

Научная статья

УДК 37.02

DOI: 10.20323/1813-145X-2025-3-144-75

EDN: ZTFRDK

Изучение русского языка как иностранного с применением корпусных технологий

Халида Нурисламовна Галимова¹, Мария Борисовна Казачкова²

¹Кандидат филологических наук, доцент кафедры иностранных языков и перевода, Казанский инновационный университет им. В. Г. Тимирязова. 420111, г. Казань, ул. Московская, д. 42/4

²Кандидат филологических наук, доцент кафедры английского языка, Одинцовский филиал Московского государственного института международных отношений. 143007, г. Одинцово, ул. Ново-Спортивная, д. 3
¹galikha@mail.ru, <https://orcid.org/0000-0003-1817-5004>

²m.kazachkova@odin.mgimo.ru, <https://orcid.org/0000-0002-0357-3010>

Аннотация. В современном обществе наблюдается увеличение интереса к изучению русского языка. Русский язык занимает пятое место среди самых распространённых языков мира и служит средством общения для 258 миллионов человек. В настоящее время наиболее востребованным и полезным видом изучения русского языка является курсовое обучение. Это обусловлено целенаправленным подходом к занятиям, который учитывает интересы, потребности и уровень языкового владения обучающихся. В учебный процесс включается социокультурная информация, отражающая особенности менталитета, культуры, традиций и образа жизни носителей языка. Для продвижения позиций русского языка необходимо разрабатывать инновационные технологии обучения и создавать современные учебники высокого качества. В настоящее время педагогическое сообщество имеет возможность интегрировать Национальный корпус русского языка в свою образовательную практику. Национальный корпус является ключевым дидактическим инструментом в области лингводидактики. С опорой на корпусные данные может быть организована работа с грамматическим, лексическим, литературным, лингвокультурным, историческим и др. явлениями. Авторы статьи на примерах компонентов лингвистического корпуса показывают, как можно обучать русскому языку инофонов. Частотный словарь, созданный сотрудниками лаборатории «Мультидисциплинарные исследования текста» на основе Национального корпуса русского языка, служит ценным лингводидактическим материалом и основой для разработки тестов на вербальный интеллект. Цель интерактивных тестов: проверить уровень языковой компетенции иностранного гражданина, необходимой для решения коммуникативных задач. В процессе тестирования проверяются умение различать значения лексических единиц и использовать их в заданном контексте. Тесты размещены на платформе-агрегаторе RuLex.

Ключевые слова: корпусная лингвистика; русский язык; лингвистический корпус; конкорданс; корпус-менеджер; частотный словарь; тесты на вербальный интеллект

Для цитирования: Галимова Х. Н., Казачкова М. Б. Изучение русского языка как иностранного с применением корпусных технологий // Ярославский педагогический вестник. 2025. № 3 (144). С. 75–87. <http://dx.doi.org/10.20323/1813-145X-2025-3-144-75>. <https://elibrary.ru/ZTFRDK>

Original article

Studying russian as a foreign language using corpus technologies

Khalida N. Galimova¹, Mariia B. Kazachkova²

¹Candidate of philological sciences, associate professor of the department of foreign languages and translation, Kazan innovative university named after V. G. Timiryasov. 420111, Kazan, Moskovskaya str., 42

²Candidate of philological sciences, associate professor of the english department, Moscow state institute of international relations (the Odintsovo branch). 143007, Odintsovo, Novo-Sportivnaya str., 3

¹galikha@mail.ru, <https://orcid.org/0000-0003-1817-5004>

²m.kazachkova@odin.mgimo.ru, <https://orcid.org/0000-0002-0357-3010>

Abstract. In modern society, there is an increasing interest in studying the Russian language. The Russian language is the fifth most widely spoken language in the world and is used as a means of communication by 258 million people. Today, the most popular type of studying the Russian language is coursework in Russian. This is due to a targeted approach to classes

that takes into account the interests and needs of students. The educational process includes socio-cultural information reflecting the characteristics of the mentality, culture, traditions and lifestyle of native speakers. To restore the position of the Russian language in the modern world, the most important thing is to develop innovative teaching technologies and create modern high-quality textbooks. The pedagogical community has the opportunity to integrate the National Corpus of the Russian Language into its educational practice. National corpus is a key tool in the field of linguodidactics. Based on corpus data, work can be organized with almost any grammatical, lexical, literary, linguacultural, historical phenomenon. The authors of the article use examples of the components of the National Corpus of the Russian Language to show how Russian can be taught as a foreign language. The frequency dictionary, created by the laboratory «Multidisciplinary Text Research» based on the NCRY, is the basis for developing tests of verbal intelligence. The purpose of interactive tests: to check the level of development of the language competence of a foreign citizen, necessary for solving communication problems in accordance with the State Standard for Russian as a Foreign Language. During testing, the ability to distinguish the meanings of lexical units and use them in a given context is checked. The tests are placed on the Ru.Lingva aggregator platform.

Key words: corpus linguistics; russian language; linguistic corpus; concordance; corpus-manager; frequency dictionary; tests of verbal intelligence

For citation: Galimova Kh. N., Kazachkova M. B. Studying russian as a foreign language using corpus technologies. *Yaroslavl pedagogical bulletin*. 2025; (3): 75-87. (In Russ.). <http://dx.doi.org/10.20323/1813-145X-2025-3-144-75>. <https://elibrary.ru/ZTFRDK>

Введение

Роль русского языка в современных условиях становится все более значимой в связи с изменениями, происходящими в нашей стране и странах, с которыми Россия непосредственно сотрудничает.

На сегодняшний день открыто более 60-ти ресурсных центров по изучению русского языка за рубежом. Например, «в Китае интерес к русскому языку стремительно растет, в Бразилии общее число владеющих русским языком бразильцев составляет около 25 тыс. человек. Русский язык преподаётся в 40 индийских вузах, важными центрами русистики являются Российские центры науки и культуры в Дели, Мумбаи, Варанаси, Калькутте» [Тимирысова, 2024, с. 175], в индийском городе Меерут на базе Университета Чаудхари Чаран Сингха. В ЮАР знающих русский язык немного (около 2,6 тыс.), но его изучение становится популярным, так как между Россией и Южной Африкой в рамках партнерства на 2023-2026 существует договоренность о создании «сети образовательных центров русского языка, которые будут открыты как в самой Африке» [Тимирысова, 2024, с. 176], так во многих регионах РФ. Наиболее сильные позиции русского языка остаются в ряде стран СНГ, в которых русский язык имеет статус официального.

