

МУЛЬТИМЕДИЙНЫЙ КОРПУС РУССКИХ ГОВОРОВ УДМУРТИИ: РАЗРАБОТКА И ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ¹

The Multimedia Corpus of Russian Dialects of Udmurtia:
Development and Possible Use

Виктор Аркадьевич Баранов
victor.a.baranov@gmail.com

Ижевский государственный технический университет (Ижевск, Россия)

Регина Александровна Верняева
anikina.regina@gmail.com

Ижевский государственный технический университет (Ижевск, Россия)

Екатерина Анатольевна Жданова
zhdanovaea@gmail.com

Ижевский государственный технический университет (Ижевск, Россия)

Victor A. Baranov
victor.a.baranov@gmail.com

Kalashnikov Izhevsk State Technical University (Izhevsk, Russian Federation)

Regina A. Vernyaeva
anikina.regina@gmail.com

Kalashnikov Izhevsk State Technical University (Izhevsk, Russian Federation)

Ekaterina A. Zhdanova
zhdanovaea@gmail.com

Kalashnikov Izhevsk State Technical University (Izhevsk, Russian Federation)

ISSN: 1698-322X ISSN INTERNET: 2340-8146

Fecha de recepción: 24.12.2019

Fecha de evaluación: 09.07.2020

Cuadernos de Rusística Española nº 16 (2020), 39 - 54

РЕЗЮМЕ

Статья посвящена представлению мультимедийного Корпуса русских говоров Удмуртии (<http://dialect.manuscripts.ru>) как электронного ресурса для изучения диалектной лексики не только методами корпусной лингвистики, но и методами лингвистической географии и электронной лексикографии. Корпус включает записи устной диалектной речи, сделанные во время диалектологических практик студентами и сотрудниками вузов республики в 70 – 80-е годы XX в. Мультимедийной составляющей корпуса являются аудиозаписи разговоров с диалектоносителями, сделанные в 1990 – 2000-е годы. Корпус имеет лексическую

1. Исследование выполнено при финансовой поддержке РФФИ и Правительства Удмуртской Республики в рамках научного проекта № 17-14-18006

разметку, позволяющую осуществлять поиск диалектных слов. В лексикографическом модуле корпуса осуществляется поиск лексемы и представление лингвистической и экстралингвистической информации о ней. В лингвогеографическом модуле можно произвести выборку всех ответов на один из вопросов программы, на которой основана разметка корпуса, и вынести полученные слова на карту Удмуртии. В статье рассмотрены некоторые результаты диалектологических исследований, полученные благодаря корпусу, приведены примеры карт и словарных статей, составленных при помощи корпусного менеджера.

Ключевые слова: лингвистический корпус, русские говоры Удмуртии, лексическая разметка, электронная лексикография, лингвистическая география.

ABSTRACT

The article is devoted to the presentation of the multimedia Corps of Russian dialects of Udmurtia (<http://dialect.manuscripts.ru>) as an electronic resource for studying dialect vocabulary not only by the methods of corpus linguistics, but also by the methods of linguistic geography and electronic lexicography. The article discusses the features of the linguistic content of the corpus and technical solutions used in its development, as well as some results of dialectological studies obtained thanks to the corpus.

Keywords: linguistic corpus, Russian dialects of Udmurtia, lexical markup, electronic lexicography, linguistic geography.

В рамках усиливающейся тенденции к созданию и применению лингвистических корпусов в диалектологии (Качинская, Сичинава 2015: 156-157; Земичева, Иванцова 2018: 193-195) в Ижевском государственном техническом университете им. М.Т. Калашникова разработан мультимедийный корпус русских говоров Удмуртии. Диалектные корпуса (как российские, так и зарубежные) ориентированы на различный материал, находящийся в распоряжении исследователей: письменные, аудио- и видеозаписи, сочетание различных видов данных (Жданова 2016: 137-141). В условиях утраты традиционных говоров в России особую ценность представляют материалы, относящиеся к возможно более раннему периоду.

Различными оказываются также разметка диалектных корпусных ресурсов и предоставляемые пользователям возможности знакомства и работы с ними, так как они обусловлены, с одной стороны, спецификой материала, с другой — целями и задачами, которые стоят перед разработчиками корпуса: анализ фонетики, грамматики, лексики, изучение диалектного дискурса, территориальной дистрибуции диалектных различий и т. д. (см. Летучий 2005: 215-232; Крючкова; Юрина 2011: 58-63).

Необходимость создания Корпуса русских говоров Удмуртии обусловлена тем, что на сегодняшний день отсутствуют научные исследования, обобщающие информацию о русском диалектном языке на территории современной Удмуртской Республики. В то же время в вузах региона накоплен обширный языковой материал, собранный в ходе многолетних полевых исследований, который может стать основой для словаря или атласа русских говоров Удмуртии. Создание современного электронного корпуса помогает сохранить диалектные записи и облегчить всем заинтересованным лицам доступ к ним, дает исследователям инструменты для разностороннего изучения русских народных говоров.

