

ЦИФРОВАЯ ЛЕКСИКОГРАФИЯ DIGITAL LEXICOGRAPHY

Научная статья
УДК 398, 81'322
doi: 10.17223/22274200/37/3

Томский фольклорный корпус: от идеи до воплощения

Светлана Владимировна Волошина¹

¹ *Томский государственный университет, Томск, Россия, vsv1304@yandex.ru*

Аннотация. В статье представлена концепция создания Томского фольклорного корпуса на основе материалов студенческих фольклорных экспедиций 1960–2010 гг. в сёла Томской, Кемеровской, Новосибирской областей, Хакасии, Красноярского края и Алтайского края. Обосновывается необходимость разработки такого ресурса, возможные направления его использования, в том числе лексикографическое, указываются предполагаемые типы разметки материалов, необходимые метаданные.

Ключевые слова: региональный фольклор, корпусная фольклористика, электронный фольклорный корпус, Томский фольклорный корпус

Благодарности. Результаты были получены в рамках выполнения государственного задания Минобрнауки России, проект № FSWM-2025-0016.

Для цитирования: *Волошина С.В.* Томский фольклорный корпус: от идеи до воплощения // Вопросы лексикографии. 2025. № 37. С. 49–73. doi: 10.17223/22274200/37/3

Original article

Tomsk Folklore Corpus: From idea to implementation

Svetlana V. Voloshina¹

¹ *Tomsk State University, Tomsk, Russian Federation, vsv1304@yandex.ru*

Abstract. The article discusses the current problem of preserving and systematizing folklore material in the digital age. An overview of the existing electronic folklore corpora created and being created in Russia is presented. Despite the growing interest in corpus linguistics, electronic folklore corpora have not yet become widespread, especially in a regional context. The aim of the article is to present the concept of creating the electronic Tomsk Folklore Corpus based on the materials of the annual summer folklore practice of students of the Faculty of Philology of Tomsk State University. The expeditions took place in 1960–2010 in villages of Tomsk, Kemerovo and Novosibirsk Oblasts, as well as in Khakasia, Krasnoyarsk and Altai Krai. Following the results of the expeditions, the students prepared diaries and handwritten notebooks, in which they noted the place and time of collecting the material, indicated information about the collectors and informants, and wrote down texts with the designation of genres. The notebooks are stored in the TSU Laboratory of General and Siberian Lexicography. A total of 756 notebooks (5 of them contain diaries) and 51 expedition diaries have been preserved. The article substantiates the need to create such a resource, emphasizes the importance of preserving and cataloging folklore materials as an important part of traditional spiritual culture. The main stages of work on the corpus are described, including creating a catalog, digitizing handwritten notebooks, typing and checking texts, developing a markup system and building an information search system. Special attention is paid to the principles of marking up folklore texts, which should enable a multidimensional analysis of the material. The types of markup considered are: the year and place of collection of the material, the gender of the informant, language, genre, subject, the presence of a media file (photos and audio and video recordings for future expeditions). It is assumed that the search in the corpus will also be possible by word or combination of words. The development of the Tomsk Folklore Corpus is a continuation of the traditions of Tomsk Dialectological School and is aimed at preserving and popularizing the unique folklore heritage of Siberia. The article describes the potential uses of the corpus, including scientific (and lexicographical), educational, and cultural. The Tomsk Folklore Corpus will be accessible to a wide range of users, including researchers, teachers, students, and anyone interested in folklore, including folklore groups.

Keywords: regional folklore, corpus folklore studies, electronic folklore corpus, Tomsk Folklore Corpus

Acknowledgments. This research was supported by the Ministry of Higher Education and Science of the Russian Federation, Project No. FSWM-2025-0016.

For citation: Voloshina, S.V. (2025) Tomsk Folklore Corpus: From idea to implementation. *Voprosy leksikografii – Russian Journal of Lexicography*. 37, pp. 49–73. (In Russian). doi: 10.17223/22274200/37/3

Введение

Проблема хранения, каталогизации фольклорного материала, его сохранения как части традиционной духовной культуры является одной из актуальных в филологии. Собираемые иногда в течение многих десятилетий экспедиционные материалы требуют их современного представления, цифровой обработки, новых методов изучения. Наиболее удобным инструментом и форматом репрезентации таких уникальных коллекций текстов наряду со словарями выступают лингвистические корпуса.

С начала XXI в. уже были созданы национальные корпуса различных языков мира, их диалектов, устной речи, текстов отдельных языковых личностей, поэтические корпуса, корпуса прозаических, публицистических текстов и т.д. Фольклорные же корпуса появляются реже, и разработка их концепций, по мнению исследователей, начинается в 2000-х гг. [1. С. 194], но пока корпусная фольклористика находится в процессе своего становления [2]. Языковые корпуса, как считает А.А. Кузьмина, «без репрезентации фольклорных текстов представляются неполноценными, неполными» [3. С. 11]. При этом констатируется недостаточная разработанность теоретической базы их создания и нехватка специалистов [3; 4].

Публикации о концепциях фольклорных корпусов, их структуре, применении в исследованиях, появляющиеся в основном с 2010-х гг., показывают, что корпуса фольклорных текстов включают в качестве подкорпусов в национальные корпуса того или иного языка [4. С. 78] или разрабатывают как самостоятельные; гораздо реже встречаются фольклорные корпуса региональных лингвокультур.

Исследователи пишут о создании фольклорных подкорпусов национального корпуса якутского языка [3], марийского языка [4], калмыцкого языка [1], о созданных подкорпусах сказок и причитаний вепсского языка [5], корпусе корякских фольклорных сказок [6], фольклорном подкорпусе тувинского языка и тувинского героического эпоса [7; 8], фольклорном подкорпусе лингвистического корпуса адыгейского языка [9]; расширении корпуса бурятского языка за счет включения в него фольклорных текстов (сказок) [10], а также отмеча-

ют самостоятельные фольклорные корпуса – башкирского языка [11; 12] и уже выделенные в его составе подкорпуса, например, подкорпус башкирских афоризмов [13], корпус нганасанских фольклорных текстов [14]. С 2014 г. сотрудниками Института филологии СО РАН и Института систем информатики имени А.П. Ершова СО РАН разрабатывается электронный портал «Фольклор народов Сибири», на котором размещены алтайский, алюторский, бурятский, корякский, хакасский, хантыйский, шорский корпуса фольклорных текстов [15], с 2011 по 2020 г. создавался электронный корпус текстов на языках малочисленных народов Сибири (ненецком, телеутском, эвенкийском, шорском) [16].

Наряду с подкорпусами национальных корпусов разных языков, фольклорные корпуса становятся частью корпусов текстов традиционных региональных лингвокультур, например, фольклорный подкорпус в структуре электронного лингвокультурологического корпуса Кубани [17], электронный текстовый корпус лингвокультуры Северного Приангарья включает в свой состав диалектный, мультимедийный и **фольклорный** подкорпуса [18]. Вместе с тем диалектные корпуса, не имеющие фольклорных подкорпусов, могут включать в свой состав тексты фольклорных жанров. Например, некоторые тексты Томского диалектного корпуса включают пословицы, поговорки, заговоры, песни, загадки, частушки, приметы и т.д., которые пользователи могут найти благодаря имеющейся жанровой разметке (например, в нём содержатся 48 частушек, 8 заговоров, 7 загадок и т.д.) [19].