Д. Е. Грибов, заместитель министра просвещения России, на презентации программы «Народы БРИКС выбирают жизнь» в Российском Университете Дружбы Народов имени Патриса Лумумбы говорил о том, что в ряде дружественных нам стран существуют программы обмена и сотрудничества в области образования, и в ближайшее время данной формой сотрудни-

чества будут охвачены все страны БРИКС и БРИКС+».

Наиболее популярным видом изучения русского языка как иностранного (РКИ) является курсовое обучение русскому языку, интерес к которому объясняется «четко выраженной целевой направленностью занятий, учетом интересов и потребностей обучающихся языку, включением в систему занятий социокультурной информации, характеризующей особенности и менталитет, культуру, традиции, образ жизни носителей языка» [Варганова, 2009, с. 143].

«Многие языковые курсы создаются как в рамках существующих высших учебных заведений, так и самостоятельных учебных подразделений, имеющих свои программы обучения, особенности организации и проведения занятий» [Тимирысова, 2024, с. 175]. «В России сегодня обучается более 350-ти тыс. иностранных студентов. Показатели растут из года в год: с 2011 число обучающихся в вузах страны иностранцев возросло на 38 тысяч» [Баранова, 2020, с. 78].

С 2021 года МГИМО (университет) МИД России активно участвует в реализации национальной государственной программы «Приоритет 2030». В рамках проекта «Russian School Abroad» университет создает сеть Русских международных школ (РМШ) (Школы полного цикла (чаще всего с 1-го по 11-й класс) за пределами России. По мнению Анатолия Торкунова, «русская международная школа – драйвер образовательной экосистемы школ МГИМО. Это не просто учебные заведения, это коммуникационные площадки, интеграторы для кристаллизации передовых методик образования для тех, кто мыслит и хочет мыслить по-русски» [Русские международные школы]. С 2022 года открыты школы в г Ташкенте (Респуб-

лика Узбекистан), в г. Бишкеке (Кыргызская Республика), г. Стамбуле и г. Силифке (Турецкая Республика), г. Фошань (Китайская Народная Республика), г. Никосии (Республика Кипр). М. Карлова, вдова российского чрезвычайного и полномочного посла А. Карлова в Турции, признается, что «появление в Турции русскоязычных государственных образовательных программ создает предпосылки для углубления сотрудничества наших стран в гуманитарной сфере, продвижении русского языка и образования, культурного диалога между нашими народами» [Русские международные школы]. До 2030 года предполагается открытие не менее 15-ти школ в Турецкой республике.

В основе проекта – совместная деятельность Горчаковского лицея МГИМО, школ-партнеров за рубежом и цифровой образовательной платформы IBLIS (International Blended Learning School). Обучение проходит в смешанном формате, сочетает в себе дистанционные модули и очные обучающие сессии с педагогами Горчаковского лицея и зарубежной школы-партнера в стране проживания обучающихся. Р. Котов, руководитель проекта, подчеркивает, что в Русской международной школе (РМШ) (Школы полного цикла (чаще всего с 1-го по 11-й класс), учебная программа которых сочетает национальный и российский образовательные компоненты) обучают не просто учителя, а мудрые и опытные наставники. «Философия каждого – постоянное повышение квалификации, ювелирный подбор программ, воспитание детей в парадигме *lifelong learning* и открытости миру» [Русские международные школы]. Программы РМШ реализуются в соответствии с ФГОС. Высокий уровень качества преподавания и соответствие Российским образовательным стандартам позволяют обучающимся подготовиться к экзаменам и по завершении обучения пройти государственную итоговую аттестацию (ГИА), получить российский аттестат.

Для восстановления позиций русского языка в мире важной задачей становится создание новых передовых технологий обучения и высококачественных современных учебников.

Результаты исследования

Лингвистический корпус

Сегодня педагогическое сообщество располагает масштабным лингвистическим и дидактическим инструментарием для обучения русскому языку иностранных студентов. Один из них – лингвистический (языковой) корпус [Захаров, 2020, с. 234], который является составной

частью передовых педагогических технологий, одной из уникальных форм обучения. Согласно А. И. Глазыриной, применение корпусных технологий в сфере лингвистического обучения рассматривается как бурно развивающаяся область, привлекающая внимание множества специалистов [Глазырина, 2012].

Языковой корпус является источником естественного языка. «A collection of naturally-occurring language text, chosen to characterize a state of variety of a language» [Sinclair, 1991, с. 17] («собрание естественных языковых текстов, чтобы охарактеризовать уровень многообразия языка» (*перевод авторов статьи*)).

По мнению П. В. Сысоева, корпус – это «информационно-справочная система, основанная на собрании текстов на некотором языке в электронной форме» [Сысоев, 2010, с. 99].

По меткому замечанию В. А. Плунгяна, «корпус – это не только мощный инструмент исследования языка, но и новая идеология, ориентирующая исследователя на текст как главный объект теоретической рефлексии» [Плунгян, 2008, с. 12].

По определению Вилсона Э. и Т. МакЭнери, корпус – «собрание языковых фрагментов, отобранных в соответствии с четкими языковыми критериями для использования в качестве модели языка» [McEnergy, 2001, с. 112]. В отличие от бумажных текстовых каталогов или же единичных текстов, «корпус текстов позволяет моментально осуществить поиск нужных элементов, а также выявить их количественные, качественные и статистические характеристики» [McEnergy, 2001].

«Современный этап развития технологий и систем обработки текстов предоставляет возможность собирать и обрабатывать массивные данные для корпусов» [Анохина, 2007, с. 155], что повышает доступность обучения и обучающих материалов для всех.

Существует множество различных типов лингвистических корпусов, определяющихся многообразием исследовательских и прикладных задач, для решения которых они и создаются. «Лингвисты-практики все чаще обращаются к языковым корпусам при обучении языкам, для создания учебников, пособий и решения своих профессиональных задач» [Палийчук, 2022, с. 74].

В качестве примера языковых корпусов русского языка можно привести следующие: Компьютерный корпус текстов русских газет конца XX-го века, созданный на Филологическом факультете МГУ в Лаборатории общей и компьютерной лексикологии и лексикографии, Уппсальский корпус русского языка, разработанный Тю-

бингенским университетом в Германии и отражающий современное состояние русского языка; Корпус русского литературного языка (<https://narusco.ru/>); Открытый корпус русского языка; Мультимедийный русский корпус (МУРКО); Генеральный Интернет-корпус Русского языка (ГИКРЯ), мегакорпус более чем 20 млрд. слов и др.

Удобство использования корпуса неоспоримо: к нему можно обращаться в любое время в режиме online, пользоваться им, не имея специальных навыков.