1. ХАРАКТЕРИСТИКА КОРПУСА РУССКИХ ГОВОРОВ УДМУРТИИ

1.1. Источники данных

Временной охват данных Корпуса русских говоров Удмуртии обусловлен материалом, имеющимся в распоряжении разработчиков. На данный момент корпус включает тексты из тетрадей с записями устной диалектной речи, осуществленными во время диалектологических практик студентами и сотрудниками Удмуртского государственного университета в 70 – 80-е годы XX в. База данных корпуса также дополнена текстами из «Хрестоматии по русской диалектологии», изданной в Глазовском государственном педагогическом институте (Хрестоматия по русской диалектологии 2014). Страницы этих источников размещены в корпусе в виде скан-копий, объединенных в условные тетради, привязанные к населенному пункту, где были сделаны записи. Сейчас база данных содержит 787 таких тетрадей, содержащих в совокупности 9300 страниц. Сегодня скан-копия страницы с рукописным (печатным) текстом выступает в качестве единицы измерения объема Корпуса русских говоров Удмуртии, так как объем тетради может составлять от 1 до 58 страниц.

Для каждой тетради корпуса указаны место и год записи, в составе тетради есть скан-копии страниц с характеристикой населенного пункта, данной собирателями, и сведениями о каждом информанте: фамилия, имя, отчество, год рождения или возраст, национальность, образование, вероисповедание, происхождение (о приездах), информация о выездах из населенного пункта и т. д. Пользователь может найти эти сведения при просмотре текстов.

Мультимедийной составляющей корпуса являются аудиозаписи разговоров с диалектоносителями, сделанные в 1990 – 2000-е годы. Записи, существующие на кассетах и дисках, оцифрованы и внесены в базу данных корпуса в виде фрагментов, соответствующих тексту или беседе с информантом на определенную тему. На данный момент в составе корпуса 64 аудиозаписи общей продолжительностью 1480 минут. Мультимедийная часть корпуса будет пополняться записями диалектной речи, хранящимися архивах кафедры лингвистики ФГБОУ ВО «ИжГТУ им. М.Т. Калашникова».

В общей сложности в корпусе представлены диалектные тексты, записанные в 166 населенных пунктах Удмуртской Республики.

1.2. Лингвистическая разметка

Корпус русских говоров Удмуртии создан на основе лингвогеографической информационной системы (ЛГИС) «Диалект» (<http://dialect.manuscripts.ru>), которая разрабатывалась как лингвогеографический Интернет-ресурс для хранения, анализа и демонстрации на лингвистической карте лексики, собранной по программе Лексического атласа русских народных говоров (ЛАРНГ) (Программа 1994) на территории Удмуртии. До начала работ над корпусом в базу данных ЛГИС «Диалект» уже был внесен снабженный метаинформацией лексический материал, собранный преподавателями, сотрудниками и студентами Удмуртского государственного университета в 1990

– 2000-х годах. При работе над корпусом структура базы данных сохранилась, в основе лексической разметки остался вопросник ЛАРНГ: разметка заключается в поиске в текстах ответов на вопросы программы ЛАРНГ и указании их начальной формы (лемматизации) для возможности поиска. В процессе разметки вопросник ЛАРНГ в ЛГИС «Диалект» дополнен новыми вопросами, которые необходимы для отражения специфических лексем русских говоров Удмуртии.

Морфологическая разметка на данном этапе отсутствует, но потенциал корпуса позволяет создать машиночитаемую версию текстов и применить существующие инструменты анализа. Основу корпуса составляют тексты, представляющие транскрипцию речи диалектоносителей, отражающую их произношение. Однако при разметке осуществляется орфографическая запись лексем для унификации произношения одного и того же слова и облегчения поиска. В целях обеспечения заполняемости лингвистической карты, с одной стороны, и во избежание ее перегруженности, с другой, было принято решение отмечать общеупотребительные слова, являющиеся ответами на вопросы программы ЛАРНГ, один раз в речи каждого диалектоносителя. В отдельных случаях, когда речь идет об особо частотных лексемах (например, хлеб), можно ограничиться одной фиксацией на каждый населенный пункт.

Мультимедийный характер корпуса обеспечивают аудиозаписи речи диалектоносителей. Оцифровка звуковых файлов, подготовка аудиозаписей для загрузки в базу данных ЛГИС «Диалект» была осуществлена с помощью свободно распространяемого программного продукта Audacity (<http://audacity-free.ru/>), являющегося редактором звуковых файлов и предоставляющего возможности по обработке звука любого качества (Audacity). Интерфейс и особенности работы данного программного продукта в применении к мультимедийной составляющей Корпуса русских говоров Удмуртии были описаны в (Верняева 2018: 104-108).

Все аудиофайлы в составе корпуса паспортизованы: указаны год и место, где была сделана запись, фамилия информанта. Также авторизованный пользователь может увидеть общие сведения о содержании записи: тип речи (диалог с собирателем или рассказ), тему разговора, длительность и качество записи, номер кассеты/диска и др.

Разметка аудиофайлов заключается в выделении курсором мыши на графике записи отрезка текста с интересующей лексемой (рис. 1), внесении данного слова в орфографическую запись в таблицу сегментов (поле «Ответ») и выборе из списка вопроса, которому соответствует ответ. В результате пользователь получает перечень сегментов, характеризующихся идентификационным номером фрагмента, временем начала и конца записи, лемматизированным ответом, вопросом-дефиницией, временем создания и изменения записи.

Выделенные и сохраненные сегменты, содержащие диалектные слова, отображаются при поиске в лексикографическом и лингвогеографическом модулях, могут быть прослушаны в качестве иллюстраций употребления диалектных слов в живой речи.

