

Научная статья

УДК 81'42 + 811.161.1.28

DOI 10.17223/18137083/79/22

Жанровая разметка Томского диалектного корпуса: от концепции – к реализации

Светлана Сергеевна Земичева

Томский государственный университет
Томск, Россия

optysmith@gmail.com, <https://orcid.org/0000-0003-3266-3696>

Аннотация

Статья выполнена на пересечении коммуникативной диалектологии и корпусной лингвистики, что обеспечивает ее актуальность. На примере Томского диалектного корпуса описан процесс и результаты внедрения жанровой разметки. Новизна реализуемого проекта обусловлена тем, что впервые на большом объеме диалектного материала маркированы не только фольклорные, но и речевые жанры. Описаны принципы и приемы осуществления жанровой разметки. Выявлены факторы, влияющие на состав жанров диалектного корпуса. Результаты разметки позволяют определить количественное соотношение разных речевых жанров в диалектной коммуникации. Достоверность выводов обеспечивается значительным объемом материала (более 16 тыс. жанровых фрагментов).

Ключевые слова

диалектный корпус, жанровая разметка, речевой жанр, русские говоры Сибири, диалектология, коммуникативная лингвистика, устная речь

Благодарности

Исследование выполнено при поддержке гранта РФФИ № 19-012-00320 «Томский диалектный корпус как новый ресурс для изучения народно-речевой культуры»

Для цитирования

Земичева С. С. Жанровая разметка Томского диалектного корпуса: от концепции – к реализации // Сибирский филологический журнал. 2022. № 2. С. 312–324. DOI 10.17223/18137083/79/22

Genre markup of the Tomsk dialect corpus: from concept to implementation

Svetlana S. Zemicheva

Tomsk State University
Tomsk, Russian Federation

optysmith@gmail.com, <https://orcid.org/0000-0003-3266-3696>

Abstract

The relevance of the study is due to the fact that it has been conducted at the intersection of two scientific fields: corpus linguistics and communicative dialectology. The paper presents

© Земичева С. С., 2022

ISSN 1813-7083

Сибирский филологический журнал. 2022. № 2. С. 312–324

Siberian Journal of Philology, 2022, no. 2, pp. 312–324

a comparative analysis of corpus practice based on the material of spoken language. Also, consideration is given to the process and results of creating a discursively annotated corpus of dialect speech with a size of more than 2 million tokens. Discursive markup implies the labeling of three parameters: topic, type, and genre of the text. The novelty of this research project is related to the fact that, for the first time, a large array of dialectal texts has been marked up according to intentional orientation: not only folklore but also speech genres have been annotated. The value of the new source is provided by the combination of archived data with current materials. A methodological advantage of the corpus is the possibility of combining qualitative and quantitative analysis. The paper describes the principles of implementing the genre marking in the Tomsk dialect corpus, with the factors influencing the composition of the genres of the dialect corpus revealed. The results make it possible to determine the quantitative ratio of different speech genres in dialectal communication. The observations are supported by examples from the spoken language of villagers. The accuracy of the conclusions is ensured by the reliance on quantitative data and a considerable amount of material (over 16 thousand genre fragments).

Keywords

dialectal corpus, genre markup, speech genre, Russian dialects of Siberia, dialectology, communicative linguistics, spoken language

Acknowledgements

The reported study was funded by RFBR according to the research project no. 19-012-00320 “Tomsk dialect corpus as a new resource for studying folk speech culture”

For citation

Zemicheva S. S. Genre markup of the Tomsk dialect corpus: from concept to implementation. *Siberian Journal of Philology*, 2022, no. 2, pp. 312–324. (in Russ.) DOI 10.17223/18137083/79/22

Введение. Обзор корпусной практики

Корпусное направление является одним из новых в лингвистике. Среди принципов данного подхода к языку – внимание к тексту (дискурсу); опора на количественные данные; внимание к синхронной и диахронной вариативности языка [Плунгян, 2008, с. 9–10]. В то же время обзор теоретических работ по использованию корпусных методов в исследованиях дискурса, сделанный К. М. Шилихиной [2014], включает преимущественно зарубежные статьи. В отечественном языкознании данное направление исследований пока не получило развития, хотя появляются отдельные публикации [Шмурак, 2020].

Существует и другая проблема, на которую не раз указывали ученые: непропорциональность представления письменной и устной речи – первая в имеющихся корпусах явно преобладает [Копотев, 2004, с. 12]. В связи с этим большое значение приобретает как включение данных устной речи в уже существующие корпуса, так и создание специальных корпусных ресурсов на разнообразном языковом материале, выходящем за рамки традиционного письменного текста. Среди проектов такого рода можно отметить Мультимедийный русский корпус, позволяющий изучать реальное звучание и жестовое сопровождение устной речи [Гришина, 2009]; корпус «Рассказы и разговоры о грушах», снабженный единой мультимедийной аннотацией, включающей жесты, интонацию и движения глаз [Кибрик и др., 2019]; корпус «Один речевой день» на материале естественной речи петербуржцев [Богданова-Бегларян и др., 2019; Шерстинова, 2015; 2018]; русский эмоциональный корпус [Котов, Будянская, 2016] и др. Создаются и диалектные корпуса (см. [Крючкова, Гольдин, 2011; Качинская, Малышева, 2019] и др.).

