

ПЕДАГОГИЧЕСКИЕ НАУКИ

УДК 004.934

DOI 10.22213/2413-1172-2018-1-99-103

Е. А. Жданова, кандидат филологических наук, доцент, ИжГТУ имени М. Т. Калашникова

ЭЛЕКТРОННЫЙ КОРПУС РУССКИХ ГОВОРОВ УДМУРТИИ: НОВЫЕ ВОЗМОЖНОСТИ*

Реализация проекта

«Корпус русских говоров Удмуртии»

В 2017 г. началась реализация проекта «Русские говоры Удмуртии: корпус диалектных текстов второй половины 20 – начала 21 века», описание которого было представлено в [1]. В рамках этой работы на основе лингвогеографической информационной системы «Диалект» (ЛГИС «Диалект») начато создание интернет-ресурса для работы с диалектными текстами, записанными в Удмуртской Республике.

Сейчас ЛГИС «Диалект» представляет собой единственную в своем роде систему хранения и обработки диалектных лексических данных отдельно взятого региона. В базе данных системы содержится материал 3 типов:

1) лексический материал, собранный в Удмуртии по Программе лексического атласа русских народных говоров (ЛАРНГ) [2] в 1990–2000-х гг. методом интервьюирования диалектоносителей;

2) скан-копии письменных текстов, являющихся записями речи диалектоносителей, сделанными в ходе бесед с ними или наблюдения за их разговором. Эти записи были произведены в Удмуртии в 1970–1980-е гг.;

3) аудиозаписи речи информантов, сделанные в ходе диалектологических экспедиций разных лет [3].

Языковой материал из всех районов республики, собранный в таком объеме и разными методами, обеспечивает полноту и объективность выборки, позволяет производить необходимый анализ и делать обоснованные выводы относительно лексического состава и диалектного членения русских говоров Удмуртии.

Сейчас ЛГИС «Диалект» позволяет анализировать диалектный лексический материал

с применением следующих лингвистических методов:

1) метод лингвистической географии – составление лингвистических карт Удмуртии, отражающих распространение диалектных слов;

2) лексикографический метод – представление записанных слов в виде словарных статей, включающих толкование, указание на территорию распространения, грамматические и стилистические пометы, примеры употребления слов;

3) методы корпусной лингвистики – получение статистических данных о распространенности слов, извлечение примеров их употребления и сочетаемости.

В ходе работы над проектом ЛГИС была перенесена на новую базу. Измененный режим доступа: <http://dialect.manuscripts.ru>. Кроме того, для осуществления задачи представления на карте лексики, отмеченной в текстах корпуса, разработчикам потребовалось изменить формат хранения скан-копий страниц из тетрадей с записями речи диалектоносителей: если на предыдущем этапе работ папки с файлами в меню «Тексты» ЛГИС «Диалект» соответствовали тетрадам с записями, сделанными во время диалектологических экспедиций, которые могли охватывать несколько населенных пунктов одного района, то на данном этапе эти папки были разделены в соответствии с количеством населенных пунктов, т. е. одна папка содержит записи, сделанные только в одном пункте.

Модернизация лингвогеографического модуля корпуса

В основе ЛГИС «Диалект» лежит Программа ЛАРНГ, по вопросам которой составляются лингвистические карты для атласа русских говоров Удмуртии. Именно к вопросам Программы ЛАРНГ «привязана» информация, содержащаяся в базе данных ЛГИС «Диалект».

Одной из целей создания корпуса диалектных текстов, записанных на территории Удмуртии, было составление карт для атласа русских говоров Удмуртии на основе расширенной базы данных. Создание корпуса представляется необходимым этапом данной работы, так как диалектный лексический материал, собранный для атласа русских говоров Удмуртии по Программе ЛАРНГ и внесенный в базу данных ЛГИС «Диалект», охватывает не все районы республики и относится к периоду 1990–2000-х гг. Лексический материал, извлеченный из корпуса в результате разметки, дополняет эти данные лексическими единицами, зафиксированными в большем количестве населенных пунктов в более ранний период, когда многие традиционные наименования еще не были утрачены.

