УДК 81’33
Науч. спец. 5.9.8
DOI: 10.36809/2309-9380-2026-52-179-184

Ирина Игоревна Абдулганеева
Казанский (Приволжский) федеральный университет, кандидат филологических наук, доцент,
доцент кафедры теории и практики перевода, Казань, Россия
e-mail: abdulganii@bk.ru

Анжелика Сергеевна Бутусова
Южный федеральный университет, кандидат филологических наук, доцент,
доцент кафедры немецкой филологии, Ростов-на-Дону, Россия
e-mail: butussova@mail.ru

Анализ языковых феноменов в свете корпусной лингвистики: сопоставительный аспект

Аннотация. Основной целью авторов настоящей статьи стало проведение параметрического анализа Национального корпуса русского языка (НКРЯ) и немецкоязычного корпуса DeReCo. Параметрический анализ двух языковых корпусов позволяет определить и выявить инструменты для комплексного исследования лингвистических феноменов языка как в диахроническом, так и в синхроническом аспектах, с учетом динамики языковых процессов в немецком и русском языках. Несмотря на наличие схожих параметров построения языковых корпусов, результаты исследования позволили показать специфические цели применения каждого корпуса. Немецкоязычный корпус DeReCo большей частью ориентирован на изучение языковых контактов и особенностей переводов между немецким и русским языками, тогда как НКРЯ представляет самый полный и богатый корпус, прежде всего, русскоязычных текстов различных типов дискурса. Тексты в корпусе структурированы и аннотированы с учетом различных языковых параметров, таких как морфологические, синтаксические и семантические характеристики. Наиболее ценным и сложным при параметрической разметке сопоставляемых корпусов является указание семантических параметров. Метаразметка корпуса DeReCo и НКРЯ предоставляет исследователям широкий выбор материала как в плане жанров и тематики, так и в статистическом плане, а значит, обеспечивает валидность данных при анализе языковых феноменов.

Ключевые слова: корпусная лингвистика, параметрический анализ, Национальный корпус русского языка, немецкоязычный корпус DeReCo, языковой феномен.

Irina I. Abdulganeeva
Kazan (Volga Region) Federal University, Candidate of Philological Sciences, Associate Professor,
Associate Professor of the Department of Theory and Practice of Translation, Kazan, Russia
e-mail: abdulganii@bk.ru

Anzhelika S. Butusova
Southern Federal University, Candidate of Philological Sciences, Associate Professor,
Associate Professor of the Department of Germanic Philology, Rostov-on-Don, Russia
e-mail: butussova@mail.ru

Analysis of Linguistic Phenomena in the Light of Corpus Linguistics: A Comparative Aspect

Abstract. This article presents a parametric analysis of the Russian National Corpus (RNC) and the German Reference Corpus (DeReCo). A comparative approach enables a comprehensive study of linguistic phenomena in both diachronic and synchronic perspectives, capturing the dynamics of linguistic change in Russian and German. Although the corpora share similar design principles, the analysis reveals their distinct research orientations. The DeReCo corpus primarily focuses on language contact and translation

© Абдулганеева И. И., Бутусова А. С., 2026

Для цитирования: Абдулганеева И. И., Бутусова А. С. Анализ языковых феноменов в свете корпусной лингвистики: сопоставительный аспект // Вестник Омского государственного педагогического университета. Гуманитарные исследования. 2026. № 3 (52). С. 179–184. DOI: 10.36809/2309-9380-2026-52-179-184

Статья доступна на условиях лицензии
Creative Commons Attribution 4.0 International (CC BY 4.0)

features between German and Russian, while the RNC serves as the most comprehensive and representative collection of Russian texts across various genres and registers. Texts in both corpora are structured and annotated according to multiple linguistic parameters, including morphological, syntactic, and semantic features. The annotation of semantic parameters stands out as one of the most complex and valuable aspects of the meta-markup. The rich metadata of both DeReCo and the RNC provides researchers with a diverse selection of materials in terms of genres, topics, and statistical representativeness, thereby ensuring the validity of linguistic analysis.

Keywords: corpus linguistics, parametric analysis, National Corpus of the Russian Language, German-language DeReCo corpus, linguistic phenomena.