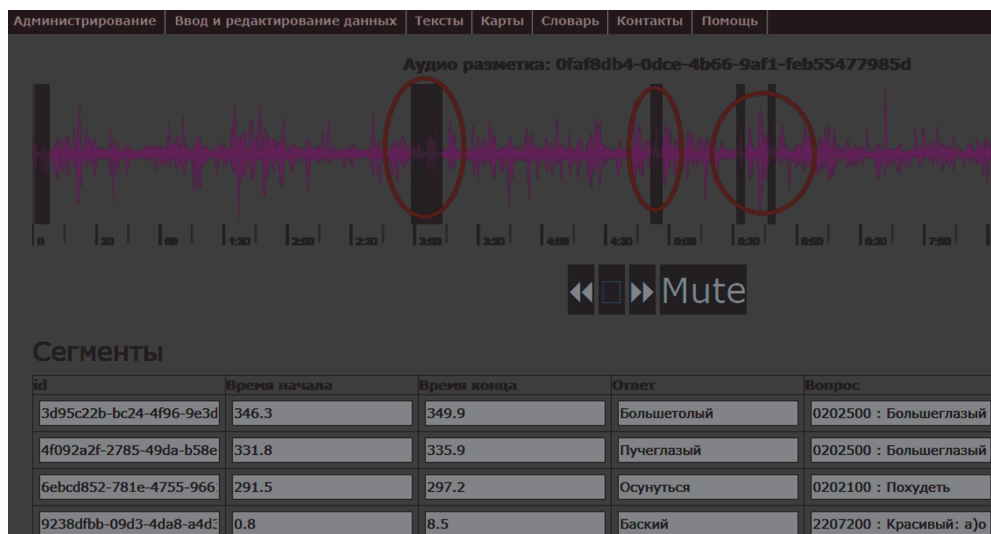


Рис. 1. Разметка аудиофайла

1.3. Корпусный менеджер

Представление материала в Корпусе русских говоров Удмуртии отражает специфику данных. Как уже упоминалось выше, ЛГИС «Диалект» была изначально ориентирована на картографирование диалектной лексики. Позже система была оснащена лексикографическим модулем, позволяющим осуществлять представление лексики русских говоров в другой востребованной диалектологами форме: в виде информации, структурно и содержательно соответствующей словарной статье (Баранов, Жданова, Кожевников, Белых 2013: 171-175).

В лексикографическом модуле (вкладка «Словарь») может быть осуществлен поиск интересующей пользователя лексемы. Результатом является выборка из всех текстов и базы данных ЛГИС «Диалект» соответствующих слов, сопровождаемых статистической информацией, толкованием, указанием на время и место фиксации, ссылками на отрывки аудио- и текстовых записей, перейдя по которым можно услышать/увидеть употребление интересующего слова в контексте, а также указанием синонимичных и омонимичных лексем.

В лингвогеографическом модуле (вкладка «Карты») можно произвести выборку всех ответов на один из вопросов программы, на которой основана разметка корпуса, и вынести полученные слова на карту Удмуртии. Отметим, что в настоящий момент лексика русских говоров Удмуртии не отражена в специальном региональном диалектном словаре или атласе. Поэтому объединение возможностей лингвистического корпуса с потенциалом ЛГИС «Диалект» является важным шагом к созданию подобных работ о русских говорах Удмуртии.

Стоит обратить внимание и на тот факт, что ЛГИС «Диалект», созданная на основе материала русских говоров Удмуртии, имеет потенциал для отражения

языковых данных любой территории и при вводе лингвистической информации из других регионов есть возможность создания лингвистических карт этих регионов или более обширной территории.

2. ИСПОЛЬЗОВАНИЕ КОРПУСА

Известно, что корпусные методы дают возможность наблюдения за поведением интересующих языковых единиц в большом массиве данных. Современные электронные диалектные корпуса ориентированы на всестороннее изучение фактического материала, отражающего явления разных уровней языка. Корпус русских говоров Удмуртии создан с целью изучения словарного состава диалектного языка различными методами. В рамках данной работы рассмотрим применение инструментов корпуса на примере вопросов дополненного раздела «Питание» программы ЛАРНГ.

2.1. Словарь

Лексикографический модуль ЛГИС «Диалект» позволяет находить в Корпусе русских говоров Удмуртии интересующую лексику и рассматривать ее с учетом всех лингвистических и экстралингвистических факторов: семантики, стилистических и грамматических особенностей, системных связей (наличия омонимов и синонимов), времени и места фиксации, характеристик говорящего. Создание мультимедийной части корпуса дает возможность услышать контекст — отрывок аудиозаписи речи информанта.

Возможности лексикографического модуля позволяют представить весь комплекс данных о слове в русских говорах Удмуртии, что помогает в исследовании лексического состава диалектного языка в нашем регионе. Например, лексема *ярушник* отмечена в корпусе в значении 'хлеб из пшеничной муки', 'хлеб из овсяной муки', 'хлеб из ржаной муки'. В такой ситуации анализ контекстов из корпуса, которыми сопровождается словарная статья, помогает понять, насколько точны данные дефиниции. В других случаях анализ контекстов помогает прийти к выводу о синкретичности семантики слова.

Корпусный подход позволяет исследовать не только семантические, но и грамматические особенности фиксируемых слов. Например, глагол *печь* в русских говорах Удмуртии отмечен в вариантах *пекчи*, *пекти*, *печти* и др. В процессе сбора данных для ЛАРНГ фиксировалась в основном начальная форма, однако интересна и парадигма этого глагола. Употребление личных форм, фиксируемых в корпусе (*пекём*, *пекошь* и т. д.), позволяет восстановить и проанализировать парадигму этого глагола.