Осуществляется создание корпусов фольклорных текстов, собранных во время фольклорных практик студентов Иркутского государственного университета с 1975 по 2011 г. [20], Уральского федерального университета [21]. Корпусные методы, на наш взгляд, используются в создании электронных версий фольклорных архивов, электронных карт фольклора отдельных регионов: так, электронная версия фольклорного архива факультета гуманитарных наук НИУ «Высшая школа экономики» создана в том числе по итогам студенческих экспедиций [22]. Экспедиционные материалы архива кафедры искусства народного пения Белгородского государственного института искусств и культуры составили основу электронной карты «Этнокуль-

турное наследие Белгородчины», включающей образцы музыкального фольклора Белгородской области [23].

Цель данной статьи – представить концепцию Томского фольклорного корпуса, создаваемого на основе текстов, собранных в 1960–2010 гг. в результате фольклорных экспедиций студентов и преподавателей Томского государственного университета. С 1960 г. студенты филологического факультета после первого курса проходили ежегодную летнюю фольклорную практику. Последние материалы с массовых выездов датируются 1992–1993 гг., менее массовые же экспедиции – 1994–2010 гг. Первая поездка была в Асиновский район Томской области, в сёла Крутоложное, Батурино и г. Асино. География всех выездов очень широкая: экспедиции осуществлялись в сёла Асиновского, Первомайского, Тегульдетского, Зырянского, Шегарского, Кожевниковского, Молчановского, Кривошеинского, Бакчарского, Чаинского, Парабельского районов и города Томской области, сёла Кемеровской, Новосибирской областей, Алтайского края, Красноярского края и Хакасии. Большой массив текстов представлен на русском языке, однако также зафиксированы тексты на хакасском, телеутском, шорском, украинском языках. По итогам практики студенты сдавали рукописные тетради, в которых отмечали место и время сбора материала, указывали данные о собирателях и информантах и записывали тексты с обозначением жанров. Тетради последние 5 лет хранятся в лаборатории общей и сибирской лексикографии ТГУ. Всего сохранилось 756 тетрадей (внутри 5-ти из них есть дневники) и 51 дневник экспедиции.

С этим материалом уже работали исследователи: были написаны научные статьи [24; 25 и др.], диссертационные работы [26–28 и др.]. Попытки каталогизировать его осуществлялись как минимум три раза. Об этом свидетельствуют номера, присвоенные тетрадям: на одной тетради может быть три номера, но все проведенные ранее работы, каталоги были утрачены. Часть собранного материала уже подвергалась корпусной систематизации: так, была создана база данных Среднеобского фольклора и ограничена рамками среднеобского говора [29]. Создаваемый же корпус, как уже отмечалось, будет включать

в том числе материалы, записанные в Новосибирской области, Алтайском крае, Красноярском крае, Хакасии.

Ранее членами Томской диалектологической школы была предложена концепция создания Томского диалектного корпуса [30; 31], который был разработан и в настоящее время постоянно пополняется и используется при разработке словарей, например [32]. Предполагается, что Томский фольклорный корпус будет создан в качестве самостоятельного. Разработка этого ресурса продолжает традиции Томской диалектологической школы, одним из ведущих направлений деятельности которой является лексикографическое. Новый корпус также выступит в качестве источника материала для лексикографов.

Актуальность такой работы обусловлена, во-первых, необходимостью сохранить собранный материал, успеть его оцифровать, поскольку уже утрачено более 200 тетрадей. Это показала каталогизация материала и присвоенные ранее номера тетрадям.

Как отмечала Т.Г. Плохотнюк, декан филологического факультета Томского государственного университета в 1994–2005 гг. и руководитель фольклорной практики, «все записанные произведения являются фактами истинной народной жизни и культуры, они должны издаваться и исследоваться, их необходимо бережно собирать и сохранить, чтобы не прервалась связь времён, чтобы не искажалось представление об истории нации. <...> Ведь знание и понимание истории помогает человеку ориентироваться и в своём времени» [33].

Во-вторых, предполагается, что корпус будет создан с возможностью его пополнения новыми материалами. М.В. Пенькова, разрабатывая концепцию фольклорного подкорпуса национального корпуса марийского языка, отмечает, что «сбор фактического “народного” живого материала особенно необходим, так как маленькими шагами, но размеренно и верно мы идем к потере универсальных старинных “вещдоков”, нередко теряем их в популярной на сегодняшний день культурной “стилизации”» [4. С. 77]. Думается, что такая ситуация наблюдается во многих регионах. Наши диалектологические экспедиции последних 5 лет показывают, что сельские жители сохраняют фольклорные традиции, используют фольклорный материал во время праздников, региональных, районных, сельских мероприятий, соблю-

дают обычаи и обряды, сожалея об утрате обрядовой культуры и указывая на необходимость её сохранения.

Например, в деревне Новосергеевка Кожевниковского района сохраняют похоронный обряд, и местные жители отмечают, что только в их деревне он еще существует и есть необходимость его сохранения: «... у нас рядом, у близлежащих сёл уже этого нету. Всё. Вот если я перестану ходить, я понятия не имею, кто сможет ходить. Понятия не имею. <...> Ну, люди хоть как-то, хоть худо-бедно, кто-то забывает, но они всё-таки вот придерживаются этого обычая при похоронах. Всё-таки они это как-то вот... устои ещё, старинные, они остались. <...> Вот у нас здесь Новосергеевка, Тека, Малиновка, Борзуновка – ни в одной деревне, только у нас здесь. <...> Вот это вот, это, извините, тоже, крайне редко, девочки, девочки и мальчики, крайне редкость, когда вот такие же обряды проводят. Крайне редкость ... то, что здесь мы вот поём не так, как в церкви. Вот чем ценно. Я говорю, а самое ценное то, что люди как деды' пели, бабки пели, так и люди до сих пор молятся так вот именно здесь, в этой местности. Я говорю, самое-то ценное что? <...> Что люди все равно свои корни, свои устои ... не забывают, об этом ведь. А остальные... у нас ни в одной деревне нет, они уже забыли. Они вышли, выпили, покурили, зашли, сели. Это пьянка. У нас такого нет. То есть очерёдность блюд, и всё строго. У нас как-то более-менее стараются вот придерживаться. Но я давно живу, я здесь с семьдесят второго года. Потом год нас не было... обратно приехали. И вот как я приехала, вот этот обряд, так он и до сих пор есть. Так мы и поём, как вот бабки пели, и так и мы сейчас поём. А скоро некому будет петь. Вот так. Обряды, ну, записывайте, если где-то хотя бы, ну вот, записывайте, собирайте, если б показали где-то, не знаю. Может, и придёт к этому время» (д. Новосергеевка, Кожевниковский район, Томская область, июль 2023 г.).

В-третьих, актуальность создания фольклорного корпуса определяется возрастающим интересом к фольклору в современном обществе и необходимостью систематизировать имеющиеся тексты. Это позволит пользователям получать более объективные данные, делать репрезентативную выборку материала при работе с текстами опреде-

ленных жанров или текстами, зафиксированными на какой-либо территории, или выбирать тексты по тематике и т. д.