Например, по разделу «Питание» Программы ЛАРНГ на территории Удмуртии собран материал в 19 из 25 районов. В каждом районе записи делались в 1-3 населенных пунктах, опрашивались по 3 информанта в каждом. Во время экспедиций 1970–1980-х гг. собиратели зачастую обходили несколько близлежащих деревень, записывая речь десятков жителей пожилого возраста. Таким образом, информация, извлеченная из текстов корпуса, значительно обогащает лингвистические карты.

Благодаря внесенным и размеченным текстам значительно расширилась география лексики тематической группы «Питание». Это произошло благодаря разметке записей, сделанных в ходе диалектологических экспедиций студентов и сотрудников Удмуртского государственного университета в населенные пункты Каракулинского, Селтинского и Сюмсинского районов, которые не были обследованы по Программе ЛАРНГ, а также за счет внесенных текстов, записанных в Глазовском и Ярском районах сотрудниками Глазовского государственного педагогического института и опубликованных в хрестоматии по русской диалектологии (авт.-сост. В. Н. Мартынова, А. О. Семакина, С. В. Шепелева) [4].

Наибольшее количество записей сделано в Каракулинском районе (экспедиции 1976, 1979 и 1986 гг.), в результате чего из размеченных текстов извлечены ответы на большинство вопросов раздела «Питание» Программы ЛАРНГ. Анализ этих ответов позволяет подтвердить сделанные ранее предположения и выводы о большом количестве общенародных слов в данном тематическом разделе, распространении литературной лексики и заимствований (*котлета* в речи неграмотной женщины 1912 г.р. (запись

1976 г.), многократно встречающиеся наименования *бульон, суп* и т. д.), незначительном проникновении удмуртских этнографизмов (*пельмени, кумышка*).

Полученные данные позволили дополнить созданные ранее карты, на которых была отражена только информация, собранная на основе вопросника ЛАРНГ. Например, данные, полученные в результате разметки текстов корпуса, дополнили карту, составленную для статьи 2016 г. [5], отражающую распространение лексем *завариха* и *заваруха*, обозначающих кушанье из муки с водой, в русских говорах Удмуртии, населенными пунктами в Селтинском и Каракулинском районах (*завариха* – Каракулинский р-н, с. Каракулино (1976); Селтинский р-н, с. Валамаз (1988); *заваруха* – Каракулинский р-н, с. Арзамасцево (1979); с. Боярка (1988)). Эти точки в целом подтверждают ареалы, отраженные на созданной ранее карте, делают ее более точной.

В базе данных ЛГИС «Диалект», основанной на вопроснике ЛАРНГ, глагол *пировать* – «пьянствовать» – отмечен в с. Какмож Вавожского р-на, с. Кукуй Воткинского р-на, с. Смольники Дебесского р-на, с. Бемыж Кизнерского р-на, с. Первомайский Киясовского р-на, с. Девятово Сарапульского р-на. Разметка диалектных текстов, проводимая в рамках проекта, дополняет этот список следующими пунктами: с. Арзамасцево, д. Ендовка Каракулинского р-на (1979); с. Кулига Кезского р-на (1979); д. Курья Красногорского р-на (1981); д. Козлово Сарапульского р-на (1972). *Она вечером пировала в деревне, брагу пила, а наутро мертва лежит.* Таким образом, ареал данной лексемы предстает совершенно иным, нежели до разметки.

Модернизация лексикографического модуля корпуса

В рамках проекта происходит расширение вопросника ЛАРНГ, лежащего в основе ЛГИС «Диалект», так как в ходе анализа текстов корпуса обнаруживаются диалектные наименования, вопросов для которых нет в программе. В ходе работы было принято решение не создавать для таких вопросов специальный раздел, а дополнять уже существующие тематические части.