Создание Корпуса русских говоров Удмуртии и введение новых вопросов для фиксации в текстах и аудиозаписях корпуса лексики, находящейся за пределами ЛАРНГ, значительно расширило круг слов, охватываемых словарным модулем. Например, раздел «Питание» дополнен 31 вопросом, благодаря чему на основе лексикографического модуля ЛГИС «Диалект» могут быть созданы словарные статьи для десятков новых слов, например:

Травяночка — сладкий напиток из солода (Воткинский р-н, с. Степаново, 1982 г.): *Бражку делали, пиво варили, травяночку. Пиво-то погорчае, а травяночка место квасу, из солода, густая она, сладенькая;*

Гречушник — хлеб из гречихи (Увинский р-н, с. Удугучин, 1982 г.): *Раньше гречушники пекли, караваи из гречихи пекли;*

Налёвка — начинка для ватрушек (Балезинский р-н, д. Киренки, 1981 г., Каракулинский р-н, д. Юньга, 1976 г., с. Арзамасцево, 1979 г., с. Аракулино, 1976 г., Кезский р-н, с. Кулига, 1978 г., Киясовский р-н, с. Киясово, 1974 г.): *Шанежки делаю небольшойё, наливаю творог, налёвка картонна али сметанна...* (Кезский р-н, с. Кулига, 1978 г.), *Налёвка — картошка, разведённая на шаньги* (Киясовский р-н, с. Киясово, 1974 г.).

Во многих случаях данные корпуса подтверждают отсутствие в русских говорах Удмуртии специфических диалектных наименований для определенных реалий, распространение общепотребительных названий: *есть, пить, хлеб, квас, пиво* и т. д. Названия некоторых блюд не встречаются в корпусе ни разу (салат, тюря (кушанье из накрошенного в квас, воду хлеба), манная каша и др.), что может свидетельствовать о нетипичности данных кушаний для местной традиционной кулинарной культуры. В некоторых случаях диалектная лексема оказывается повсеместно распространенной (*пировать* 'пьянствовать', *скать* 'раскатывать тесто', *болтушка* 'суп с мукой'), и ее нецелесообразно представлять на карте. Такие слова отражены в виде словарных статей с указанием всех мест, где они были записаны. Отдельные слова отмечены лишь в одном населенном пункте или в двух-трех, расположенных рядом, что делает их локализмами, которые целесообразно фиксировать лишь в лексикографическом плане. Например, название ржаного хлеба *челтан / чолтан* отмечено в трех населенных пунктах на севере республики (Балезинский р-н, с. Андреевцы, 1981 г.; Кезский р-н, с. Кулига, 1977 г., д. Юклята, 1978 г.), поэтому при условии отсутствия в других говорах диалектных противопоставлений данную лексему можно зафиксировать лишь в лексикографическом источнике.

2.2. Карты

Лингвогеографический модуль ЛГИС «Диалект» позволяет создавать карты русских говоров Удмуртии, отражающие лексику, отмеченную в корпусе и выбранную пользователем. Существующая разметка дает возможность строить прежде всего лексические, семантические и словообразовательные карты. Пользователь указывает параметры, на основании которых происходит выбор картографируемых единиц, цвет и форму значка, после чего выбранные слова значками автоматически отражаются на масштабируемой карте (Яндекс-карты), сопровождаемой легендой, указывающей значения символов.

Изначально лингвогеографический модуль ЛГИС «Диалект» был ориентирован на картографирование лексики, собранной для ЛАРНГ. Этот материал, зафиксированный в 1990 – 2000-х годах, отражает состояние русских говоров Удмуртии на рубеже XX – XXI веков. Однако в этот период значительная часть традиционной

диалектной лексики уже была утрачена под влиянием известных лингвистических и экстралингвистических факторов, таких как изменчивость словарного состава языка, повышение уровня образованности сельского населения, уход в прошлое многих традиционных крестьянских видов деятельности и т. д.

Представление на картах лексических данных, извлеченных из Корпуса русских говоров Удмуртии, который содержит материал 1970 – 1980-х годов, значительно обогащает лингвистические карты, показывает иную картину диалектного членения на данной территории. Яркой иллюстрацией этого служит сравнение карт, составленных по одному и тому же вопросу программы ЛАРНГ без учета данных корпуса и с представлением этих данных.

Так, карта, составленная по вопросу «Пареные овощи» программы ЛАРНГ с привлечением данных корпуса, в целом подтверждает распространение на всей территории Удмуртии обозначений *парёнка* / *парёнки* и преобладание первого наименования в западной, а второго в восточной части республики (рис. 2, 3). Стоит заметить также, что материал корпуса не фиксирует ни одного употребления наименований *парево*, *паренина*, *пареница*, *паренье*, записанных студентами-собирающими материала для ЛАРНГ в различных населенных пунктах региона без контекстов. На этом основании такие данные можно считать недостоверными и не представлять на карте.



Рис. 2. Карта к вопросу «Пареные овощи» программы ЛАРНГ, составленная с привлечением данных из корпуса русских говоров Удмуртии 2-й пол. XX – нач. XXI в.