Этапы разработки Томского фольклорного корпуса и его структура

Разработка Томского фольклорного корпуса предполагает несколько этапов.

Первым этапом работы стало создание каталога имеющихся тетрадей. Каждой тетради присваивался номер, отмечались год и место записи, собиратели материала, жанры, количество страниц, язык и наличие фотографий и дневников. Всего в каталоге, таким образом, описано 807 тетрадей, из которых 51 дневник экспедиции. Первые 200 тетрадей были систематизированы и описаны сотрудником лаборатории общей и сибирской лексикографии В.В. Галаниной в 2023 г. под руководством М.А. Толстой, остальные – И.А. Ореховой и С.В. Волошиной.

Одновременно с первым этапом начался и до сих пор продолжается второй этап – работа по сканированию тетрадей. Сотрудники отдела электронной библиотеки Научной библиотеки ТГУ сканируют тетради и размещают электронную версию в формате PDF на платформе PRO Сибирь. Уже отсканированы 440 тетрадей.

На третьем этапе делается набор отсканированного материала, к настоящему моменту с 2024 г. набраны тексты 140 тетрадей. К концу 2025 г. будут набраны тексты ещё 160 тетрадей. Набор осуществляется сотрудниками лаборатории общей и сибирской лексикографии (И.А. Ореховой, Н.А. Зюзьковой, А.Е. Рулевой и А.Д. Ставровой), студентами-волонтерами и студентами направления «Издательское дело», проходящими учебную практику.

В тетрадях и набранных текстах встречаются ошибки, опечатки, соответственно, требуется четвертый этап – проверка информации и текстов, их редактирование.

На пятом этапе предполагается внесение материала в корпус и его разметка.

Безусловно, необходимо взаимодействие с программистом, который разработает программную платформу, информационно-поисковую систему. Она, как представляется, позволит сохранять тексты, делать их разметку, осуществлять поиск по критериям, которые лягут в основу корпуса. В настоящее время десять набранных, проверенных текстов разных жанров переданы программисту для разработки демо-версии Томского фольклорного корпуса.

Фольклорные тексты – это тексты особого типа, и исследователи отмечают, что в отличие от создания других корпусов создание фольклорных корпусов отличается особой сложностью и многоаспектностью [3; 20], и «специфика фольклорного текста как семиотического кода обуславливает необходимость разработки особых подходов к выбору критериев структурирования фольклорного дискурса» [17. С. 140].

Существующие фольклорные корпуса структурируются по-разному, что зависит от поставленных исследователями целей создания корпуса, технических возможностей, от материала, с которым работают и типов разметки, которые продиктованы во многом спецификой материала. Так, это может быть единый фольклорный корпус с одинаковой разметкой для всех текстов. Согласно другой точке зрения, устное народное творчество представлено большим количеством жанров, каждый из которых требует введения в корпус дополнительных, принципиально разных разметок [34], «работающих» для одного типа текстов и совершенно «не работающих» для другого, поэтому необходимо внутри фольклорного корпуса создание подкорпусов, объединяющих тексты одного жанра: подкорпус сказок, подкорпус песен, подкорпус загадок и т.п.

Проблемы создания фольклорного подкорпуса в НКРЯ, информация о предполагаемых типах разметки, целесообразности их включения, а также принципы, которые лягут в основу концепции подкорпуса, обсуждались в научном альманахе «Традиционная культура» в 2018 г. Исследователи-фольклористы представили ответы на вопросы о паспортизации материала – о необходимости указания времени записи, региона, возраста и пола исполнителя, фамилии собирателя. Один из вопросов касался классификации материала, каким образом его представлять в корпусе: по жанрам, по сюжетам/сюжетным ситу-

ациям, по мотивам, по образам, по персонажам, по функциям, по предметному миру или по какому-то другому принципу [35].

Так, большинство созданных и создаваемых корпусов включают как экстралингвистические метаданные, так и лингвистическую разметку. Данные о месте, времени сбора материала, собирателях и информантах, жанрах включают почти все рассмотренные выше корпуса.

В отдельные корпуса введены также наряду с названными определенные параметры, присутствующие не во всех корпусах. Фольклорный архив факультета гуманитарных наук НИУ ВШЭ, включающий собственно архив и лингвистический корпус, предполагает поиск информации по ключевому слову, к корпусу прилагается таблица ключевых слов. В фольклорном подкорпусе электронного текстового корпуса лингвокультуры Северного Приангарья применяется к таким жанрам устной прозы, как быличка и бывальщина, сюжетно-мотивная разметка. В марийском фольклорном подкорпусе и подкорпусе национального корпуса якутского языка предлагается отразить вид фиксации и характер аутентичности фольклорного материала.

Тексты Томского фольклорного корпуса, как представляется, могут быть классифицированы по году, месту сбора материала, полу информантов, языку, жанру, тематике, наличию медиафайлов. Такая классификация материала обусловлена наличием указанных данных и достаточной информативностью для пользователей.

Введение в корпус данных о годе записи позволит пользователям изучать фольклорные тексты за какой-либо год или период, например, тексты 1960–70-х гг. В собранном материале отражается эпоха, это даст возможность исследователям определить, с какого времени начинает в фольклоре отражаться то или иное явление, событие, и до какого времени оно осмысливается, творчески перерабатывается. Например, частушка, записанная в 1989 г. в с. Сарафановка Молчановского района Томской области, отражает эпоху Перестройки:

*Ой, не всё погорело,
Ой, не всё пойдем тушить.
Перестройка началась,
Еще лучшие будем жить!*

Подобный материал отражает активные процессы, происходящие в русском языке, соответственно, он может включаться в качестве иллюстративного в создаваемые словари языковых изменений, словари новых слов или электронные лексикографические ресурсы, отражающие новое в русской лексике. Например, в экспедиции 2025 г. были зафиксированы частушки, в которых отражается ситуация и характерная лексика 2020–2021 гг.:

*От ковида привилась,
Значит, долго буду жить.
Это убыток государству,
Надо ж пенсию платить.*

(2025 г., с. Кривошеино, Кривошеинский район, Томская область).

Параметр «место» сбора материала позволит сделать выборку текстов, записанных в определенной области, районе, селе (посёлке, деревне), указать на распространение в той или иной местности какой-либо тематики. К примеру, тексты разных жанров о ссыльных, записанные на севере, северо-западе Томской области: песня, записанная в с. Нарым Парабельского района – месте ссылки:

*По диким степям, по Нарыму
Идёт заготовка лесов,
Там сослано много народу,
Неизвестно на сколько годов.
Они сосланы все без причины,
И безвинно страдают они. <...>*

Или частушка, зафиксированная в с. Подгорном Чаинского района Томской области:

*Кулаки мы, кулаки,
Спецпереселенцы.
За капустные листки
Меняем полотенцы.*

Интересным представляется пополнение фольклорными текстами словарей, созданных или создаваемых на материале, собранном в одном селе или районе. Например, Вершининский словарь как диалектный словарь полного типа, включающий зафиксированную в 1940–1990-е гг. XX в. лексику и фразеологию одного из типичных русских

старожильческих говоров Сибири – говора с. Вершинино (включая с. Ярское и с. Батурино) Томского района Томской области, может быть дополнен лексическими единицами и иллюстративным материалом из записанных там частушек, песен, заговоров. Так, словарь включает слово ЗАЛЁТА (ЗАЛЁТА, ы, м. Любимый, возлюбленный. – Раньше девчонок медунками звали, а ребят звали залётами. Красивые медунки были, цветочки голубенькие [36. С. 257]). В зафиксированном в селе Ярское фольклорном материале содержится слово *залётка*, оно может быть внесено в словарь и проиллюстрировано следующим образом:

*А у меня залётка был,
Звали его Васею,
Он работает всегда
По своему согласию*

(1996 г., с. Ярское, Томский район, Томская область).