Задача маркирования типов и жанров устного текста решается в корпусной практике по-разному. В устном подкорпусе НКРЯ в разделе «Устная непубличная речь» выделены такие типы текста, как «беседа», «микродиалог», «разговор», «пересказ», «рассказ», «спор». На следующем уровне разметки они маркируются по месту записи, например: «микродиалог» – «в аптеке», «в библиотеке», «в магазине», «в транспорте» и т. п. [Гришина, 2005, с. 104]. Предполагается, что за маркировкой типа текста в этом случае стоит конкретная ситуация с легко опознаваемыми участниками и стереотипным представлением об их социальных ролях. При этом подкорпус не включает аудиозаписей, пользователю доступны только расшифровки речи.

Особая разметка типов и жанров речи представлена в мультимедийном корпусе русского языка, где единицей выдачи является кликст – видеоклип (фрагменты из кинофильмов, интервью), сопровождаемый расшифровкой. В корпусе проводится разметка по орфоэпической и вокалической структуре, типу коммуникативного действия или речевого состояния, выражаемого жестом и др. Для наших целей наиболее интересна разметка по типу речевого действия. Данная классификация разработана весьма подробно [Гришина, 2009, с. 193–199]. В качестве основных типов выделены: «вопросы», «утверждения», «этикетные высказывания», «императивы», «модальные высказывания», «перформативы», «шутливые высказывания», «мелиоративные высказывания», «пейоративные высказывания» и некоторые другие. Внутри типов выделяются подтипы, например, в рамках этикетных высказываний – «благодарность», «извинение», «пожелание», «поздравление» и др. [Там же]. Одним из основных критериев классификации в данном случае является интенция говорящего.

В корпусе «Один речевой день» для маркирования типа текста используется параметр, соответствующий понятию «речевой акт», под которым понимается «целенаправленное речевое действие, рассматриваемое в контексте прагматической ситуации и обладающее определенной иллюкутивной силой» [Шерстинова, 2015, с. 3–4]. При разметке корпуса авторы исходят из предположения, что «любое высказывание может быть интерпретировано как речевой акт определенного типа», и опираются на типологию речевых поступков, предложенную И. Н. Борисовой [Шерстинова, 2018; Борисова, 2009]. Добавим, что корпус находится в ограниченном доступе, что не позволяет познакомиться с реализацией его разметки более детально. В других известных нам корпусах устной речи жанровая разметка отсутствует.

Несколько иначе дискурсивная разметка реализуется в диалектных корпусах. Возможность когнитивно-дискурсивного анализа диалектных материалов была обоснована В. Е. Гольдиным [1997]. Диалектный корпус рассматривался как модель диалектной коммуникации и новый источник исследования [Крючкова, Гольдин, 2010]. В Саратовском диалектологическом корпусе планировалось маркирование таких типов текста, как повествование, описание, рассуждение, сказка, песня, частушка, пословицы и поговорки, произведения религиозного содержания [Там же, с. 12], однако данный ресурс не представлен в открытом доступе, поэтому оценить его современное состояние нет возможности.

В диалектном подкорпусе Национального корпуса русского языка представлены 2 типа текстов: устные фольклорные и устные нефольклорные (диалог, ответы по вопроснику, пересказ, пересказ разговора, рассказ). Конкретные разновидности текста создатели корпуса определяют как «жанры» [Качинская, Малышева, 2019, с. 108].

В корпусе диалектной лингвокультуры Северного Приангарья на материале речи жителей Красноярского края разметка по типу текста организована более сложным образом. Разработчики используют понятия «вид текста» (монолог, диалог, полилог); «тип текста» (повествование, описание, рассуждение) и «жанр» (устный рассказ, разговор / беседа, интервью, реплика). Различия между ними описываются на странице ресурса «Диалектный подкорпус»¹. Выделены также фольклорные жанры (заговоры, загадки, легенды, частушки и др.).

Обзор практических разработок в сфере корпусной диалектологии позволяет констатировать, что чаще всего жанровая разметка сводится к маркированию в корпусе фольклорных жанров. Разметка диалектных текстов по интенциональной направленности не была осуществлена ни в одном из существующих русских диалектных корпусов, чем обусловлена новизна настоящей работы.

1. Принципы и приемы жанровой разметки в Томском диалектном корпусе

Томский диалектный корпус (далее – ТДК) создается на основе диалектного архива кафедры русского языка Томского государственного университета и включает материалы, собранные в период с 1947 г. до настоящего времени в Среднем Приобье (по современному административному делению – Томская область и прилегающий к ней север Кемеровской области). Общий объем ресурса на сегодняшний день – свыше 2 000 000 словоупотреблений. В ТДК реализованы разные виды разметки: паспортная, дискурсивная, морфологическая². Дискурсивная разметка подразумевает маркирование трех параметров: тема, тип, жанр текста.

В первоначальной концепции ТДК, разработанной Е. А. Юриной, предлагалось выделение макрожанров (биографический рассказ, сюжетный рассказ, описание, рассуждение, интервью) и собственно жанров по интенции говорящего: информативные, оценочные, ритуальные, императивные с последующей дифференциацией [Юрина, 2011].

