

RU

Анализ пересечений в регионально окрашенной лексике территориальных единиц Западной Сибири

Матвеева И. В.

Аннотация. Цель исследования – выявить пересечения и характерность употребления регионализмов в пяти регионах Западной Сибири. В статье рассматривается феномен регионального варьирования языка, а именно тенденции в изучении региональных особенностей языка на территории страны, а также результаты исследований регионально окрашенной лексики, полученные с применением современных технологий в выявлении и анализе регионально ограниченных единиц. Исследователями активно изучается регионально ограниченная лексика различных регионов, но вопрос пересечения подобной лексики в соседних регионах часто является малоизученным вопросом, так как требует большого количества данных для сопоставления. Региональная лексика регионов Западной Сибири, а именно: Новосибирской, Омской, Томской, Кемеровской областей и Алтайского края, подвергалась изучению, но зачастую исследовалась только одна область. Вопрос о совпадении регионализмов в соседних регионах неоднократно поднимался в современных исследованиях как требующий более детальной проработки. Но анализ пересекающейся лексики в перечисленных регионах в существующих на данный момент работах отсутствует, в связи с чем научная новизна исследования состоит в проведении сравнительного анализа регионально ограниченной лексики пяти регионов Западной Сибири на предмет пересечений и совпадений, в выявлении особенностей и представлении статистики подобных пересечений, а также в рассмотрении такого показателя, как характерность использования слов, на материале, полученном в результате глобального лингвистического исследования компании Яндекс. В результате исследования установлено, что количество «широких» регионализмов (использующихся в двух и более регионах) в соседних регионах может быть достаточно высоким.

EN

Analysis of intersections in the regionally marked vocabulary of the territorial units of Western Siberia

Matveeva I. V.

Abstract. The aim of the research is to identify the intersections and characteristics of the use of regionally specific words in five regions of Western Siberia. The paper examines the phenomenon of regional variation of language, namely, trends in the study of regional peculiarities of language in the country, as well as the results of studies of regionally marked vocabulary obtained using modern technologies in identifying and analyzing regionally limited units. Researchers are actively investigating the regionally limited vocabulary of various regions, but the question of the intersection of such vocabulary in neighboring regions is often a little-studied issue, since it requires a large amount of data for comparison. The regional vocabulary of the regions of Western Siberia, namely, the Novosibirsk, Omsk, Tomsk, Kemerovo regions and the Altai Territory, has been examined, but often only one area has been investigated. The question of the coincidence of regionally specific words in neighboring regions has been repeatedly raised in modern studies as requiring more elaboration. But there is no analysis of intersecting vocabulary in the listed regions in the currently existing works, and, therefore, the scientific novelty of this research lies in carrying out a comparative analysis of the regionally limited vocabulary of the five regions of Western Siberia to find intersections and coincidences, in identifying features and presenting statistics for these intersections, as well as in considering such an indicator as the specificity of the use of words involving the material obtained during a global linguistic research by Yandex. As a result of the research, it was found that the number of “broad” regionally specific words (used in two or more regions) in neighboring regions can be quite high.

Введение

Актуальность данного исследования обусловлена тем, что изучение языкового разнообразия в регионах развивается на протяжении многих лет, и традиционные научные термины заменяются новыми, подчеркивая современные реалии. Регионально окрашенная лексика подвергается лексикографированию, становится предметом лингвистических дискуссий, а ареалы регионализмов вызывают вопросы и интерес не только у ученых-лингвистов, но и у жителей регионов, употребляющих регионально окрашенные лексемы. Отдельные области, регионы и географические ареалы достаточно давно подвергаются исследованию с точки зрения региональной специфики. Исследователи собирают материал у носителей языка, проживающих в конкретных регионах, классифицируют его, типологизируют, лексикографируют, присваивают фонетические и грамматические характеристики. Зачастую результаты исследования представляют собой актуальную региональную картину мира, переданную в том числе с помощью региолекта. Но вопрос в том, насколько лексемы, считающиеся регионализмами в одном регионе, могут считаться регионализмами и в соседних регионах. Насколько пересекаются данные слова, что они означают в соседних регионах и можно ли считать слово регионализмом, если оно распространено не только в одном, а в нескольких соседних регионах? Проблема пересечений и совпадений регионально окрашенной лексики, которая была отнесена исследователями к одному конкретному региону, вызвала и ставила перед учеными новые вопросы. Исследование нескольких территорий одновременно представлялось трудоемким процессом из-за необходимости использовать большой объем данных. В текущих условиях, когда стало возможным задействовать искусственный интеллект в анализе большого количества данных, есть шанс ответить на вопросы о том, насколько же пересекается регионально окрашенная лексика в соседних регионах и насколько характерны те или иные лексемы в конкретных регионах.