Пол информантов – параметр, который может быть введён в корпус. Встречаются мужские и женские частушки:

*Вижу Танино окошко,
А у Тани огонёк.
Подошёл к её окошку,
Она кушает чаёк*

(1989 г., с. Молчаново, Молчановский район, Томская область).

В тексте может быть повествование от лица женского пола, а записан он от мужчины и наоборот. Некоторые фольклорные жанры можно назвать типично женскими – колыбельные, гадания.

В корпусе предполагается введение текстовой разметки – жанровой и тематической.

Согласно каталогу тетрадей, в имеющемся материале представлены тексты более 25 жанров. Среди них загадки, причитания, заговоры, песни, сказки, частушки, были, былички, страшилки, считалки, потешки, приметы, поверья, анекдоты, пословицы, поговорки, дразнилки, колыбельные, легенды, прибаутки, колядки, плачи, скороговорки, заклички и т. д. Репрезентативность материала обеспечивается таким богатым репертуаром жанров. Однако отметим, что не у всех текстов указан жанр, соответственно, при внесении их в корпус по-

требуется идентификация жанров. Кроме этого, наблюдается жанровая стратификация, что актуализирует проблему распределения текстов в пределах одного жанра. Например, в тетрадах встречаются различные типы песен: лирические, шуточные, свадебные, обрядовые, рекрутские (солдатские), военные, исторические, бытовые, застольные, любовные, сиротские, хороводные, плясовые и др. Очевидно, что они могут быть классифицированы по разным основаниям, что может быть осуществлено впоследствии, при усовершенствовании созданного корпуса. В материале также широко представлены варианты частушек, песен и т.д., записанных в разных местах или в пределах одного села, района:

*Дроби бей, дроби бей,
Под ногою воробей,
Под другою – галка,
Трепача не жалко.*

*Дроби бей, дроби бей,
Под ногою воробей,
Под другою – серый гусь,
Я измены не боюсь*

(частушки, 1988 г., с. Вавиловка, Бакчарский район).

В этих текстах варьируется какая-либо часть, это, как представляется, может быть важным для исследователя при изучении какого-либо типа текстов и должно быть учтено при разработке корпуса.

Кроме этого, выделяются частушки-отвечалки (частушки-отзовушки), которые могут существовать как самостоятельно, так и в связке с другими частушками, потому что представляют своеобразный диалог их исполнителей:

*Ой, подружка, моя Зина,
Пойдем садики садить,
Чтобы нашим ухажерам
Было весело ходить.*

*Ой, подружка, моя Оля,
Не пойду я засадить.*

*У меня нет ухажеров,
Ко мне некому ходить.*

*Ой, подружка, моя Зина,
Пойдем садики садить.
Заведи ты ухажёра,
Будет кто к тебе ходить.*

*Ой, подружка, моя Оля,
Как я буду заводить?
Подошла к одному парню,
Он не хочет говорить.*

Такие типы частушек, как представляется, требуется размещать в корпусе как единый текст, не по отдельности.

Одной из главных, на наш взгляд, станет тематическая разметка текстов в корпусе. Разножанровые тексты, разного объема и содержания могут быть упорядочены по параметру «тема». В настоящее время определяется круг тем, среди которых можно уже назвать такие, как религия, нечистая сила, любовь, дружба, измена, разлука, ревность, убийство, война, армия, оружие, тюрьма, страна, Родина, космос, атмосферные явления, растения (деревья, ягоды, цветы, травы, фрукты, дикоросы), животные (дикие, домашние), птицы, насекомые, мужчина/женщина, части тела, физическое состояние человека (физиология), внешность, болезни, вредные привычки, характер, психологическое состояние человека, национальность, профессии, возраст, богатство/бедность, семья, дом (домашняя утварь, части дома, быт), еда и напитки, праздники, досуг, свадьба, сватовство, искусство и культура, одежда, обувь, цвет, сиротство, музыкальные инструменты, речь и язык, пространство (ландшафт: лес, водные объекты, горы и др.), время (сутки, сезон и др.), политика (эпоха), исторические личности, колхозы, совхозы, транспорт, дорога, смерть, кладбище и др.

Необходимым видится поиск в корпусе по слову, а также по сочетанию слов, чтобы пользователь мог найти слово, используемое в том или ином контексте, жанре.

Все размещаемые в корпусе материалы предполагается снабдить информацией об источнике – указать ссылки на тетради и номера тетрадей, откуда взят текст. Таким образом пользователь сможет обратиться к первоисточнику.

Кроме этого, студенты с 1960 по 1978 г. вели дневники экспедиций, в которых делились своими впечатлениями о практике, начиная от отъезда из Томска и заканчивая возвращением в город, в них они описывали встречи с информантами, условия, в которых записывали фольклорные тексты и т. д. Эти дневники экспедиций обладают большой социокультурной информативностью: раскрывают процесс сбора материала, исторические сведения, данные о сёлах и деревнях, в том числе ныне не существующих, показывают особенности эпохи, а также дают информацию о собирателях, их восприятию сельской культуры и об информантах: *«Населяют Горск латыши и латгальцы, русские и украинцы. Нужно сказать, что к концу практики к нам немножко привыкли и даже был случай, когда сами предложили записать у них частушки. Нам кажется, что если бы практика продлилась ещё несколько дней, то мы бы собрали значительно больше материала. С наибольшей охотой население исполняло частушки. Поздние частушки иллюстрируют то, что народ определенным образом отзывается на вещи, происходящие в селе:*

*«Как на том берегу
Брякала гитара,
А ильинские ребята
По копейке пара»*

(д. Ильинка, Асиновский район, Томская область, 1973 г.).

Представляется важным интегрировать дневники с текстами, извлеченными из тетрадей. В некоторых случаях дневники уже помогли установить даты сбора материала, не отмеченные в тетрадях.

Во многих тетрадях содержатся рисунки, в более поздних появились фотографии, что также можно представить вместе с размещаемым в корпусе текстом.

Параметр «наличие медиафайла» позволит в будущем после проведенных экспедиций дополнять записанный материал фотографиями и видео-, аудиофайлами.

Представленные типы разметки и десять размеченных текстов разных жанров уже переданы программисту.

