

ЕЛЕНА РАФХАТОВНА ГУСЕВА

кандидат филологических наук, старший преподаватель кафедры русского языка Института филологии, Петрозаводский государственный университет (Петрозаводск, Российская Федерация)
rafhatovna@mail.ru

АЛЕКСЕЙ ГЕОРГИЕВИЧ МАРАХТАНОВ

заместитель директора Регионального центра новых информационных технологий, Петрозаводский государственный университет (Петрозаводск, Российская Федерация)
marahthanov@mail.petrso.ru

ИНФОРМАЦИОННАЯ СИСТЕМА «ДИАЛЕКТНАЯ ЛЕКСИКА И ОНОМАСТИКА КАРЕЛИИ» КАК ДИНАМИЧЕСКИЙ МЕДИАРЕСУРС ПО КРАЕВЕДЕНИЮ*

Описывается опыт создания информационной системы (медиатеки) «Диалектная лексика и ономастика Карелии» / *Dialect lexis and onomastics of Karelia* (ДЛОК / DLOK), фундаментального цифрового архива данных. В базы данных вводится лексика, зафиксированная начиная с XVI века по настоящее время, что дает возможность изучить вопросы формирования лексической системы говоров Карельского Поморья, проследить динамику лексической и ономастической системы региона, выявить архаичные и вновь появившиеся элементы, определить пласт слов субстратного происхождения, а также установить изменения социального, экономического характера в истории региона, отраженные в слове. Медиатека представляет собой динамический электронный информационный ресурс «с открытым входом», позволяющий описывать материал в рамках различных лингвистических подходов: в дальнейшем предполагается создание не только лексических, но и текстовых баз данных, включение информации на аудио- и видеоносителях. Информационная система снабжена инструментами редактирования, поиска и визуализации данных в привязке к электронной карте Карелии и сопредельных областей. Входящие в медиатеку ДЛОК / DLOK базы данных могут служить примером для создания подобных систематизированных структур по другим регионам России. С 2017 года ведется разработка сектора «Карельское Поморье».

Ключевые слова: цифровая гуманитаристика, медиатека, диалектология, ономастика, история, Карелия

С целью сохранения и изучения историко-культурного наследия, представленного в речевой культуре Русского Севера – архаической зоне русских народных говоров и русского фольклора, территории тысячелетних контактов русского и прибалтийско-финского населения в прошлом и настоящем, Лаборатория лингвистического краеведения и языковой экологии (Лаборатория ЛКЭ) Института филологии и Исследовательская лаборатория локальной и микроистории Карелии (ИЛЛМИК) Института истории, политических и социальных наук ПетрГУ начали работу над созданием информационной системы (медиатеки) «Диалектная лексика и ономастика Карелии» / *Dialect lexis and onomastics of Karelia* (ДЛОК / DLOK), которая представляет собой цифровой архив данных, включающий материалы по лексике Карелии и сопредельных областей (Обонежье, Мурманская, Архангельская области). Информационная система может использоваться не только в исследовательских, но и в образовательных, просветительских целях. Ее создание обусловлено необходимостью интеграции научных парадигм в образовательные и просветительские процессы: ссылки на цифровой архив

данных будут размещены на информационном портале «Русская диалектология» (Институт русского языка имени В. В. Виноградова, г. Москва), сайтах Национального архива Финляндии, Центра поморской культуры (г. Беломорск) и др. Работа по созданию информационной системы началась в 2017 году в рамках проекта «Медиатека: русские говоры Карелии» по Программе развития опорного университета ФГБОУ ВО «Петрозаводский государственный университет».

1. МЕДИАТЕКА – ФУНДАМЕНТАЛЬНЫЙ ЦИФРОВОЙ АРХИВ ДАННЫХ

Информационная система (медиатека) ДЛОК / DLOK включает языковые факты, извлеченные из памятников письменности XVI–XIX веков, фольклорных текстов, записи диалектной речи, выполненные в экспедициях за последние 60 лет. Медиатека снабжена инструментами редактирования, поиска и визуализации данных в привязке к электронной карте Карелии и сопредельных областей. При проектировании архитектуры цифрового архива учитывается возможность расширения системы за счет добавления новых типов

информационных объектов (например, топонимические данные, текстовые базы данных), новых объектов хранения (мультимедиа, фотографии артефактов и т. д.), новых полей электронных карточек. Информационная система разработана на базе YiiFramework 2 и СУБД MySQL и доступна по адресу <https://lexis.petrsgu.ru/>; в период активного информационного наполнения доступна в режиме ограниченного доступа; размещена на серверном оборудовании Петрозаводского государственного университета. Техническое обслуживание осуществляют специалисты Регионального центра новых информационных технологий (РЦНИТ) ПетрГУ.