За основу нашего исследования взяты компоненты Национального корпуса русского языка (НКРЯ), который является результатом масштабного проекта Российской академии наук, начатого в 2001 году [Шаров, 2003; Плунгян, 2005]. НКРЯ представляет собой информационно-справочную систему, интернет-сервис, «вещающий обширный массив филологически компетентных речевых данных. Все данные собраны в единую систему по определённым критериям (автору текстов, жанру, времени создания текста и т. п.) и дополнены поисковой системой, основанной на метаразмётке исследуемого текста» [Захаров, 2020, с. 239].

НКРЯ – это «подвижный» корпус, который постоянно обновляется и пополняется новыми текстами. Это позволяет отслеживать диахронические и синхронические изменения в языке, развитие того или иного слова в различные исторические периоды, выявлять неологизмы и устаревшую лексику, анализировать тенденции в употреблении языка с течением времени.

Данный корпус был создан группой ученых-лингвистов Института русского языка им. В. В. Виноградова для изучения современного русского языка, его структуры, лексики, грамматики и стиля.

Основные цели НКРЯ:

1. *Исследование языка.* Корпус предоставляет исследователям возможность анализировать частотность слов, фраз и грамматических конструкций.

2. *Образование.* Он может быть использован в качестве учебного материала для преподавателей и обучающихся при изучении языка. «Дидактические материалы НКРЯ позволяют создавать задания, основу которых составляют предложения из реальных текстов корпуса» [Анохина, 2007, с. 154]. Обычные учебно-методические пособия зачастую не могут полностью отразить черты, характерные для естественного разговорного дискурса, поскольку чаще всего они не яв-

ляются примерами живого использования языка. Кроме того, удобный интерфейс дает возможность пользователям легко взаимодействовать с данными и выполнять необходимые операции.

3. *Лексикография.* Данные корпуса могут служить основой для создания различных языковых словарей и других лексикографических ресурсов.

В настоящее время полный объём НКРЯ составляет более 2 млрд. словоупотреблений. Около 40 % объёма основного корпуса – художественная литература; 60 % – нехудожественные разножанровые тексты.

Рассмотрим возможности корпусной лингвистики как актуального и перспективного направления в современной языковой науке.

Корпусное обучение (corpus-based language learning) становится все более актуальным в области языкового обучения и обработки естественного языка. Данный метод ориентирован на практическое освоение языка, его использование в реальных ситуациях и текстах, что подразумевает изучение иностранного языка (в частности, русского) через практику речи этого языка [Беляев, 2004, с. 7].

В современной отечественной методике преподавания РКИ только делаются первые шаги в этом направлении, разрабатываются учебные материалы, опирающиеся на корпусные технологии. Методы корпусной лингвистики представляются наиболее оптимальными, так как предполагают обработку большого количества текстов с целью решения самых разнообразных лингвистических задач. Для пользователей корпусов «на первый план выдвигается не только содержание текстов, но и их метатекстовая информация (метаразмётка, паспортизация, аннотация), примеры употребления определенных языковых элементов и конструкций» [Анохина, 2007, с. 153].

Метаразмётка текстов корпуса с полным прописыванием определенных атрибутов, характеризующих обстоятельства создания текста, его автора (авторов), тематику, жанровые особенности, дату создания и т. п., позволяет произвести социолингвистическую и семантическую диверсификацию текстов в зависимости от уровня сложности, изучаемой темы и т. д. Таким образом, метаразмётка необходима для исследователя, прежде всего, для поиска выборки текстов с определенными параметрами.

Существует большое число разновидностей разметки. Например, «при лексико-семантической разметке большинству слов в тексте приписывается один или несколько семантических и словообра-

зовательных признаков, например, “вещество”, “животные”, “профессия”, “признак”, “характер человека”, “отглагольное существительное” и т. п. Используется также фасетная классификация, при которой одно слово может попадать в несколько тематических групп» [Лонцкая, 2023, с. 4].

Морфологическая метаразметка представляет в корпусе информацию о морфологических формах и значениях (часть речи, род, число, падеж, время, вид и т. п.).

Темпоральная же разметка охватывает слова и словосочетания с временной семантикой, например, *сегодня, вечером, 5 ноября, ночью, с детства*, и отражает их вклад в формирование смысла предложения.

Отбор текстов и коллокатов «дает возможность сделать обучение индивидуальным и позволяет преподавателю варьировать, адаптировать дидактический материал, тем самым, оптимизируя процесс обучения» [Анохина, 2007, с. 155].

Использование НКРЯ в процессе обучения РКИ

НКРЯ – важнейший инструмент любого исследователя, который занимается русским языком и вообще русской словесной культурой.

Рассмотрим некоторые возможности использования Корпуса для обучения инофонов русскому языку с использованием базы НКРЯ.

Одним из педагогических инструментов в процессе обучения языку является создание строк программы-конкорданса как «упорядоченного списка всех употреблений словоформы или сочетания в контексте и с ссылкой на источник» [Казачкова, 2022, с. 34] и дату. А. С. Герд описы-

вает конкорданс как «упорядоченный список словоформ с указанием всех вхождений того или иного слова в заданный массив текстов, при этом для каждой словоформы приведен соответствующий контекст, иллюстрирующий данное словоупотребление» [Герд, 1981, с. 75]. В. П. Захаров определяет конкорданс как список контекстов, где «искомая единица представлена в её лексическом окружении и характеризуется набором статистических данных» [Захаров, 2005, с. 55].

Конкордасы иногда называют Корпусными менеджерами (корпус-менеджерами). Однако есть некоторые определенные различия между этими двумя понятиями. Корпус-менеджеры дают возможность не только найти слово или сочетание в контексте, но также получить из корпуса дополнительную информацию исходя из его метаданных, что может включать информацию о частях речи, синтаксической структуре, семантике и других аспектах; частотность употребления слов, а также гипертекстовую разметку документа для просмотра веб-страниц в браузере (HTML).

Конкорданс необходим для быстрой и лучшей ориентации в материале. Указав в «поиске точных форм» в основном корпусе или в одном из подкорпусов слово или словосочетание, соответствующее определенному типу, преподаватель получит подборку в сотни предложений. Например, введем словосочетание «родительский дом», между компонентами которого установились атрибутивные отношения. Полученные результаты отражены на рисунке (см. рис. 1).

Конкорданс
KWIC
График
Статистика
Частотность
2-граммы
3-граммы
4-граммы
5-граммы

Скачать

7. Д. П. Дроздов. Большая Ордынка. Прогулка по Замоскворечью от Москворецкого моста до Серпуховской площади (2017)

После смерти Меланьи Захаровны **родительский дом** достался семье младшего брата Давида Ивановича.

В конце жизни он вовсе разорился, продал **родительский дом** и перебрался в маленький домик на окраине города.