Введение (Introduction)

За последние десятилетия корпусная лингвистика укрепилась как отдельный раздел науки исследования языка, а также как основная эмпирическая парадигма современного научного знания о языке. Неоспорим тот факт, что изыскания корпусной лингвистики внесли огромный вклад в изучение функционирования отдельных языковых феноменов в русле когнитивного, сопоставительного, корпусно-дискурсивного направлений в лингвистике. Семантический вектор современного языкознания, ознаменованный антропоцентричностью научного знания, выдвигает на повестку дня необходимость исследования новых форматов функционирования языкового материала, появление и формирование которых детерминировано развитием цифровых технологий, — традиционный языковой текст зачастую является сегодня комбинированным продуктом различных семиотических систем. В связи с этим языковые корпусы стали осознаваться как комплексные специализированные инструменты, позволяющие на основе обширной разметки данных исследовать различные лингвистические феномены с учетом представленности статистических показателей при изучении отдельных языковых явлений, которые повышают верифицируемость научных выводов по проведенным исследованиям.

Целью настоящего исследования стало осуществление сопоставительного анализа исследовательского потенциала Национального корпуса русского языка (НКРЯ) и DeReCo — крупнейших национальных лингвистических корпусов России и Германии — на основании параметрического анализа корпусов.

Задачи исследования подчинены заявленной цели и направлены на рассмотрение русско- и немецкоязычного корпусов с позиции анализа языковых феноменов на основании выделенных параметрических характеристик. Мы полагаем, что наш сопоставительный обзор может послужить воспроизводимости лингвистических изысканий с использованием определенных инструментов и тем самым способствовать как развитию цифрового инструментария, так и совершенствованию методологии исследования определенных языковых феноменов. Описывая перспективы научно-исследовательского потенциала корпусов, мы не претендуем на исчерпывающее изложение, но фокусируемся на наиболее продуктивных направлениях русских и немецких корпусных исследований.

Методы (Methods)

Материалом для исследования послужили Национальный корпус русского языка (http://ruscorpora.ru) и немецкоязычный корпус DeReCo (https://www.ids-mannheim.de/digspra/kl/projekte/korpora). В настоящей работе применялись эмпирические методы, включая сравнительный. В работе также представлены результаты эмпирического обзора русскоязычной (С. О. Савчук, Т. А. Архангельский, В. П. Захаров, С. Ю. Богданова, В. И. Заботкина и др.) и немецкоязычной (M. Купиц (M. Kupietz), Х. Кайбель (H. Keibel), С. Прелль (S. Pröll), Т. Вебер (T. Weber) и др.) литературы по корпусной лингвистике.

Результаты и обсуждение (Results and Discussion)

Национальный корпус русского языка представляет собой масштабную базу данных, содержащую собрание текстов не только на русском, но и на иностранных языках. Корпус был создан с целью обеспечения исследователей, филологов и лингвистов, инструментом для глубинного анализа языковых явлений — сегодня НКРЯ служит фундаментальной платформой для исследования самых разных языковых аспектов. Охватывая широкий временной диапазон, начиная от древнерусских текстов и заканчивая современными материалами, НКРЯ позволяет осуществлять как диахронический, так и синхронический анализ языковых явлений. НКРЯ включает миллионы текстов различных жанров и стилей, в том числе художественную литературу, научные работы, газетные статьи, разговорные высказывания и другие виды текстовой продукции. Тексты в корпусе структурированы и аннотированы с учетом языковых параметров, таких как морфологические, синтаксические и семантические характеристики. Такая многоуровневая аннотация обеспечивает возможность проведения комплексного анализа и способствует повышению точности и достоверности при исследовании определенных языковых феноменов. Следует отметить, что параметрический анализ НКРЯ представляет собой систематизированный подход к исследованию языковых данных, основанный на выявлении и измерении различных параметров языка. Данный метод позволяет анализировать не только отдельные элементы языка, такие как лексический и грамматический уровни, но и их взаимосвязи, а также динамику изменения в определенный период развития языка.

В отличие от НКРЯ, Немецкий справочный корпус (DeReCo и COSMAS II) является крупнейшей моноязычной национальной коллекцией электронных текстов. Открытый доступ к подкорпусам DeReCo, в том числе онлайн, предоставляет широкие исследовательские перспективы. В данной статье мы представим направления корпусных исследований на базе DeReCo, как они видятся нам и самим сотрудникам Института немецкого языка (IDS). НКРЯ и DeReCo, которые архивируют и накапливают многочис-

ленные языковые материалы, необходимо также рассматривать как репрезентативные источники языкового выражения концептуальной картины мира определенной лингвокультуры, в основе которой лежит когнитивная база, понимаемая нами как структура «для хранения и использования информации, накопленной в процессе эволюции и деятельности человечества» [1, с. 71].