2. ЭЛЕКТРОННЫЕ БАЗЫ ДАННЫХ КАК СТРУКТУРНЫЕ КОМПОНЕНТЫ МЕДИАТЕКИ

Использование компьютерных технологий как особого способа представления языковых единиц (нарицательных и проприальных) позволяет классифицировать материалы по разным основаниям: 1) лингвистическим (ономасиологический, семантический, формальный – словообразовательный, морфемный, структурный); 2) экстралингвистическим (география, хронология, источник). Данный прикладной аспект работы должен обеспечить в последующем эффективность лексикографической обработки материалов наряду с традиционными методами (описательным, компонентным, сопоставительным, сравнительным). В настоящий момент информационная система (медiateка) ДЛОК / DLOK включает три базы данных:

- 1) Русская диалектная лексика Карелии (автор – Е. Р. Гусева, исполнители – И. Н. Дьячкова, Л. П. Михайлова, Н. Г. Урванцева);
- 2) Историческая антропонимия Карелии по документам XVI–XVII веков (автор, исполнитель – И. А. Кюршунова);

3) Ономастикон сказок Карелии (автор, исполнитель – О. В. Семенова).

Базы данных, входящие в информационную систему ДЛОК / DLOK, могут служить примером для создания подобных систематизированных структур по другим регионам России. В рамках проекта «Медiateка: русские говоры Карелии» в 2017 году начата разработка сектора «Карельское Поморье»: в базы данных введены диалектные и ономастические сведения, связанные с этим регионом.

2.1. Электронная база данных «Русская диалектная лексика Карелии»

База данных предназначена для хранения информации о русской диалектной лексике, зафиксированной начиная с 60-х годов XX столетия в экспедициях по Карелии и сопредельным областям; включает материалы, хранящиеся в архиве Лаборатории ЛКЭ [1]: сведения тетрадью с диалектными текстами и картотек Карельского государственного педагогического института / университета / академии и ПетрГУ. Материалы до 2004 года вошли в «Словарь русских говоров Карелии и сопредельных областей» [4] (далее – СРГК), в базу данных они вводятся с пометой СРГК.

Важной составляющей работы является определение темы, к которой относится лексема, номера вопроса и номера района по Федеральной программе собирания сведений для «Лексического атласа русских народных говоров» (ЛАРНГ)¹. В базе данных имеются соответствующие поля. Звездочкой (*) отмечены поля, по которым можно осуществлять выборку материала. Информация о лексеме сопровождается сканом карточки с диалектным словом (см. карточку 1).

Карточка 1

Поля электронной базы данных

заголовочное слово (ударная гласная обозначена прописной буквой)* кЕчкара
грамматическая помета* -ы, ж. р.
фонетическая транскрипция [кёчкара]
значение слова морское дно, которое обнажается во время отлива
иллюстрация На отливе по Белому морю, по кЕчкаре бродят медведи. Камбала
закапывается в грунт, они откапывают. [Что такое кЕчкара?] Грунт на отливе.
информатор* Семенова Надежда Сергеевна
населенный пункт с. Нюхча
район Беломорский
регион* Карелия
год записи* 2018
ссылка на тетрадь с диалектными текстами* Т. 1, С. 15
фольклорное*
фразеологическая единица*
примечания аудиозапись
тема (по программе ЛАРНГ) 3
номер вопроса (по программе ЛАРНГ) -
номер района (по программе ЛАРНГ) 11
место хранения Картотека ПетрГУ
собиратель Гусева Е. Р.

В течение 2017 года в рамках проекта «Медиаатекта: русские говоры Карелии» выполнены следующие работы:

- методом сплошной выборки извлечены карточки из картотеки «Словаря русских говоров Карелии и сопредельных областей» по Беломорскому, Кемскому, Лоухскому районам Карелии – 5512 ед.;
- сведения карточек СРГК внесены в электронные таблицы заданного формата (каждая единица описана через набор параметров, см. выше: поля базы данных) в программе редактирования таблиц Microsoft Office Excel, карточки СРГК сфотографированы – 2932 ед.;
- осуществлен импорт данных из электронных таблиц заданного формата в базу данных «Русская диалектная лексика Карелии» – 1000 ед.

