



The PhytoLex plant name database: results and perspectives

Valeria B. Kolosova ¹, Kira I. Kovalenko ^{1,2*}, Georgy A. Molkov ¹

¹*Institute for Linguistic Studies, Russian Academy of Sciences, Tuchkov pereülok 9, 199053 Saint Petersburg, Russia*

²*European University at Saint Petersburg, Gagarinskaya st. 6/1A, 191187 Saint Petersburg, Russia*

Article info

History:

Received September 17, 2021

Accepted October 7, 2021

Published December 12, 2021

Key words:

historical lexicology

historical semantics

old literature

religious literature

ethnolinguistics

Abstract

Russian plant names are a semantic group poorly represented in historical dictionaries. During the project “Russian phytonyms in diachronic aspect (11th–17th c.)” (2017–2020), the PhytoLex database was created, which contains plant names recorded in Russian texts of the 11th–17th centuries. The population of the database was accompanied by research work devoted to various issues relating to the representation of plant names in the Russian language in general, and in particular in relation to some specific genres: herbals, documents of the Apothecary Chancery, business documents, church literature and lexicons. Some earlier unknown plant names were identified. These database materials will be the basis for the dictionary of Old Russian plant names, draft word entries for which have been compiled. Technical solutions will be used during the creation of a further database representing plant names recorded in later periods of the Russian language.

1. Introduction

The PhytoLex database (phytolex.iling.spb.ru) was created as a result of the project “Russian phytonyms in diachronic aspect (11th–17th c.)”. The aim of the project—fixing the genesis, chronology of appearance, ways of borrowing plant names during various stages of development of the Russian language—required the creation of a database of Old Russian plant names.

At the beginning of the project, the principles of creation and population of the database were worked through on the text of “Materials for the Old Russian Dictionary by Written Monuments” by Izmail I. Sreznevsky. The processing of the work demonstrated a number of problems: 1) identification of plants (doubtful or no translation); 2) identification of the source; 3) numerous misprints in quotations; 4) the time of appearance of a plant name in the language; 5) questions of graphical representation (peculiarities of the writing of words and characters: titlos, letters written over the line, etc.); 6) automatic word search in existing databases; 7) absence of the reverse (Russian – Old Russian) dictionary (Kolosova, 2017). It became clear that for the representation of plant names in the database, profound research work is needed to investigate both sources for the database and the principles of technical implementation. By the end of the project, participants had produced about 20 articles devoted to these issues; some observations remained unpublished. The main results of this research work will be discussed in this paper.

2. Plant names in various genres

Russian texts of the 11th–17th centuries are represented by a number of genres. In the database, each text is marked with a genre tag, and the classification of genres was designed by the members of the project specially for our purposes. The distribution of records can be seen in Table 1.

The largest number of records came from business documents, especially of the Apothecary Chancery, and from lexicons, the compilers of which used a wide range of sources as material for the word entries.

*Email address: kira.kovalenko@gmail.com.

Genre	Number of records
Canonical Bible	832
Service books	9
Theological writings	83
Hagiographical texts	127
Apocrypha	13
Encyclopedic books	9530
— agricultural books	442
— botanical books (mostly herbals)	632
— geographical books	657
— lexicography	7255
— lexicons	5803
— phrase books	1452
— books on medicine	201
Business documents	5086
— Apothecary Chancery	3297
— Embassy Chancery	107
Historical writings	92
Letters	34
Newspapers	11
Fiction	208
Varia	602

Table 1: Distribution of records by genre

Original Russian herbals were created beginning only in the 17th century, but although they are not numerous in the database, they contain extremely interesting information about plants used in medicine. Biblical and service books, theological and hagiographical text give us examples of the oldest records. Some of them can be traced to the 9th–10th centuries—the time when the earliest Church Slavonic books are supposed to have been created¹. In the PhytoLex database they are represented by the ancient translation of the Four Gospels, Cyril of Jerusalem's *Catechetical and Mystagogic Lectures*, John the Exarch's *Hexameron*, *Paralipomena of Jeremiah*, and *The Life of the Great Martyr Irene of Macedonia*.