Например, в ходе разметки текстов по разделу «Питание» Программы ЛАРНГ в Каракулинском районе, ранее не входившем в число территорий, где проходил сбор материала по указанному разделу, был обнаружен ряд диалектных слов, относящихся к теме «Питание»,

но не соответствующих вопросам данного раздела вопросника ЛАРНГ.

Кула́га – кушанье из ржаной муки или солода с квасом или водой (с. Боярка (1986); с. Каракулино, д. Юньга (1976)). *Стряпали кулагу* – замешают из ржаной муки с квасом, жиденькая, вроде киселя.

Салма́ – похлебка из клецок с капустой (с. Каракулино, д. Юньга (1976)). *Салму стряпали из кислого теста: разоскут, нарежут катышками.., варят и едят со сметаной с капустой*. По данным Словаря русских народных говоров, это слово зафиксировано в вятских говорах и на территории Башкирии [6, вып. 36, с. 63]. Таким образом, это слово в целом характерно для нашего региона.

Мали́нки – небольшие фигурные мучные изделия, которые готовили на Масленицу (с. Каракулино (1976)). *В Масленку стряпали малинки-те: масла положат, яичек побьют, рассекают сочни, режут ленточками, гнут, как надо, вот и малинки*.

Налёвка – начинка для ватрушек (с. Арзамасцево (1979); с. Каракулино, д. Юньга (1976)). *Толкли налёвку на шаньги: истолчешь картошку, потом сметаны нальешь; налёвку делали: протокваши нальешь, муки клали, яиц бьешь, потом масла*. По данным СРНГ, в таком значении слово *налёвка* зафиксировано в Свердловской области [6, вып. 20, с. 9].

Ватрушки с *налёвкой* в Удмуртии называют *налёвные шаньги* (с. Арзамасцево (1979); с. Каракулино, с. Первомайское, д. Юньга (1976)). Слово *шаньга* (*шанежка*) в качестве обозначения ватрушки широко распространено на территории Удмуртии. А словом *ватрушка* в Каракулинском районе называют булку, выпечку из пресного теста (с. Арзамасцево (1979); д. Боярка (1986)).

Анализируя приведенные диалектные слова, можно отметить, что в основном это названия блюд традиционной народной или местной кухни, которые не имеют территориально закреплённых различий или соответствий в других регионах, поэтому не представляют интереса с точки зрения лингвогеографического подхода, на который ориентирована Программа ЛАРНГ. Однако в лексикографическом аспекте фиксация такой лексики является важным шагом к доскональному словарному описанию диалектного языка нашего региона.

Также в диалектном языке XX столетия нашли отражение реалии, связанные с тяжёлыми периодами войн и голода. Такие слова существуют в языке короткое время, совре-

менные сельские жители вряд ли помнят эти обозначения, однако для полной картины функционирования русских говоров Удмуртии такие лексемы также представляют большую ценность.

Потры́нки – лепешки из мороженой картошки (с. Арзамасцево (1979); с. Каракулино, д. Юньга (1976)). *Потрынки* – лепешки из картошки гнилой.

Олялюшки (*люлюшки*) – лепешки из травы или гнилой картошки (Каракулинский р-н, д. Боярка (1986); д. Юньга (1976); Селтинский р-н, с. Малый Валамаз (1988)). *Лебедные люлюшки* делали.

Лебедной год – голодный год (с. Арзамасцево (1979); с. Каракулино (1976)). *В войну потрынки пекли, стряпали в лебедной год*.

Ряд аналогичных примеров из других районов Удмуртии приведен в [7].

Дефиниции данных слов введены в качестве дополнительных вопросов в программу, лежащую в основе ЛГИС «Диалект». В результате такой работы по данным, относящимся к этим вопросам, возможно составление лингвистических карт. Кроме того, ответы на указанные вопросы и сопутствующую информацию (дата и место фиксации, сведения об информантах, примеры употребления) можно увидеть в лексикографическом модуле ЛГИС «Диалект» (вкладка «Словарь»), данные лексемы наравне с остальными будут включены в словарь русских говоров Удмуртии.