Для достижения вышеуказанной цели исследования необходимо решить следующие задачи:

- собрать и проанализировать теоретическую базу современных исследований регионально окрашенной лексики;
- рассмотреть понятия «широкий регионализм» и «узкий регионализм» по отношению к регионально окрашенной лексике пяти регионов Западной Сибири;
- выявить и описать существующие пересечения в регионально окрашенной лексике регионов и сделать выводы относительно имеющихся пересечений;
- рассмотреть и сравнить характерность слов для рассматриваемых регионов.

Материалом для работы послужил словарь региональной лексики, составленный с помощью исследования Яндекс, с привлечением лингвистов Института русского языка им. В. В. Виноградова Российской академии наук, а именно лексемы, отнесенные к регионализмам в исследуемых регионах (567 регионализмов (<https://yastatic.net/s3/milab/2021/local-words/data/RussianLocalWordsDataset.zip>)). К исследованию также привлекались сведения из словаря «Языки русских городов» (<https://web.archive.org/web/20120705132446/http://community.lingvo.ru/goroda/dictionary.asp>).

Теоретическую базу исследования составляют труды в области лингвистической регионалистики, а именно работа А. С. Герда (2001), в которой описана часть используемой терминологии (региолект); концептуальные вопросы регионального варьирования и причины возникновения такого явления представлены в статьях Т. И. Ерофеевой, Ф. Л. Скитовой (1990) и Е. В. Ерофеевой (2006; Ерофеева, Ерофеева, 2020); работы В. И. Беликова (2004а; 2004б; 2005; 2013; 2016; 2023; Беликов, Дубяга, Рванова и др., 2021) являют собой исследования функционирования регионализмов в словарях и в Интернете, также в них описаны возможности первых инструментов сети Интернет, позволяющих осуществлять исследования характерности лексем для того или иного региона. Теоретические аспекты исследования региолекта и функционирования регионализмов на примере дальневосточного региолекта отражают статьи Е. А. Оглезневой (2008; 2013; 2019); типологизация регионализмов представлена в книге А. П. Майорова (2010); теоретические обоснования использования сети Интернет в работе с регионально ограниченной лексикой рассмотрены в статьях таких авторов, как Г. И. Романий (2014) и М. В. Ахметова (2014).

Методы исследования: описательно-аналитический (при сборе и анализе теоретического материала по региональному варьированию языка, классификации и обобщении полученных результатов), сравнительно-сопоставительный (при выявлении лексических регионализмов и их дифференциации), количественно-статистический (для количественного анализа полученных данных).

Практическая значимость исследования состоит в возможности использования его результатов в дальнейшем изучении вопросов регионального варьирования языка, а также для понимания широкого и узкого подходов в употреблении регионализмов. Данное исследование позволяет выявить точки пересечения регионально ограниченной лексики пяти регионов и уточнить характерность тех или иных лексем для определенных регионов. Полученный материал и результаты исследования можно использовать в процессе преподавания таких дисциплин, как «Лексикография», «Лексикология», «Русский язык как иностранный», «Лингвокультурология», спецкурсов по лингвистической регионалистике и литературному краеведению.

Обсуждение и результаты

Вопрос о статусе локально окрашенных типов литературной речи неоднократно становился предметом лингвистической дискуссии, но до сих пор не получил однозначного решения. В то же время сегодня ни у кого

не вызывает сомнения тот факт, что русский язык как язык общения относительно однородного коллектива, проживающего на огромной территории, подвержен территориальной и социальной дифференциации. Изучение языкового разнообразия в регионах развивается на протяжении многих лет, и традиционные научные термины уже не способны охватить новые аспекты этого процесса. Естественным образом возникают новые понятия, целью которых является полное описание уникальных языковых особенностей каждого конкретного региона. Одним из таких терминов является «региолект», предложенный петербургским лингвистом В. И. Трубинским и активно использованный А. С. Гердом. Этот термин отражает особенности региональной городской речи, а не только сельской, связанной с диалектами и полудиалектами (Герд, 2001).

Региональная специфика может проявляться на разных уровнях языка: целесообразно говорить о фонетических, грамматических и собственно лексических регионализмах. Так, о региональных вариантах разноразноуровневых языковых единиц пишет А. П. Майоров (2010, с. 12) в книге «Введение в лингвокраеведение Бурятии». На региональную специфику в фонетике, грамматике, лексике указывает Е. А. Оглезнева (2013, с. 31-32). Исследователи также ставят вопрос об ареале регионализмов, выделяя узкие регионализмы (зачастую известны только в одном-двух регионах) и широкие (известные в двух и более регионах), которые в эпоху всепроникновения Интернета становятся более распространенными. Большинство исследований посвящено изучению лексических регионализмов, а также их лексикографированию. В фокусе нашего исследовательского внимания также находятся лексические регионализмы, под которыми мы понимаем **лексические единицы языка, бытующие на определенных территориях**.