Предполагается, что корпус будет доступен широкому кругу пользователей: как исследователям, так и всем, интересующимся фольклором (школьникам, учителям, фольклорным коллективам, краеведам и т.д.) после регистрации на сайте, где будет размещаться корпус.

Такой ресурс может внести существенный вклад в развитие практической лексикографии. Материал, систематизированный по указанным параметрам, может быть использован при создании словарей, переиздании и дополнении как лексическими единицами, ранее не представленными в словаре, так и новыми значениями, и новым иллюстративным материалом.

Например, в Словаре русских старожильческих говоров средней части бассейна реки Оби не зафиксировано слово *ва́толя*. Однако оно встречается в частушках, записанных в Асиновском и Колпашевском районах:

*Колечко моё
Золото', венчалное,
Мне **ва́толя** приказал –
Не ходи печальная!*

(г. Асино, Асиновский район, Томская область),

*У меня **ва́толи** (милёнка) два,
Как и полагается,
У обоих волоса
Кудрями завиваются*

(г. Колпашево, Колпашевский район, Томская область).

Словарные статьи Словаря синонимов сибирского говора также могут быть дополнены лексическими единицами. Например, словарный ряд «ЛЮБИМАЯ, МЕДУНКА, МИЛАЯ, ПОДРУГА» может быть дополнен лексической единицей *задрыжина*, *матаня* (=возлюбленная):

*Я свою задрыжину
Посажу на лыжину,
Ты езжай, задрыжина,
Куда поедет лыжина*

(г. Томск, Томская область),

*У кого какая баня,
У меня – осинова.
У кого кака **матаня**,
У меня красивая*

(с. Вавиловка, Бакcharский район, Томская область, 1988 г.).

А словарный ряд «ЗАЛЁТА, ДРУЖОК, ЛЮБИМЫЙ, МИЛОЙ, МИЛЫЙ» может быть дополнен словами *ватоля, залётка, матаня, милёнок (=возлюбленный)*:

*Пойдём сплясанём,
Четырёх вспомняем:
Колю, Петю, Ванечку
И старого **матанечку**.*

*У кого какая баня,
У меня – из кирпичей.
У кого какой **милёнок**,
У меня из трепачей*

(с. Вавиловка, Бакcharский район, Томская область, 1988 г.).

Лексикографы на репрезентативном материале за короткое время смогут проследить функционирование нужной лексической единицы в текстах разных жанров, определить ее сочетаемость, контекстуальное окружение, что в свою очередь позволит уточнить семантику и наличие символического компонента в значении.

Заключение

Итак, Томский фольклорный корпус разрабатывается с учетом принципов полноты (максимальный охват имеющегося материала), репрезентативности (богатство жанрового состава), доступности. Поиск материала в нём будет осуществляться по году, месту сбора материала, полу информантов, языку, жанру, тематике, наличию медиа-файлов, слову. В результатах поиска предполагается указание наличия вариантов у того или иного текста и ссылки на источник материала (тетрадь, дневник практики).

Материалы корпуса могут быть востребованы в лексикографической практике: при создании словарей разных типов; в исследовательской практике: фольклористами при изучении региональной специфики фольклора, жанрового разнообразия и т. п., этнографами, лингвистами; в образовательной деятельности: при использовании корпусных материалов в курсах по фольклористике, этнолингвистике, этнографии и т. п.; в культурно-просветительской работе: при популяризации фольклорного наследия Сибири, в организации праздников и фестивалей народного творчества, расширении репертуара фольклорных коллективов, в создании музейных выставок и т. п.

Разработка Томского фольклорного корпуса продолжает традиции создания фольклорных корпусов и Томского диалектного корпуса. Его создание послужит сохранению уникального архива текстов, собираемых в течение полувека. Корпус станет удобным инструментом для исследователей, который позволит делать выборку материала по обозначенным в статье параметрам. Этот ресурс может быть интегрирован в будущем в другие корпуса фольклорных материалов, в частности зафиксированных на территории Сибири, встроен в различные типы создаваемых электронных справочников и словарей, дополнен изданными сборниками сибирского фольклора.

Список источников

1. Куканова В.В. Фольклорный подкорпус: проблемы, структура и перспективы использования // Участие калмыков в укреплении Российской государственности : материалы Региональной науч.-практ. конф., посвящ. 1150-летию Российской государственности и Году российской истории (Элиста, 29 ноября 2012 г.). Элиста : КИГИ РАН, 2012. С. 192–197.
2. Тучкова Н.А. Корпусный метод в фольклористике и его возможности для этнографических исследований (опыт использования на селькупском фольклорном материале) // Томский журнал лингвистических и антропологических исследований. 2018. Вып. 2 (20). С. 100–111. doi: 10.23951/2307-6119-2018-2-100-111
3. Кузьмина А.А. Концепция создания фольклорного подкорпуса Национального корпуса якутского языка: постановка проблемы, структура, технология // Филология: научные исследования. 2020. № 11. С. 11–18. doi: 10.7256/2454-0749.2020.11.34168

4. Пенькова М.В. Фольклорный подкорпус Национального корпуса марийского языка: к вопросу создания // Финно-угроведение. 2023. № 1 (64). С. 75–80. doi: 10.51254/2312-0312-2023-64-75-80
5. Зайцева Н.Г., Крижановская Н.Б. Корпусная лингвистика в прибалтийско-финском исследовательском пространстве (на материале Корпуса вепсского языка и Открытого корпуса вепсского и карельского языков) // Альманах северо-европейских и балтийских исследований. 2018. № 3. С. 263–273.
6. Сафонова Д.П. Продуктивные модели аффиксального словообразования существительных по данным корпуса фольклорных корякских сказок // Известия Российского государственного педагогического университета им. А.И. Герцена. 2012. № 146. С. 120–128.
7. Ондар М.В. База данных текстов тувинского героического эпоса: первый этап // Новые исследования Тувы. 2016. № 4. С. 66–75.
8. Монгуш Ч.М., Ондар М.В. База данных и средства создания контекстов для представления и анализа тувинского героического эпоса // Программные продукты, системы и алгоритмы. 2017. № 3. С. 1–6.
9. Унарокова Р.Б., Соколова А.Н., Хакунова Э.Х. Ритуальные приговоры и импровизированные возгласы-призывы джегуако в контексте свадебной обрядности адыгов (в рамках корпусных исследований адыгейского языка) // Сибирский филологический журнал. 2023. № 1. С. 45–61.
10. Бадмаева Л.Д. О расширении корпуса бурятского языка // Банзаровские чтения : материалы междунар. науч. конф., посвященной 200-летию со дня рождения Д. Банзарова и 90-летию БГПИ-БГУ (Улан-Удэ, 30–31 марта 2022 г.): в 2 ч. Ч. 1. Улан-Удэ : Бурятский государственный университет имени Доржи Банзарова, 2022. С. 227–229.
11. Сиразитдинов З.А., Бускунбаева Л.А., Ишмухаметова А.Ш., Шамсутдинова Г.Г. О структуре и содержании фольклорного корпуса башкирского языка // Электронная письменность народов Российской Федерации-2021 & IWCLUL 2021 : материалы междунар. науч.-практ. конф. (Сыктывкар, 23–24 сентября 2021 г.). Сыктывкар : Коми Республиканская академия государственной службы и управления, 2022. С. 133–139.
12. Сиразитдинов З.А. Корпусные проекты лаборатории лингвистики и информационных технологий ИИЯЛ УНЦ РАН // Известия Уфимского научного центра РАН. 2013. № 4. С. 104–111.
13. Шамсутдинова Г.Г., Ишмухаметова А.Ш., Бускунбаева Л.А. Структура и состав базы данных загадок в подкорпусе текстов афористических жанров башкирского фольклора // Вестник Калмыцкого института гуманитарных исследований РАН. 2017. Т. 10, № 4 (32). С. 146–153.
14. Корпус нганасанских фольклорных текстов. М., 2003–2005. URL: <https://iling-ran.ru/gusev/Nganasan/texts/index.php> (дата обращения: 12.02.2025).
15. Электронный портал «Фольклор народов Сибири». Новосибирск, 2025. URL: <https://folk-old.philology.nsc.ru/#> (дата обращения: 12.02.2025).

