

Арктика и Север. 2024. № 55. С. 243–252.

Краткое сообщение

УДК 81'282(470.11)(045)

DOI: <https://doi.org/10.37482/issn2221-2698.2024.55.243>

Электронный словарь архангельских говоров

Ненашева Лариса Викторовна^{1✉}, доктор филологических наук, доцент, профессор

Шурыкина Людмила Сергеевна², ассистент

^{1,2} Северный (Арктический) федеральный университет имени М.В. Ломоносова, набережная Северной Двины, 17, Архангельск, Россия

¹ l.nenasheva@narfu.ru ✉, ORCID: <https://orcid.org/0000-0001-8580-5980>

² l.shurykina@narfu.ru, ORCID: <https://orcid.org/0009-0004-8547-1967>

Аннотация. Исследовательская группа Северного (Арктического) федерального университета имени М.В. Ломоносова реализует проект «Тематический словарь архангельских говоров с электронной поддержкой», поддержанный Российским научным фондом. В задачи проекта входит издание трёх выпусков словаря, которые будут включать в себя лексику, отражающую традиционную картину мира архангельского крестьянина: в первом выпуске помещена лексика, связанная с наименованиями одежды, обуви, головных уборов; во втором выпуске — наименования жилых и нежилых построек; в третьем выпуске собрана лексика, называющая пищу и напитки архангельского жителя. Вторая задача проекта — создание электронного диалектного корпуса, посвящённого архангельским говорам, которого в настоящее время нет в университетах Северо-западного региона. В статье представлен опыт реализации данного проекта, который находится на стыке диалектологии, этнографии, культурологии и корпусной лингвистики. Результаты данного исследования и электронная база смогут найти применение в диалектологических, фольклорных и этнографических экспедициях, в преподавании русской словесности в вузе и в школе, в организации краеведческой работы, в просветительских проектах, направленных на популяризацию северной духовной культуры, в подготовке культурно-массовых и научных мероприятий, посвящённых языку и культуре Русского Севера.

Ключевые слова: архангельские говоры, корпусная лингвистика, диалектный словарь

Благодарности и финансирование

Исследование выполнено за счёт гранта Российского научного фонда № 23-28-01380, «Тематический словарь архангельских говоров с электронной поддержкой» (<https://rscf.ru/project/23-28-01380/>).

Electronic Dictionary of Arkhangelsk Dialects

Larisa V. Nenasheva^{1✉}, Dr. Sci. (Phil.), Associate Professor, Professor

Lyudmila S. Shurykina², Assistant

^{1,2} Northern (Arctic) Federal University named after M.V. Lomonosov, Naberezhnaya Severnoy Dviny, 17, Arkhangelsk, Russia


¹ l.nenasheva@narfu.ru ✉, ORCID: <https://orcid.org/0000-0001-8580-5980>

² l.shurykina@narfu.ru, ORCID: <https://orcid.org/0009-0004-8547-1967>

* © Ненашева Л.В., Шурыкина Л.С., 2024

Для цитирования: Ненашева Л.В., Шурыкина Л.С. Электронный словарь архангельских говоров // Арктика и Север. 2024. № 55. С. 243–252. DOI: <https://doi.org/10.37482/issn2221-2698.2024.55.243>

For citation: Nenasheva L.V., Shurykina L.S. Electronic Dictionary of Arkhangelsk Dialects. *Arktika i Sever* [Arctic and North], 2024, no. 55, pp. 243–252. DOI: <https://doi.org/10.37482/issn2221-2698.2024.55.243>

 Статья опубликована в открытом доступе и распространяется на условиях лицензии [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/)

Abstract. The research group of the Northern (Arctic) Federal University is implementing the project “Thematic dictionary of Arkhangelsk dialects with electronic support” supported by the Russian Science Foundation. The project aims to publish three issues of the dictionary, which will include the vocabulary reflecting the traditional world picture of the Arkhangelsk peasant: the first issue contains the vocabulary related to the names of clothes, shoes, hats; the second issue contains the names of residential and non-residential buildings; the third issue contains the vocabulary naming the food and drinks of the Arkhangelsk inhabitant. The second task of the project is to create an electronic dialect corpus devoted to the Arkhangelsk dialects, which is currently not available in universities of the North-West region. The article presents the experience of the project, which is at the intersection of dialectology, ethnography, cultural studies and corpus linguistics. The results of this research and the electronic database can be used in dialectological, folklore and ethnographic expeditions, in teaching Russian literature at universities and schools, in organization of local history work, in educational projects aimed at promoting northern spiritual culture, in preparation of cultural and scientific events devoted to the language and culture of the Russian North.