Один из важнейших вопросов, который представляется важным разрешить в исследованиях, посвященных регионально ограниченной речи: может ли считаться слово лексическим регионализмом, если оно известно или используется жителями нескольких регионов? Использование одного и того же регионально ограниченного слова в том же значении, что и в соседнем регионе, а возможно и в регионе, находящемся на большом расстоянии от исследуемого, справедливо порождает мысль, что это слово не может считаться регионализмом вообще, и тем более регионализмом конкретного региона. Тем не менее, если на территории страны существуют хотя бы несколько регионов, в которых слово не используют или не знают в контексте, в котором оно распространено в другом регионе, оно может считаться регионализмом. В связи с этим исследователями были предприняты попытки разграничения регионализмов, имеющих широкий ареал, и регионализмов, распространенных в одном или двух соседних регионах. М. А. Бородина, определяя территорию распространения регионализмов, выделяет узкие и широкие регионализмы. Узкие регионализмы – это те слова, которые представлены только в определенном регионе, широкие – имеют более обширный ареал. Автор определяет функционирование регионализмов в нашей стране следующим образом: «Для русской истории, равно как и для истории народов СССР, узкий регионализм, замыкающийся рамками какого-либо одного региона, не является характерным явлением. В истории народов СССР господствует сочетание развития местных традиций (esp. языков, диалектов) с обще- и межнациональными тенденциями развития. Именно поэтому у нас отсутствует ярко выраженный регионализм, который наблюдается в современной Франции» (Бородина, 1982, с. 35). Здесь вполне очевидно указывается на то, что использование того или иного слова только в одном месте не является обязательным условием для определения регионализма. Об этом пишут и другие исследователи. По мнению И. С. Зварыкиной, «важным является то, насколько характерна та или иная реалья для данного региона, насколько активно жители региона используют этот регионализм, насколько они воспринимают ту или иную единицу как “свою”, свойственную своему региону. Иными словами, регионализм должен являться частью региональной культуры» (2013, с. 31).

Главным условием для выделения регионализмов, следовательно, является то, что таковыми могут считаться «языковые единицы, имеющие пространственно-ограниченную сферу функционирования» (Климкова, 2008, с. 53), независимо от того, каково количество регионов, в которых они употребляются. «Региолекты возникают в результате стирания диалектных различий, на фоне миграционного перенесения языковых фактов из одного региона в другой, влияния литературного языка и т. д., и это мотивирует возможность употребления одного и того же слова не в одном, а в нескольких региолектах, но не на всей территории распространения русского языка» (Теркулов, Курмакаева, Мозговой и др., 2018, с. 45).

Другим не менее спорным вопросом является факт наличия в различных региональных лексикографических изданиях, а также в изданиях, материалы которых исследуются в нашей работе, лексем, которые отражены в словарях нормативной и разговорной речи, молодежного сленга и жаргона. Предполагается, что слова, содержащиеся в словарях нормативной лексики без соответствующих помет, не могут считаться регионализмами, но на сегодняшний день это утверждение подвергается сомнению. В. И. Беликов (2004а, с. 23-24) обращает внимание филологов на то, что некоторые особенности речи, которые большинство ученых считает выходящими за пределы литературной нормы, так или иначе попадают даже в самые авторитетные нормативные словари, например в МАС («Словарь русского языка» под ред. А. П. Евгеньевой) или в БТС («Большой толковый словарь русского языка» под ред. С. А. Кузнецова). В качестве примеров исследователь приводит следующие лексемы: «кура» (в БТС у данной лексемы отсутствует помета «просторечное») и «водогрей» – газовая колонка (данная лексема зафиксирована в БТС, изданном в 1998 г., без помет, но, по мнению В. И. Беликова, имеет локально ограниченное употребление в Петербурге); «латка» (вид посуды) и «поробрик», которые включены в МАС без каких-либо помет, хотя, как убежден ученый, они «также имеют строгую географическую прикреплённость» (2004а, с. 33-34). Автор подчеркивает необходимость и в академической лексикографии учитывать региональную норму, так как «многие из такого рода единиц оказываются единственным используемым в повседневной практике (а иногда и единственно известным) способом обозначения определенного понятия для тех, кого никак нельзя исключить из числа носителей литературного языка» (Беликов, 2004а, с. 27-28).

Таким образом, изучение региолектов, ареала употребления регионализмов и прочих нюансов регионально окрашенной лексики сопряжено с рядом сложностей и проблем:

1. Фрагментарность данных. Информация о региолектах может быть разрозненной и недостаточной для полного охвата всех их аспектов. Исследователям приходится собирать данные из различных источников и контекстов.

