries*

Dictionary According to the Manuscripts of the 10th–11th centuries). Because this dictionary was sold out shortly after it was published, it saw an unaltered reprint in 1999. For the purposes of the *GORAZD* project, this dictionary was given a more accurate Czech name that equates to *Dictionary of the Oldest Old Church Slavonic* (*Slovník nejstarších staroslověnských památek*).

From the beginnings of work upon it, the *DOOCSM* was intended as a dictionary primarily for study purposes that was, through its simplified structure, to offer a basic orientation in the oldest Old Church Slavonic vocabulary, and it thus is a more accessible alternative for ordinary users than the monumental *OCSD*. The explanatory material within it is minimized and is primarily used for the more complex prepositions and verbs. The meanings of its headwords are given in Czech and Russian; Greek and Latin equivalents are only listed as a summary in headwords' headings. Unlike the much more broadly conceived *OCSD*, the *DOOCSM* could show the number of documents pertaining to its individual headwords, and so users can easily deduce which words appear often in the texts of canonical Old Church Slavonic, and which infrequently. For study purposes, detailed morphological tables with declension samples for all flective parts of speech were added to this dictionary.

The *DOOCSM* is based on the lexical material from 18 manuscripts. However, even before this dictionary was published, another six manuscripts that demanded classification among Old Church Slavonic canonical texts due to their archaic features were discovered or newly published. This led the authors to prepare a second, revised and supplemented edition of the dictionary. Its drafting at the Institute of Slavonic Studies began after the year 2000. However, the original five-member authorial collective then narrowed down to only Emilie Bláhová (†2016) and Ludmila Pacnerová (†2008). The second edition is supplemented with material from newer manuscripts, and it also includes the innovation of adding English equivalents and changing the dictionary's language for explanations from Russian to Czech, thereby broadening its applicability within the Czech environment. In 2003, the *Indexes for the Old Church Slavonic Dictionary* (*Indexy k Staroslověnskému slovníku*) were published as a supplement to the *DOOCSM*. This work, the fruit of the labors of Zdenka Ribarova, contains seven types of indexes (e.g. a retrograde index).

Before her death, Emilie Blahová managed to review the entire planned edition of the dictionary and comprehensively fix its minor errors. At present the team is performing a final review of the individual headwords, double-checking the integration of all corrections, and unifying the Czech wordings for the explanatory material. The headwords for the letter ꙁ will be published before the culmination of the *GORAZD* project. The provision of access to the entire electronic version of the dictionary is planned to arrive after the publishing of its printed version.

The Greek-Old Church Slavonic Index

Another of the electronic databases planned within the *GORAZD: An Old Church Slavonic Digital Hub* project is the *Greek-Old Church Slavonic Index* database. It aims to present the lexical material covered in the first volume of the *Greek-Old Church Slavonic Index* (*Řecko-staroslověnský index*), a lexicographic work published in eight parts by the Institute of Slavonic Studies from 2008 to 2014.

The drafting of a Greek-Old Church Slavonic Dictionary-Index was intended to provide, alongside the Latin-Old Church Slavonic Index and the Old High German Dictionary-Index, a necessary supplement for exploring the vocabulary of Old Church Slavonic; this intention dates back to the beginnings of work on the *OCSD* in the 1940s, even though efforts to process and present Greek and Old Church Slavonic lexical equivalents naturally occurred earlier as well, e.g. in the form of Greek-Old Church Slavonic listings of words for certain Old Church Slavonic texts translated from Greek. Thus in connection and in parallel with the excerpting of texts for the *OCSD*, briefer Greek-Old Church Slavonic sheets were also produced for use in a future dictionary-index; unlike those in the sheet-based card index for the *OCSD*, these did not register a given lexeme's broader context. All of the source texts in the *OCSD* translated from Greek originals and all of the biblical citations from original Old Church Slavonic texts are processed in this way. The first conception for the drafting of the Greek-Old Church Slavonic Dictionary-Index, or more precisely the material in the Greek-Old Church Slavonic card index, was presented to fellow experts by Vladimír Kyas in 1978. No further work with this index occurred until the 1980s, when Ilona Páclová picked up this project. After that
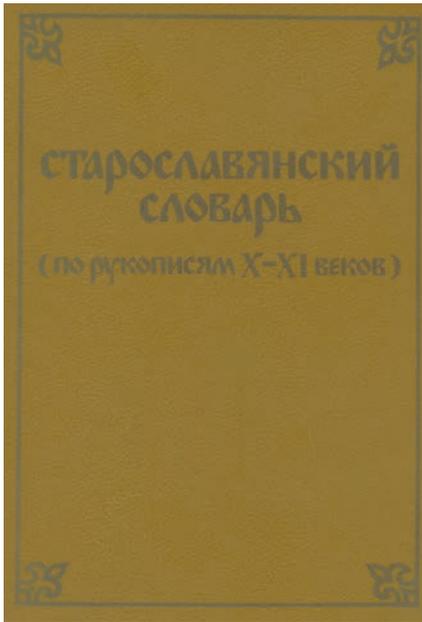
Example of the headword **богатити сѧ** from the original edition of the *Old Church Slavonic Dictionary According to the Manuscripts of the 10th–11th Centuries*