Keywords: *Arkhangelsk dialects, corpus linguistics, dialect dictionary*

Введение

На кафедре русского языка и речевой культуры Высшей школы социально-гуманитарных наук и международной коммуникации Северного (Арктического) федерального университета имени М.В. Ломоносова хранится картотека архангельских говоров, записанных в диалектологических экспедициях по Архангельской области, начиная с шестидесятых годов XX в. Этот интересный обширный материал является неисчерпаемым источником для изучения живого северного слова. В настоящее время назрела необходимость систематизировать этот богатейший материал и сделать его доступным для исследователей и специалистов, занимающихся вопросами диалектологии, фольклористики, этнографии, культуры Русского Севера.

Наряду с Национальным корпусом русского языка¹, где представлена коллекция текстов на русском языке, размещён подкорпус исторических и диалектных текстов, в последние годы учёными отдельных вузов создаются электронные диалектные корпуса, которые размещаются на страницах университетов. Такие электронные базы данных имеются в Волгоградском государственном социально-педагогическом университете, Омском государственном университете имени Ф.М. Достоевского, Томском государственном университете и др.

В настоящее время в картотеке кафедры русского языка и речевой культуры собранные диалектные материалы сохраняются в рукописном виде на недолговечных бумажных носителях и к ним затруднён доступ исследователей и специалистов, поэтому назрела необходимость создания электронного корпуса и информационной системы для того, чтобы продублировать традиционный способ хранения и обеспечить удобную работу с данными. В САФУ имени М.В. Ломоносова на кафедре русского языка и речевой культуры впервые разрабатываются веб-приложение и мобильное приложение, предоставляющие удобный доступ к диалектным материалам. Мобильное приложение, кроме того, позволит работать с материалами в полевых условиях и будет способствовать систематизации и классификации собранных материалов непосредственно на месте.

¹ Национальный корпус русского языка. URL: <https://ruscorpora.ru> (дата обращения: 25.07.2023).

Основная часть

А.С. Герд в восьмидесятые годы XX в. предложил сформировать цифровой фонд исторических и диалектных текстов. Его идею поддержал и аргументировал В.Е. Гольдин [1]. Сегодня существует широкий ряд корпусов, как зарубежных, так и отечественных, которые хранят в себе диалектные тексты и демонстрируют элементы диалектной речи, например, зарубежные — The Nordic Dialect Corpus², The Freiburg English Dialect Corpus³, Helsinki Corpus of British English Dialects⁴, российские — диалектный подкорпус в составе Национального корпуса русского языка (НКРЯ)⁵, Саратовский диалектный корпус⁶, Томский диалектный корпус⁷, Кубанский диалектный корпус⁸, диалектный корпус лингвокультуры Северного Приангарья⁹, корпус народной речи Среднего Прииртышья¹⁰, Волгоградский лексический атлас¹¹.

Говоры Русского Севера, в частности, архангельские говоры, из-за своей удалённости от центра и архаичности, не одно десятилетие привлекали внимание лингвистов, филологов, этнографов. В последние десятилетия отечественными учёными серьёзно осваивается лексика Русского Севера, результатом исследований стали публикации таких словарей, как «Архангельский областной словарь», издаваемый МГУ имени М.В. Ломоносова под редакцией О.Г. Гецовой и Е.А. Нефедовой, «Словарь говоров Русского Севера» под редакцией А.К. Матвеева, издаваемый Уральским университетом, «Словарь русских говоров Карелии и сопредельных областей» под редакцией А.С. Герда, изданный Санкт-Петербургским государственным университетом, «Словарь пинежских говоров» А.Н. Левичкина и С.А. Мызникова (в 2014 г. вышел в свет первый выпуск словаря с пробными статьями).

В настоящее время также активизировалась работа по созданию тематических словарей, например, таких как «Тематический словарь говоров Тверской области» (2002–2006) [2], «Ловецкое слово: Словарь рыбаков Волго-Каспия» Э.В. Копыловой (1984) [3], словарь костромского краеведа А.В. Громова «Лексика льноводства, прядения и ткачества в костром-

² Nordic Dialect Corpus. URL: <http://www.tekstlab.uio.no/nota/scandiasyn/> (дата обращения: 14.07.2023).