2.1. Church literature

In various genres of church literature, the names of plants are found sporadically, but collectively they form a varied and extensive body of material. In some cases, church text provides the earliest records of plant names—both known from other sources, and exotic and rather rare in writing. In particular, in various Hexamers the names *диктамонъ* (*Origanum dictamnus* L.), *ластовникъ* (*Chelidonium majus* L.), *брестие* (*Ulmus glabra* Huds.), *дряныни*, *дръныни*, *дрънь* (*Cornus mas* L.) and others were found. The translation of *Life of Andrew the Fool* contains the earliest record of *фасуль* (*Phaseolus* spp.), the translation of Flavius Josephus's *Books of the History of the Jewish War against the Romans*—*кунръ* (*Lawsonia inermis* L.) and *муровалъ* (*Balanites aegyptiaca* (L.) Delile).

Of particular interest are the plant names mentioned in the Gospel. The history of the Slavic translation of this text, which included several stages of editing in the South Slavic region before it entered Russia, led to the appearance in it of a large number of lexical variants (Slavova, 1989, p. 15–16), including variants of the plant names. The PhytoLex database contains the materials of the majority of the oldest and most significant copies of the Gospel, representing the text of the Cyril and Methodius translation in various

¹Each source is supplied with two dates: the year of the source creation, sometimes hypothetical, and the year of the copy which was used for the records extraction.

modifications: the Four Gospels (or *Tetra*, in which the readings of the evangelists follow the numbering the chapters) and the Gospel Lectionary (or *Aprakos*, in which the readings are given in the order of their readings at the liturgy throughout the year); the Lectionary, in turn, was divided into short and full types (Zhukovskaya, 1968; Temčinas, 1989, p. 12–13). Comparison of the Gospel of various types shows that the ancient translation into Church Slavonic was a compilation of texts translated from Greek at different points of time (Temčinas, 2013, p. 71).

The PhytoLex materials reflect this heterogeneous character of the Gospel text: such plants as *Ficus sycomorus* L., *Ficus carica* L., *Aquilaria* spp., *Anethum graveolens* L., *Mentha longifolia* (L.) Huds. are represented in two or more different ways. So, *Ficus sycomorus* L. (sycamore) has two main designations—*сикамина*, which is based on the transliterated Greek word *συκάμινος*, and Slavic *ягодичина*; their connection with different layers of editing in the Gospel has already been indicated in the research of Slavova (1989). The database material allows us to see that this distribution is preserved not only for *сикамина* and *ягодичина* but for their modifications as well: *сикамина*, *сюкамина* and *сикамня* are opposed to *ягодичина* and *ягодина*, which are represented only in full Gospel Lectionaries—Mstislav Gospel (State Historical Museum, Sin. 1203), Museum Gospel (Russian State Library, Rum. 104), Simonov Gospel (Russian State Library, Rum. 105) and Dobrilov Gospel (Russian State Library, Rum. 103). Words denoting fig tree (*Ficus carica* L.) are found in a number of Gospel readings. In most of them, variants *смоковъница* and *смокы* have a specific use: *смокы* is typical for the full Gospel Lectionaries (*Mt*, 21, 20; 24, 32; *Mk*, 11, 21; 13, 28; *Lk*, 13, 6; 13, 7); in *Jn*, 1, 48, the variant *смокы* was also found in the Ostromir Gospel (Russian National Library, Ф.п. I.5), containing the text of the short Gospel Lectionary. At the same time, the word *смоковъница* is neutral with respect to the type of text—it is found both in Four Gospels as well as in Lectionaries. For mint (*Mentha longifolia* (L.) Huds.), the initial distribution of names can be seen in *Lk*, 11, 42, where the lexical variant *вонялица* was recorded only in copies of the full Gospel Lectionary; in *Mt*, 23, 23 the distribution cannot be traced.