Рис. 3 Карта к вопросу «Пареные овощи» программы ЛАРНГ, составленная только на основе данных, собранных в Удмуртии в к. 1990 – нач. 2000-х гг. методом опроса по программе ЛАРНГ

Сравнение карт, составленных по вопросу «Кисель из муки», созданных на основе разных данных — материалов для ЛАРНГ (рис. 4) и сведений из корпуса (рис. 5), показывает, насколько данные корпуса могут уточнить и дополнить сведения о диалектных обозначениях в русских говорах Удмуртии. Наименования с основой *мучн-*, записанные для ЛАРНГ, являются единичными, не подтверждены контекстами и, скорее всего, взяты недобросовестными собирателями из числа примеров, приведенных в вопроснике ЛАРНГ в качестве возможных ответов. Обозначение *зырет* (*зылет*) является заимствованием из удмуртского языка, в котором имеет несколько иное значение, и не может считаться характерным для традиционных русских говоров Удмуртии. В то же время данные корпуса подтверждают бытование на северо-востоке республики лексемы *повалиха*, а также расширяют ареал обозначения *заваруха*.



Рис. 4. Карта по вопросу «Кисель из муки» программы ЛАРНГ, составленная только на основе данных, собранных в Удмуртии в к. 1990 – нач. 2000-х гг. методом опроса



Рис. 5. Карта по вопросу «Кисель из муки» программы ЛАРНГ, составленная только на основе данных корпуса русских говоров Удмуртии 2-й пол. XX – нач. XXI в.

Еще одним ярким примером важности материалов корпуса для картографирования лексики русских говоров Удмуртии являются карты, составленные по вопросу «Приспособление для раскатывания теста». Вариант, визуализирующий данные, полученные только в ходе экспедиций по программе ЛАРНГ (рис. 6), показывает большое разнообразие в основном единичных лексмах, разбросанных по всей территории региона. В то же время карта, дополненная материалами корпуса (рис. 7), указывает на распространение слова *пирожник* обширным ареалом в восточной части республики. Возможно, к 1990-м годам XX века, когда начался сбор материала для ЛАРНГ, эта лексема во многих говорах была утрачена, вытеснена общеупотребительным *скалка*, но для изучения традиционных народных говоров наличие такого ареала является показательным. Отметим также, что в корпусе нет ни одной фиксации лексемы *каталка*, *каток*, *валик*, *валёк*, в то же время эти слова приведены в качестве примерных ответов на данный вопрос в программе ЛАРНГ. Это, к сожалению, дает нам еще один повод усомниться в достоверности материалов, собранных в 1990 – 2000-х годах. В корпусе находим обозначения *скальник*, *сучильник* и *сочник*, которые во время экспедиций последних десятилетий не фиксировались.

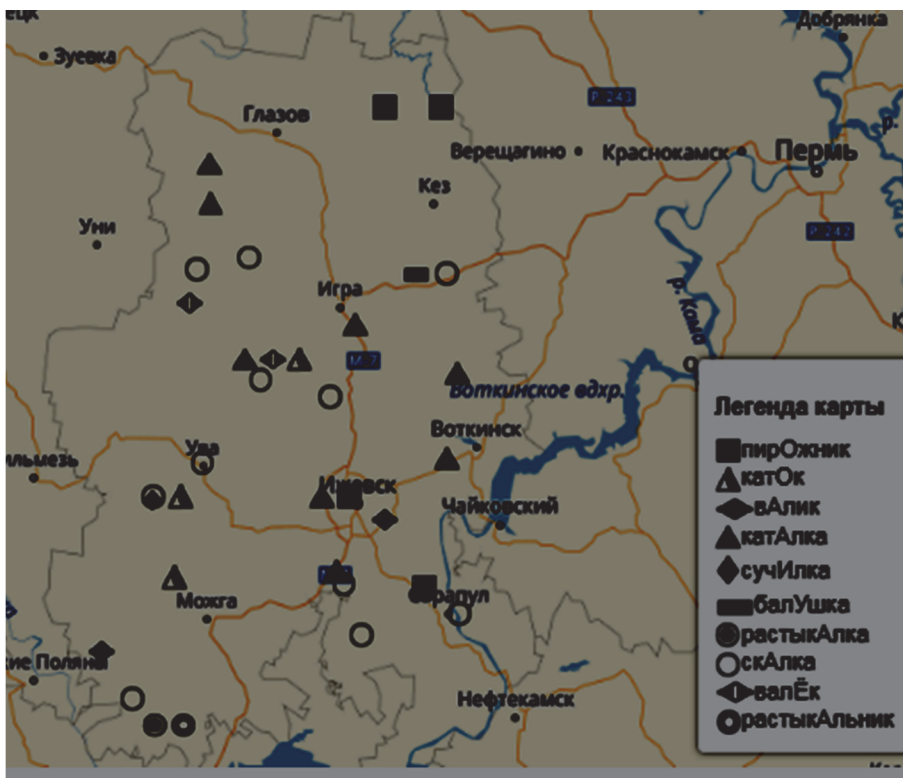


Рис. 6. Карта по вопросу «Приспособление для раскатывания теста» программы ЛАРНГ, составленная только на основе данных, собранных в Удмуртии в к. 1990 – нач. 2000-х гг. методом опроса