В ходе реализации проекта данная концепция была несколько трансформирована. Вместо макрожанров выделяются типы текста, отражающие неоднородность имеющихся материалов. Всего выделено 10 типов текста, которые можно объединить в 4 группы:

I) спонтанные тексты: диалог диалектоносителей, полилог, ситуативные вкрапления, реплики;

II) воспроизводимые тексты: религиозные, прецедентные, пересказ разговора;

III) спровоцированные тексты: ответы на вопросники;

IV) письменные тексты: официальные документы, школьные тетради (эти типы встречаются в имеющихся материалах очень редко).

Как и в первоначальной концепции, за основу жанрового деления нефольклорной речи взят перечень речевых жанров, предложенный Т. В. Шмелёвой [1997]. Список жанров был значительно изменен. Часть жанров, предложенных в перечне Е. А. Юриной (например, поручение), не использовалась при разметке в связи с их не востребованностью в имеющихся материалах; другие жанры, напротив,

¹ См.: Электронный текстовый корпус лингвокультуры Северного Приангарья. URL: <http://angara.sfu-kras.ru/?page=dialect#> (дата обращения 03.05.2021).

² Томский диалектный корпус v. 2. // Лаборатория общей и сибирской лексикографии НИ ТГУ. URL: http://losl.tsu.ru/losl_search (дата обращения 05.12.2021). Режим доступа: свободный.

в ходе работы были введены в перечень (возражение, угроза, поучение, обещание, разрешение, отказ, злопожелание и др.). При разработке классификации диалектных речевых жанров учитывались результаты работы О. А. Казаковой [2007], где подробно охарактеризован каждый жанр, описаны их отличия друг от друга.

Итоговый список жанров выглядит следующим образом.

Среди императивных жанров в ТДК выделены предложение, просьба, распоряжение, приказ, совет, поучение, запрет, угроза, разрешение, отказ, согласие, обещание. Отличия между предложением, просьбой, распоряжением – в степени категоричности. Приказ противопоставлен другим жанрам на основе дополнительного критерия принадлежности высказывания лицу, облеченному властью. Разрешение, отказ, согласие представляют собой реакции на побуждения собеседника. Совет и обещание противопоставлены на основе актора, потенциально осуществляющего действия (в первом случае это слушающий, во втором – говорящий). В отличие от совета, поучение касается не конкретных действий, а общих правил поведения, принципов жизни.

Среди ритуальных жанров в ТДК маркируются следующие: приветствие, прощание, благодарность, извинение, поздравление, благопожелание, злопожелание, приглашение, угощение, связанные с определенными этикетными ситуациями.

При классификации оценочных жанров на первом уровне разделяются жанры с негативной и позитивной оценкой. Далее обобщенно характеризуется объект оценки (говорящий, другие люди, явления – растения, животные, предметы, события, жизнь в целом). В результате выделяется 7 оценочных речевых жанров: негативная самооценка, порицание (негативная оценка другого человека, высказываемая заочно), упрек (негативная оценка другого человека в глаза), неодобрение (негативная оценка предмета, явления); позитивная самооценка, похвала (позитивная оценка другого человека), одобрение (позитивная оценка предмета, явления).

При классификации информативных жанров учитывался, во-первых, тип речи (монологическая / диалогическая), во-вторых, общее содержание информации (мысли / действия / чувства), в-третьих, модальность высказывания. В ТДК в монологической речи маркируются следующие информативные жанры: объяснение (цель – сообщить собеседнику новую информацию), намерение (цель – сообщить собеседнику о своих планируемых действиях), жалоба (цель – сообщить о своих негативных эмоциях, получить сочувствие собеседника). Кроме того, выделяются предположение, сообщение о желании, сообщение о необходимости, связанные с передачей дополнительных модальных смыслов (неуверенность, желательность, необходимость). На основе интенциональной направленности размечаются по жанру также диалогические реплики: вопрос (цель – получить информацию от собеседника); подтверждение (сообщает, что выраженное в вопросе собеседника предположение верно), отрицание (сообщает, что предположение собеседника неверно), возражение (оспаривание позиции собеседника в ответ на его утверждение).

Среди фольклорных жанров выделены следующие: обрядовый фольклор; заговор; лирическая песня; фольклорная проза; пословица / поговорка; примета; загадка; детский фольклор; частушка; анекдот.

Как видно из перечня, в некоторых случаях представлены группы жанров (обрядовый фольклор, лирические песни, фольклорная проза, детский фольклор). Это связано с тем, что, во-первых, фольклорные материалы в имеющихся архивных данных представлены эпизодически; во-вторых, команда исследователей не

включала фольклориста, поэтому разграничить смежные жанры в некоторых случаях было трудно.

Краткое описание каждого жанра и примеры соответствующих текстов даны в пользовательской инструкции³.

Жанровая разметка ТДК осуществлялась в 2020–2021 гг. сотрудниками Лаборатории общей и сибирской лексикографии ТГУ А. А. Васильченко, С. С. Земичевой и доцентом кафедры русского языка Л. А. Дубцовой.