2. Нормативная неоднородность. Региолекты могут содержать элементы, которые не укладываются в стандартные грамматические или лексические нормы. Это создает проблемы в определении их правильности и описании (например, в формировании определения регионализма и правильной постановки ударения в нем).

3. Динамичность и изменчивость. Региональные особенности языка могут изменяться со временем под воздействием множества факторов, включая миграцию, глобализацию, технологический прогресс и медиа. Изучение этих изменений и тенденций является сложной задачей.

4. Непостоянство границ. Границы, определяющие региональные особенности языка, не всегда четкие и также подвергаются изменениям с течением времени. Лингвистические черты могут переплетаться с соседними регионами, что создает проблемы идентификации и анализа.

5. Комплексность межъязыковых влияний. В региональном контексте могут влиять несколько языков и диалектов, что усложняет анализ и идентификацию тех элементов, которые принадлежат именно региональному влиянию.

6. Отсутствие единой методологии. Не существует универсального метода исследования региолектов и регионализмов, что делает сложным сопоставление результатов различных исследований.

7. Сложность определения ареала. Получить точную информацию о том, в каких регионах и насколько распространен тот или иной регионализм, крайне сложно, так как для этого необходимо обработать огромное количество данных, что невозможно сделать вручную.

Несмотря на эти сложности, изучение региолектов важно для понимания разнообразия языковых культур и сохранения языкового многообразия.

В данной работе остановимся на проблеме определения границ использования регионализмов и наличия пересечений в региолектах на разных территориях на примере анализа функционирования регионализмов в пяти регионах Западной Сибири.

Несмотря на то, что изучение регионально окрашенной лексики отдельных регионов началось уже достаточно давно, глобальный и обширный поиск, сбор и анализ регионализмов во всех регионах России впервые были начаты с запуска онлайн-словаря «Языки русских городов». Электронный словарь «Языки русских городов» разрабатывался с 2005 года при участии группы социолингвистов и обычных граждан под руководством В. И. Беликова. Каждое слово, включенное в этот словарь, представляет собой лексическое явление, характерное для определенного города, области или региона. Чтобы слово было включено в электронный формат этого проекта, оно должно быть не типичным для общего употребления, а быть характерным именно для определенной местности. В 2005 году технологические возможности не позволяли лингвистам провести исчерпывающий анализ, поэтому сбор слов осуществлялся практически вручную. Во время составления словаря авторы разработали свой метод выявления региональных слов. Процесс был довольно прост: участники обсуждений предлагали свои варианты, которые затем проверялись профессиональными редакторами с использованием базы региональных СМИ «Интегрум» и архивов Яндекс.Блогов. Это привело к созданию уникального массива данных за несколько лет функционирования форума, включающего несколько тысяч региональных слов с высоким индексом цитирования в региональных СМИ и социальных сетях. Это позволило сделать вывод о глобальной природе лингвистического разнообразия в нормативном уровне языка. Публикация словаря «Языки русских городов» стала одним из высших достижений в этой работе. Внедрение технологий должно было начать новую эру в изучении региональной лексики, учитывая сложность и невозможность рассматривания всего спектра региональной лексики всех регионов России одним исследователем. Однако идеологи словаря «Языки русских городов» не думали, что автоматизация может заменить труд лингвиста. Присутствие общего интереса к тщательной лексикографической работе уступало место автоматизации, стремлению к максимальному охвату и скорости (и низкой стоимости), но в ущерб качеству (Беликов, Дубяга, Рванова и др., 2021, с. 80).

В настоящее время интернет-технологии и их возможности существенно способствуют исследованиям в области лингвистики, в частности в изучении лингвистических особенностей конкретных регионов, благодаря точному и оперативному анализу доступной информации. Г. И. Романий (2014, с. 174) считает, что Интернет имеет ряд важных преимуществ для лексикологических исследований регионализмов:

- 1) охват различных форм языка, как территориальных, так и стилистических;
- 2) богатые поисковые возможности, позволяющие производить показательные выборки примеров различной территориальной и стилевой принадлежности, ограничивать область поиска (например, искать лексемы определенной тематики на тематических сайтах, искать в заданном регионе или в заданном языковом регистре – в архивах публицистики, художественной литературы, блогах и т. п.), вести комбинированный поиск (искать определенные синтагмы/фразы, включающие релевантную лексику или ее комбинации с ключевыми словами на заданном расстоянии, в пределах одного предложения и др.);
- 3) возможность датировать словоупотребление и извлекать сведения об авторе (место проживания, возраст, образование и др.).