Сегодня функционирование текстовых массивов НКРЯ охватывает более 17 млрд словоупотреблений. Корпус представляет собой собрание 16 независимых корпусов, специфика каждого определяется изучением отдельных языковых феноменов и постановкой в соответствии с этим определенных исследовательских задач. Разнообразием жанров представленных текстов отмечен Основной корпус НКРЯ. Тексты, включенные в Основной корпус, представляют своего рода языковой стандарт, т. е. литературный язык определенной эпохи. Самым крупным по репрезентативности материала считается Газетный корпус НКРЯ, он охватывает материалы средств массовой информации с 1983 по 2021 г. Поэтический корпус отличается тем, что, наряду с семантической и морфологической разметкой, в нем предусмотрена специальная стиховедческая разметка, которая позволяет конструировать запрос строк, написанных определенным метром. Ряд авторов отмечают наличие новых компонентов в НКРЯ как ресурсный функциональный комплекс статистики и визуализации данных [2].

В отношении анализа аспектов синтаксического уровня языка НКРЯ также представляет возможности для решения исследовательских задач. По мнению Г. Я. Мартыненко, «в последние годы достигнуты определенные успехи в области разработки формализованных синтаксических моделей, элементы метаязыка которых легко поддаются количественному измерению» [3, с. 173]. Синтаксическая разметка представлена в НКРЯ таким образом, что оказываются доступны конструкции, т. е. сложные синтаксические единицы со своим значением, часто фиксированным набором и порядком лексических переменных, заданным грамматическим оформлением и лексическим наполнением разной степени свободы — от почти застывших фразеологизмов до свободных сочетаний с минимальными ограничениями и составляющими их количественные показатели в зависимости от типа текста.

Данные НКРЯ открывают широкий спектр возможностей для изучения семантических аспектов лексического уровня языка в синхроническом и диахроническом срезе. Согласно М. Н. Конновой, компьютерно-корпусный подход весьма плодотворен для комплексного когнолингвокультурологического изучения тех изменений, которые затрагивают аксиологический пласт значения лексем. Приводящие к трансформации ценностной отнесенности слова, эти изменения, как правило, отражают сдвиги более глубокого, ментального, мировоззренческого характера, затрагивающие сознание носителей языка [4, с. 271]. НКРЯ позволяет анализировать частотный профиль различных лексем, функционирование в определенных типах текста которых охватывает отдельный период. Целесообразность применения НКРЯ в рамках изучения семантики языковых единиц определяется также тем, что данные НКРЯ позволяют раскрыть сложные случаи полисемии, проанализировать возможные способы переводов идиоматизированных реплик и клише, провести анализ семантики эмоционального лексикона с опорой на корпусные данные. Например, исследователи используют НКРЯ при анализе передачи дискурсивных маркеров, рассматриваемых в качестве адвербиалов, которые модифицируют иллокутивность высказывания и могут употребляться в качестве инструмента вежливого тона [5, с. 121]. В основном корпусе НКРЯ доступен поиск по словообразовательной структуре (морфемному составу) слова. Если анализировать материалы ранней версии НКРЯ, то в «основном корпусе противопоставлялись тексты со снятой (вручную) морфологической омонимией (около 6 млн словоупотреблений) и основная масса текстов, где были представлены все возможные разборы; в большинстве корпусов со специальной разметкой присутствовала только неснятая омонимия. В результате использования нейросетевой модели для основного и газетных корпусов это противопоставление фактически снято» [2, с. 12]. Очевидно, что для изучения многих языковых явлений и этого объема недостаточно. Поэтому сейчас создаются корпусы, где счет словоупотреблениям идет на миллиарды.