В ходе работы выработаны следующие **принципы** жанровой разметки в ТДК:

- 1) жанры маркируются вручную на основе заранее составленного перечня;
- 2) разметка несплошная, маркируются лишь те фрагменты текста, которые соответствуют одному из жанров перечня;
- 3) жанр может быть частью высказывания, совпадать с отдельным высказыванием или представлять собой совокупность высказываний;
- 4) жанры маркируются как в актуальной коммуникации, так и внутри цитирования чужой речи. Такой подход позволяет отчасти компенсировать отсутствие или слабую представленность в корпусе некоторых интенциональных фрагментов (например, упрека), скрытых от постороннего наблюдателя;
- 5) жанр рассматривается как автономный от содержания конструкт.

Кратко охарактеризуем **методику** жанровой разметки. За основу взят подход А. Вежицкой [1997], основанный на идее семантических примитивов. Выявление коммуникативного смысла речевого поступка в данной методике основывается на экспликации его имплицитно-перформативного модуса. В качестве рабочих приемов можно использовать переформулирование высказывания, а также подстановку перформативного глагола, например, фразу *не торопитесь, деуки, замуж* легко квалифицировать как совет, поскольку она допускает подстановку глагола *советую*, в то время как другие глаголы (*запрещаю, *распоряжаюсь, *предлагаю) не могут быть включены в ее структуру.

Другой прием – отграничение от смежных жанров с опорой на имеющиеся теоретические сведения и групповые договоренности. Например, фраза, приведенная выше, не может быть отнесена к запрету, так как запрет предполагает, что «выгодополучателем» от несовершения и прекращения действия будет говорящий (а в случае совета – слушающий).

В техническом плане жанровая разметка основывается на приписывании каждому фрагменту текста числовых идентификаторов – тегов.

Для пользователя корпуса доступен поиск по конкретному жанру, при этом возможны комбинации с другими типами поиска (выбор темы, определенного периода или места записи, пола, возраста, уровня образования информанта). Уникальность созданного ресурса обусловлена возможностью сочетать при поиске тематические и жанровые параметры (например, можно найти песни о Великой Отечественной войне или одобрительные высказывания о еде). Сочетание жанрового поиска с пословным позволяет также частично нейтрализовать явления омографии и полисемии (например, поиск слова *пожалуйста* в рамках жанра «благодарность» или слова *пока* только в жанре «прощание»).

³ Текстовая разметка Томского диалектного корпуса // Томский диалектный корпус. Инструкция для пользователя. URL: https://losl.tsu.ru/sites/default/files/instructions/Topics_result.docx (дата обращения 02.05.2021). Режим доступа: свободный.

2. Жанровая разметка в Томском диалектном корпусе: количественные результаты

В результате реализации жанровой разметки в ТДК выделено более 16 000 размеченных фрагментов. Наиболее частотны в имеющихся материалах информативные речевые жанры (8 379 фрагментов). Достаточно частотны также оценочные (3 475), императивные (2 276) речевые жанры, реже встречаются жанры фольклорного текста (527) и ритуальные речевые жанры (467).

Количественный анализ речевых жанров позволяет сделать следующие наблюдения. Среди информативных диалогических реплик более частотно отрицание, имеющее реактивный характер, по сравнению с возражением, порождаемым внутренним несогласием, без вопроса (782 и 261 фрагмент соответственно).

Соотношение негативно оценочных и позитивно оценочных речевых жанров примерно равное (1 539 и 1 570 фрагментов соответственно на момент подготовки статьи). При этом оценка чаще направлена на внешний мир (1 697), других людей (1 080), реже актуализируется самооценка (332 фрагмента).

Среди негативно оценочных речевых жанров чаще встречается порицание, которое дается заочно, чем упрек, высказываемый «в глаза» (437 и 131 фрагмент соответственно).

Среди ритуальных речевых жанров благопожелания (122) используются значительно чаще злопожеланий (10), что отчасти связано с ситуацией общения (запись речи в присутствии посторонних), а также свидетельствует о доброжелательности как коммуникативной доминанте народно-речевой культуры.

Среди фольклорных жанров к наиболее востребованным относятся пословицы и поговорки (178). В отличие от других фольклорных текстов, эти малые жанры часто используются в речи спонтанно.

В группе текстов, маркированных в корпусе как «Обрядовый фольклор», значительная часть (83 % всех) приходится на свадебные причитания и приговоры; тексты, связанные с другими обрядами (крестильным, похоронным, календарными), отмечаются значительно реже. Частотность данных текстов, как представляется, обусловлена не только значимостью свадебного обряда, но и влиянием собирателей (вопрос о свадьбе в экспедициях задается очень часто).

3. Факторы, влияющие на жанровую структуру диалектного корпуса

На жанровую структуру коммуникации влияют такие факторы, как степень знакомства и характер отношений, возраст и пол коммуникантов, ситуация общения, тип речевой культуры и др. На данный момент определено влияние некоторых из них на жанровый состав ТДК.