16. Корпусы ИЭА РАН. М., 2011–2020. URL: <https://corpora.ica.ras.ru/corpora/> (дата обращения: 12.02.2025).
17. Литус Е.В. К вопросу о разметке фольклорных текстов в структуре электронного лингвокультурологического корпуса Кубани (на материале народных песен кубанского свадебного обряда) // Ресурсы региона: культурно-историческое развитие в контексте науки и образования : материалы Всерос. науч.-практ. конф. (Славянск-на-Кубани, 10–14 сентября 2015 г.). Славянск-на-Кубани : Филиал ГГБОУ ВПО «Кубанский государственный университет» в г. Славянске-на-Кубани, 2015. С. 138–142.
18. Фельде О.В. Электронный текстовый корпус лингвокультуры Северного Приангарья: принципы создания, структура и возможности использования // Вестник Кемеровского государственного университета. 2021. Т. 23, № 4 (88). С. 1086–1095.
19. Томский диалектный корпус // Лаборатория общей и сибирской лексикографии Томского государственного университета. Томск, 2025. URL: <https://losl.tsu.ru/?q=corpus> (дата обращения: 12.02.2025).
20. Васильева Е.В., Тиллабаева А.А., Шеметова Д.И. Корпус фольклорных текстов: от идеи до разметки // «Цифра» в социально-гуманитарных исследованиях: метод, поле, реальность : материалы науч. конф. молодых ученых (Иркутск, 15–18 ноября 2023 г.). Иркутск : Иркутский государственный университет, 2024. С. 45–51.
21. Маслова М.Д. Особенности создания морфологической разметки диалектного корпуса на примере собрания текстов исчезающих деревень Алапаевского и Верхотурского районов Свердловской области // Четвертая зимняя школа по гуманитарной информатике : сб. ст. (Калининград, 10–11 декабря 2020 г.). Калининград : БФУ им. И. Канта, 2020. С. 85–91.
22. Фольклорный архив Факультета гуманитарных наук НИУ «Высшая школа экономики». М., 2025. URL: <https://folklore.linghub.ru/> (дата обращения: 12.02.2025).
23. Жиров М.С., Жирова О.Я., Кузнецова Н.С. Электронный фольклорный архив в современном социокультурном пространстве // Самарский научный вестник. 2021. Т. 10, № 1. С. 283–286.
24. Тубалова И.В. «Свои» и «чужие» в фольклорном коллективе: лингвокогнитивные модели социализации в среднеобском фольклоре // Евроазиатский межкультурный диалог: «Свое» и «чужое» в национальном самосознании культуры. Томск : ТГУ, 2007. С. 398–404.
25. Тубалова И.В., Эмер Ю.А. Современная сибирская деревня в фольклорном отражении (лингвокультурологическое описание) // Вестник Томского государственного университета. 2007. № 294. С. 86–92.
26. Бохонная М.Е. Эстетическая интерпретация «вещного» мира в языке среднеобского фольклора (на материале лирической песни и частушки) : автореф. дис. ... канд. филол. наук. Томск, 2006. 29 с.

27. Эмер Ю.А. Миромоделирование в современном песенном фольклоре (когнитивно-дискурсивный анализ): автореф. дис. ... д-ра филол. наук. Томск, 2011. 39 с.
28. Тубалова И.В. Полифонический текст в устных личностно-ориентированных дискурсах: автореф. дис. ... д-ра филол. наук. Томск, 2015. 43 с.
29. Мишанкина Н.А., Тубалова И.В., Эмер Ю.А. Филология и информатика: специфика электронного представления региональных фольклорных текстов // Гуманитарная информатика. 2004. № 1. С. 102–113.
30. Юрина Е.А. Томский диалектный корпус: в начале пути // Вестник Томского государственного университета. Филология. 2011. № 2 (14). С. 58–63.
31. Земичева С.С., Иванцова Е.В. Проект создания Томского диалектного корпуса в свете тенденций развития корпусной лингвистики // Сибирский филологический журнал. 2018. № 3. С. 192–205.
32. Аганова Н.А., Волошина С.В., Демешкина Т.А. и др. Словарь детства: говоры Среднего Приобья (с лингвокультурологическим комментарием) / под ред. М.М. Угрюмовой. Томск: Изд-во Том. ун-та, 2018. 200 с.
33. Плохотнюк Т.Г. Народная культура остается востребованной: студенты ФилФ изучают фольклор во время учебной практики // Alma Mater. № 2387 21.06.2005. URL: http://almamater.tsu.ru/show_story.phtml?nom=2387&s=1861 (дата обращения: 28.03.2025).
34. Николаев Д.С. Создание электронного корпуса фольклорных текстов на русском языке. URL: https://prof1.wciom.ru/fileadmin/file/nauka/grusha2015/s2_6/Nikolaev.pdf (дата обращения: 28.03.2025).
35. О фольклорном подкорпусе национального корпуса русского языка // Традиционная культура. 2012. № 1 (45). С. 14–25.
36. Вершининский словарь. Т 2. Г–З. / гл. ред. О.И. Блинова. Томск: Изд-во Том. ун-та, 1999. 319 с.

References

1. Kukanova, V.V. (2012) [Folklore Subcorpus: Problems, Structure and Prospects of Use]. *Uchastiye kalmykov v ukreplenii Rossiyskoy gosudarstvennosti* [Participation of the Kalmyks in Strengthening Russian Statehood]. Proceedings of the Regional Conference dedicated to the 1150th Anniversary of Russian Statehood and the Year of Russian History. Elista. 29 November 2012. Elista: Kalmyk Institute for Humanities of the Russian Academy of Sciences. pp. 192–197. (In Russian).
2. Tuchkova, N.A. (2018) *Korpusnyy metod v fol'kloristike i yego vozmozhnosti dlya etnograficheskikh issledovaniy (opyt ispol'zovaniya na sel'kupskom fol'klornom materiale)* [Corpus Method in Folklore Studies and Its Potential for Ethnographic Research (Experience of Using Selkup Folklore Material)]. *Tomskiy zhurnal lingvisticheskikh i antropologicheskikh issledovaniy*. 2 (20). pp. 100–111. doi: 10.23951/2307-6119-2018-2-100-111