При выявлении межъязыковых аспектов, а также сопоставлении языковых феноменов различных языков с целью определения универсальных и специфичных свойств языка применяется параллельный подкорпус НКРЯ. Перспективы формирования новых разметок в НКРЯ связаны с антропоцентричностью языка, его способностью меняться и обогащаться. Вопрос о представленности этнокультурного своеобразия языков в НКРЯ также видится актуальным и перспективным, значимым ввиду необходимости сохранения своеобразия и уникальности национальных языков народов и этносов, проживающих на территории нашей страны. Расширение диалектных текстов и их метаданных, а также параллельного подкорпуса НКРЯ в паре «русский и национальный язык» — задачи для НКРЯ на ближайшие годы. Сегодня в параллельном подкорпусе НКРЯ представлены такие языки нашей страны, как башкирский, карельский, бурятский, хакасский, чувашский. Следующим направлением для разметки и пополнения НКРЯ является представленность текстовых данных на основе выделения участников определенной профессиональной группы — социолектов. Эти данные необходимы, так как корпус может быть не в полной мере репрезентативным для языкового употребления, определенного лингвосообщества (см. об этом [6]).

Немецкий справочный корпус DeReCo, как и НКРЯ, предоставляет исследователям широкий выбор материала в плане жанров и тематики, в статистическом плане, а значит, обеспечивает валидность потенциальных исследований. Поскольку в отечественном научном дискурсе рассматриваемый немецкий корпус не был рассмотрен так детально, как НКРЯ, остановимся на его устройстве подробнее.

В отличие от НКРЯ, немецкоязычный корпус DeReCo определяет сбалансированность и репрезентативность только в отношении статистики, поскольку релевантность подборки текстов диктуется каждым конкретным исследованием [7].

Как и в НКРЯ, в DeReCo представлены различные типы текстов: художественная проза, научная и научно-популярная литература, периодика и др. На сегодняшний день DeReCo насчитывает 600 подкорпусов. В общей сложности корпус содержит 68,9 млрд словоформ, что соответствует примерно 172,2 млн книжных страниц). Размечены 70 корпусов. Неразмеченные корпуса содержат подборки периодических изданий немецких медиа различных временных отрезков с 1998 по 2023 г. Все тексты распределены по 18 архивам.

В DeReCo тексты классифицируются по тематике и по жанрам. Благодаря этой классификации и метаданным пользователи могут получить представление о контенте корпуса, тематики и жанры доступны для выбора при осуществлении поиска. Корпус предлагает зарегистрированным пользователям возможность создать свой собственный подкорпус внутри DeReCo, предварительно связавшись с администратором. Создание пользовательского подкорпуса возможно благодаря наличествующей жанровой и тематической разметке контента. Большинство классификаций типов текста автоматически извлекаются из исходных метаданных, но требуют проверки.

Все газетные тексты и пресс-релизы корпуса DeReCo размечены по темам, т. е. им присуждается тематическая категория согласно тематической таксономии DeReCo. Тематическая таксономия корпуса включает 12 основных категорий и 52 подкатегории, которые изначально были основаны на категориях Open Directory [8]. Архивы C и C2 размечены морфологически и синтаксически. Морфологическая аннотация в DeReCo предполагает, помимо частеречной, разметку следующих признаков: время и аспект, неспрягаемые глаголы, наклонение, сокращение, имя собственное, число, степени сравнения, координационные маркеры, маркеры предложений. Синтаксическая аннотация фиксирует цепочку глаголов (один главный и остальные вспомогательные), номинальную фразу; наречную фразу; предложную фразу; структуру предложения.

Как показало исследование, мейнстримом корпусных исследований на базе DeReCo являются исследования лексикографические и, прежде всего, новейших неологизмов [9; 10]. На базе платформы DeReCo пользователям доступен ряд онлайн-словарей (OWID), связанных с корпусом и другими ресурсами. Теоретические подходы и описание онлайн-словаря современных немецких неологизмов представляет Д. Штеффанс [11]. Онлайн-словарь допускает поиск по ключевым словам, которые сгруппированы в алфавитном порядке. Пользователям предоставляется исчерпывающая информация о данных словарных статей и возможность подключения к ним. Помимо неологизмов интерфейс OWID открывает доступ к онлайн-словарям паронимов, пословиц, устойчивых коллокаций, коммуникационных глаголов, процессуальных форм, иностранных слов и elexiko.