8. В. Г. Глушкова. Путешествие из Москвы в Ярославль. Москва -- Сергиев Посад -- Переславль-Залесский -- Ростов Великий -- Ярославль (2015)

Их женившийся в 1893 г. старший сын Александр Евгеньевич с молодой и очень привлекательной 19-летней женой Инессой Теодоровной (Федоровной) Стефан (1874–1920) поселились в их подмосковной усадьбе Ельдиго и нечасто приезжали в **родительский дом**-усадьбу в Пушкине (см. стр. 22–25).

9. А. Б. Сальников. Отдел // «Волга», 2015

Он оценил, что жена не бросила его, пока у него не было работы, потому что без этой работы на это государство он сам вернулся в **родительский дом** и со стороны выглядел бы даже более жалко, чем сисадмин.

10. Алексей Горшенин. Многостаночник // «Сибирские огни», 2013

Негасимым Божественным светом был освещен и **родительский дом** Плитченко, на создание которого положил он всю свою творческую жизнь.

Активация Windows

Чтобы активировать Windows, перейдите в «Параметры».

Страницы: 1 2 3 4 5 ... 22

Рисунок 1. Результаты поиска словосочетания «родительский дом» в базе НКРЯ

Контекст использования словосочетания даёт возможность наиболее точно и конкретно выявить смысл и значение отдельного входящего в него слова или фразы.

Изучение русских фразеологизмов представляет собой особенную трудность и часто вызывает непонимание их семантики. Изучение идиом, «встроенных» в конкретный контекст, оказывается особенно полезным при их исследовании (см. рис. 2).

Конкорданс KWIC График Статистика Частотность 2-граммы 3-граммы 4-граммы 5-граммы

1. Юрий Давыдов. Синие тюльпаны (1988-1989) Ⓜ

Кавторанг Карпов не давал *бить баклуши*, дважды в неделю — шлюпочные занятия. Ⓜ Ⓜ

2. М. Б. Бару. Второй сон Любви Александровны // «Волга», 2015 Ⓜ

Он же первым стал *бить баклуши* — чурбачки-заготовки для изготовления ложек. Ⓜ Ⓜ

3. В. В. Лорченков. Бездна (2012) // «Волга», 2013 Ⓜ

Хватит уже *бить баклуши* за мамин счет! — восклицает она. Ⓜ Ⓜ

4. Виктор Слипенчук. Зинзивер (2001) Ⓜ

На ходу спит, а *бить баклуши* — первая! Ⓜ Ⓜ

Рисунок 2. Результаты поиска идиомы «бить баклуши» в базе НКРЯ

НКРЯ играет значительную роль при изучении синонимов в русском языке. Как правило, обучающиеся допускают ошибки в их употреблении. Например, при различении синонимов «болезнь» и «заболевание» вы можете обратиться к толковым словарям, но примеров употребления данных синонимичных лексических единиц в них будет недостаточно. Например, «*Болезнь. Расстройство здоровья, нарушение деятельности организма. Детские болезни. Заразная болезнь. Болезни растений. Болезни роста.*» «*Заболевание. Процесс действия по гл. заболеть. Аллергическое заболевание. Момент заболевания*» (Толковый словарь русского языка под ред. С. И. Ожегова, Н. Ю. Шведова).

Подобрать примеры функционирования синонимов поможет использование материалов НКРЯ. Из полученных данных можно сделать вывод о том, что существительные «болезнь» и «заболевание» — близки по значению, констатируют наличие расстройства здоровья. Но отглагольное существительное «заболевание», означает еще и начало процесса недомогания, момент возникновения недуга. Кроме того, используя материалы НКРЯ, можно узнать, в каких сферах функционирования чаще употребляются слова (см. рис. 3, 4).

№	Значение атрибута ▼	Тексты ▼	Вхождения ▼	IPM ▼
1	художественная	2049	5691 (36,26%)	35,74
2	публицистика	2500	5598 (35,67%)	39,91
3	учебно-научная	548	2039 (12,99%)	46,15
4	бытовая	939	1951 (12,43%)	57,88
5	церковно-богословская	161	405 (2,58%)	76,56
6	электронная коммуникация	55	105 (0,67%)	31,05
7	официально-деловая	50	72 (0,46%)	13,39
8	реклама	17	22 (0,14%)	26,06
9	производственно-техническая	9	12 (0,08%)	7,32

Рисунок 3. Сферы функционирования употребления слова «болезнь» (по данным НКРЯ)

№	Значение атрибута ▼	Тексты ▼	Вхождения ▼	IPM ▼
1	учебно-научная	239	630 (41,58%)	14,26
2	публицистика	439	565 (37,29%)	4,03
3	художественная	139	175 (11,55%)	1,1
4	бытовая	62	66 (4,36%)	1,96
5	электронная коммуникация	21	41 (2,71%)	12,12
6	реклама	7	10 (0,66%)	11,85
7	официально-деловая	7	10 (0,66%)	1,86
8	церковно-богословская	7	10 (0,66%)	1,89
9	производственно-техническая	4	8 (0,53%)	4,88

Рисунок 4. Сферы функционирования употребления слова «заболевание» (по данным НКРЯ)

Эти небольшие фрагменты показывают, «как быстро можно найти иллюстративный материал, проверить его употребление на реальных примерах благодаря технологии поиска в лингвистическом корпусе, более того, произвести отбор материала для анализа и составления упражнений в дифференциации изучаемого лексического и грамматического явления» [Сысоев, 2010, с. 100].

НКРЯ имеет различные подкорпусы. «Подкорпус параллельных текстов, размещенный на сайте НКРЯ, имеет наиболее функциональный интерфейс среди всех общедоступных корпусов (подкорпусов) и представляет особый интерес. Он состоит из текстов на русском языке с их профессиональными параллельными переводами на иностранные языки» [Тимирясова, 2024, с. 180]. На сайте НКРЯ представлены аутентичные тексты с выровненным переводом на английский, немецкий, китайский, испанский, португальский, хинди, турецкий, арабский, узбекский и др.»

Незаменимым помощником для изучающих русский язык на всех уровнях обучения может стать и подкорпус устной речи, «в который включены аудиозаписи публичной и частной устной речи, транскрипты кинофильмов (подкорпус МУРКО), а также подкорпус диалектных текстов, включающий запись диалектной речи различных регионов России с сохранением их фонетической, семантической и грамматической специфики» [Захаров, 2020, с. 131].

Нетрадиционным способом освоения русского языка может стать «использование Газетного корпуса (корпуса современных СМИ), который также представлен в составе НКРЯ. Газетный подкорпус представляет собой публицистику 2000-х гг. Этот

компонент корпуса интересен и полезен иностранцам, желающим ознакомиться с политической или экономической ситуацией в России» [Анохина, 2007, с. 154].