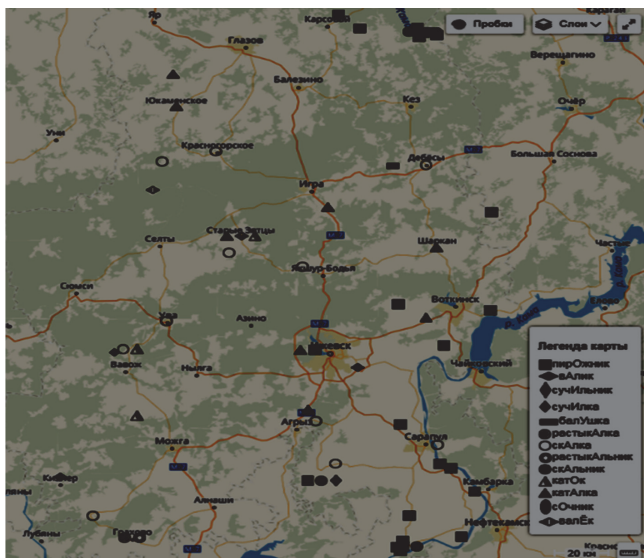


Рис. 7. Карта по вопросу «Приспособление для раскатывания теста» программы ЛАРНГ, составленная с привлечением данных из корпуса русских говоров Удмуртии 2-й пол. XX – нач. XXI в.

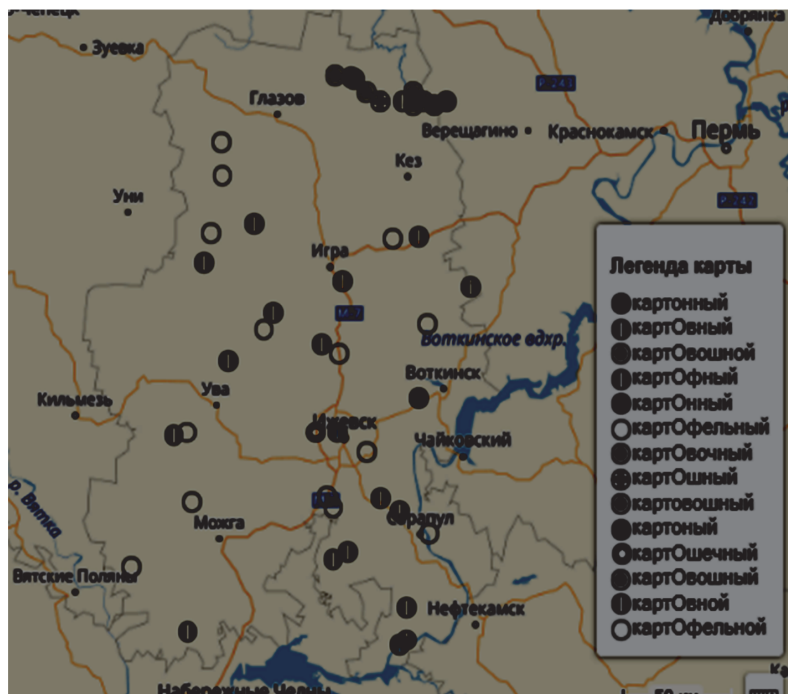


Рис. 8. Словообразовательная карта к вопросу «Приготовленный из картофеля» программы ЛАРНГ

Примером достаточно показательной словообразовательной карты может служить карта к вопросу «Приготовленный из картофеля» (рис. 8). Данные корпуса и материалы, собранные для ЛАРНГ, в совокупности показывают, что наиболее распространенными в Удмуртии оказываются образования с усечением последнего слога производящей основы *картофель* (*картовный*, *картофный*), на северо-востоке отмечено образование с полным отсутствием (или полной ассимиляцией) согласного перед суффиксом *-н-* (*картонный*).

Вопросы, внесенные разработчиками Корпуса русских говоров Удмуртии в ЛГИС «Диалект» для расширения программы ЛАРНГ с целью фиксации большего количества диалектных наименований, существующих в говорах данного региона, также могут лечь в основу лингвистической карты. Так, была составлена карта к вопросу «Большая ложка для вынимания пельменей», внесенному в раздел ЛАРНГ «Питание» (рис. 9). Этот вопрос позволил не только отметить в словарном модуле ряд новых лексем, но и показать на карте их территориальное разграничение. На картах могут быть визуализированы и ответы на другие вопросы (например, названия лепешек из гнилой мороженой картошки, травы и т. д. (рис. 10)), которые повсеместно составляли основу рациона крестьян в годы Великой Отечественной войны. Конечно, картографирование этих слов не имеет отношения к исследованию традиционных русских говоров, но является показательным в плане изучения истории и быта населения Удмуртии в XX веке.