Original edition of the *Old Church Slavonic Dictionary According to the Manuscripts of the 10th–11th Centuries*

Example of the same headword finally adjusted in the *GORAZD* interface

Example of the headword γέννησις in the printed edition of the *Greek-Old Church Slavonic Index*

Example of the same headword processed in the *GORAZD* interface

scholar's untimely death in 1988, Emilie Bláhová took over the project; however, work on the project did not truly begin until 2004. In the meantime, the conception for work on the dictionary-index was changed and adjusted, primarily so as to make it easier for users to understand the headwords and navigate among them. The Old Church Slavonic equivalents for Greek lexemes were thus alphabetized, the efforts to depict the translation techniques for verb forms were abandoned, and the register of variant readings in Greek and Old Church Slavonic manuscripts was fleshed out in more detail. The selection of texts for inclusion into the excerpting base was also revised, and certain newly discovered manuscripts that could not be incorporated into the *OCSD* until the *Supplements to Volume I* were also filled in. The first volume of the *GOCSI* presented headwords for the letters alpha, beta, and gamma (α′–γωνιαῖος), processed on the basis of the Greek-Old Church Slavonic card index. Its introductory parts contain, among other things, synopsis tables for biblical texts, using which one can find specific biblical places in Old Church Slavonic source texts.

At present, the entire headword portion of the first volume of the *GOCSI* is included in the *Greek-Old Church Slavonic Index* database. After digitization, all headwords were checked and errors from the beginning of the volume that had only been uncovered later in the course of the work were corrected. Brief *Supplements to Volume I* supplements and corrections noted at the end of the first volume were also worked directly into the database text. Besides these enhancements, the database also enables users to search the dictionary based on Greek headings and subheadings, Old Church Slavonic headings, and the grammatical genders and parts of speech to which the Greek lexemes belong, and to perform full-text searches. In the future we expect to gradually supplement the database with further headwords in connection with the ongoing processing of material for the further letters of the alphabet, and to link up the *GOCSI* database to the *OCSD* database.

What will the future bring?

The *GORAZD: An Old Church Slavonic Digital Hub* project offers a broad range of possibilities both for supplementing the dictionary databases it contains and for expanding its contents. In the electronic *Old Church Slavonic Dictionary*, the *Supplements to Volume I of the Dictionary of Old Church Slavonic* is to be incorporated, and this will naturally oblige the authors to similarly rework the three remaining volumes. However, the headwords will not be added in alphabetical order as they have been so far: since digital content may be added in any order, digitization will enable a much more effective workflow based on the individual source texts. Also, the perspective of lexicographically processing Old Church Slavonic texts that could not previously be incorporated into the *OCSD* has now opened up. This is relevant both for part of the manuscripts discovered in 1975 at the St. Catherine monastery in the Sinai (the *Psalter of Dimitri* and the *Sinai Missal*) – as these were inaccessible to the authors of the *Supplements to Volume I* during their work – and for certain newly identified and published texts of Czech Church Slavonic origin (e.g. the *Legend of St. Anastasia* and the St. Gregory prayers). Advanced digital search tools also open up the theoretical possibility of extracting the Czech Old

Emilie Bláhová (1931–2016), long termed author of the *Old Church Slavonic Dictionary*, co-editor of the *Old Church Slavonic Dictionary According to the Manuscripts of the 10th–11th Centuries* and until her death, the main editor of the *Greek-Old Church Slavonic Index*

Church Slavonic material from the *OCSD* and processing it as a separate dictionary. Other sets of texts incorporated into the *OCSD* may now see gradual revision and supplementation as well. The project team will also strive to interconnect the *OCSD* headwords with the contents of the Old Church Slavonic card index and the headwords of the *Greek-Old Church Slavonic Index*. The electronic processing of the Latin-Old Church Slavonic and Old High German-Old Church Slavonic indexes is also a possibility. While their material is currently stored only in the form of index sheets, these have in past years also been scanned.

The future development of the *GORAZD* project need not, however, be restricted solely to supplementing and expanding lexicographic databases, even though these are – and undoubtedly will remain – its core. It can also offer other content, primarily multimedia content for use in popularizing Old Church Slavonic and teaching students at universities and high schools. This can include e.g. maps, voice recordings for the most important Old Church Slavonic texts, and other teaching materials. Another natural extension could be the creation of a text bank with a direct link between individual words and the headwords of electronic dictionaries, including automatic analysis of grammatical forms. However, these are challenges to be faced in other, future projects.

should mention the *Encyclopedia of Byzantium* (*Encyklopedie Byzance*) by Vladimír Vavřínek and Pavel Balcárek (2011) and *Cyril and Methodius Between Constantinople and Rome* by Vladimír Vavřínek (*Cyril a Metoděj mezi Konstantinopolí a Římem*; 2013). Also in preparation are monographs on the subjects of the Byzantine nobility and of Kherson in the Early Middle Ages.