Образ России у большинства иностранцев ассоциируется с русской литературой, с великими писателями, такими как Фёдор Достоевский, Лев Толстой, Антон Чехов и Александр Пушкин. Русская литература не только помогает улучшить языковые навыки, но, прежде всего, несет в себе культурный код, ведет к пониманию сущности русского народа и его истории, обладает ярко выраженным национальным характером и представляет менталитет русского человека через все его общенациональные формы: литературную, разговорную, диалектную, просторечную. Именно в русской литературе сконцентрировано многообразие уникальных стилистических приемов и других ярких экспрессивных средств языка: метафор, эпитетов, фразеологизмов, образных сравнений, метонимий, олицетворений и др. Все это создает неповторимую атмосферу произведений и углубляет понимание русской культуры. Этот корпус более подходит студентам продвинутого и среднего этапов обучения.

Частотный словарь русского языка

НКРЯ также предлагает словарь самых частотных слов русского языка, построенный на его основе. Частотный список слов даёт возможность получить представление о тематике и основных понятиях текста. Полный частотный словарь создан в рамках Федеральной целевой программы Федерального агентства по образованию РФ «Русский язык» (Госконтракт П66).

Словарь частотности современного русского языка НКРЯ под редакцией О. Н. Ляшевской и С. А. Шарова – это лексикографический источник, который содержит информацию как о частотности употребления, так и статистическом распределении словоформ по текстам и жанрам, по времени создания текстов. Он был составлен на основе анализа текстов и может использоваться для различных целей, таких как лексикогра-

фия, обучение языку и исследование языковых тенденций. Объем выборки, на которой строится большинство разделов словаря, составляет более 92 млн. словоупотреблений. Он включает тексты художественной литературы, средств массовой информации, технические, деловые документы, монографии и т. д., а также записи разговорной речи. Баланс текстов, представленных в частотном словаре, показан на рисунке (см. рис. 5).

Функциональный стиль	Доля	Размер подкорпуса, токенов	Размер подкорпуса, орф. слов	Кол-во текстов
Художественная литература	39.04%	45 150 317	35 150 521	2 418
Публицистика	42.21%	48 818 173	39 739 644	27 390
Другая нехудожественная литература,	16.96%	19 618 518	15 478 151	7 495
в т. ч. учебно–научная	11.30%	13 067 152		3 994
официально–деловая	1.62%	1 872 482		1 075
электронная коммуникация	1.49%	1 727 363		133
церковно–богословская	1.44%	1 664 804		488
реклама	0.57%	659 707		1 232
бытовая	0.48%	556 291		439
производственно–техническая	0.26%	295 206		134
Устная непубличная речь	0.88%	1 017 568	758 407	1 005
Другое (в т.ч. смешанный стиль)	0.90%	1 037 468	827 580	61
Итого	100%	115 642 044	91 954 303	38 369

Рисунок 5. Соотношение текстов разных функциональных стилей в частотном словаре

Частотный словарь может быть полезен на практике для изучающих русский. Он также интересен для разработки различных тестов. В словаре можно получить следующую информацию об употреблении слова в корпусе: порядковый номер в общем частотном списке, ранг слова, то есть эмпирическую закономерность распределения частотности слов (чем чаще слово встречается, тем выше его ранг в словаре), часть речи; общее количество употреблений словоформы (абсолютная нормализованная частот-

ность слова на миллион, ipm), распределение употребления слова в текстах, созданных в разные периоды времени. В словарях значимой лексики можно также получить информацию о сравнительной частотности слова в общем корпусе и в подкорпусе текстов определенного функционального стиля (художественная литература, научная литература публицистика и т. д.) [Ляшевская, 2021]. Пример оформления частотного словаря представлен на рисунке (см. рис. 6).

Частотный список лемм

Лемма	Лемма	Часть речи	Частота (ipm)	Худ.лит.1950-60-е	1970-80-е	1990-2000-е	Публ. 1950-60-е	1970-80-е	1990-2000-е
1	и	conj	35801.8	37546.0	39187.7	36514.7	36703.6	38352.1	35539.5
2	в	pr	31374.2	23943.5	25005.7	26185.7	35295.3	33264.3	36372.5
3	не	part	18028.0	21154.3	22921.5	22349.1	16127.4	18796.9	16853.2
4	на	pr	15867.3	17169.9	16411.7	16663.0	15249.8	15961.9	16675.2
5	я	spro	12684.4	17022.7	17090.5	18257.0	13881.5	16386.7	10447.1
6	быть	v	12160.7	12467.9	13492.8	12778.1	14458.2	15091.8	12646.0
7	он	spro	11791.1	19708.0	18587.1	16882.0	11721.6	12799.3	8811.7
8	с	pr	11311.9	11186.1	10974.4	11655.3	11060.5	11662.5	11462.7
9	что	conj	8354.0	8155.2	8313.9	8938.8	7634.5	9702.4	8743.1
10	а	conj	8198.0	10298.0	10772.2	9897.5	5748.9	7643.3	7298.8
11	по	pr	5786.7	4506.0	4616.7	4840.0	5059.6	5353.8	6656.1
12	это	spro	5707.8	6038.4	6137.5	5946.6	5378.3	6614.3	5831.8
13	она	spro	5592.4	8375.6	8348.8	9474.0	5302.7	4673.6	3733.1
14	этот	apro	5414.0	5140.8	4817.3	4723.1	6032.0	6689.8	5919.4

Рисунок 6. Отрывок Частотного словаря под редакцией О. Н. Ляшевской и С. А. Шарова

Сотрудниками научно-исследовательской лаборатории «Мультидисциплинарные исследования текста» Института филологии и межкультурной коммуникации Казанского федерального университета под руководством доктора филологических наук, профессора М. И. Солнышкиной на основе данного словаря был реализован масштабный проект «по разработке тестов на вербальный интеллект (лексические тесты). Данные тесты оценивают словарный запас иностранных граждан и определяют их уровень владения РКИ в соответствии с российской государственной системой тестирования» [Тимирысова, 2024, с. 177].

Цель настоящих лексических интерактивных тестов – проверка уровня сформированности языковой компетенции иностранного гражданина, необходимой для решения определённых коммуникативных задач [Солнышкина, 2015] в соответствии с Государственным стандартом по русскому языку как иностранному. Тесты помещены на платформе-агрегаторе RuLingva.

Разрабатываемый инструмент позволяет измерить уровень владения русским языком, а также проводить сопоставительные анализы в синхронии и диахронии, осуществлять кросс-культурные исследования.

Частотный словарь, на основе которого были разработаны тесты, был подвергнут редактированию. Из готового частотного списка были исключены стоп-слова (личные местоимения, частицы), эмоционально-оценочная лексика (собачка, березка и т. д.); некоторые слова-термины

(авиафобия, канифора и т. д.). В результате был сформирован словник, состоящий из стилистически нейтрально окрашенных слов, принадлежащих к знаменательным частям речи, включающий 25 000 лексических единиц. Таким образом была сформирована выборка содержания теста.