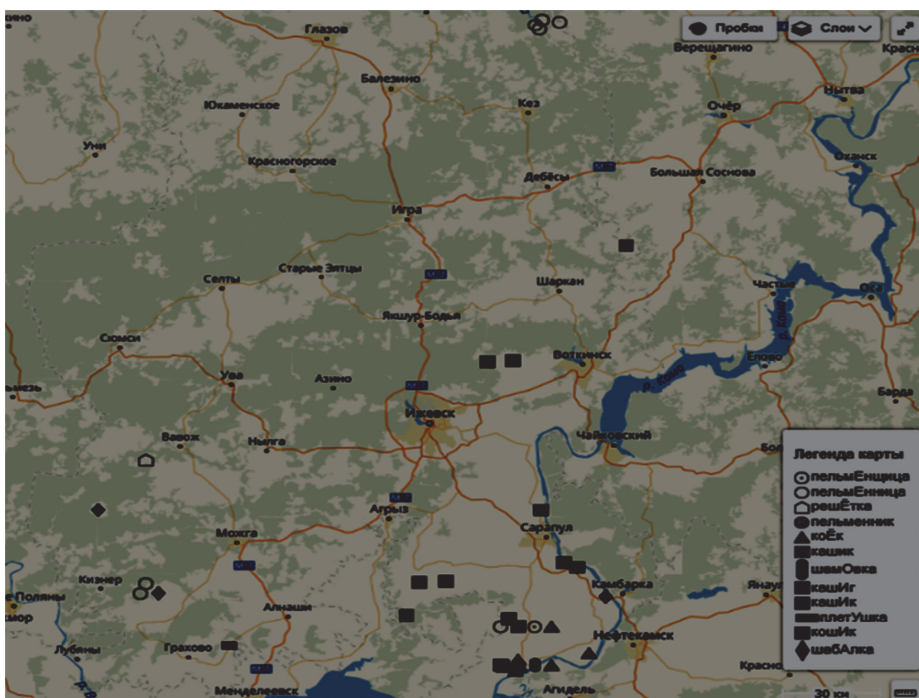


Рис. 9. Карта к вопросу «Большая ложка для вынимания пельменей» вне программы ЛАРНГ (только материалы корпуса)

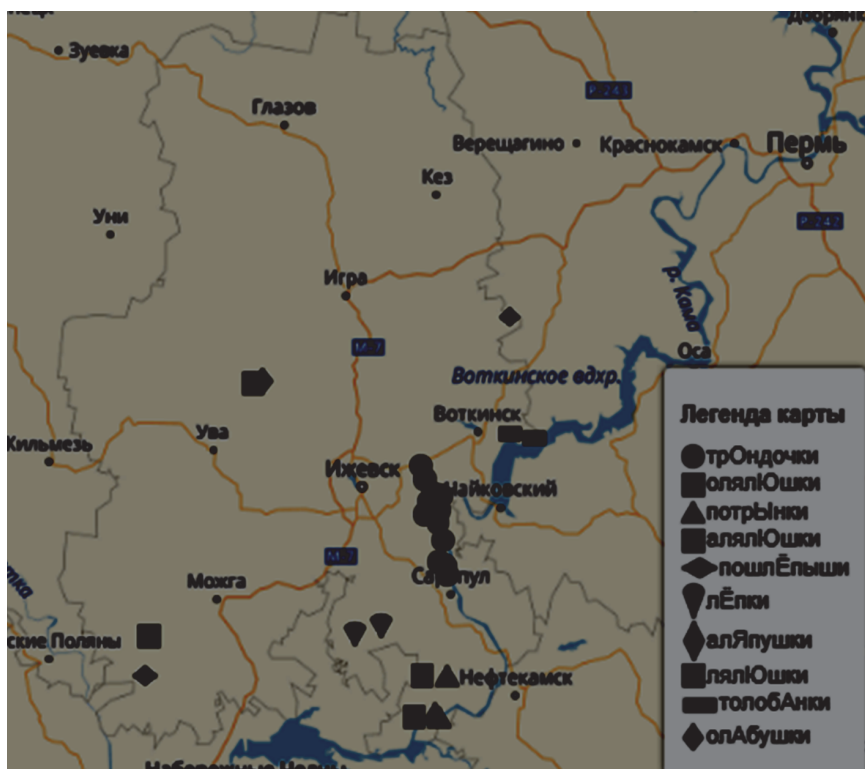


Рис. 10. Карта к вопросу «Лепешки из гнилой мороженой картошки, травы и т. д.» вне программы ЛАРНГ (только материалы корпуса)

3. ВЫВОДЫ

Мультимедийный корпус русских говоров Удмуртии представляет собой электронный лингвистический ресурс, направленный на решение проблем региональной диалектологии. Корпус содержит языковой материал, собранный более чем в полутораста населенных пунктах в течение нескольких десятилетий разными методами: прямым опросом, полевой транскрипционной записью спонтанной речи, с помощью звукозаписывающей аппаратуры, — что обеспечивает репрезентативность данных. При создании корпуса были применены современные методы обработки письменных текстов и аудиозаписей, что позволило внести в базу данных материалы в их исходной форме.

Разметка корпуса ориентирована на изучение лексического состава русских говоров Удмуртии. И в письменных, и в аудиотекстах отмечены слова, соответствующие вопросам программы ЛАРНГ, а также вопросам, сформулированным разработчиками корпуса. Такой подход позволил зафиксировать в материалах корпуса диалектные слова, относящиеся к разнообразным темам, в том числе лексемы, ранее не фиксировавшиеся в диалектологических источниках.

Корпус русских говоров Удмуртии создан прежде всего для лексикологических исследований, в связи с этим корпусный менеджер ориентирован не только на выполнение стандартных задач — представление слов в контексте в сопровождении статистических данных, но и на применение таких востребованных современной диалектологией методов исследования, как лексикографический и картографический. Совокупность этих методов позволяет дать полную картину русских говоров Удмуртии.