³ Freiburg English Dialect Corpus. URL: <https://fred.ub.uni-freiburg.de/> (дата обращения: 14.07.2023).

⁴ Helsinki Corpus of British English Dialects. URL: <https://varieng.helsinki.fi/CoRD/corpora/Dialects/> (дата обращения: 14.07.2023).

⁵ Корпус диалектных текстов: Национальный корпус русского языка. URL: <https://ruscorpora.ru/corpus/dialect> (дата обращения: 24.07.2023).

⁶ Крючкова О.Ю., Гольдин В.Е. Мультимедийный диалектологический корпус — важнейший ресурс сохранения и изучения народно-речевой культуры. URL: <http://sarteorlingv.narod.ru/dialekt/kru4kova-goldin.html> (дата обращения: 24.07.2023).

⁷ Томский диалектный корпус. URL: <https://losl.tsu.ru/?q=corpus> (дата обращения: 24.07.2023).

⁸ Кубанский диалектный корпус. URL: <https://ethnolex.ru/kubdk/> (дата обращения: 20.07.2023).

⁹ Электронный текстовый корпус лингвокультуры Северного Приангарья. URL: <http://angara.sfu-kras.ru> (дата обращения: 20.07.2023).

¹⁰ Словарь констант народной речи Среднего Прииртышья. URL: <http://dict.univer.omsk.su/> (дата обращения: 20.07.2023).

¹¹ Слово на карте. Лексический атлас Волгоградской области. URL: <http://dialekt.vspu.ru/?q=node/2> (дата обращения: 20.07.2023).

ских говорах по реке Унже» (1992) [4], «Словарь географической терминологии в русской речи Пермского края» Е.Н. Поляковой (2007) [5].

Богатый диалектный материал, хранящийся на кафедре русского языка и речевой культуры, позволяет создать тематический словарь, в котором слова группируются по тематическому принципу. В словаре содержится лексика той или иной сферы жизни людей (одежда, пища, постройки, природа, животный и растительный мир и т. д.) в пределах Архангельской области. Материалы словаря могут послужить хорошим источником для реконструкции традиционной картины мира архангельского крестьянина. В связи с развитием компьютерных технологий назрела необходимость создания электронного варианта «Тематического словаря архангельских говоров».

В рамках гранта, поддержанного Российским научным фондом, предполагается подготовить три выпуска тематического словаря архангельских говоров, включающих в себя лексику, связанную с названиями одежды, жилых и нежилых построек, пищи и напитков, т. е. лексику, отражающую традиционную картину мира архангельского крестьянина.

Первый выпуск тематического словаря «Одежда, обувь, головные уборы, украшения, ткани» включает в себя такие подгруппы, как «общее название одежды», «качественное название одежды», «женская верхняя одежда», «мужская верхняя одежда», «сарафаны, платья, женские кофты, рубашки, юбки», «мужские рубашки, штаны», «материалы, из которых делалась одежда», «рукавицы, варежки, перчатки», «головные уборы, женские и мужские», «обувь, женская и мужская», «носки, чулки», «детская одежда», «лексика, связанная с ношением одежды и обуви», «лексика, связанная с изготовлением и починкой одежды и обуви», «украшения».

На данный момент диалектный материал, записанный преподавателями и студентами пединститута (ныне САФУ) в ходе диалектологических экспедиций, хранится в виде полевых записей на бумажных носителях (в тетрадях и на карточках) и на магнитофонных кассетах. Этот материал очень ценен, поскольку сохраняет самобытную, уникальную, протяжную, певучую северную речь. Диалектный материал в тетради записан с указанием персональных данных информанта: фамилия, имя, отчество; год рождения (возраст); образование; также отмечается, является ли информант коренным жителем — эти сведения особенно важны, так у коренных жителей сохраняется подлинная северная речь. Диалектные слова в тетради выделяются подчеркиванием, над словом обязательно ставится ударение. Слово записывается в контексте, опираясь на который можно сформулировать значение. Часто в тетрадях можно встретить рисунки бытовых предметов, которые наглядно показывают, как выглядит предмет, о котором идёт речь. На рис. 1, 2 представлены образцы полевых записей, зафиксированных в диалектологических экспедициях в деревне Ухта Каргопольского района в июне-июле 1991 г. (рис. 1) и в Ошевенске Каргопольского района в июне-июле 1993 г. (рис. 2).