На следующем этапе списки слов были ранжированы по уровням владения языком по шкале уровней Российской системы сертификационно-го тестирования по РКИ.

За основу классификации списков по сложности для распределения по шести уровням был взят «Лексический минимум по русскому языку» под редакцией Н. П. Андрияшина и Т. В. Козлова [Андрияшина, 2020].

Таким образом, тест состоит из лексических заданий, правил их применения, оценок за выполнение блока заданий и интерпретации тестовых результатов [Копылова, 2004]. При составлении тестов к каждому тестовому заданию подбирается один правильный ответ и три правдоподобных неправильных вариантов ответа (дистракторы). Они используются для отвлечения внимания от правильного ответа тех, кто либо совсем не знает правильный ответ, либо пытается угадать его во время тестирования.

Для диагностики этого этапа привлекались только такие лексические задания, которые соответствуют программным требованиям [Хамраева, 2023]. Интуитивно понятный интерфейс, экономичность тестов (небольшие затраты на выполнение тестов) и несложность их организа-

ции позволяют легко проходить тесты как взрослым, так и детям.

Тест представляет систему заданий возрастающей сложности. В настоящее время разработаны тесты уровня A1, A2, B1.

При разработке тестов элементарного уровня преимущественно использовались картинки для правильного ответа и дистракторов, поскольку студенты с данным уровнем владения языком могут взаимодействовать с другими людьми лишь в несложном разговоре, способны читать и понимать короткие надписи. Таким образом, основную часть тезауруса составляют слова, являющиеся для языка «базовыми», отражающие фундаментальные понятия. При создании тестов для диагностики элементарного уровня в качестве дистракторов могут использоваться как картинки, так и простые слова (см. рис. 7, 8).

Через реку был ... мост.
мостовой
мягкий
водный
широкий

Рисунок 7. Фрагмент лексического теста для диагностирования уровня A1

Базовый уровень русского языка достаточен для туристических путешествий по стране и общения с носителями языка. На данном уровне владения языком человек может читать несложные рассказы, статьи, адаптированные тексты, письма, сообщения в мессенджере на повседневные темы, рекламные объявления, указатели, знаки и т. п. При составлении тестов на базовый уровень также могли использоваться картинки, более сложные слова, а также контексты для вставки слов (см рис. 8.)

Книги автора ... популярностью.
смотрят
кончаются
читают
пользуются

Рисунок 8. Фрагмент лексического теста для диагностирования уровня A2

Первый (I) уровень B1 (пороговый уровень) – первый сертификационный уровень владения русским языком – позволяет инофону справлять-

ся с большинством ситуаций в социально-бытовой, социально-культурной и учебно-профессиональной сферах общения, предполагает чтение и понимание рекламных буклетов, несложных статей в журналах и газетах, а также умение писать личные письма, описывая свои впечатления и события.

Сертификат данного уровня дает возможность поступление в российский вуз с обязательным последующим изучением русского языка.

Здесь в качестве заданий к тестам могут использоваться дефиниции (см. рис. 9).

Становится (стать) зрелым, вышедшим из детского возраста
взрослеть
стареть
растить
рождаться

Рисунок 9 Фрагмент лексического теста для диагностирования уровня B1

Разработанные тестовые задания на данный момент проходят этап эмпирической апробации. Тесты последующих уровней находятся в стадии разработки. Предстоит также огромная работа над надежностью, сопряженной с прогностической валидностью, стандартизацией тестов, и полный процесс их психометрической проверки, проведения лицензирования, сертификации в соответствующих органах.

Лингвистический Анализатор RULEX

Специалистами лаборатории создан лингвистический анализатор RULEX, предназначенный для оценки сложности текстов на русском языке для изучающих его как иностранный.

RULEX предоставляет пользователю качественные и количественные параметры текстов, определяет читательский, терминологический диапазон текста и индексы сложности тестовых материалов (см. рис. 10), включая «индекс удобочитаемости Флеш-Кинкейда и индекс лексического разнообразия (TTR). Анализатор также классифицирует словоформы по уровням владения русским языком и генерирует список частотности употребления лексических единиц» [Солнышкина, 2015, с. 88].

	Показатель	Значение	Пояснение
1	Средняя длина предложения	26	①
2	Средняя длина слова (в слогах)	3.23	①
3	Средняя длина слова (в буквах)	7.31	①

Индекс читабельности текста ①			
	Показатель	Значение	Пояснение
1	FKGL (PSIS)	5.8	①

Лексические параметры			
	Показатель	Значение	Пояснение
	Лексическое разнообразие	0.88	①

Рисунок 10. Лингвистический анализатор RULEX

Оценка сложности текста является важным инструментом в образовании, коммуникации и исследованиях. В образовательной среде информация о сложности текста помогает преподавателям сделать вывод о целесообразности использования того или иного текста, а также адаптировать материалы в соответствии с уровнем подготовки учащихся.

Анализ сложности текста позволяет оценить уровень читательских навыков у студентов или целевой аудитории. Это может помочь в диагностике и планировании дальнейшего обучения.

Определение сложности текста может быть полезным в исследованиях восприятия и понимания языка, а также в изучении когнитивных процессов, связанных с чтением.

Заключение

Подводя итоги, отмечаем, что в современном мире русский язык является связующим звеном для миллионов людей, объединяя их в одну языковую общность, сохраняя культурные ценности, традиции и историческое наследие. Русский язык есть основной инструмент науки, передачи знаний, идей и развития критического мышления.

Национальный корпус русского языка – неопенимый лингводидактический материал и для преподавателей, с помощью него можно составлять задания, организовывать самостоятельную и исследовательскую работу, проверять данные учебников и словарей, и для обучающихся: корпус дает им возможность формировать и развивать лексико-грамматические навыки, осуществлять сбор фактического материала при выполнении исследовательских проектов.

Тесты на вербальный интеллект, созданные сотрудниками НИЛ «Мультидисциплинарные

исследования текста» на основе частотного словаря НКРЯ, также как и сам словарь, являются полезными инструментами для лингвистов, переводчиков, преподавателей языка и всех, кто интересуется языком.