Коммуникативная ситуация сбора материала предполагает общение малознакомых людей, чаще одного пола (женщины), различных по возрасту (молодые собиратели, пожилые информанты) и уровню образования (собиратели – студенты или преподаватели университета, информанты – малограмотные крестьяне). Цель собирателя – сбор информации о традиционной культуре, о жизни в прошлом. Эта установка порождает востребованность речевого жанра объяснения (всего в ТДК выделено 618 фрагментов): [*А что такое полати?*] *Полати? А это русская печка была, вот так, а здесь были отстойники такие деревянные и там доски настелены. Там спать можно было. Как печка, допустим, моя стоит, а сбоку туда вот сюда как бы сбоку полати, деревянные. Там спали* (Большая Саровка, 2016). Наличие жанра «объяснение» как части биографического расска-

за диалектоносителя отмечалось С. В. Волошиной [2016, с. 50–51]. Некоторые наблюдения о функционировании данного жанра в речи диалектоносителя В. П. Вершининой сделаны в работе О. А. Казаковой [2007, с. 54–55].

Коммуникативная ситуация экспедиционного сбора материала актуализирует также речевой жанр специализированного совета, в котором говорящий рекомендует других информантов: *Председатель колхозу у нас был старый активист. На двух войнах был. Да вы к ему сходите, там через дом живёт* (Улановка, 1968). Подобные высказывания составляют 20 % от всех, относящихся к речевому жанру «совет».

Спонтанный характер ситуации общения, неподготовленность информанта объясняет актуализацию речевого жанра извинения (28 примеров): *Извините, что я... сижу как эта, ага. С огорода. Я на огороде была* (Вознесенка, 2018). Распространено извинение за использование сниженной или нецензурной лексики: *И вот она говорит, я тебя всю жизнь вспоминаю, как же ты могла видеть, что это говно [о плохом человеке]. Простите меня за выражение* (Колпашево, 2016). Наличие подобных высказываний связано с тем, что собиратель воспринимается как представитель другой культуры. Сходные наблюдения сделаны О. А. Казаковой [2007, с. 70–71].

Разница в возрасте актуализирует речевой жанр поучения (105 фрагментов), направленный от старшего поколения младшему: *Вот так, девочки, живите, свою жизнь берегите. Выйдете замуж, живите, цените всё, чтоб как попало не жить, и не пить, и стараться для себя всё делать. Вот тогда будет очень... учётесь, учётесь, старайтесь, всю силу отдавайте. Шчас учёным только жизнь. А такі [без образования], будут только работать* (Рыбалово, 2015).

Добавим, что влияние культурного фактора проявляется не столько в номенклатуре жанров, сколько в содержании соответствующих фрагментов, определенных жанровых стратегиях или «сценариях» (по А. Вежбицкой). Материалы ТДК позволяют сделать новые наблюдения в этом направлении.

Так, специфика народно-речевой культуры проявляется в способе выражения позитивной самооценки, которая часто реализуется через цитирование чужой речи или ссылку на мнение других членов сельского социума: *Они [дети] меня все слушались, они до сих пор, у них мама на первом плане. Не хочу хвалиться... вот можете у любого из соседей спросить, вам хоть кто скажет: мама на первом плане до сих пор. Со всем советуются, со всеми бедами бегут ко мне, даже внуки... с мамой поссорятся, тут как тут. У меня девять внуков, десять правнуков: вот это баушка. Четверо детей* (Новосельцево, 2012). Такое поведение соответствует этикетному правилу, отраженному в пословице: «не хвали себя сам (жди, когда люди похвалят)». Важно отметить, что данная стратегия сохраняется в речи сельских жителей старшего поколения до настоящего времени.

Региональный и темпоральный факторы, судя по всему, не оказывают существенного влияния на жанровую структуру диалектной коммуникации. В отдельных случаях можно отметить динамику языковых средств, например использование междометия *давайте* в качестве прощальной реплики, что является относительно новым явлением, характерным для разговорной речи [Ермолов, Богданова-Бегларян, 2019]: *Чё, всё, наверно? Ну, давайте* (Петрово, 2008).

Влияние на состав и функционирование речевых жанров таких социолингвистических факторов, как пол и уровень образования информанта на данном этапе также не очевидно и требует дальнейших исследований.

Выводы

Проведенный анализ показал, что проблема формализации устного дискурса решается по-разному в существующих корпусах русского языка. При этом нередко используется понятие «речевого акта» или «речевого жанра».

Материалы диалектных архивов представляют собой ценный ресурс, на основе которого может быть создан дискурсивно размеченный корпус устной речи, однако эта задача до сих пор не была решена в полной мере. Осуществление жанровой разметки диалектного корпуса позволяет отразить структуру соответствующего дискурса, исследовать на его основе выделенные жанры диалектной речи с учетом коммуникативных, социолингвистических и лингвокультурологических аспектов, а также «анализировать речевые жанры количественно», что является «давней мечтой жанроведов» [Дементьев, Степанова, 2016, с. 25].

Жанровая разметка в ТДК, как и в других диалектных корпусах, осуществляется вручную. При этом перечень включает не только фольклорные, но и речевые жанры. Методологическую основу процедуры определения речевого жанра составляют работы А. Вежбицкой, О. А. Казаковой.

Количественные результаты позволяют сделать вывод о том, что в диалектной коммуникации преобладают информативные речевые жанры, оценочные и императивные речевые жанры используются реже, наименьшую частотность имеют жанры фольклорного текста и ритуальные речевые жанры. Достоверность выводов обусловлена значительным объемом проанализированного материала (более 16 тыс. фрагментов).