Электронная база лексики русских говоров Карелии, включающая диалектные слова, функционирующие в речи жителей региона с середины XX века по настоящее время, дает возможность исследователю проследить динамику лексической системы, изменения социального, экономического характера в истории региона, отраженные в слове.

2.2. Электронная база данных «Историческая антропонимия Карелии по документам XVI–XVII вв.»

База данных, включающая материалы ИЛЛИМИК и Лаборатории ЛКЭ, предназначена для хранения информации по антропонимии Карелии XVI–XVII веков; для включения представленных в ней материалов в научный оборот с целью сопоставления способов номинации лица и мотивов его именования в делопроизводственных документах разных регионов Карелии; для выявления универсальных и специфических принципов номинации человека, обусловленных особым географическим положением Карелии в указанное историческое время и ее этническим составом; для использования основной единицы базы данных – антропонимической карточки – в вузовском процессе обучения (см. карточку 2).

В 2017 году в электронную базу данных введено 1237 ед. из картотеки антропонимов по Выгозерскому погосту из Писцовой книги Обонежской пятины 1496 и 1563 гг. (ПКОП)².

Карточка 2

Поля базы данных

заголовочное слово (антропоним) Ярыгин
 статус именуемого крестьянин
 иллюстрация деревня на Суме реки Гридинская Шилникова: в ней Иванко Ростовцев да сын его Иванко, в ней Федко Малгин, в ней Истомка Ярыгин, в ней Сенка Гаврилов Попов, в ней Ульянко Семенов Попов...
 география:
 погост Выгозерский
 населенный пункт дер. на Суме реки Гридинская Шилникова
 источник и хронология:
 документ ПКОП
 дата 1583
 страница в источнике 160
 этимология апеллятивная основа
 мотив именования именование по отцу
 анализ структуры:
 реализация в тексте Истомка Ярыгин
 статус единицы прозвище в основе патронима
 модель двухкомпонентная
 формула Н+Н
 реконструкция производящей базы Ярыга
 форма патроним
 формант модификата -ин
 примечание

Ранее методика обработки и систематизации антропонимических материалов апробирована И. А. Кюршуновой в соавторстве с О. В. Черняковым при составлении электронной базы данных «Историческая антропонимия Карелии по документам XVI–XVII веков (по материалам Олонецкого погоста)» [3].

2.3. Электронная база данных «Ономастикон сказок Карелии»

База данных предназначена для хранения антропонимической информации (личных имен, прозвищ, фамилий и проч.), представленной в сказках, записанных на территории Карелии (см. карточку 3). В 2017 году в базу данных

введено 1000 ед. из беломорских сказок, рассказанных М. М. Коргуевым³. Последующее введение ономастического материала в научный оборот позволит исследовать проблемы имен собственных не только в номинативном аспекте

(анализ принципов, способов, мотивов номинации лиц), но и в этнолингвистическом (картина мира, отраженная в именнике), в идиостилистическом (особенности представления сказочных сюжетов разными сказителями и т. д.).

Карточка 3

Поля базы данных

имя в тексте чаревна Татьяна
официальное имя Татьяна
статус единицы антропоним
модель 2-компонентное
формула А+К
формант
контекст Усадил он на стул свою чаревну Татьяну и стал бабушке обьеснять, кто оне и как едут.
примечания
синтаксические особенности
сказитель Коргуев
регион Беломорье
год фиксации 1936
название сказки Как чарьского сына унес комнатный ероплан
выходные данные Т. 1, Кн. 1, С. 96

3. МЕДИАТЕКА – ДИНАМИЧЕСКИЙ ЭЛЕКТРОННЫЙ ИНФОРМАЦИОННЫЙ РЕСУРС

В 2018 году получен грант РФФИ на реализацию проекта «Карельское Поморье: лексика и ономастика (XIV–XXI вв.)», целью которого является дальнейшее наполнение медиатеки контентом. Продолжается работа по введению диалектных сведений из картотек Карельского государственного педагогического института и Петрозаводского государственного университета, фольклорных источников – сказок М. М. Коргуева, Ф. Н. Свиньина, исторических источников – «Соловецкого летописца»⁴. Осуществляется чтение, транслитерация и набор в современной кириллице архивных текстов деловой письменности в скорописи XVII века из фонда «Олонецкая воеводская изба»⁵, хранящегося в Научном архиве Санкт-Петербургского института истории РАН на базе ИЛЛМИК, создается полнотекстовая база данных из транслитерированных текстов XVII века, отражающих историческую лексику северной Карелии, в том числе Карельского Поморья.