Библиографический список

1. Андрияшина Н. П. Лексический минимум по русскому языку как иностранному. Элементарный уровень. Базовый уровень. Первый сертификационный уровень / Н. П. Андрияшина, Т. В. Козлова. 6-е изд. Санкт-Петербург : Златоуст, 2020. 116 с.
2. Анохина С. А. Использование материалов Национального корпуса русского языка при изучении лексических категорий // Национальный корпус русского языка и проблемы гуманитарного образования. Москва : ИД ГУ-ВШЭ, 2007. С. 150–159.
3. Баранова И. И. Особенности дистанционного обучения русскому языку иностранных учащихся в современных условиях российского вуза / И. И. Баранова, М. В. Виноградова, М. Ю. Доценко // ПНИО. 2020. № 6 (48). URL: <https://cyberleninka.ru/article/n/osobennosti-distantionnogo-obucheniya-russkomu-yazyku-inostrannyh-uchaschihsya-v-sovremennyh-usloviyah-rossiyskogo-vuza> (дата обращения: 25.04.2024).
4. Беляев Б. А. Применение принципа сознательности в обучении иностранному языку // Психологические основы обучения неродному языку: хрестоматия / сост. А. А. Леонтьев. Москва : МПСИ, 2004. С. 7–13.
5. Варганова О. В. Курсовое обучение в системе овладения русским языком иностранных учащихся // Международная научно-практическая конференция «Славянская культура: истоки, традиции, взаимодействие. X Юбилейные Кирилло-Мефодиевские чтения». Москва : Ремдер, 2009. С. 140–146.
6. Герд А. С. Автоматизация в лексикографии и словари-конкордансы // Филологические науки. 1981. № 1. С. 72–78.
7. Глазырина А. И. Использование конкордансов при обучении английскому языку для специальных

целей // Педагогическое образование в России. 2012. № 1. С. 178–182.

8. Захаров В. П. Корпусная лингвистика : учебник / В. П. Захаров, С. Ю. Богданова. 3-е изд., перераб. Санкт-Петербург : Изд-во С.-Петерб. ун-та, 2020. 161 с.

9. Казачкова М. Б. Создание лингвистического корпуса учебников английского языка / М. Б. Казачкова, Х. Н. Галимова // Иностранные языки в школе: научно-методический журнал. 2022. № 2. С. 32–38.

10. Копылова С. С. Конструирование лексических тестов для контроля языковой компетенции // Гаудеамус. 2004. № 6. С. 108–115.

11. Лонцкая М. Применение инструментов корпусной лингвистики в обучении русскому языку как иностранному // МНИЖ. 2023. № 2 (128). URL: <https://cyberleninka.ru/article/n/primeneniye-instrumentov-korpusnoy-lingvistiki-v-obuchenii-russkomu-yazyku-kak-inostrannomu> (дата обращения: 25.04.2024).

12. Ляшевская О. Н. Новый частотный словарь русской лексики / О. Н. Ляшевская, С. А. Шаров. URL: <http://dict.ruslang.ru/freq.php> (дата обращения: 20.05.2024).

13. Национальный корпус русского языка. URL: <https://ruscorpora.ru> (дата обращения: 26.11.2024).

14. Палийчук Д. А. Корпусные технологии в лингвистических исследованиях // Гуманитарные исследования. История и филология. 2022. № 6. С. 72–79.

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

16. Российское образование для иностранных граждан. URL: <http://www.russia.edu.ru/> (дата обращения: 26.11.2024).

17. Русские международные школы. URL: <https://russianabroad.school/about/rws/> (дата обращения: 26.11.2024).

18. Солнышкина М. И. Сложность текста: этапы изучения в отечественном прикладном языкознании / М. И. Солнышкина, А. С. Кисельников // Вестн. Том. гос. ун-та. Филология. 2015. № 6 (38). С. 86–99.

19. Сысоев П. В. Лингвистический корпус в методике обучения иностранным языкам. 2010. № 1. С. 99–111.

20. Тимирясова А. В. БРИКС: российский взгляд: пролегомены в научное исследование феномена БРИКС / А. В. Тимирясова, И. Е. Дискин, И. И. Бикеев [и др.] ; под научн. ред. проф. И. Е. Дискина. Казань : Изд-во «Познание» Казанского инновационного университета, 2024. 224 с.

21. Толковый словарь русского языка: 72500 слов и 7500 фразеологических выражений / С. И. Ожегов, Н. Ю. Шведова. Москва : АЗЪ, 1993. 960 с.

22. Факультет государственного управления Московского государственного университета имени М. В. Ломоносова. URL: <http://spa.msu.ru/> (дата обращения: 26.11.2024).

23. Хамраева Е. А. Комплексная диагностика уровня владения русским языком обучающихся в полиэтнических классах РФ / Е. А. Хамраева, Н. И. Кытина, Е. В. Болмазова // Русистика. 2023. № 3. С. 370–386.

24. Aref'ev A. The Russian language in the world: Past, present, and future // Herald of the Russian Academy of Sciences. 2014. № 84. p. 357–364.

25. McEnery T., Wilson A. Corpus Linguistics. Edinburgh: Edinburgh University Press, 2001. 235 p.

26. Sinclair J. Developing Linguistic Corpora: a Guide to Good Practice Corpus and Text. Oxford: Tuscan Word Centre, 2005.

27. The Most Spoken Languages 2023. URL: <https://statisticsanddata.org/data/the-most-spoken-languages-2023/> (дата обращения: 03.02.2025).

Reference list

1. Andrijushina N. P. Leksicheskiy minimum po russkomu yazyku kak inostrannomu. Jelementarnyj uroven'. Bazovyj uroven'. Pervyj sertifikacionnyj uroven' = Lexical minimum in Russian as a foreign language. Elementary level. Basic level. First certification level / N. P. Andrijushina, T. V. Kozlova. 6-e izd. Sankt-Peterburg : Zlatoust, 2020. 116 s.

2. Anohina S. A. Ispolzovanie materialov Nacional'nogo korpusa russkogo yazyka pri izuchenii leksicheskikh kategorij = Use of materials of the National Corps of the Russian language in the study of lexical categories // Nacional'nyj korpus russkogo yazyka i problemy humanitarnogo obrazovaniya. Moskva : ID GU-VShJe, 2007. S. 150–159.

3. Baranova I. I. Osobennosti distancionnogo obucheniya russkomu yazyku inostrannyh uchashhihsja v sovremennyh usloviyah rossijskogo vuza = Features of distance learning in the Russian language of foreign students in modern conditions of a Russian university / I. I. Baranova, M. V. Vinogradova, M. Ju. Docenko // PNiO. 2020. № 6 (48). URL: <https://cyberleninka.ru/article/n/osobennosti-distantsionnogo-obucheniya-russkomu-yazyku-inostrannyh-uchashhihsja-v-sovremennyh-usloviyah-rossijskogo-vuza> (дата обращения: 25.04.2024).