Данные жанрово размеченного текстового корпуса позволяют выявить факторы, влияющие на жанровую структуру коммуникации. Так, коммуникативная ситуация беседы с собирателем способствует появлению речевых жанров объяснения, совета, извинения. Разница в возрасте актуализирует речевой жанр поучения, направленный от старшего поколения к младшему. Эти выводы в целом согласуются с наблюдениями других исследователей.

Представляется, что жанровая разметка, в отличие от тематической, не имеет региональной специфики, а выявленные закономерности (например, в количественном соотношении и сочетании жанров) типичны не только для диалекта, но и для других разновидностей устной речи. Своеобразие традиционной народно-речевой культуры выявляется не через номенклатуру жанров, а через их содержательный анализ, описание стратегий и тактик говорящего. В частности, в речевом жанре позитивной самооценки выявляется специфическая тактика ссылки на общественное мнение.

Проведенная работа вносит вклад в решение масштабной проблемы формализации дискурса, что определяет ее теоретическую значимость. Внедрение жанровой разметки Томского диалектного корпуса будет стимулировать развитие таких научных направлений, как коммуникативная диалектология, диалектная жанрология, лингвоперсонология и др. В то же время отсутствие в новых материалах ярко выраженной региональной специфики дает возможность использовать имеющиеся материалы в преподавании русского языка как иностранного. Представленный опыт может быть востребован при создании других корпусных ресурсов.

Список литературы

Богданова-Бегларян Н. В., Блинова О. В., Зайдес К. Д., Попова Т. И., Шерстинова Т. Ю. Корпус естественной речи: проблемы ручного аннотирования прагма-

тических маркеров и пути их решения // Анализ разговорной русской речи (АР³-2019). Труды восьмого междисциплинарного семинара. 2019. С. 5–10.

Борисова И. Н. Русский разговорный диалог: структура и динамика. М.: ЛИБРОКОМ, 2009. 320 с.

Вежбицкая А. Речевые жанры // Жанры речи. 1997. № 1. С. 99–112.

Волошина С. В. Речевой жанр автобиографического рассказа в диалектной коммуникации // Портреты речевых жанров: разные дискурсивные практики / Под ред. Т. А. Демешкиной. Томск: Изд-во Том. ун-та, 2016. С. 37–96.

Гольдин В. Е. Теоретические проблемы коммуникативной диалектологии: Дис. ... д-ра филол. наук в виде науч. докл. Саратов, 1997. 52 с.

Гришина Е. А. Мультимедийный русский корпус (МУРКО): проблемы аннотации // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. СПб., 2009. С. 175–214.

Гришина Е. А. Устная речь в Национальном корпусе русского языка // Национальный корпус русского языка: 2003–2005. М., 2005. С. 94–110.

Дементьев В. В., Степанова Н. Б. Корпусные методы в исследовании речевых жанров: проблема ключевых фраз // Жанры речи. 2016. № 3. С. 24–41.

Ермолов О. Б., Богданова-Бегларян Н. В. Языковое оформление прощания в современной разговорной речи (на материале звукового корпуса «Один речевой день») // Коммуникативные исследования. 2019. Т. 6, № 2. С. 307–331.

Казакова О. А. Диалектная языковая личность в жанровом аспекте. Томск: Изд-во Том. политехн. ун-та, 2007. 200 с.

Качинская И. Б., Малышева А. В. Народная речь в Национальном корпусе русского языка // Русская речь. 2019. № 4. С. 103–118.

Кибрик А. А., Кортаев Н. А., Федорова О. В., Евдокимова А. А. Единая мультимедийная аннотация как инструмент анализа естественной коммуникации // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной международной конференции «Диалог» (Москва, 29 мая – 1 июня 2019 г.). М., 2019. Вып. 18 (25). С. 265–280.

Коптев М. В. Введение в корпусную лингвистику: Учеб. пособие для студентов филологических и лингвистических специальностей университетов. Прага: Animedia Company, 2014. 195 с.

Котов А. А., Будянская Е. М. Видеокорпус обращений граждан по вопросам оплаты коммунальных услуг // Вестник Ярослав. гос. ун-та им. П. Г. Демидова. Серия: Гуманитарные науки. 2016. № 2 (36). С. 93–99.

Крючкова О. Ю., Гольдин В. Е. Корпус русской диалектной речи: концепция и параметры оценки // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Беласово, 25–29 мая 2011 г.). М., 2011. Вып. 10 (17). С. 359–367.

Крючкова О. Ю., Гольдин В. Е. Саратовский диалектный корпус: новый научный и образовательный ресурс. Концепция, методические материалы. Саратов, 2010. 39 с.

Плунгян В. А. Корпус как инструмент и как идеология: о некоторых уроках современной корпусной лингвистики // Русский язык в научном освещении. 2008. № 16 (2). С. 7–20.

Шерстинова Т. Ю. Прагматическое аннотирование коммуникативных единиц в корпусе ОРД: микроэпизоды и речевые акты // Корпусная лингвистика – 2015. Тр. Междунар. конф. / Отв. ред. В. П. Захаров, О. А. Митрофанова, М. В. Хохлова. СПб., 2015. С. 451–459.

Шерстинова Т. Ю. Структура повседневного диалога как последовательность речевых актов // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной международной конференции «Диалог» (Москва, 30 мая – 2 июня 2018 г.). М., 2018. Вып. 17 (24). С. 637–651.