Интернет как средство коммуникации носителей разных региолектов видится благоприятной средой для выявления региональных лексических особенностей, в частности если учитывать явление «неосознанного регионализма» (когда использующий в повседневной речи ту или иную лексику человек не осознает, что это

слово может быть неизвестно в другом регионе) (Беликов, 2016). Один из инструментов для проверки ограниченности слов по региону с использованием поисковой системы Яндекс был впервые описан и использован В. И. Беликовым и М. В. Ахметовой. Суть применения инструмента состоит в том, что, введя любой поисковый запрос (лексему) в строчку поиска Яндекс или в дополнительный сервис Яндекс.Вордстат, возможно отследить частотность данного запроса по регионам поиска (Ахметова, 2014; Беликов, 2004b; 2005). Таким образом, с помощью поиска слова в Интернете можно очертить его ареал.

В 2021 году технологии компании Яндекс позволили провести глобальную работу в этом направлении: были взяты тексты публикаций и комментариев, появившихся в Яндекс.Дзене с января 2019 года по апрель 2021 года (15 млн публикаций, 217 млн комментариев, 11 млрд словоупотреблений, 6 млн пользователей), и составлен список всех встретившихся в них слов. Каждому слову была присвоена доля пользователей из определенного региона, которые хотя бы раз использовали это слово. Также был вычислен показатель, отражающий характерное употребление слова в этом регионе: отношение доли местных пользователей, использующих данное слово, к средней доле пользователей во всех регионах. Затем для каждого региона составили список из 500 слов с показателем характерности не менее трех, и совместно с лингвистами из Института русского языка им. В. В. Виноградова Российской академии наук эти слова были проверены вручную. Слова, которые употребляли менее десяти человек в данном регионе, были исключены. Кроме того, из списков убрали топонимы, названия организаций, фамилии, ругательства и грубые слова, которые могли казаться оскорбительными, а также обычные слова, которые попали в список из-за местных особенностей, таких как, например, промышленная специализация региона (Исследования Яндекса. Как говорится: местные слова в разных регионах России. 06.06.2021. <https://yandex.ru/company/researches/2021/local-words>). Таким образом, мы видим, что наличие технологий даже в нынешних условиях не способно заменить ручную работу лингвиста, осуществляющего финальный отбор слов. Несмотря на наличие погрешностей, благодаря появлению такого словаря с распределенными по регионам лексемами впервые стало возможно провести сравнительный анализ регионализмов смежных регионов и ответить на вопросы, насколько схожи слова в соседних регионах и можем ли мы говорить о том, что существуют слова, которые используются только в одном из ряда смежных регионов.

Проведем такой сравнительный анализ лексем, отнесенных исследованием к регионализмам Западной Сибири: а именно регионализмов Новосибирской области и четырех граничащих с ней регионов – Омской области, Томской области, Кемеровской области и Алтайского края. В словарь, составленный в рамках исследования Яндекса, в качестве регионализмов вошло 84 лексемы из Новосибирской области, 83 – из Томской области, 61 – из Омской области, 150 – из Кемеровской области и 105 – из Алтайского края. Если мы попробуем провести корреляцию количества регионализмов и количества жителей в областях, то здесь мы не найдем взаимосвязи, так как это скорее связано со спецификой территории, с преобладающей причиной формирования регионализмов. В Кемеровской области, например, это может быть осязаемое влияние профессионального жаргона шахтеров на формирование регионально окрашенной лексики, что способствует увеличению регионально ограниченных лексем. Лексемы, отнесенные исследованием Яндекса к региональным, присутствуют в списках регионализмов и других областей, но на данном этапе остановимся на исследовании представленности лексем только в обозначенных пяти регионах. С помощью сервиса для сравнения текстов Copyleks и сервисов Microsoft Excel удалось провести сравнительный анализ слов во всех пяти регионах. 84 регионализма (из 483 во всех пяти регионах) имеют минимум одно совпадение с регионализмами других регионов, это 17,3% от общего количества слов, данные по регионам представлены в Таблице 1.

Таблица 1. Количество регионализмов, имеющих совпадения с другими регионами

	Всего регионализмов	Количество регионализмов, имеющих хотя бы одно совпадение с регионализмами других регионов	Процент совпавших регионализмов от их общего количества в регионе
Омская область	61	21	34%
Новосибирская область	84	48	40%
Томская область	83	35	29%
Кемеровская область	150	71	47%
Алтайский край	105	57	54%

Регионализмов, совпавших во всех пяти регионах, оказалось всего четыре (даны определения, представленные в словаре):

- расповаживать – избаловать, разбаловать;
- лупень – определение в словаре отсутствует;
- горошница – гороховая каша;
- божничка (на божничку садить, ставить) – превозносить.

В Таблицах 2 и 3 представлены регионализмы, совпавшие в четырех (16 лексем) и трех (17 лексем) регионах из пяти. Плюсами в таблицах обозначены регионы, в которых слово считается регионализмом.