4. Beljaev B. A. Primenenie principa soznatel'nosti v obuchenii inostrannomu yazyku = Applying the principle of conscientiousness in teaching a foreign language // Psihologicheskie osnovy obucheniya nerodnomu yazyku: hrestomatija / sost. A. A. Leont'ev. Moskva : MPSI, 2004. S. 7–13.

5. Varganova O. V. Kursovoe obuchenie v sisteme ovladeniya russkim jazykom inostrannyh uchashhihsja = Coursework in the system of mastering the Russian language by foreign students // Mezhdunarodnaja nauchno-prakticheskaja konferencija «Slavjanskaja kul'tura: istoki, tradicii, vzaimodejstvie. X Jubilejnye Kirillo-Mefodievskie chteniya». Moskva : Remder, 2009. S. 140–146.

6. Gerd A. S. Avtomatizacija v leksikografii i slovarkonkordansy = Automatisatation in lexicography and concordance dictionaries // Filologicheskie nauki. 1981. № 1. S. 72–78.

7. Glazyrina A. I. Ispol'zovanie konkordanserov pri obuchenii anglijskomu jazyku dlja special'nyh celej = Use of concordancers in teaching English for special purposes // Pedagogicheskoe obrazovanie v Rossii. 2012. № 1. S. 178–182.
8. Zaharov V. P. Korpusnaja lingvistika : uchebnik = Corpus linguistics: a textbook / V. P. Zaharov, S. Ju. Bogdanova. 3-e izd., pererab. Sankt-Peterburg : Izd-vo S. Peterb. un-ta, 2020. 161 s.
9. Kazachkova M. B. Sozdanie lingvisticheskogo korpusa uchebnikov anglijskogo jazyka = Creating a linguistic corpus of English textbooks / M. B. Kazachkova, H. N. Galimova // Inostrannye jazyki v shkole : nauchno-metodicheskij zhurnal. 2022. № 2. S. 32–38.
10. Kopylova S. S. Konstruirovanie leksicheskikh testov dlja kontrolja jazykovoj kompetencii = Constructing lexical tests to control language competence // Gaudeamus. 2004. № 6. C. 108–115.
11. Lonckaja M. Primenenie instrumentov korpusnoj lingvistiki v obuchenii russkomu jazyku kak inostrannomu = Application of corpus linguistics tools in teaching Russian as a foreign language // MNIZh. 2023. № 2 (128). URL: <https://cyberleninka.ru/article/n/primenenie-instrumentov-korpusnoj-lingvistiki-v-obuchenii-russkomu-jazyku-kak-inostrannomu> (data obrashhenija: 25.04.2024).
12. Ljashevskaja O. N. Novyj chastotnyj slovar' russkoj leksiki = New frequency dictionary of Russian vocabulary / O. N. Ljashevskaja, S. A. Sharov. URL: <http://dict.ruslang.ru/freq.php> (data obrashhenija: 20.05.2024).
13. Nacional'nyj korpus russkogo jazyka = National Corps of the Russian Language. URL: <https://ruscorpora.ru> (data obrashhenija: 26.11.2024).
14. Palijchuk D. A. Korpusnye tehnologii v lingvisticheskikh issledovanijah = Corpus technologies in linguistic studies // Gumanitarnye issledovaniya. Istorija i filologija. 2022. № 6. S. 72–79.
15. Plungjan V. A. Korpus kak instrument i kak ideologija: o nekotoryh urokah sovremennoj korpusnoj lingvistiki = Corpus as a tool and as an ideology: about some lessons of modern corpus linguistics // Russkij jazyk v nauchnom osveshhenii. 2008. № 2. (16). S. 7–20.
16. Rossijskoe obrazovanie dlja inostrannyh grazhdan = Russian education for foreign citizens. URL: <http://www.russia.edu.ru/> (data obrashhenija: 26.11.2024).
17. Russkie mezhdunarodnye shkoly = Russian International Schools. URL: <https://russianabroad.school/about/rws/> (data obrashhenija: 26.11.2024).
18. Solnyshkina M. I. Slozhnost' teksta: jetapy izuchenija v otechestvennom prikladnom jazykoznanii = Complexity of the text: stages of study in domestic applied linguistics / M. I. Solnyshkina, A. S. Kisel'nikov // Vestn. Tom. gos. un-ta. Filologija. 2015. № 6 (38). C. 86–99.
19. Sysoev P. V. Lingvisticheskij korpus v metodike obuchenija inostrannym jazykam = Linguistic corpus in the methodology of teaching foreign languages. 2010. № 1. S. 99–111.
20. Timirjasova A. V. BRIKS: rossijskij vzgljad: prolegomeny v nauchnoe issledovanie fenomena BRIKS = BRICS: Russian view: prolegomena in the scientific study of the BRICS phenomenon / A. V. Timirjasova, I. E. Diskin, I. I. Bikeev [i dr.] ; pod nauchn. red. prof. I. E. Diskina. Kazan' : Izd-vo «Poznanie» Kazanskogo innovacionnogo un-teta, 2024. 224 s.
21. Tolkovyj slovar' russkogo jazyka: 72500 slov i 7500 frazeologicheskikh vyrazhenij = Explanatory dictionary of the Russian language: 72500 words and 7500 phraseological expressions / S. I. Ozhegov, N. Ju. Shvedova. Moskva : AZ, 1993. 960 s.
22. Fakul'tet gosudarstvennogo upravlenija Moskovskogo gosudarstvennogo universiteta imeni M. V. Lomonosova = Faculty of Public Administration, Lomonosov Moscow State University. URL: <http://spa.msu.ru/> (data obrashhenija: 26.11.2024).
23. Hamraeva E. A. Kompleksnaja diagnostika urovnja vladenija russkim jazykom obuchajushhihsja v politnicheskikh klassah RF = Comprehensive diagnostics of the level of proficiency in the Russian language of students in multi-ethnic classes of the Russian Federation / E. A. Hamraeva, N. I. Kytina, E. V. Bolmazova // Rusistika. 2023. № 3. S. 370–386.
24. Aref'ev A. The Russian language in the world: Past, present, and future // Herald of the Russian Academy of Sciences. 2014. № 84. r. 357–364.
25. McEnery T., Wilson A. Corpus Linguistics. Edinburgh: Edinburgh University Press, 2001. 235 p.
26. Sinclair J. Developing Linguistic Corpora: a Guide to Good Practice Corpus and Text. Oxford: Tuscan Word Centre, 2005.
27. The Most Spoken Languages 2023. URL: <https://statisticsanddata.org/data/the-most-spoken-languages-2023/> (data obrashhenija: 03.02.2025).

Статья поступила в редакцию 14.03.2025; одобрена после рецензирования 24.04.2025; принята к публикации 15.05.2025.

The article was submitted 14.03.2025; approved after reviewing 24.04.2025; accepted for publication 15.05.2025.