Помимо этого разметка текстов дает возможность показать в контексте слова, уже отмеченные в Удмуртии при сборе материала по Программе ЛАРНГ, но записанные без примеров употребления, что ставило под сомнение точность и достоверность фиксации. Например, в Каракулинском р-не отмечены слова *жучить* – «есть жадно и много» (вопрос № 19032 Программы ЛАРНГ): *«Детдомовски ребятишки вовсе голодны ходят. Сыру капусту жучат»*, и *парёнки* – «пареные овощи» (общее название) (вопрос № 19109 Программы ЛАРНГ): *«Парёнки ели: раньше калегу порежешь в вольную печь, свёклу нарежешь, поставишь сваришь»*. До фиксации приведенных примеров в ЛГИС «Диалект» не было контекстов, подтверждающих употребление этих диалектных наименований. Примеры использования слова в речи являются важной частью словарной статьи, необходимой составляющей его лексикографического описания. Особенно существенно это для диалектных слов, значение которых во многом определяется по контексту, так как исследова-

тель зачастую не имеет других данных о семантике и особенностях бытования диалектных наименований. Поэтому извлечение контекстов при разметке записей диалектной речи значительно расширяет возможности ее лексикографического представления и наши знания о диалектной лексике Удмуртии.

Использование данных корпуса русских говоров Удмуртии в научной и методической работе

В проекте корпуса было указано и на возможность более широкого использования содержащейся в нем информации. Уже в 2017 г. эта перспектива была реализована.

На кафедре «Лингвистика» ИжГТУ имени М. Т. Калашникова было опубликовано 2 электронных учебно-методических пособия по дисциплинам «Лингвистическая география в России» [8] и «Электронные издания» [9], в которых ЛГИС «Диалект» и созданный на ее основе корпус выступают в качестве одного из средств обучения студентов направления «Фундаментальная и прикладная лингвистика». Также ЛГИС «Диалект» используется при преподавании дисциплины «Русский язык и культура речи» студентам нелингвистических направлений [10].

В июне 2017 г. на той же кафедре написана и успешно защищена выпускная квалификационная работа студентки направления «Фундаментальная и прикладная лингвистика» Е. Н. Мосовой «Диалектная лексика русских говоров Удмуртии в речи школьников и студентов», для которой лексикографический модуль ЛГИС «Диалект» послужил источником информации о лексическом составе русских говоров Удмуртии [11]. В предыдущие годы было написано 3 дипломных работы студентов специальности «Теоретическая и прикладная лингвистика» на материале лингвогеографического модуля ЛГИС «Диалект». Модернизация системы значительно расширила возможности ее использования в учебно-методической работе.

Как было указано в проекте, корпус текстов, представляющих собой записи живой речи жителей сельской местности, родившихся преимущественно в конце XIX – начале XX века, является важным источником не только лингвистической, но и исторической информации. В текстах содержатся многочисленные рассказы о быте крестьян начала XX века (традиционной кухне, одежде, предметах обихода, обычаях), тяжелой крестьянской жизни в период коллективизации, Гражданской и Великой Отечественной войн, а также воспоминания о важных исторических событиях, например, об освобождении

«баржи смерти» около с. Гольяны на Каме в 1918 г. В Завьяловском районе записан рассказ Анны Андреевны Пушиной, сестры героя ВОВ Федоры Пушиной, о ее семье и жизни знаменитой сестры.

Таким образом, создаваемый корпус представляет интерес не только для лингвистов, в том числе педагогов, но и для историков, культурологов, краеведов.

К перспективам развития корпуса относится разметка аудиоматериалов с возможностью их подключения к лингвогеографическому и лексикографическому модулям ЛГИС «Диалект», а также разработка инструментария для фонетического и грамматического анализа письменных и аудиозаписей.