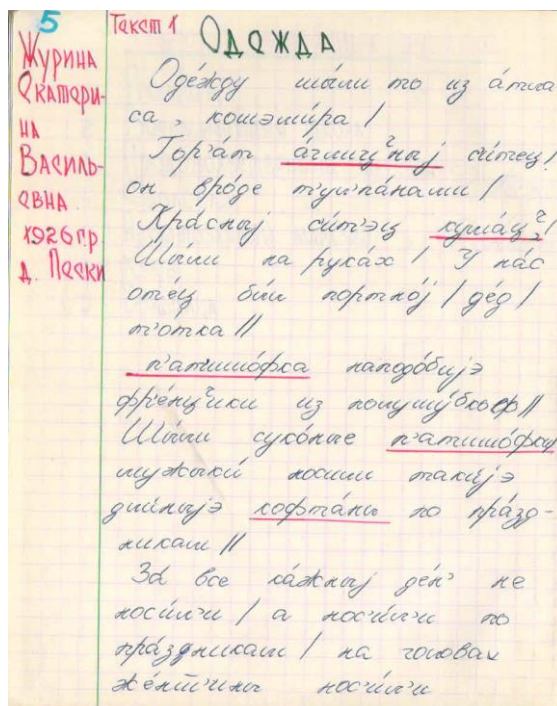


Рис. 1. Диалектная запись.

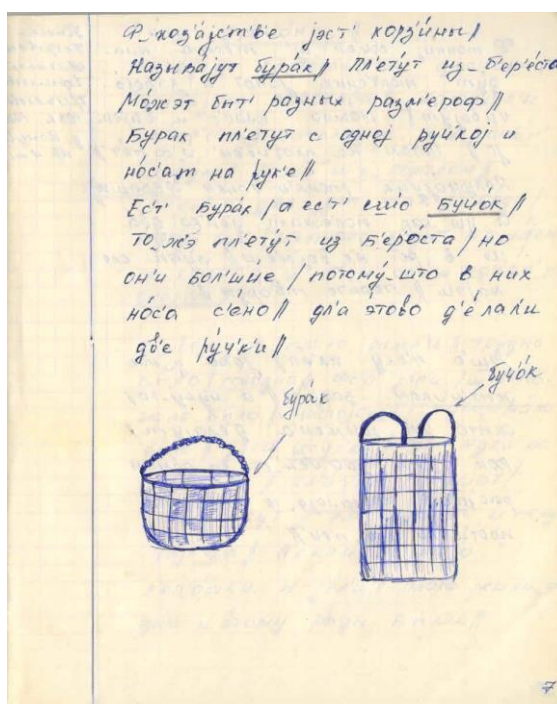


Рис. 2. Диалектная запись с иллюстрациями.

Из собранных записей выписывали на карточки диалектные слова в исходной форме: имя существительное — в именительном падеже единственного или множественного числа; имя прилагательное в именительном падеже, мужском роде, единственном числе; глагол — в инфинитиве. На карточке формулировалось толкование слова (определение значения слова) и записывался контекст, раскрывающий значение слова (иллюстративный материал). Также фиксировался населённый пункт и район Архангельской области, где данное слово было записано, и инициалы информанта.



Рис. 3. Слово «Казачина».

На данный момент к публикации подготовлен первый выпуск тематического словаря «Одежда, обувь, головные уборы, украшения, ткани». Этот же материал загружается в электронный диалектный корпус. Сформированные карточки заносятся в базу данных диалектных слов при помощи специально разработанного приложения «Копилка слов».

«Копилка слов» позволяет формировать списки диалектных слов. Для этого информация о каждом слове заносится в специальную электронную карточку. Записи можно редактировать или при необходимости удалять. Полученный список слов можно экспортировать в формате .docx как набор словарных статей или в формате .csv как пригодный для анализа датасет.

Внешний вид приложения представлен на рис. 4.