Elexiko — это онлайн-информационная система («Словарь») современного немецкого языка, которая документирует, объясняет и научно комментирует словарный запас немецкого языка на основе текущих языковых данных (до 2013) в отдельных модулях. Пункт меню «Отредактированные ключевые слова» дает обзор всех ключевых слов. Расширенный поиск осуществляется по всем обработанным ключевым словам, имеющим общий признак, например все ключевые слова, все сокращения, все наречия. Использование пункта меню «К газетным статьям» позволяет перейти либо ко всем словарным статьям, либо только к ключевым словам, отредактированным редакцией. Здесь также есть дополнительная возможность для доступа к словарным статьям elexiko в определенных подмножествах. Каждый из словарей снабжен детальным пояснением терминов, описанием проекта и инструкциями для пользователя.

Весьма полезной мы считаем методику Х. Кейбеля [12], направленную на полуавтоматическое обнаружение неологизмов в газетных подкорпусах DeReCo. Она состоит из шести этапов:

– автоматическое определение исходного списка с частотами встречаемости слов. Частотность анализируется в газетных публикациях за каждый рассматриваемый год. Отмечается, что в идеале объем текстов за каждый анализируемый год частотности употребления слов должен быть приблизительно равным;

– отсеивание слов, которые не являются неологизмами, а представляют собой необычные символы, например URLs. Весьма логично применяется количественный критерий, т. е. учитываются только те слова, которые встречаются в корпусе не менее 20 раз;

– отсеивание имен собственных и их нестандартных орфографических номинаций;

– отсеивание регионализмов;

– ручная фильтрация: черновая проверка полученного списка;

– экспертный анализ: проверка полученного списка экспертами.

Нам представляется важным подчеркнуть, что данная методика полуавтоматического обнаружения неологизмов, разработанная на базе описываемого корпуса DeReCo, заслуживает внимания и может оказаться полезной для применения и в других корпусах.

Следующей сферой корпусных исследований является корпусная грамматика с использованием корпусного аналитического инструмента KoGra-R [13; 14]. KoGra-R позволяет выполнять исследования качественного и количественного функционирования морфосинтетических явлений в дискурсе. Таким образом, KoGra-R является эффективным инструментом актуальных полиаспектных исследований с синтаксической, семантической и метаязыковой точек зрения. KoGra-R предназначен также для статистического анализа. Он содержит списки частотности одного или нескольких лингвистических феноменов, встречающихся в DeReCo.

Исследовательский сравнительно-статистический функционал KoGra-R раскрывается посредством региональных предпочтений в использовании отдельных словоформ, например генитивных алломорфов и относительных местоимений *was* и *das* [15].

Вышеописанные идеи корпусных исследований базируются на образцах письменных текстов. Однако ими исследовательский потенциал DeReCo не ограничивается. Грамматические структуры современного немецкого языка исследуются также на материале устных подкорпусов. Как уже было замечено, объем корпуса позволяет проводить мас-

штабные исследования и получать валидные результаты. На материале устных текстов представляется возможным изучать особенности употребления определенных лексических и фонетических феноменов немецкого языка как в отдельности или совокупности речи [16], так и в сопоставлении [17]. Новизна НКРЯ в этом отношении, в отличие от DeReCo, состоит в том, что он располагает также мультимедийным корпусом. Перспективными являются исследования корпусного стандарта современной немецкой устной речи, в частности его диатопические (региональные) и диастрические (социальные) фонетические и морфосинтаксические вариации в рамках одного национального языка [16].

Немецкоязычный корпус, как и НКРЯ, располагает и к проведению сравнительно-сопоставительных исследований. На материале анализируемого корпуса возможна реализация такого амбициозного проекта, как описание типологии немецкого языка в контрастивном аспекте по отношению к другим европейским языкам, к таким как нидерландский, английский, французский, испанский, венгерский и/или турецкий [18]. Наряду с этими контрастивными в узком смысле европейскими языками рассматриваются контрастивные в широком смысле европейские языки, такие как скандинавские, кельтские, кавказские. Сопоставление ведется по так называемым функциональным доменам, присваиваемым формально определенным крупным языковым сферам. Например, к сфере объектов номинальных конструкций могут быть отнесены такие функциональные домены, как классификация референтного диапазона (например, личный, безличный, собирательный, индивидуальный). Для объектной сферы глагола или глагольных конструкций приписываются следующие функциональные домены: темпоральность, аспектуальность, модальность, диатеза (синтаксические + семантические роли).