Как мы можем видеть, наименьшее количество общих совпадений имеет Омская область, что может быть связано с ее более удаленной позицией от остальных регионов. Наибольшее количество совпадений имеют Новосибирская область, Кемеровская область и Алтайский край. Интерес для нас представляют также данные о том, каким образом пересекаются регионализмы соседних регионов с регионализмами Новосибирской

области – региона с крупнейшим из представленных областным центром, в который чаще всего происходит миграция из исследуемых регионов. Отметим, сколько лексем Новосибирской области совпадает с регионализмами соседних регионов. Первая цифра – количество совпавших лексем, вторая – процент совпавших лексем от всего количества регионализмов Новосибирской области (Таблица 4).

Таблица 2. Регионализмы, имеющие совпадения в четырех сибирских регионах

	Новосибирская область	Томская область	Омская область	Кемеровская область	Алтайский край
Верхонка	+	+		+	+
Догоняшка	+	+		+	+
Если и еслиф	+	+		+	+
Запашистый	+	+		+	+
Каралька	+		+	+	+
Мордушка	+		+	+	+
Мультифора	+	+		+	+
Огребаться	+	+		+	+
Перепираться	+		+	+	+
Попроведовать	+		+	+	+
Повдоль	+	+		+	+
Разновес	+	+		+	+
Строжиться	+	+		+	+
Халда	+	+		+	+
Шмутье	+		+	+	+
Фотать	+	+		+	+

Таблица 3. Регионализмы, имеющие совпадения в трех сибирских регионах

	Новосибирская область	Томская область	Омская область	Кемеровская область	Алтайский край
Толченка			+	+	+
Следи́к	+	+			+
Свешать		+		+	+
Пурхаться	+			+	+
Постряпать	+	+		+	
Полоротый	+			+	+
Пожиться	+		+	+	
Пимы		+		+	+
Околка	+		+		+
Однешка	+	+		+	
Насмеливаться	+			+	+
Матерка	+			+	+
Куржак	+			+	+
Движняк	+		+		+
Вехотка	+	+			+
Бухануть	+			+	+
Авторазбор			+	+	+

Таблица 4. Количество регионализмов Новосибирской области, совпадающих с соседними регионами

	Томская область	Омская область	Кемеровская область	Алтайский край
Новосибирская область	21 (25%)	14 (17%)	39 (46%)	33 (39%)

Таким образом, было выявлено, что наибольшее число совпадений, а именно практически половина регионализмов, отнесенных к Новосибирской области, являются также регионализмами Кемеровской области, а меньше всего совпадений у Новосибирской области с Омской областью.

Несмотря на то, что слово может являться регионализмом в нескольких регионах, такой показатель, как **характерность слова для региона**, может продемонстрировать, насколько это слово можно считать регионально ограниченным и специфичным именно для этого региона, а также сузить его ареалы. Чем этот показатель выше, тем характернее это слово для конкретного региона (о методах подсчета характерности слова см. выше). Сравним слова, которые считаются регионализмами во всех пяти регионах, и определим, для каких регионов они наиболее характерны. Первая цифра в таблице – это количество уникальных пользователей, употребивших регионализм хотя бы раз, вторая цифра – показатель характерности слова (чем выше показатель, тем характернее слово для региона). Слова с показателем характерности менее трех вычеркивались из списка при отборе (Таблица 5).

Таблица 5. Показатели характерности употребления слов по регионам

	Новосибирская область	Томская область	Омская область	Кемеровская область	Алтайский край
божничка	92 (4,4)	23 (6,8)	31 (5,0)	42 (5,1)	45 (6,7)
горошница	32 (3,03)	15 (8,8)	16 (5,1)	27 (6,5)	16 (4,7)
распознаживать	34 (3,6)	10 (6,7)	11 (4,03)	15 (4,1)	14 (4,7)
лупень	69 (4,2)	10 (3,8)	18 (3,7)	47 (7,3)	17 (3,2)

Исходя из этих данных, мы можем увидеть, что три слова из четырех, считающихся регионализмами во всех пяти регионах, наиболее характерны для Томской области.

Самыми же характерными словами из всего списка предполагаемых регионализмов для указанных пяти регионов являются лексемы, представленные в Таблице 6.