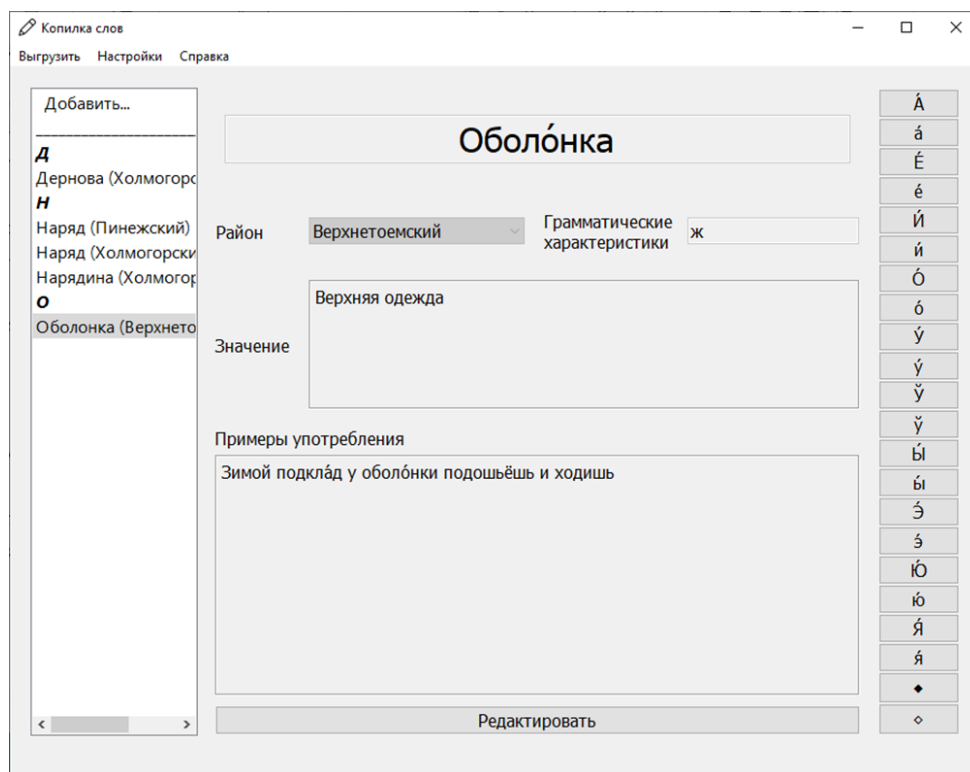


Рис. 4. Внешний вид приложения.

Строка меню позволяет получить доступ к возможностям программы. В подменю «Выгрузить» находятся две строки: 1) «csv», которая позволяет выгрузить собранные словарные карточки в формате .csv для анализа машинными средствами, и 2) «word», которая позволяет выгрузить собранные словарные карточки в формате .docx в виде словарика, отсортированного по алфавиту. Подменю «Настройки» содержит строку «Изменить размер шрифта», посредством нажатия на которую вызывается диалоговое окно настроек шрифта текстовых элементов. Подменю «Справка» содержит соответствующую строку, которая вызывает окно с краткой информацией о приложении и его функционале и с контактами разработчика.

В центре расположена словарная карточка, в поля которой заносится соответствующая информация с бумажных карточек.

Слева от словарной карточки находится органайзер слов. Слова в нём отображаются без ударений, чтобы избежать лишнего визуального шума, но с указанием районов, в которых эти слова используются. Слова в органайзере отсортированы по алфавиту и разбиты на разделы по первой букве, каждый раздел предваряется инициалом — соответствующей заглавной буквой.

Нажатие на элемент «Добавить...» позволяет переключиться в режим добавления нового слова. Посредством нажатия на выбранное слово показывается соответствующая словарная карточка. Нажатие на кнопку «Редактировать» позволяет переключить словарную карточку в режим редактирования. Для того, чтобы продолжить работу со списком слов, необходимо сохранить или отменить изменения в текущей карточке, при этом для сохранения должно быть заполнено хотя бы первое поле карточки.

Справа от области словарной карточки находится панель дополнительных символов. Они используются для постановки ударений (´), обозначения краткости (ˇ), добавления прочей необходимой информации (♦, ◇). Использовать их можно, если курсор мыши поставить в поле «слово», «значение слова» или «примеры употребления» и затем щёлкнуть на нужном символе левой кнопкой мыши. Символ будет вставлен на позиции курсора. Для постановки ударений также можно использовать апострофы: по окончании редактирования поля все гласные буквы с апострофами будут заменены на гласные буквы с ударениями.