2.2. Lexicons

Exotic plants and their names were of a particular interest for the lexicographers of the 17th c. They included them in handwritten lexicons with descriptions of the appearance and use in everyday life, in medicine and religious practices. Taking previously compiled lexicons as a source, compilers copied their materials and added new word entries, among which plant names took their place. One of the popular sources became the Russian-Greek phrase book of the 14th–15th centuries. Although it contained transliterated Greek words which were not among those borrowed into the Russian language, the compilers actively took its materials into their lexicons, so that Russian readers of the 17th c. could know the Greek equivalents of the Russian *виноградъ* ‘vine’, *репа* ‘turnip’, *тыква* ‘pumpkin’, *редька* ‘radish’, *чеснок* ‘garlic’, *лукъ* ‘onion’, *укропъ* ‘dill’, *шафранъ* ‘saffron’, etc. Other plant names (sometimes with descriptions) were included mostly from texts well known at that period of time: sacred scripture, hagiographic and patristic literature, historical and geographical texts. A monk of the Solovky monastery, Sergiy Shelonin, enlarged his lexicon mostly by adding word entries taken from the literary works he edited at the time (Kovalenko, 2018, p. 199–234). As a result, descriptions of some real and some mysterious plant names were included:

- coconuts (“Indian nuts”) – from *Christian Topography* by Cosmas Indicopleustes,
- cedars – from *Chronicle* by George Hamartolos,
- *istiri* tree “which shines like the sun” and peppers – from the *Russian Chronograph*,
- *skorbut* grass (*Smilax?*) – from *Skete Patericon*,
- cloves, cardamom, cinnamon, peppers, maize, figs, *Myristica fragrans*, and some unidentified plants – from *Cosmography* (Russian translation of the *Atlas* by Gerardus Mercator) (Kovalenko, 2020).

After the third edition of the lexicon was finished, Sergiy included a great number of plant names from a translation of Hieronymus Brunschwig’s *Medicinarius*. He wrote all the plant names he found there into the free space which was typically left after letter subdivisions (the whole list of word entries is given

as a supplement to the article by [Kovalenko, 2017](#)), so that Russian readers became acquainted with the European botanical terminology of the time.

2.3. Business documents

Another group of important sources containing a great number of plant names is the business documents, especially documents of the Apothecary Chancery of the 16th and 17th centuries. They demonstrate that, during this period, words denoting plants were significantly replenished. Preference was given to the Latin pharmaceutical names that appeared long before scientific botanical nomenclature. The lexical group of plant names was formed in a short period and became a kind of preparatory material ([Olekhovich, 2018a](#)). In addition, medical texts in Latin had a significant impact on the formation of the medical prescription genre, in particular, its structure ([Olekhovich, 2018b](#)).

Apart from members of Russian authorities, high-ranking foreigners also used the services of the Apothecary Chancery. Kirill S. Khudin studied the corpus of medical prescriptions issued for Prince Valdemar of Denmark during his trip to Moscow in 1644–45. The article contains an alphabetical list of plants used by the Apothecary Chancery with their attribution to modern botanical terminology and an analysis of the physical status and completeness of the source materials studied ([Khudin, 2013](#)).

Documents related to the establishment of boundaries between land plots are also represented in the PhytoLex database in a significant quantity, since trees and bushes often served as places for drawing boundary marks. Boundary documents are accurately dated and geographically attributed to a specific region, which allows us to determine the area of both the word and the plant designated by this word. Thus, a comparison of two documents from the Uglich district showed the distribution of some plant names, as well as adjectives derived from them ([Shchekin, 2018](#)).

2.4. Herbals

Plants that appear in herbals were analysed by Alexandra B. Ippolitova, among them: *трава с откушенным корнем* (“grass with a bitten off root”), *плакун* (*plakun*), *адамова голова* (“Adam’s head”), *коровий язык* (“cow’s tongue”) and *чертогрыз* (“devil’s gnaw”) ([Ippolitova, 2018a,b, 2019](#)). In particular, based on texts about plants *трава с откушенным корнем* (“grass with a bitten off root”) and their Western European analogues, it was possible to identify this plant as the *Succisa pratensis*. Also, the primary source of some of these legends was found to be the incunabula of 1492 “Gaerde der suntheit” ([Ippolitova, 2018a](#)). In the case of *plakun* grass, the picture is much more complicated: contenders for this role are *Lythrum salicaria*, *Lythrum virgatum* L., *Epilobium angustifolium*, *Adenophora liliifolia* Ledb., *Iris sibirica* L., *Succisa pratensis* Moench., and *Veronica longifolia* L. ([Ippolitova, 2019](#)).