Библиографические ссылки

1. *Жданова Е. А.* Проект Корпуса диалектных текстов русских говоров Удмуртии // Интеллектуальные системы в производстве. Ижевск : Изд-во ИжГТУ имени М. Т. Калашникова, 2016. № 4. С. 137–141.
2. Программа собирания сведений для лексического атласа русских народных говоров / отв. ред. И. А. Попов. СПб., 1994.
3. *Верняева Р. А., Жданова Е. А.* Цифровое представление звучащей диалектной речи в лингвогеографической информационной системе «Диалект» // Гуманитарное образование и наука в техническом вузе : сб. докл. Всерос. науч.-практ. конф. с международ. участием (Ижевск, 24–27 октября 2017 г.) [Электронный ресурс]. Ижевск : Изд-во ИжГТУ имени М. Т. Калашникова, 2017. С. 457–462.
4. Хрестоматия по русской диалектологии / авт.-сост.: В. Н. Мартыянова, А. О. Семакина, С. В. Шепелева. Глазов : ГГПИ, 2014.
5. *Жданова Е. А.* Субстантивная лексика тематической группы «Питание» в русских говорах Удмуртии (лингвогеографический и системный анализ) // Вестник Удмуртского университета. 2016. Т. 26. Серия «История и филология». Вып. 5. С. 11–20.
6. Словарь русских народных говоров / ИЛИ РАН. 2002–2016. URL: <https://iling.spb.ru/vocabula/srng/srng.html> (дата обращения: 16 декабря 2017 г.).
7. *Жданова Е. А.* Лексика тематической группы «Питание» в Корпусе диалектных текстов русских говоров Удмуртии // И. А. Бодуэн де Куртенэ и мировая лингвистика : международ. конф. «VI Бодуэновские чтения» (Казан. федер. ун-т, 18–21 окт. 2017 г.) : в 2 т. / под общ. ред. К. Р. Галиуллина, Е. А. Горобец, Д. А. Мартыянова, Г. А. Николаева. Казань : Изд-во Казан. ун-та, 2017. Т. 1. С. 91–93. URL: <https://kpfu.ru/ baudouin2017/sbornik-trudov-konferencii> (дата обращения: 16 декабря 2017 г.).
8. *Жданова Е. А.* Лингвистическая география в России : метод. указ. Ижевск, 2017. URL: <https://drive.google.com/open?id=18a6Vtp784u-L6Wuza5ee3VQcS nvGhwud> (дата обращения: 16 декабря 2017 г.).

9. Верняева Р. А. Электронные издания. Практические работы и рекомендации к их выполнению : метод. рекомендации к вып. практ. работ для студ. напр. «Фундаментальная и прикладная лингвистика» профиля «Теоретическая и прикладная лингвистика». Ижевск : Изд-во ИжГТУ имени М. Т. Калашникова, 2017. 28 с. URL: <https://drive.google.com/open?id=11BPWHBMY4zBRA0L0okYOik7oFYzZZx5Q> (дата обращения: 16 декабря 2017 г.).

10. Жданова Е. А. Использование современных информационных технологий при преподавании дисциплины «Русский язык и культура речи» студентам нелингвистических направлений ИжГТУ имени М. Т. Калашникова // Гуманитарное образование и наука в техническом вузе : сб. докл. Всерос. науч.-практ. конф. с междунар. участием (Ижевск, 24–27 октября 2017 г.) [Электронный ресурс]. Ижевск : Изд-во ИжГТУ имени М. Т. Калашникова, 2017. С. 699–702.

11. Жданова Е. А., Мосова Е. Н. Диалектная лексика русских говоров Удмуртии в речи школьников и студентов // Гуманитарное образование и наука в техническом вузе : сб. докл. Всерос. науч.-практ. конф. с междунар. участием (Ижевск, 24–27 октября 2017 г.) [Электронный ресурс]. Ижевск : Изд-во ИжГТУ имени М. Т. Калашникова, 2017. С. 383–387.