В рамках тестирования приложение было предложено для работы группе из восьми студентов направления «Филология» в качестве одного из заданий по производственной практике. В результате тестирования было установлено, что приложение работает корректно и выполняет заявленные функции. Сформированные в результате обработки файлы в формате .csv (датасеты) пригодны для анализа.

Был собран и проанализирован датасет на основе материалов, вошедших в первый выпуск тематического словаря «Одежда, обувь, головные уборы, украшения, ткани». Так, на рис. 5 показано облако слов, которое отражает наиболее частотные слова. Чем чаще слово употребляется, тем крупнее оно показывается на рисунке.



Рис. 5. Облако слов.

На рис. 6 показана сложенная столбчатая диаграмма, отражающая распределение по районам самых частотных слов в этом датасете. На рисунке приведены относительные частоты — доли, которые занимает каждое из частотных слов в каждом конкретном районе. Диаграмма демонстрирует неравномерность употребления слов: так, например, слово «ранний» встречается в Виноградовском районе чаще прочих частотных слов и занимает лидирующую позицию среди прочих районов, а слова «сарафан» или «шить» вовсе не нашли отражения в словаре для этого района. На диаграмме видно также, что слово «носить» является абсолютным лидером как среди всех частотных слов, так и среди всех районов.

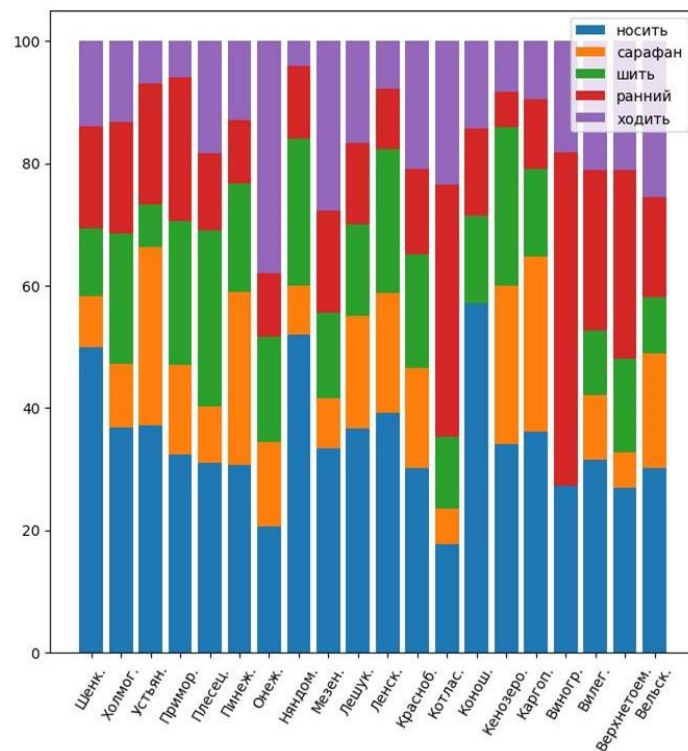


Рис. 6. Распределение по районам самых частотных слов

Заключение

Таким образом, на базе Северного (Арктического) федерального университета имени М.В. Ломоносова готовится к изданию многотомный тематический словарь говоров Архангельской области и впервые создаётся самостоятельный диалектный корпус, посвящённый архангельским говорам. Материалы тематического словаря архангельских говоров станут доступными широкому кругу читателей, всем любителям живого «родникового», как писал Ф.А. Абрамов, слова, а наполненный этими материалами диалектный корпус поможет специалистам в исследованиях русского фольклора, в работах по этнографии, социальной истории, народному костюму и быту, а также в организации музейно-этнографической деятельности. Результаты данного исследования смогут найти применение в преподавании русской словесности в вузе и в школе; в организации краеведческой работы; в просветительских проектах, направленных на популяризацию северной духовной культуры; в подготовке культурно-массовых мероприятий, посвящённых языку Русского Севера; в организации музейно-этнографической деятельности.