На уровне предложения речь идет о таких функциональных областях, как «конституция фактов / участие», «изменение фактов». На уровне фраз также рекомендуется отталкиваться от существующих в немецком языке типов конструкций в качестве методологической основы, например именное словосочетание, относительное предложение, инфинитивная конструкция, сослагательное предложение и т. д. [18].

Одной из перспективных задач является изучение функционирования современного немецкого языка как в Германии, так и за ее пределами, в районах проживания языковых меньшинств, например на Южном Кавказе или в Техасе. Перспективная цель проекта состоит в том, чтобы систематически производить и исследовать голосовые записи в двух районах проживания языковых меньшинств, которые до сих пор были мало изучены. Статистический инструментарий подкорпусов DeReCo позволяет проводить исследования по квантитативной лингвистике для уточнения формул частотности употребления слов и выражений, вероятностного распределения абсолютной и относительной частотности по подкорпусам, коэффициента различия и классов частотности употребления [7]. Под абсолютной частотностью понимается величина, показывающая, сколько раз искомая единица употреблена во всех текстах корпуса. Относительная частотность показывает, какую долю составляет в данном корпусе искомая единица. Эта доля обычно указывается в процентах или так называемых промилях (на миллион слов). Относительная частотность выводится из абсолютной. Валидной данная величина может считаться только при условии работы с национальными или сопоставимыми по объему корпусами. В результате исследования делается вывод о целесообразности параллельного использования обеих величин.

Заключение (Conclusion)

Несмотря на наличие схожих параметров построения языковых корпусов, параметрический анализ позволил показать их специфические стороны и цели применения. Использование DeReCo наиболее релевантно для исследований в сферах лексикографии, грамматической вариативности немецкого языка (на материале письменных текстов) в топологическом и социологическом аспектах, фонетической вариативности в современном немецком языке в сопоставительном аспекте, а также в области корпусной и квантитативной лингвистики. Параметризация НКРЯ также подтверждает его исследовательский потенциал при анализе и изучении самого широкого круга лингвистических феноменов на основе разножанровых языковых материалов, позволяя на основе статистических данных анализировать разноуровневые параметры языка как в синхроническом, так и в диахроническом срезе. Сложно переоценить ценность анализируемых корпусов в отношении представленности статистических данных по всем параметрам в отношении единиц языка.

Анализ корпусных данных русско- и немецкоязычного корпусов включает в себя слуховую транскрипцию, аннотирование фонетических, синтаксических, морфологических, лексических параметров, диалектометрию и статистику. В обоих анализируемых корпусах каждому языковому уровню соответствуют аннотации дискурсивных стратегий и референциальных отношений. При этом немецкоязычный корпус DeReCo позволяет компоновать так называемые виртуальные корпуса, предоставляет сложные варианты поиска, например лемматизацию, операторы близости, поиск по границам предложений, логические операторы; способен выполнять сложный анализ несмежной коллокации более высокого порядка; имеет различные представления результатов поиска и разнообразных интерфейсных клиентов: Web, программный клиент MS Windows, скрипт-клиент. Корпусный поиск и анализ видятся достаточно валидными вследствие поддержки в принципе неограниченного числа слоев лингвистической аннотации. Оба корпуса ставят перед собой цель дополнять все аннотации, которые представляют интерес для пользователей или поставщиков инструментов, и сделать их доступными.

1. Каколири И. К., Лучинская Е. Н. Концептуализация научного знания в лингвистической терминологии // Вестн. Адыг. гос. ун-та. Сер. 2 : Филология и искусствоведение. 2024. № 2 (337). С. 71–75. DOI: 10.53598/2410-3489-2024-2-337-71-75

2. Национальный корпус русского языка 2.0: новые возможности и перспективы развития / С. О. Савчук, Т. А. Архангельский, А. А. Бонч-Осмоловская [и др.] // Вопросы языкознания. 2024. № 2. С. 7–34. DOI: 10.31857/0373-658X.2024.2.7-34

3. Мартыненко Г. Я. Методы математической лингвистики в стилистических исследованиях. СПб. : Нестор-История, 2019. 296 с.

4. Коннова М. Н. Когнитивные сдвиги в семантике темпоральных лексем: корпусный аспект // Методы когнитивного анализа семантики слова: компьютерно-корпусный подход / под общ. ред. В. И. Заботкиной. М. : Языки славянской культуры, 2015. С. 269–311.