Шилихина К. М. Использование корпусов в исследованиях дискурса // Вестник ВГУ. Серия: Лингвистика и межкультурная коммуникация. 2014. № 3. С. 21–26.

Шмелёва Т. В. Модель речевого жанра // Жанры речи: Сб. науч. ст. Саратов, 1997. С. 91–96.

Шмурак Р. И. К уточнению понятия упрёка с помощью корпусных инструментов // Изв. РАН. Серия литературы и языка. 2020. Т. 79, № 3. С. 24–48.

Юрина Е. А. Томский диалектный корпус: в начале пути // Вестник Том. гос. ун-та. Филология. 2011. № 2 (14). С. 58–63.

Список источников

Томский диалектный корпус v 2 // Лаборатория общей и сибирской лексикографии НИ ТГУ. URL: http://losl.tsu.ru/losl_search (дата обращения 05.12.2021). Режим доступа: свободный.

Текстовая разметка Томского диалектного корпуса // Томский диалектный корпус. Инструкция для пользователя. URL: https://losl.tsu.ru/sites/default/files/instructions/Topics_result.docx (дата обращения 02.05.2021). Режим доступа: свободный.

References

Bogdanova-Beglaryan N. V., Blinova O. V., Zaydes K. D., Popova T. I., Sherstinoва T. Yu. Korpus estestvennoy rechi: problemy ruchnogo annotirovaniya pragmaticheskikh markerov i puti ikh resheniya [Natural speech corpus: problems and solutions of manual pragmatic markers annotation]. In: *Analiz razgovornoy russkoy rechi (ARz-2019). Trudy vos'mogo mezhdistsiplinarnogo seminara* [Analysis of colloquial Russian speech (ARz-2019). Proceedings of the eighth interdisciplinary seminar]. St. Petersburg, 2019, pp. 5–10.

Borisova I. N. *Russkiy razgovornyy dialog: struktura i dinamika* [Russian conversational dialogue: structure and dynamics]. Moscow, LIBROKOM, 2009, 320 p.

Dement'ev V. V., Stepanova N. B. Korpusnye metody v issledovanii rechevykh zhanrov: problema klyuchevykh fraz [Corpus genristics: a problem of key phrases]. *Speech Genres*. 2016, no. 3, pp. 24–41.

Ermolov O. B., Bogdanova-Beglaryan N. V. Yazykovoe oformlenie proshchaniya v sovremennoy razgovornoy rechi (na materiale zvukovogo korpusa “Odin rechevyden”) [Linguistic design of “goodbye” situation in modern colloquial speech (based on the material of the speech corpus “one speaker’s day”)]. *Communication Studies*. 2019, vol. 6, no. 2, pp. 307–331.

Gol'din V. E. *Teoreticheskie problemy kommunikativnoy dialektologii* [Theoretical problems of communicative dialectology]. Abstract of Dr. philol. sci. diss. Saratov, 1997, 52 p.

Grishina E. A. Mul'timediynny russkiy korpus (MURCO): problemy annotatsii [Multimedia Russian Corpus (MURCO): problems of annotation]. In: *Natsional'nyy korpus russkogo yazyka: 2006–2008. Novye rezul'taty i perspektivy* [National corpus of the Russian language: 2006–2008. New results and prospects]. St. Petersburg, 2009, pp. 175–214.

Grishina E. A. Ustnaya rech' v Natsional'nom korpuse russkogo yazyka [Spoken language in the National Corpus of the Russian language]. In: *Natsional'nyy korpus russkogo yazyka: 2003–2005* [National corpus of the Russian language: 2003–2005]. Moscow, 2005, pp. 94–110.

Kachinskaya I. B., Malysheva A. V. Narodnaya rech' v Natsional'nom korpuse russkogo yazyka [Folk speech in Russian National Corpus]. *Russkaya Rech' (Russian Speech)*. 2019, no. 4, pp. 103–118.

Kazakova O. A. *Dialektnaya yazykovaya lichnost' v zhanrovom aspekte* [Dialectal linguistic personality in the genre aspect]. Tomsk, TSPU Publ., 200 p.

Kibrik A. A., Korotaev N. A., Fedorova O. V., Evdokimova A. A. Edinaya mul'tichanal'naya annotatsiya kak instrument analiza estestvennoy kommunikatsii [Unified multichannel annotation: a tool for analysing natural communication]. In: *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: Po materialam ezhegodnoy mezhdunarodnoy konferentsii "Dialog" (Moskva, 29 maya – 1 iyunya 2019 g.)* [Computer linguistics and intellectual technologies: On the materials of the annual international conference "Dialogue" (Moscow, May 29 – June 1, 2019)]. Moscow, 2019, iss. 18 (25), pp. 265–280.

Kopotev M. V. *Vvedenie v korpusnuyu lingvistiku: Ucheb. posobie dlya studentov filologicheskikh i lingvisticheskikh spetsial'nostey universitetov* [Introduction to the corpus linguistics] Praga, Animedia Company, 2014, 195 p.