3. Kuz'mina, A.A. (2020) Kontseptsiya sozdaniya fol'klornogo podkorpusa Natsional'nogo korpusa yakutskogo yazyka: postanovka problemy, struktura, technologya [The Concept of Creating a Folklore Subcorpus of the National Corpus of the Yakut Language: Problem Statement, Structure, Technology]. *Filologiya: nauchnye issledovaniya*. 11. pp. 11–18. doi: 10.7256/2454-0749.2020.11.34168
4. Pen'kova, M.V. (2023) Fol'klorny podkorpus Natsional'nogo korpusa mariyskogo yazyka: k voprosu sozdaniya [Folklore Subcorpus of the National Corpus of the Mari Language: On the Issue of Creation]. *Finno-ugrovedeniye*. 1 (64). pp. 75–80. doi: 10.51254/2312-0312-2023-64-75-80
5. Zaytseva, N.G. & Krizhanovskaya, N.B. (2018) Korpusnaya lingvistika v pribaltiysko-finskom issledovatel'skom prostranstve (na materiale Korpusa vepsskogo yazyka i Otkrytogo korpusa vepsskogo i karel'skogo yazykov) [Corpus Linguistics in the Baltic-Finnish Research Space (Based on the Veps Language Corpus and the Open Corpus of Veps and Karelian Languages)]. *Al'manakh severoyevropeyskikh i baltiyskikh issledovaniy*. 3. pp. 263–273.
6. Safonova, D.P. (2012) Produktivnye modeli affiksial'nogo slovoobrazovaniya sushestvitel'nykh po dannym korpusa fol'klornykh koryakskikh skazok [Productive Models of Affixal Noun Formation According to the Corpus of Koryak Folklore Tales]. *Izvestiya Rossiyskogo gosudarstvennogo pedagogicheskogo universiteta im. A.I. Gertsena*. 146. pp. 120–128.
7. Ondar, M.V. (2016) Baza dannykh tekstov tuvinskogo geroicheskogo eposa: pervyy etap [Database of Texts of the Tuvan Heroic Epic: The First Stage]. *Novye issledovaniya Tuvy*. 4. pp. 66–75.
8. Mongush, Ch.M. & Ondar, M.V. (2017) Baza dannykh i sredstva sozdaniya kontekstov dlya predstavleniya i analiza tuvinskogo geroicheskogo eposa [Database and Tools for Creating Contexts for the Representation and Analysis of the Tuvan Heroic Epic]. *Programmnye produkty, sistemy i algoritmy*. 3. pp. 1–6.
9. Unarokova, R.B., Sokolova, A.N. & Khakunova, E.Kh. (2023) Ritual'nye prigovery i improvizirovannye vozzglasy-prizvyvy dzheguako v kontekste svadebnoy obryadnosti adygov (v ramkakh korpusnykh issledovaniy adygeyskogo yazyka) [Ritual Sayings and Improvised Calls of Jeguako in the Context of Adyghe Wedding Rituals (Within the Framework of Corpus Studies of the Adyghe Language)]. *Sibirskiy filologicheskii zhurnal*. 1. pp. 45–61.
10. Badmaeva, L.D. (2022) [On the Expansion of the Buryat Language Corpus]. *Banzarovskiy chteniye* [Banzarov Readings]. Proceedings of the International Conference dedicated to the 200th Anniversary of D. Banzarov's Birth and the 90th Anniversary of BSPI-BSU. Ulan-Ude. 30–31 March 2022. In 2 parts. Part 1. Ulan-Ude: Banzarov Buryat State University. pp. 227–229. (In Russian).
11. Sirazitdinov, Z.A., Buskunbayeva, L.A., Ishmukhametova, A.Sh. & Shamsutdinova, G.G. (2022) [On the Structure and Content of the Folklore Corpus of the Bashkir Language]. *Elektronnaya pis'mennost' narodov Rossiyskoy Federatsii-2021 & IWCLUL 2021* [Electronic Writing of the Peoples of the Russian Federation-2021 & IWCLUL 2021]. Proceedings of the International Conference.

- Syktvykar. 23–24 September 2021. Syktvykar: Komi Republican Academy of Public Service and Administration. pp. 133–139. (In Russian).
12. Sirazitdinov, Z.A. (2013) Korpusnye proyekty laboratorii lingvistiki i informatsionnykh tekhnologiy IYYaL UNTs RAN [Corpus Projects of the Laboratory of Linguistics and Information Technologies of the IHL Ufa SC RAS]. *Izvestiya Ufimskogo nauchnogo tsentra RAN*. 4. pp. 104–111.
 13. Shamsutdinova, G.G., Ishmukhametova, A.Sh. & Buskunbayeva, L.A. (2017) Struktura i sostav bazy dannykh zagadok v podkorpuse tekstov aforisticheskikh zhanrov bashkirskogo fol'klora [Structure and Composition of the Riddle Database in the Subcorpus of Aphoristic Genre Texts of Bashkir Folklore]. *Vestnik Kalmytskogo instituta gumanitarnykh issledovaniy RAN*. 10 (4 (32)). pp. 146–153.
 14. *Korpus nganasanskikh fol'klornykh tekstov* [Corpus of Nganasan Folklore Texts]. (2025) [Online] Available from: <https://iling-ran.ru/gusev/Nganasan/texts/index.php> (Accessed: 12.02.2025).
 15. *Elektronnyy portal "Fol'klor narodov Sibiri"* [Electronic Portal "Folklore of the Peoples of Siberia"]. (2025) [Online] Available from: <https://folk-old.philology.nsc.ru/#> (Accessed: 12.02.2025).
 16. *Korpusy IEA RAN* [Corpora of the IEA RAS]. (2025) [Online] Available from: <https://corpora.iea.ras.ru/corpora/> (Accessed: 12.02.2025).
 17. Litus, E.V. (2015) [On Marking Folklore Texts in the Structure of the Electronic Linguistic and Cultural Corpus of Kuban (Based on Folk Songs of the Kuban Wedding Ritual)]. *Resursy regiona: kul'turno-istoricheskoye razvitiye v kontekste nauki i obrazovaniya* [Resources of the Region: Cultural and Historical Development in the Context of Science and Education]. Proceedings of the All-Russian Conference. Slavyansk-on-Kuban, 10–14 September 2015. Slavyansk-on-Kuban: Branch of the Kuban State University in Slavyansk-on-Kuban. pp. 138–142. (In Russian).
 18. Fel'de, O.V. (2021) Elektronnyy tekstovyy korpus lingvokul'tury Severnogo Priangar'ya: printsipy sozdaniya, struktura i vozmozhnosti ispol'zovaniya [Electronic Text Corpus of the Linguistic Culture of the Northern Angara Region: Principles of Creation, Structure and Possibilities of Use]. *Vestnik Kemerovskogo gosudarstvennogo universiteta*. 23 (4 (88)). pp. 1086–1095.
 19. TSU Laboratory of General and Siberian Lexicography. (2025) *Tomskiy dialektnyy korpus* [Tomsk Dialect Corpus]. [Online] Available from: <https://losl.tsu.ru/?q=corpus> (Accessed: 12.02.2025).
 20. Vasil'yeva, E.V., Tillabayeva, A.A. & Shemetova, D.I. (2024) [Corpus of Folklore Texts: From Idea to Markup]. *"Tsifra" v sotsial'no-gumanitarnykh issledovaniyakh: metod, pole, real'nost'* ["Digital" in Socio-Humanitarian Research: Method, Field, Reality]. Proceedings of the Conference of Young Scientists. Irkutsk. 15–18 November 2023. Irkutsk: Irkutsk State University. pp. 45–51. (In Russian).
 21. Maslova, M.D. (2020) Osobennosti sozdaniya morfologicheskoy razmetki dialectnogo korpusa na primere sobraniya tekstov ischezayushchikh dereven' Alapayevskogo i Verkhoturskogo rayonov Sverdlovskoy oblasti [Features of Creating Mor-