Проект «Карельское Поморье: лексика и ономастика (XIV–XXI вв.)» предполагает полевое обследование. Летом 2018 года состоялась комплексная экспедиция, цель которой – сбор диалектного, фольклорного, ономастического материала (села Беломорского района РК) [2]. Имеющиеся в лексикографических источниках сведения требуют вторичного обследования для выявления степени сохранности говора и обнаружения новых языковых фактов. В задачи экспедиции входили сбор лексики русских говоров по программе «Лексического атласа рус-

ских народных говоров»³; запись живой речи на аудио- и видеоносители с целью фиксации вербальных и невербальных средств общения; камеральная обработка записанной диалектной речи; фиксация русской локальной микротопонимии и антропонимии; фотографирование этнографических реалий и диалектоносителей; запись рассказов поморов о личной, семейной, трудовой жизни и событиях, проходивших в разные исторические периоды (после революции 1917 года, коллективизации, в годы Великой Отечественной войны, в послевоенный период и т. д.). Особое внимание было обращено на обрядовую лексику, связанную с мореплаванием, рыбным промыслом и морскими походами в Скандинавию. Проведена работа по каталогизации музейного фонда этнокультурного центра «Хламной сарай» в пос. Нюхча, что позволяет визуализировать диалектные лексемы и представить в виде фотографий материальных предметов быта в электронной базе данных «Русская диалектная лексика Карелии».

Итак, информационная система «Диалектная лексика и ономастика Карелии» / *Dialect lexis and onomastics of Karelia* (ДЛОК / DLOK) представляет собой динамический электронный ресурс «с открытым входом», то есть возможностью постоянного дополнения и корректировки лексикографического материала в базах данных с последующим картографированием – созданием электронных карт. В нее включаются языковые факты, извлеченные из исторических, фольклорных источников, записи диалектной речи за последние 60 лет, представленные как на бумажных (предполагается создание не только лексических баз данных, но и текстовых),

так и аудио- и видеоносителях, что позволяет описывать материал в рамках различных лингвистических подходов (системно-структурный; антропоцентрический: когнитология, прагма-

лингвистика, теория речевых жанров, лингвоперсоналогия и т. д.) и использовать в междисциплинарных исследованиях (лингвогеография, этнолингвистика, контактология).

* Исследование выполнено при финансовой поддержке РФФИ, проект № 18-012-00810 «Карельское Поморье: лексика и ономастика (XVI–XXI вв.)».

ПРИМЕЧАНИЯ

- ⁴ Программа собирания сведений для «Лексического атласа русских народных говоров»: В 2 ч. / Отв. ред. И. А. Попов. СПб.: Изд-во ИЛИ РАН, 1994.
- ⁵ Писцовые книги Обонежской пятины: 1496 и 1563 гг.: материалы по истории Карельской АССР / Акад. наук СССР, Археогр. комис.; Под общ. ред. М. Н. Покровского; Сост. А. М. Андрияшев. Л.: Изд-во Акад. наук СССР, 1930. IV. 268 с. (Материалы по истории народов СССР / Акад. наук СССР, Археогр. комис. Вып. 1).
- ⁶ Беломорские сказки, рассказанные М. М. Коргуевым / Под ред. А. Н. Нечаева. М.: Сов. писатель, 1938. 257 с.
- ⁷ Соловецкий летописец (текст по списку конца XVIII в. из фондов НБ РК) / Подгот. к публ. и текстол. коммент. Е. Н. Кутковой. Петрозаводск: Национальная библиотека РК, 2003 [Электронный ресурс]. Режим доступа: http://library.karelia.ru/Resursy/Rukopisnye_knigi/Letopise_Soloveckij/ (дата обращения 01.10.2018).
- ⁸ Чернякова И. А. Полнотекстовая электронная документальная коллекция «Олонецкая воеводская изба» [Электронный ресурс]. Режим доступа: http://illmik.petsu.ru/illmik/Archive_2.html (дата обращения 12.10.2018).