Список источников

1. Гольдин В.Е. Диалектологический текстовый машинный фонд говора и исследование диалектных изменений // Современные процессы в русских народных говорах. Саратов, 1991. С. 17–28.
2. Тематический словарь говоров Тверской области. Вып. 1 / Под ред. Т.В. Кирилловой, Л.Н. Новиковой. Тверь: ТвГУ, 2002. 240 с.
3. Тематический словарь говоров Тверской области. Вып. 2 / Под ред. Т.В. Кирилловой, Л.Н. Новиковой. Тверь: ТвГУ, 2003. 184 с.
4. Тематический словарь говоров Тверской области. Вып. 3 / Под ред. Т.В. Кирилловой, Л.Н. Новиковой. Тверь: ТвГУ, 2004. 228 с.
5. Тематический словарь говоров Тверской области. Вып. 4 / Под ред. Т.В. Кирилловой, Л.Н. Новиковой. Тверь: ТвГУ, 2005. 190 с.
6. Тематический словарь говоров Тверской области. Вып. 5 / Под ред. Т.В. Кирилловой, Л.Н. Новиковой. Тверь: ТвГУ, 2006. 150 с.
7. Копылова Э.В. Ловецкое слово: Словарь рыбаков Волго-Каспия. Волгоград: Нижне-Волжское книжное издательство, 1984. 128 с.
8. Громов А.В. Лексика льноводства, прядения и ткачества в костромских говорах по реке Унже: словарь: учебное пособие. Ярославль: Ярославский государственный педагогический университет им. К.Д. Ушинского, 1992. 118 с.
9. Полякова Е.Н. Словарь географических терминов в русской речи Пермского края. Пермь, 2007. 423 с.

References

1. Goldin V.E. Dialektologicheskiy tekstovyy mashinnyy fond govora i issledovanie dialektnykh izmeneniy [Dialectological Text-Machine Fund of the Vernacular and the Study of Dialectal Changes]. In: *Sovremennyye protsessy v russkikh narodnykh govorakh* [Modern Processes in Russian Folk Dialects]. Saratov, 1991, pp. 17–28. (In Russ.)
2. Kirillova T.V., Novikova L.N. *Tematicheskiy slovar' govorov Tverskoy oblasti. Vyp. 1* [Thematic Dictionary of Dialects of the Tver Region. Issue 1]. Tver, TvGU Publ., 2002, 184 p. (In Russ.)
3. Kirillova T.V., Novikova L.N. *Tematicheskiy slovar' govorov Tverskoy oblasti. Vyp. 2* [Thematic Dictionary of Dialects of the Tver Region. Issue 2]. Tver, TvGU Publ., 2003, 240 p. (In Russ.)

4. Kirillova T.V., Novikova L.N. *Tematicheskiy slovar' govorov Tverskoy oblasti. Vyp. 3* [Thematic Dictionary of Dialects of the Tver Region. Issue 3]. Tver, TvGU Publ., 2004, 228 p. (In Russ.)
5. Kirillova T.V., Novikova L.N. *Tematicheskiy slovar' govorov Tverskoy oblasti. Vyp. 4* [Thematic Dictionary of Dialects of the Tver Region. Issue 4]. Tver, TvGU Publ., 2005, 190 p. (In Russ.)
6. Kirillova T.V., Novikova L.N. *Tematicheskiy slovar' govorov Tverskoy oblasti. Vyp. 5* [Thematic Dictionary of Dialects of the Tver Region. Issue 5]. Tver, TvGU Publ., 2006, 150 p. (In Russ.)
7. Kopylova E.V. *Lovetskoe slovo: Slovar' rybakov Volgo-Kaspiya* [Lovetskoe Slovo: Dictionary of Volgo-Caspian Fishermen]. Volgograd, Nizhne-Volzhsкое knizhnoe izdatel'stvo Publ., 1984, 128 p. (In Russ.)
8. Gromov A.V. *Leksika l'novodstva, pryadeniya i tkachestva v kostromskikh govorakh po reke Unzhe: slovar': uchebnoe posobie* [The Vocabulary of Flax Growing, Spinning and Weaving in Kostroma Dialects along the Unzha River]. Yaroslavl, Yaroslavl State Pedagogical University named after K.D. Ushinsky Publ., 1992, 118 p. (In Russ.)
9. Polyakova E.N. *Slovar' geograficheskikh terminov v russkoy rechi Permskogo kraya* [Dictionary of Geographical Terms in the Russian Speech of the Perm Region]. Perm, 2007, 423 p. (In Russ.)

Статья поступила в редакцию 31.07.2023; принята к публикации 07.08.2023

Вклад авторов: все авторы внесли эквивалентный вклад в подготовку публикации

Авторы заявляют об отсутствии конфликта интересов