- phological Markup for a Dialect Corpus Using the Example of a Collection of Texts from Disappearing Villages of the Alapaevsk and Verkhoturye Districts of the Sverdlovsk Region]. In: *Chetvertaya zimnyaya shkola po gumanitarnoy informatike: sb. st.* [Fourth Winter School on Humanities Informatics: Collection of Articles]. (Kaliningrad, 10–11 December 2020). Kaliningrad: Immanuel Kant Baltic Federal University. pp. 85–91.
22. HSE Faculty of Humanities. (2025) *Fol'klornyy arkhiv Fakul'teta gumanitarnykh nauk NIU "Vysshaya shkola ekonomiki"* [Folklore Archive of the Faculty of Humanities of the National Research University "Higher School of Economics"]. [Online] Available from: <https://folklore.linghub.ru/> (Accessed: 12.02.2025).
23. Zhirov, M.S., Zhirova, O.Ya. & Kuznetsova, N.S. (2021) Elektronnyy fol'klornyy arkhiv v sovremennom sotsiokul'turnom prostranstve [Electronic Folklore Archive in the Modern Sociocultural Space]. *Samarskiy nauchnyy vestnik*. 10 (1). pp. 283–286.
24. Tubalova, I.V. (2007) "Svoi" i "chuzhie" v fol'klornom kollektive: lingvokognitivnye modeli sotsializatsii v sredneobskom fol'klore ["Own" and "Others" in the Folklore Collective: Linguocognitive Models of Socialization in the Middle Ob Folklore]. In: *Yevroaziatskiy mezhkul'turnyy dialog: "Svoe" i "chuzhoye" v natsional'nom samosoznanii kul'tury* [Eurasian Intercultural Dialogue: "Own" and "Other" in the National Self-Consciousness of Culture]. Tomsk: Tomsk State University. pp. 398–404.
25. Tubalova, I.V. & Emer, Yu.A. (2007) Sovremennaya sibirskaya derevnya v fol'klornom otazhenii (lingvokul'turologicheskoye opisaniye) [The Modern Siberian Village in Folklore Reflection (Linguocultural Description)]. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. 294. pp. 86–92.
26. Bokhonnaya, M.E. (2006) *Esteticheskaya interpretatsiya "veshchnogo" mira v yazyke sredneobskogo fol'klora (na materiale liricheskoy pesni i chastushki)* [Aesthetic Interpretation of the "Material" World in the Language of the Middle Ob Folklore (Based on Lyrical Songs and Chastushkas)]. Abstract of Philology Cand. Diss. Tomsk.
27. Emer, Yu.A. (2011) *Miromodelirovaniye v sovremennom pesennom fol'klore (kognitivno-diskursivnyy analiz)* [World Modeling in Modern Song Folklore (Cognitive-Discursive Analysis)]. Abstract of Philology Dr. Diss. Tomsk.
28. Tubalova, I.V. (2015) *Polifonicheskiy tekst v ustnykh lichnostno-orientirovannykh diskursakh* [Polyphonic Text in Oral Person-Oriented Discourses]. Abstract of Philology Dr. Diss. Tomsk.
29. Mishankina, N.A., Tubalova, I.V. & Emer, Yu.A. (2004) Filologiya i informatika: spetsifika elektronnoy predstavleniya regional'nykh fol'klornykh tekstov [Philology and Informatics: Specifics of Electronic Representation of Regional Folklore Texts]. *Gumanitarnaya informatika*. 1. pp. 102–113.
30. Yurina, E.A. (2011) Tomsk Dialect Corpus: At the Beginning of the Journey. *Vestnik Tomskogo gosudarstvennogo universiteta. Filologiya – Tomsk State University Journal. Philology*. 2 (14). pp. 58–63. (In Russian).

31. Zemicheva, S.S. & Ivantsova, E.V. (2018) *Proyekt sozdaniya Tomskogo dialekt-nogo korpusa v svete tendentsiy razvitiya korpusnoy lingvistiki* [The Project of Creating the Tomsk Dialect Corpus in the Light of Trends in the Development of Corpus Linguistics]. *Sibirskiy filologicheskii zhurnal*. 3. pp. 192–205.
32. Agapova, N.A. et al. (2018) *Slovar' detstva: govory Srednego Priob'ya (s lingvo-kul'turologicheskimi kommentariami)* [Dictionary of Childhood: Dialects of the Middle Ob Region (with Linguocultural Commentary)]. Tomsk: Tomsk State University.
33. Plokhotnyuk, T.G. (2005) *Narodnaya kul'tura ostayotsya vstrebovannoy: studenty FilF izuchayut fol'klor vo vremya uchebnoy praktiki* [Folk Culture Remains in Demand: PhilF Students Study Folklore During Educational Practice]. *Alma Mater*. 2387. 21.06.2005. [Online] Available from: http://almamater.tsu.ru/show_story.phtml?nom=2387&s=1861 (Accessed: 28.03.2025).
34. Nikolayev, D.S. (2025) *Sozdaniye elektronnoy korpusa fol'klornykh tekstov na russkom yazyke* [Creation of an Electronic Corpus of Folklore Texts in Russian]. [Online] Available from: https://profi.wciom.ru/fileadmin/file/nauka/grusha2015/s2_6/Nikolaev.pdf (Accessed: 28.03.2025).
35. *Traditsionnaya kul'tura*. (2012) *O fol'klornom podkorpuse natsional'nogo korpusa russkogo yazyka* [On the Folklore Subcorpus of the National Corpus of the Russian Language]. 1 (45). pp. 14–25.
36. Blinova, O.I. (ed.) (1999) *Vershininskiy slovar'* [Vershininsky Dictionary]. Vol. 2. Tomsk: Tomsk State University.

Сведения об авторе:

Волошина Светлана Владимировна – д-р филол. наук, профессор кафедры русского языка Томского государственного университета (Томск, Россия). E-mail: vsv1304@yandex.ru

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Svetlana V. Voloshina, Dr. Sci. (Philology), professor, National Research Tomsk State University (Tomsk, Russian Federation). E-mail: vsv1304@yandex.ru

The author declares no conflicts of interests.

*Статья поступила в редакцию 14.04.2025;
одобрена после рецензирования 20.08.2025; принята к публикации 04.09.2025*

*The article was submitted 14.04.2025;
approved after reviewing 20.08.2025; accepted for publication 04.09.2025*