References

1. Zhdanova E. A. (2016). *Intellectualnye sistemy v proizvodstve* [Intelligent Systems in Manufacturing], no. 4, pp. 137-141 (in Russ.).

2. Popov I. A. (1994). *Programma sobiraniya svedeni idla lexicheskogo atlasa russkih narodnyh govorov* [The program of gathering information for the Lexical atlas of Russian folk dialects]. St. Petersburg (in Russ.).

3. Vernyaeva R. A., Zhdanova E. A. (2017). *Cifrovoe predstavlenie zvuchashchei rechi v lingvogeograficheskoi informacionnoi sisteme "Dialect"* [Digital representation of sounding dialect speech in the linguistic-geographical information system "Dialect"]. Proceedings of the *Gumanitarnoe obrazovanie i nauka v tehničeskom vuze* (ed. V. A. Baranov), pp. 457-462 (in Russ.).

4. Martyanova V. N., Semakina A. O., Shepeleva S. V. (2014). *Hrestomatia po russkoy dialektologii* [Collection of Russian dialectology]. Glazov (in Russ.).

5. Zhdanova E. A. (2016). *Vestnik Udmurtskogo universiteta* [Bulletin of Udmurt University], vol. 26: History and Philology, no. 5, pp. 11-20 (in Russ.).

6. *Slovar russkih narodnyh govorov* [Dictionary of Russian folk dialects], available at <https://iling.spb.ru/vocabula/srng/srng.html> (accessed December 16, 2017) (in Russ.).

7. Zhdanova E. A. (2017). *Lexika tematicheskoy gruppy "Pitanie" v korpuse dialektnykh textov russkih govorov Udmurtii* [Vocabulary of the theme group "Nutrition" in the corpus of Russian dialects of Udmurtia]. Proceedings of the *I. A. Baudouin de Courtenay i mirovaya lingvistika* (eds. K. R. Galiullin, E. A. Gorobec, D. A., Martyanov G. A. Nikolaev), pp. 91-93, available at <https://kpfu.ru/ baudouin2017/sbornik-trudov-konferencii> (accessed December 16, 2017) (in Russ.).

8. Zhdanova E. A. (2017). *Lingvisticheskaya geografija v Rossii* [Linguistic geography in the Russia], available at <https://drive.google.com/open?id=18a6Vtp784u-L6Wuza5ee3VQcSnnGhwud> (accessed December 16, 2017) (in Russ.).

9. Vernyaeva R. A. (2017). *Elektronnye izdania. Prakticheskie raboty i rekomendacii k ih vypolneniu* [Electronic publications. Practical work and recommendations for their implementation], available at <https://drive.google.com/open?id=11BPWHBMY4zBRA0L0okYOik7oFYzZZx5Q> (accessed December 16, 2017) (in Russ.).

10. Zhdanova E. A. (2017). *Ispolzovanie sovremennykh informacionnykh tehnologii pri prepodavanii discipliny "Russkii yazyk i kultura rechi" studentam nelingvisticheskikh napravlenii IzhGTU imeni M. T. Kalashnikova* [Use of modern information technologies in teaching of "Russian and speech culture" to non-linguistic fields students]. Proceedings of the *Gumanitarnoe obrazovanie i nauka v tehničeskom vuze* (ed. A. V. Baranov), pp. 699-702 (in Russ.).

11. Zhdanova E. A., Mosova E. N. (2017). *Dialektnaya lexika russkih govorov Udmurtii v rechi shkolnikov i studentov* [The vocabulary of Russian dialect of Udmurtia in the school children and students' speech]. Proceedings of the *Gumanitarnoe obrazovanie i nauka v tehničeskom vuze* (ed. V. A. Baranov), pp. 383-387 (in Russ.).