Таблица 6. Список наиболее характерных регионализмов в исследуемых регионах

Регион	Лексема	Количество уникальных пользователей, использовавших слово	Показатель характерности слова для региона
Новосибирская область	задуряться	88	58.8781265
Томская область	просрота	14	44.6523835
Омская область	бульжить	25	68.9866591
Кемеровская область	углярка	75	29.8028111
Алтайский край	приспрашивать	23	39.6687777

Считаем, что информация о характерности слова для региона представляется очень важной при обсуждении феномена регионализма, потому как именно эта характеристика позволит сделать вывод, что слово наиболее характерно для какого-либо региона (-ов), несмотря на то, что известно и употребляется в некоторых других. При исследовании слова важно учитывать коэффициент характерности в разных регионах, так, например, слово «фотать» (фотографировать) исследованием Яндекса представлено как регионализм большого количества областей (в скобках указан коэффициент характерности для региона), среди них: Хабаровский край (4,0), Томская область (3,9), Сахалинская область (3,0), Республика Хакассия (3,5), Республика Саха (3,4), Республика Бурятия (3,1), Республика Алтай (3,1), Приморский край (3,4), Новосибирская область (6,6), Красноярский край (3,4), Кемеровская область (3,4), Камчатский край (3,1), Иркутская область (3,2), Забайкальский край (3,0), Амурская область (3,2), Алтайский край (3,0). Помимо широкого ареала, в большинстве случаев характерность слова для региона не превышает 4, а при условии, что слова, характерность которых была ниже 3, удалялись из списка, мы можем сделать вывод, что региональная особенность этого слова слабо выражена. Решение об отнесении его к регионализмам, вероятно, может быть продиктовано тем, что в упомянутых областях просторечный вариант «фотать» распространен чуть шире литературного «фотографировать» или другого просторечного аналога «фоткать». Вопрос причисления к регионализмам широко распространенных слов с низким региональным коэффициентом наряду с отнесением к регионализмам слов, зафиксированных в словарях, является спорным в лингвистической регионалистике.

Заключение

Таким образом, проанализировав базу актуальных исследований региональной лексики, мы приходим к следующим выводам: принимая во внимание факт существования погрешности даже в исследованиях, где применяются искусственный интеллект и современные технологии, мы в какой-то мере можем опираться на эти результаты, так как они основаны на обработке большого количества данных, что является весомым аргументом в пользу исследования. Выявив и описав существующие пересечения регионально ограниченной лексики пяти исследуемых регионов и узнав, что количество хотя бы раз совпавших регионализмов в соседних пяти регионах равно 17,3% от числа всех исследуемых лексем, мы допускаем, что в реальной языковой среде оно может быть еще больше и слова, которые зачастую неизвестны жителям других регионов, известны жителям соседних регионов, а различаться будет только процент характерности использования слова в регионе – в каком-то из них оно будет более популярно, а где-то – менее. Для каждого из регионов мы сравнили процент характерности совпадающих слов, а также выявили наиболее характерные лексемы.

Подводя итог анализа пересечений слов в соседних регионах Западной Сибири, в которых миграционные и коммуникационные процессы проходят очень плотно, мы можем подчеркнуть, что достаточно большое количество лексем совпадают друг с другом в процентном соотношении (и в сравнении с более удаленными областями), особенно это касается Новосибирской и Кемеровской областей, тем не менее большое количество регионализмов являются характерными только для какого-то одного региона из пяти, не имея пересечений даже с соседними областями. Исходя из результатов исследования, мы предполагаем, что понятие «широкий регионализм» может быть также разделено на подтипы: «широкий» регионализм межрегионального

характера – распространенный на территории большого количества регионов, в том числе и удаленных друг от друга, и «широкий» регионализм межобластного характера, который имеет распространение, например, только в соседних друг с другом областях, зачастую объединенных в один регион (Сибирь, Урал, Дальний Восток и так далее). Кроме того, при анализе региональной лексики стоит учитывать, что региональная особенность слов может быть слабо выражена (когда лексема распространена в большом количестве областей и имеет низкий коэффициент характерности).

Дальнейшее изучение пересечений в использовании регионализмов на территории различных регионов страны даст больше понимания о специфических ареалах распространения той или иной лексемы и позволит построить языковые и внеязыковые гипотезы об исторических причинах подобного явления, что можно отнести к перспективам исследования данного вопроса.