In addition to information about the use of plants in traditional medicine, herbals provide rare information about how their authors imagined the appearance of the described plants. On the basis of more than 2000 illustrations, it was shown that the images of plants in translated works and in “folk” herbals were based on fundamentally different approaches. While the translated herbals depicted some kind of “portraits” of individual plants, in the “folk” herbals plants were shown in the natural environment. Illustrators of translated works copied European engravings, sometimes adding their own, more familiar manner of depiction. The artists of the “folk” herbals worked on the illustrations independently, but they also acted in their usual style typical for the “common” manuscripts ([Ippolitova, 2018c](#)).

3. Lexicographical representation

Particular attention is paid in the project to the issues of lexicographic representation of plant names and the shortcomings in historical dictionaries of the Russian language, as in some cases it is possible to make clarifications in the dating and semantics of some words thanks to the PhytoLex database. Additionally, there are quite a significant number of newly revealed plant names which are not represented in the dictionaries ([Kovalenko et al., 2018](#); [Shchekin, 2019](#)). That shows that the creation of the dictionary of Old Russian plant names (11th–17th c.) is the next logical step after research work made by members

of the team. The dictionary will be compiled in Lexonomy—a special dictionary writing system. The Old Russian–Latin part will contain words and their meanings and provide forms, gender, phonological and orthographic variants, etymology, definition(s), illustrations, and sources. Materials may contain geodata, plants functions (medicinal, magical, religious, etc.), metaphorical meaning, etc. The Latin–Old Russian part will give the opportunity to find all the names of a particular plant recorded in the Old Russian language. The electronic format of the dictionary will allow several people to work simultaneously, which is very important in a joint project (Kovalenko & Kolosova, 2019).

4. Technical implementation

The PhytoLex database was created in the Django framework; for the Web interface, the Bootstrap framework was used. The data are stored in an open-source object-relational database system, PostgreSQL. It is a free, open, constantly updatable resource, currently containing more than 16 300 records of plant names.

Technical issues included such aspects as: data modelling, creation and normalization of controlled vocabularies, the development of a database and web application for the project's data curators and anonymous users on the web, and data visualization (Kolosova *et al.*, 2018). Complexity in the database structure was caused by the variety of information types: this includes data necessary for both linguists (etymology, quotations in original orthography, information about language of manuscript and its foreign original in case it is relevant, detailed source descriptions and references to the publications of the manuscripts) and ethnobotanists (quotations in simplified orthography, Latin terminology, Russian scientific names, links to the botanical catalogues). Although the field “plant name” is obligatory, it is possible to include citations, which contain only plant description, supplying metalinguistic information. For the convenience of users who populate the database and for the unification of the data, a number of controlled vocabularies were created, such as: a list of languages, lists of medical forms, plant use, plant parts, botanical classification terms, and some others. This makes it easy to search data and to structure information according to the user's requirements.

5. Conclusions

The PhytoLex database contains multidimensional information about plants recorded in Russian texts of the 11th–17th c. The wide coverage of textual material, including printed books, parchment and paper manuscripts, and birchbarks as well as genres variety made it possible to collect and partly identify a great number of plant names. Non-original lexemes are represented in variant forms that show the inclusion of borrowings into the Russian language. The database also gives many examples of polysemantic plant names that are rarely found in the historical dictionaries. Plenty of citations from various genres give a wide view of how a plant was used in everyday life, religion, medicine, cuisine, and other spheres. This makes the PhytoLex database a valuable resource for a Dictionary of Old Russian Plant Names (11th–17th c.). At the moment, the database provides opportunities for cross-genre comparison of phytonymic vocabulary of the 11th–17th centuries. In the future, these data can be used to compare the Old Russian material with the later states of the Russian language.