БИБЛИОГРАФИЯ

- БАРАНОВ, В.А., ЖДАНОВА, Е.А., КОЖЕВНИКОВ, Д.Б., БЕЛЫХ, А.А. (2013): «Лингвогеографическая система “Диалект”: история создания, новые возможности, технологические решения, демонстрация данных». *Интеллектуальные системы в производстве*. № 1 (21). С. 171-175.
- ВЕРНЯЕВА, Р.А. (2018): «Использование лингвогеографической системы “Диалект” в учебном процессе (на примере курса “Технологии обработки текста и звучащей речи”)». *Вестник ИжГТУ им. М.Т. Калашникова*. Т. 21. №1. С. 104-108.
- ЖДАНОВА, Е.А. (2016): «Проект Корпуса диалектных текстов русских говоров Удмуртии», *Интеллектуальные системы в производстве*, Изд-во ИжГТУ им. М.Т. Калашникова, Ижевск, 4. С. 137-141.
- ЗЕМИЧЕВА, С.С., ИВАНЦОВА, Е.В. (2018): «Проект создания Томского диалектного корпуса в свете тенденции развития корпусной лингвистики», *Сибирский филологический журнал*, 3. С. 192-205.
- КАЧИНСКАЯ, И.Б., СИЧИНАВА, Д.В. (2015): *Диалектный подкорпус сегодня*. Труды Ин-та рус. яз. им. В.В. Виноградова. Т. 6. Москва.
- КРЮЧКОВА, О.Ю. Электронный корпус русской диалектной речи и принципы его разметки [Электронный ресурс]. Режим доступа http://sarteorlingv.narod.ru/dialekt/elektr_korpus.html#ftn0 (дата обращения 01.12.2019).
- ЛЕТУЧИЙ, А.Б. (2005): *Корпус диалектных текстов: задачи и проблемы*. Национальный корпус русского языка: 2003–2005. Индрик, Москва.
- МАРТЬЯНОВА, В. Н., СЕМАКИНА, А. О., ШЕПЕЛЕВА, С. В. (2014): *Хрестоматия по русской диалектологии*. Изд-во Глазовского государственного педагогического института. Глазов.
- ПОПОВ, И. А. (1994): *Программа собирания сведений для лексического атласа русских народных говоров*. Санкт-Петербург.
- Руководство по программе Audacity [Электронный ресурс]. Режим доступа <http://beginwithsoftware.com/videoguides/audacity-guide-rus.html> (дата обращения 01.12.2019).
- ЮРИНА, Е.А. (2011): «Томский диалектный корпус: в начале пути», *Вестник Томского государственного университета. Филология*, 2 (14). С. 58-63.

BIBLIOGRAPHY

BARANOV, V.A., ZHDANOVA, E.A., KOZHEVNIKOV, D. B., BELYKH, A. A. (2013): “Lingvogeograficheskaya sistema “Dialekt”: istoriya sozdaniya, novyye vozmozhnosti, tekhnologicheskiye resheniya, demonstratsiya dannykh”. *Intellektual'nyye sistemy v proizvodstve*. # 1 (21). pp. 171-175.

KACHINSKAYA, I.B., SICHINA, D.V. (2015): *Dialectnyi podcorpus segodnya*. Tr. In-ta rus.yaz. im. V. V. Vinogradova. T. 6. Moskva.

KRYUCHKOVA, O.Yu. Elektronnyy korpus russkoy dialektnoy rechi i printsipy yego razmetki [Elektronnyy resurs]. Rezhim dostupa: http://sarteorlingv.narod.ru/dialekt/elektr_korpus.html#ftn0 (data obrashcheniya 01.12.2019).

LETUCII, A.B. (2005): *Corpus dialectnykh tekstov: zadachi i problemy*. Nacional'nyi korpus russkogo yazyka: 2003–2005. Indrik, Moskva.

MARTYANOVA, V.N., SEMAKINA, A.O., SHEPELEVA, S.V. (2014): *Khrestomatiya po russkoy dialektologii*. Izdatel'stvo Glazovskogo gosudarstvennogo pedagogicheskogo instituta. Glazov.

POPOV, I.A. (1994): *Programma sobiraniya svedeniy dlya leksicheskogo atlasa russkikh narodnykh govorov*. Sankt-Peterburg.

Rukovodstvo po programme Audacity [Elektronnyy resurs]. Rezhim dostupa <http://beginwithsoftware.com/videoguides/audacity-guide-rus.html> (access date 01.12.2019).

VERNYAEVA, R.A. (2018): “Ispol'zovaniye lingvogeograficheskoy sistemy “Dialekt” v uchebnom protsesse (na primere kursa “Tekhnologii obrabotki teksta i zvuchashchey rechi)””. *Vestnik IzhGTU im. M.T. Kalashnikova*. T. 21. #1. pp. 104-108.

YURINA, E.A. (2011): “Tomskiy dialektnyy korpus: v nachale puti”, *Vestnik Tomskogo gosudarstvennogo universiteta, Filologiya*, 2 (14), pp. 58-63.

ZEMICHEVA, S.S., IVANOVA, E.V. (2018): “Proect sozdaniya Tomskogo dialectnogo korpusa v svete tendencii razvitiya korpusnoi ligvistiki”, *Sibirskii philologicheskii zhurnal*, 3, pp. 192-205.

ZHDANOVA, E.A. (2016): “Proect Korpusa dialectnykh tekstov russkikh govorov Udmurtii”, *Intellektual'nye sistemy v proizvodstve*, Izd-vo IzhGTU im. M.T. Kalashnikova, Izhevsk, 4, pp. 137-141.