Источники | References

1. Ахметова М. В. Лексические регионализмы и локализмы в русскоязычном Интернете // Русский язык и новые технологии / сост. Г. Ч. Гусейнов; под ред. М. В. Ахметовой, В. И. Беликова. М.: Новое литературное обозрение, 2014.
2. Беликов В. И. Избранные статьи / ред.-сост. С. Г. Татевосов. М., 2023.
3. Беликов В. И. О словарях региональной лексики: вступительное слово // Ливинская И. В., Матвеев А. И. Словарь региональной лексики и народных топонимов города Новосибирска. Изд-е 2-е, стер. Новосибирск: Свиньин и сыновья, 2016.
4. Беликов В. И. Регионализация русской орфоэпической нормы и данные словарей // Слово и словарь – Vocabulumetvocabulary: сб. науч. тр. по лексикографии. Гродно: Изд-во Гроднен. гос. ун-та им. Янки Купалы, 2005.
5. Беликов В. И. Региональные варианты русского языка и традиционные сельские диалекты // Региональные варианты национального языка: мат. всерос. (с междунар. участием) науч. конф. / науч. ред. А. П. Майоров. Улан-Удэ: Изд-во Бурят. гос. ун-та, 2013.
6. Беликов В. И. Сравнение Петербурга с Москвой и другие соображения по социальной лексикографии // Русский язык сегодня: сб. ст. М.: Изд-во Ин-та русского языка РАН, 2004а. Вып. 3. Проблемы русской лексикографии.
7. Беликов В. И. Yandex как лексикографический инструмент // Компьютерная лингвистика и интеллектуальные технологии: тр. междунар. конф. «Диалог'2004». М.: Наука, 2004b.
8. Беликов В. И., Дубяга А. О., Рванова Л. Ю., Селегей В. П. Корпусная региональная лексикография: принципы, методы и предварительные результаты // Компьютерная лингвистика и интеллектуальные технологии: по мат. ежегод. междунар. конф. «Диалог». М.: РГГУ, 2021. Вып. 20.
9. Бородин М. А. Диалекты или региональные языки? (К проблеме языковой ситуации в современной Франции) // Вопросы языкознания. 1982. № 5.
10. Герд А. С. Несколько замечаний касательно понятия «диалект» // Русский язык сегодня: сб. ст. М.: Наука, 2001. Вып. 1.
11. Ерофеева Е. В. Региолект как промежуточный идиом //...Слово отзовется: памяти Аллы Соломоновны Штерн и Леонида Вольковича Сахарного: сб. ст. Пермь: Изд-во Перм. гос. ун-та, 2006.
12. Ерофеева Е. В., Ерофеева Т. И. Социальное варьирование в региолекте // Филологические заметки. 2020. Т. 1. № 18.
13. Ерофеева Т. И., Скитова Ф. Л. Локальные элементы в литературной речи горожан // Языковой облик уральского города: сб. науч. тр. Свердловск: Изд-во УрГУ, 1990.
14. Зварыкина И. С. К вопросу о соотношении диалектного и регионального в русском языке (на примере лексики Астраханского края) // Гуманитарные исследования. 2013. № 2 (46).
15. Климкова Л. А. Нижегородская микропонимия в языковой картине мира: дисс. ... д. филол. н. М., 2008.
16. Майоров А. П. Введение в лингвокраеведение Бурятии: учеб. пособие для студентов вузов. Улан-Удэ: Изд-во Бурят. гос. ун-та, 2010.
17. Оглезнева Е. А. Дальневосточный региолект русского национального языка: к вопросу о его формировании и специфике // Региональные варианты национального языка: мат. всерос. (с междунар. участием) науч. конф. / науч. ред. А. П. Майоров. Улан-Удэ: Изд-во Бурят. гос. ун-та, 2013.
18. Оглезнева Е. А. Дальневосточный региолект русского языка: особенности формирования // Русский язык в научном освещении. 2008. Вып. 16.
19. Оглезнева Е. А. Региональные заимствования как идентификаторы региолектов // Язык в различных сферах коммуникации: мат. III междунар. науч. конф. Чита: Изд-во Забайкал. гос. ун-та, 2019.
20. Романий Г. И. Региональная городская лексика в русском языке: происхождение, типология и ареализация по данным узуса в сети Интернет // Русский язык и новые технологии / сост. Г. Ч. Гусейнов; под ред. М. В. Ахметовой, В. И. Беликова. М.: Новое литературное обозрение, 2014.
21. Теркулов В. И., Курмакаева Н. П., Мозговой В. И., Першина К. В., Кудрейко И. А., Матвиенко О. В., Мельникова Л. В., Евсеева М. Г., Коробов-Латынцев А. Ю., Юрьева Е. В., Бровец А. И., Коробова-Латынцева В. С., Куприянова К. А. Донецкий региолект: монография / под ред. В. И. Теркулова. Донецк: Фолиант, 2018.

Информация об авторах | Author information



Матвеева Ирина Владимировна¹, к. филол. н.

¹ Новосибирский государственный технический университет



Matveeva Irina Vladimirovna¹, PhD

¹ Novosibirsk State Technical University

¹ *Irina.matveeva1@mail.ru*

Информация о статье | About this article

Дата поступления рукописи (received): 06.10.2023; опубликовано online (published online): 17.11.2023.

Ключевые слова (keywords): региолект; регионализм; регионально ограниченная лексика; региональная лексикография; regional dialect; regionally specific word; regionally limited vocabulary; regional lexicography.