Bibliography

- Iropolitova, A. B. (2018a). *Легенда о траве с откушенным корнем в русской рукописной традиции XVI–XIX вв.* [Legend about a herb with a bitten root in the Russian manuscript tradition of the 16th–19th centuries], in “Живая старина”, (1), p. 6–10, [online].
- Iropolitova, A. B. (2018b). *Растения с названием Адамова глава в русских рукописных травниках* [Plant name *Adamova glava* in the Russian manuscripts herbals of the 18th–19th centuries], in “Полетата на этнологията: между знанието и познанието. 25 години асоциация «Онгъл»”, 16, p. 50–76.
- Iropolitova, A. B. (2018c). *К истории русской ботанической иллюстрации: «ученые» и «народные» травники XVI–XVIII веков* [Towards a History of Russian Botanical Illustrations: “Learned” and “Folk” Herbals in the Sixteenth to Eighteenth Centuries], in “ВИВЛЮОИКА: E-Journal of Eighteenth-Century Russian Studies”, 6, p. 13–45, Crossref.

- Ippolitova, A.V. (2019). *Плаун в русских рукописных травниках XVII – начала XX в. [Плаун (*Lythrum Salicaria*) in Russian herbal manuscripts 17th–20th century]*, in *Гора калинова (Бильни свет у традиционалној култури словена)*, Уред. З. Карановић, Belgrade / Vilnius, p. 217–249.
- Khudin, K.S. (2017). *Номенклатура лекарственных растений в медицинской практике Аптекарского приказа (на материале рецептов для датского королевича Вольдемара 1644–1645 гг.) [Nomenclature of medicinal plants used in the Apothecary Chancery's practices (based on prescriptions issued for Prince Valdemar of Denmark in 1644–1645)]*, in “Acta Linguistica Petropolitana. Труды Института лингвистических исследований”, XIII (2), Этноботаника 2: растения в языке и культуре, Наука, Saint Petersburg, p. 395–415, [online].
- Kolosova, V.B. (2017). *Проблемы анализа фитонимов в «Материалах для словаря древнерусского языка по письменным памятникам» И.И. Срезневского [Problems of phytonym analysis in “Materials for the Old Russian dictionary by written monuments” by I.I. Sreznevsky]*, in Осипова, Е.П. (ed.), *И.И. Срезневский и русское историческое языкознание: опыт и перспективы. К 205-летию со дня рождения И.И. Срезневского: сборник статей Международной научно-практической конференции, 21–23 сентября 2017 г.*, Ryazan State University, Ryazan, p. 229–235.
- Kolosova, V., Zaytseva, K. & Kovalenko, K. (2018). *PhytoLex – the Database of Russian Phytonyms: from Idea to Implementation*, in *SlaviCorp 2018. 24–26 September 2018, Charles University, Prague. Book of Abstracts*, p. 88–90, [online].
- Kovalenko, K.I. (2017). “Сказание о пропущении вод” как лексикографический источник [“Das New Distilleir Buch“ as a lexicographical source], in “Acta Linguistica Petropolitana. Труды Института лингвистических исследований”, XIII (2), Этноботаника 2: растения в языке и культуре, Наука, Saint Petersburg, p. 416–472.
- Kovalenko, K.I. (2018). *Азбучовник Давида Замарага как источник по русской лексикографии XVII века [David Zamaray's lexicon as a source for the Russian lexicography of the 17th century]*, PhD Dissertation, Saint Petersburg.
- Kovalenko, K.I. (2020). «Трава есть нькая тако нарицаема»: названия растений в *Азбучовнике* Сергия Шелонина [“There is some grass, the so-called”: plant names in *Lexicon* by Sergius Shelonin], in Сапожникова, О.С. (ed.), *Сергий Шелонин и Древнерусская книжность XVI–XVII вв.*, Альянс-Архео, Moscow / Saint Petersburg, p. 10–41.
- Kovalenko, K. & Kolosova, V. (2019). *Old Russian plant names dictionary (the 11th–17th cc.): Word entries drafts* in Gürlek, M., Çiçekler, A.N. & Taşdemir, Y. (eds), *ASLALLEX 2019. Proceedings of the 13th International Conference of the Asian Association for Lexicography*, Asos Publisher, Istanbul, p. 316–328, [online].