Kotov A. A., Budyanskaya E. M. Videokorpus obrashcheniy grazhdan po voprosam oplaty kommunal'nykh uslug [Video corpus of dialogues on issues of utility bills payments]. *Vestnik Yaroslavskego gosudarstvennogo universiteta im. P. G. Demidova. Seriya gumanitarnye nauki*. 2016, no. 2 (36), pp. 93–99.

Kryuchkova O. Yu., Gol'din V. E. Korpus russkoy dialektnoy rechi: kontseptsiya i parametry otsenki [A corpus of Russian dialectal speech: the concept and parameters of evaluation]. In: *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: Po materialam ezhegodnoy mezhdunarodnoy konferentsii "Dialog" (Moskva, 25–29 maya, 2011 g.)* [Computer linguistics and intellectual technologies: On the materials of the annual international conference "Dialogue" (Moscow, May 25–29, 2011)]. Moscow, 2011, iss. 10 (17), pp. 359–367.

Kryuchkova O. Yu., Gol'din V. E. *Saratovskiy dialektnyy korpus: novyy nauchnyy i obrazovatel'nyy resurs. Kontseptsiya, metodicheskie materialy* [Saratov dialect corpus: a new scientific and educational resource. Concept, teaching materials]. Saratov, 2010, 39 p.

Plungyan V. A. Korpus kak instrument i kak ideologiya: o nekotorykh urokakh sovremennoy korpusnoy lingvistiki [Corpus as a tool and as an ideology: on some lessons of modern corpus linguistics]. *Russian Language and Linguistic Theory*. 2008, no. 16 (2), pp. 7–20.

Sherstinova T. Yu. Pragmaticheskoe annotirovanie kommunikativnykh edinit v korpuse ORD: mikroepizody i rechevye akty [Approaches to pragmatic annotation in the ORD corpus: micro episodes and speech acts]. In: *Korpusnaya lingvistika – 2015. Tr. Mezhdunar. konf.* [Corpus linguistics – 2015. Proceedings of the intern. conf.] V. P. Zakharov, O. A. Mitrofanova, M. V. Khokhlova (Eds). St. Petersburg, 2015, pp. 451–459.

Sherstinova T. Yu. Struktura povsednevnogo dialoga kak posledovatel'nost' rechevykh aktov [The structure of everyday dialogue as the sequence of speech acts]. In: *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: Po materialam ezhegodnoy mezhdunarodnoy konferentsii "Dialog" (Moskva, 30 maya – 2 iyunya 2018 g.)* [Com-

puter linguistics and intellectual technologies: On the materials of the annual international conference “Dialogue” (Moscow, May 30 – June 2, 2018)]. Moscow, 2018, issue 17 (24), pp. 637–651.

Shilikhina K. M. Ispol'zovanie korpusov v issledovaniyakh diskursa [Using corpora in discourse research]. *Proceedings of Voronezh State University. Series: Linguistics and intercultural communication*. 2014, no. 3, pp. 21–26.

Shmeleva T. V. Model' rechevogo zhanra [The model of speech genre]. In: *Zhanry rechi: Sb. nauch. st.* [Genres of speech: Coll. of sci. papers]. Saratov, 1997, pp. 91–96.

Shmurak R. I. K utochneniyu ponyatiya upreka s pomoshch'yu korpusnykh instrumentov [Clarifying the concept of reproach using corpus tools]. *The Bulletin of the Russian Academy of Sciences: Studies in Literature and Language*. 2020, vol. 79, no. 3, pp. 24–48.

Voloshina S. V. Rechevoy zhanr avtobiograficheskogo rasskaza v dialektnoy kommunikatsii [Speech genre of autobiographical story in dialect communication]. In: *Portrety rechevykh zhanrov: raznye diskursivnye praktiki* [Portraits of speech genres: different discursive practices]. T. A. Demeshkina (Ed.). Tomsk, TSU Publ., 2016, pp. 37–96.

Wierzbicka A. Rechevye zhanry [Speech genres]. *Speech Genres*. 1997, no. 1, pp. 99–112.

Yurina E. A. Tomskiy dialektnyy korpus: v nachale puti [Tomsk dialectal corpus: the starting point]. *Tomsk State University Journal of Philology*. 2011, no. 2 (14), pp. 58–63.

List of sources

Tekstovaya razmetka Tomskogo dialektного korpusa [Textual markup of the Tomsk dialect corpus]. In: *Tomskiy dialektnyy korpus. Instruktsiya dlya pol'zovatelya* [Tomsk dialect corpus. Instructions for the user]. URL: https://losl.tsu.ru/sites/default/files/instructions/Topics_result.docx (accessed: 15.06.2021). Access mode: free.

Tomskiy dialektnyy korpus v 2 [Tomsk dialect corpus]. In: *Laboratoriya obshchey i sibirskoy leksikografii NI TGU* [Laboratory of General and Siberian Lexicography of NR TSU]. URL: http://losl.tsu.ru/losl_search (accessed: 15.06.2021). Access mode: free.

Информация об авторе

Светлана Сергеевна Земичева, кандидат филологических наук

Information about the author

Svetlana S. Zemicheva, Candidate of Philology

Статья поступила в редакцию 21.10.2021;
одобрена после рецензирования 21.12.2021; принята к публикации 21.12.2021
The article was submitted 21.10.2021;
approved after reviewing 21.12.2021; accepted for publication 21.12.2021