5. Шилихина К. М. Изучение дискурсивных маркеров методами корпусной лингвистики // Вестн. Воронеж. гос. ун-та. Сер. : Лингвистика и межкультурная коммуникация. 2015. № 3. С. 120–125.

6. Колпакова Г. В. Корпусная лингвистика и лексикография // Грани познания. 2011. № 2 (12). С. 42–50.

7. Kupietz M., Keibel H. The Mannheim German Reference Corpus (DeReKo) as a Basis for Empirical Linguistic Research // Working Papers in Corpus-based Linguistics and Language Education. 2009. No. 3. P. 53–59.

8. Weiß Ch. Die thematische Erschließung von Sprachkorpora // OPAL — Online publizierte Arbeiten zur Linguistik. Mannheim : Institut für Deutsche Sprache, 2005. No. 1. 14 S. Elektronische Version. URL: https://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/11804 (date of access: 20.09.2024).

9. Steffens D., Nikitina O. Deutsch-russisches Neologismenwörterbuch. Neuer Wortschatz im Deutschen 1991–2010 : in 2 Bd. Mannheim : Institut für Deutsche Sprache, 2014. 598 S.

10. Steffens D., Al-Wadi D. Neuer Wortschatz. Neologismen im Deutschen 2001–2010 : in 2 Bd. Mannheim : Institut für deutsche Sprache, 2013. 577 S.

11. Steffens D. Vom Print- zum Onlinewörterbuch — Zur Erfassung, Beschreibung und Präsentation von Neologismen am IDS // Deutsch: lokal — regional — global / hrsg. von J. Dąbrowska-Burkhardt, L. M. Eichinger, U. Itakura. Tübingen : Narr Francke Attempto, 2017. S. 281–294.

12. Keibel H., Hennig S., Perkuhn R. Effiziente halbautomatische Detektion von Neologismuskandidaten. Technical Report IDS-KL-2010-01. Mannheim : Institut für deutsche Sprache, 2011. 16 S. Elektronische Version. URL: http://ids-mannheim.de (date of access: 20.09.2024).

13. Bubenhofer N., Konopka M., Schneider R. Präliminarien einer Korpusgrammatik / unter Mitwirkung von C. Brinckmann, K. Hein, B. Strecker. Tübingen : Narr Francke Attempto, 2013. 245 S. (Korpuslinguistik und interdisziplinäre Perspektiven auf Sprache (CLIP) ; Bd. 4).

14. Grammatik im Korpus. Korpuslinguistisch-statistische Analysen morphosyntaktischer Variationsphänomene / hrsg. von E. Fuß, M. Konopka, A. Wöllstein. Tübingen : Narr Francke Attempto, 2019. 357 S.

15. Hansen-Morath S., Wolfer S. Standardisierte statistische Auswertung von Korpusdaten im Projekt "Korpusgrammatik" (KoGra-R) // Grammatische Variation: Empirische Zugänge und theoretische Modellierung / hrsg. von M. Konopka, A. Wöllstein. Berlin ; Boston : de Gruyter, 2017. S. 345–356.

16. Deppermann A., Kleiner S., Knöbl R. 'Standard usage': Towards a realistic conception of spoken standard German // Language Variation — European Perspectives IV : selected papers from the Sixth International Conference on Language Variation in Europe (ICLaVE 6), Freiburg, June 2011 / ed. by G. Kaufmann, J. Caro Reina, P. Auer. Amsterdam ; Philadelphia : John Benjamins, 2013. P. 83–116.

17. Pröll S., Weber T. Die Denaturierung oraler Varietäten — Bairisch, Niederdeutsch, Färöisch // Struktur von Variation zwischen Individuum und Gesellschaft. Akten der 14. Bayerisch-Österreichischen Dialektologietagung, Salzburg, 7–9.11.2019 / hrsg. von P. C. Vergeiner, S. Elspaß, D. Wallner. Stuttgart : Franz Steiner Verlag, 2022. S. 337–360.

18. Zifonun G. Grammatik des Deutschen im europäischen Vergleich — Verbgrammatik (GDE-V), Aspektualität (GDE-A) und Fragesätze (GDE-F) // Studia Linguistica XX / hrsg. von L. Cirko. Wrocław : Wydawnictwo Uniwersytetu Wrocławskiego, 2001. S. 171–186.