- Kovalenko, K.I., Kolosova, V.B. & Shekin, A.S. (2018). *Исторические словари как источники базы данных русских фитонимов (XI–XVII вв.) [Historical dictionaries as sources of the PhytoLex – Russian phytonyms database (11th–17th cc.)]*, in Крылова, О.Н., Мызников, С.А., Приёмышева, М.Н. & Пурицкая, Е.В. (eds), *Российская академическая лексикография: современное состояние и перспективы развития. Сборник научных статей по материалам Международной научной конференции, посвященной 70-летию выхода первого тома академического «Словаря современного русского литературного языка»*, Нестор-История, Saint Petersburg, p. 253–262.
- Olekhovich, O.G. (2018a). *Денотативно-семантические корреляции при формировании фитонимической лексики (на примере «Документов Аптекарского приказа» (XVII в.) [Denotative-semantic correlations in the formation phytonymic vocabulary (on the example of the documents of the Pharmaceutical Order of the 17th century)]*, in “Современная наука: актуальные проблемы теории и практики. Серия «Гуманитарные науки»”, 12, p. 174–179, [online].
- Olekhovich, O.G. (2018b). *Становление и развитие медицинского жанра «сказка» на примере опубликованных «Документов Аптекарского приказа» (XVII в.) [Establishment and development of medical genre «tale» in the documents of the Pharmaceutical Order of the 17th century]*, in Castellví, J., Zainouldinov, A., Garcia, I. & Ruiz-Zorrilla, M. (eds), *Proceedings of the International Conference on Russian Studies at the University of Barcelona*, Trialba Ediciones, Barcelona, p. 1391–1401.
- Shchekin, A.S. (2018). *Межевые документы XVI–XVII вв. как источник по истории фитонимической лексики русского языка [Documents of land surveying of the 16th–17th centuries as sources for the history of plants denominations in the Russian language]*, in “Славянская историческая лексикология и лексикография”, 1, p. 195–201, Crossref.
- Shchekin, A.S. (2019). *Об одном малоизученном памятнике русской переводной историко-географической литературы [About one understudied monument of Russian translated historical and geographical literature]*, in “Русский язык в школе”, 80 (4), p. 72–77, [online].
- Slavova, T. (1989). *Преславская редакция Кирилло-Методиевского староболгарского перевода Евангелия [Preslav edition of the Cyril and Methodius Old Bulgarian translation of the Gospel]*, in “Кирилло-Методиевские студии”, 6, p. 15–129.
- Temčinas, S. (1991). *Дистрибуция глагольных разностей в древнейших славянских списках евангелия и объем первоначального перевода [Distribution of verb variations in the oldest Slavic copies of the Gospel and the volume of the original translation]*, in *Исследования по глаголу в славянских языках. История славянского глагола*, Moscow, p. 9–41.
- Temčinas, S. (2013). *Кирилло-методиевское протоапракосное Евангелие: структура, состав и богослужебная функция [Cyril and Methodius Proto-Lectinary of the Gospel: Structure, Composition and Liturgical Function]*, in *Krakowsko-wileńskie Stidia Slawistyczne*, 8, *Cyrylomethodejski komponent kultury chrześcijańskiej Słowian w regionie karpackim. Historia, tradycje, odwołania*, Kraków, p. 49–72.
- Zhukovskaya, L.P. (1968). *Типология рукописей древнерусского полного апракоса XI–XIV вв. в связи с лингвистическим изучением их [Typology of manuscripts of the Old Russian full type Lectionary of the 11th–14th centuries in connection with their linguistic study]*, in *Памятники древнерусской письменности. Язык и текстология*, Moscow, p. 199–332.