СПИСОК ЛИТЕРАТУРЫ

1. Гусева Е. Р. Об архиве лаборатории лингвокраеведения и языковой экологии Карельского государственного педагогического университета // Методика полевых работ и архивное хранение фольклорных, этнографических и лингвистических материалов: Материалы науч.-практ. семинара, г. Петрозаводск, 23–24 марта 2009 г. / Отв. ред. В. П. Кузнецова. Петрозаводск, 2009. С. 22–34.
2. Гусева Е. Р., Андрианова А. Ю. Комплексная экспедиция «Карельское Поморье: лексика и ономастика (XVI–XXI вв.)» // CARELiCA. 2018. № 2 (20). С. 52–61.
3. Кюршуннова И. А., Черняков О. В. База данных по исторической антропонимии как ономастический тезаурус // Громовские чтения. Вып. 3. Живое народное слово и костромской край: Сб. материалов и исследований междунар. науч. конф. Кострома, 7–9 нояб. 2016. Кострома, 2016. С. 401–407.
4. Словарь русских говоров Карелии и сопредельных областей / Гл. ред. А. С. Герд. Вып. 1–6. СПб., 1994–2005.

Guseva E. R., Petrozavodsk State University (Petrozavodsk, Russian Federation)
Marakhtanov A. G., Petrozavodsk State University (Petrozavodsk, Russian Federation)

INFORMATION SYSTEM “DIALECT LEXIS AND ONOMASTICS OF KARELIA” AS A DYNAMIC MEDIA RESOURCE FOR LOCAL STUDIES*

The article describes the experience of creating an information system (media library) “Dialect Lexis and Onomastics of Karelia” (DLOK), a fundamental digital data archive. DLOK media library includes language facts extracted from written records of the period between the XVI and the XIX centuries, folklore texts, dialect recordings made during the expeditions over the past 60 years – data stored in the archives of the Laboratory of Regional Linguistic Studies and Language Ecology and the Research Laboratory for Local and Microhistory of Karelia at Petrozavodsk State University. The information system is equipped with tools for editing, searching and visualizing data in relation to the electronic map of Karelia and the adjacent areas. Currently, the system includes three databases: Russian dialect vocabulary of Karelia, historical anthroponymy of Karelia according to the documents from the XVI–XVII centuries, and onomasticon of fairy tales of Karelia. The databases included in the DLOK information system can serve as an example for creating similar systematized structures in other regions of Russia. The media library is a dynamic electronic information resource: in the future it is planned to create not only lexical, but also textual databases, including information on audio and video media, which will enable to describe the materials using various linguistic approaches.

Key words: digital humanities, media library, dialectology, onomastics, history, Karelia

* The study was funded by the Russian Foundation for Basic Research as part of the project No 18-012-00810 “Karelian Pomorye: lexis and onomastics (XVI–XXI centuries)”.

REFERENCES

1. Guseva E. R. The archive of the Laboratory of Regional Linguistic Studies and Language Ecology of Karelian State Pedagogical University. *Metodika polevykh rabot i arkhivnoe khranenie fol'klornykh, etnograficheskikh i lingvisticheskikh materialov: Materialy nauchno-prakticheskogo seminar, Petrozavodsk, 23–24 marta 2009*. Petrozavodsk, 2009. P. 22–34. (In Russ.)
2. Guseva E. R., Andrianova A. Yu. Complex expedition “Karelian Pomorye: lexis and onomastics (XVI–XXI centuries)”. *CARELiCA*. 2018. No 2 (20). P. 52–61. (In Russ.)
3. Kyurshunova I. A., Chernyakov O. V. Historical anthroponymy database as an onomastic thesaurus. *Gromovskie chteniya. Vypusk 3. Zhivoe narodnoe slovo i kostromskoy kray: Sb. materialov i issledovaniy mezhdunar. nauch. konf. Kostroma, 7–9 noyabrya. 2016 g.* Kostroma, 2016. P. 401–407. (In Russ.)
4. Dictionary of Russian dialects of Karelia and adjacent areas. Issues 1–6. St. Petersburg, 1994–2005. (In Russ.)

Поступила в редакцию 16.10.2018