The **Department of Slavonic Linguistics and Lexicography** focuses on the vocabularies, grammatical systems, and dialects of Slavic languages. The East Slavic languages have long been its primary focus. Its lexicographic activities have had as their primary output translation dictionaries for Russian (among them the six-volume *Large Russian-Czech Dictionary*) and Ukrainian; it has also produced dictionaries of neologisms for both mentioned languages. This department is gradually providing online access to the *Russian-Czech Dictionary Database* (*Rusko-česká slovníková databáze*), and in parallel it is preparing the digitization of the unpublished *Large Czech-Russian Dictionary* (*Velký česko-ruský slovník*) that was produced in the 1970s and 1980s. It is likewise digitally processing the *Database of the Dialects of Carpathian Ruthenia* (*Databáze dialektů Podkarpatské Rusi*), which is based primarily on the archive material in the dialect-studies card index of prof. Ivan Pankevych from the 1920s and 1930s. In the area of South Slavic languages, the publishing of a Slovenian--Czech dictionary is being prepared. The department also performs research into neologisms in Bulgarian, as well as analyses of calques in Czech and Serbian. In the area of grammatical research, we can highlight the department's research into verb bonds (valency) in Slavic languages. One of its results is the publication *Research into Verb Valency in the Slavic Countries* (*Výzkum slovesné valence ve slovanských zemích*, 2016).

The **Department of the History of Slavonic Studies and Literature** is active both within the literary-studies area of domestic Slavic studies tradition, founded on a comparative conception of Slavic literatures, and to a growing extent within the literary--studies area of European Slavic studies, exploring the dialogue among European cultures and literatures. The first sphere of its activities includes the history of Slavic studies and inter-Slavic literary and cultural relations and literature before the start of the 20th century. Here let us mention the unique bibliographical project *Slavica in the Czech Tongue* (*Slavica v české řeči*), whose four volumes map the translations of belles-lettres from other Slavic languages into Czech from the early 1800s until 1918. And among its book-form publications, let us note e.g. its anthology of the studies of Karel Krejčí, a Czech scholar of Polish Studies and comparative literature (2014). The second sphere is research into the Slavic literatures of the 20th century, and especially relationships in the Czech-German and Russian-German cultural contexts. Among the department's projects in this sphere, we can mention *Literature and Knowledge* (*Literatura a vědění*), which was realized in 2012–2016 in cooperation with colleagues from Germany and Sweden, or the newly begun project *Crimea as a Crossroads of Cultural Memory* (*Krym jako křižovatka kulturní paměti*, 2016–2019) supported within the *Strategy AV21 – Europe and the State: Between Barbarism and Civilization* research program. The third sphere is research into the Russian diaspora in interwar Czechoslovakia. Here it publishes archive materials, for example extensive memoirs and diaries; meanwhile a *Lexicon of the Russian Emigration in the Czechoslovak Republic* (*Lexikon ruské emigrace v ČSR*) has recently been produced. Of its book-form outputs, let us name e.g. the monograph devoted to the scholarly legacy of the historian Nikolai Okunev (2012). Its attentions are also devoted towards Russian art in Czech collections. The publication *Russian Paintings, Drawings, and Graphic Art from the 19th Century to the mid-20th Century from the Collections of the Art Gallery in Náchod* (*Ruská malba, kresba a grafika od 19. do poloviny 20. století ze sbírky Galerie výtvarného umění v Náchodě*, 2015) by Julie Jančárková was awarded the 2016 Jury Prize of the Academia publishing house.

The *GORAZD: An Old Church Slavonic Digital Hub* project, accessible at www.gorazd.org, opens up the results of long years of Czech work in Old Church Slavonic studies for use by both experts and the public. These results include the monumental four-volume *Old Church Slavonic Dictionary* (*Slovník jazyka staroslověnského*, 1966–1997) including the *Supplements to Volume I of the Dictionary of OId Church Slavonic* (*Dodatky k I. dílu Slovníku jazyka staroslověnského*, 2016), a sample portion of the expanded new edition of the *Dictionary of the Oldest Old Church Slavonic Manuscripts* (*Slovník nejstarších staroslověnských památek*, first published in 1994), the first published volume of the *Greek-Old Church Slavonic Index* (*Řecko-staroslověnský index*, 2014), and a unique Old Church Slavonic sheet-based card index comprising over a million sheets. The project's outputs will also come to include special software tools for the digitization of explanatory and encyclopedic multilingual dictionaries.