



СЕМАНТИКА И СЕМИОТИКА SEMANTICS AND SEMIOTICS

DOI: 10.22363/2313-2299-2023-14-2-285-304

EDN: NCULEN

УДК 811.161.1'373'374

Научная статья / Research article

Репрезентативность фразеологических зон «Большого универсального словаря русского языка»

Л.Г. Золотых¹  , Е.Н. Барышникова² , О.Ю. Космачёва³ 

¹Сычуаньский университет,
610065, Китайская Народная Республика, г. Чэнду, ул. Ванцзянлу, 29

²Российский университет дружбы народов,
117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, 6

³Астраханский государственный университет им. В.Н. Татищева,
414056, Российская Федерация, г. Астрахань, ул. Татищева 20а

 scu1959rus@mail.ru

Аннотация. Фразеологический массив и универсальное содержание фразеологических зон словарных статей «Большого универсального словаря русского языка», где многократно предложены фразеологические единицы в зависимости от грамматически опорного компонента соответственно алфавитному расположению заголовочного слова, представляют интерес в аспекте формирования исследовательского корпуса для создания словаря электронного фразеологического словаря русского языка (АСПОТ). Цель исследования — определить показатели репрезентативности фразеологических зон в структуре толкового словаря и представить фразеостатистические модели частотного соотношения заголовочных слов как структурно-семантического компонента фразеологических единиц в алфавитном ряду. Материалом исследования стали 2545 фразеологических зон, которые сгруппированы в соответствии с буквами русского алфавита. В рамках корпусного подхода использовалась комбинация лингвостатистического метода, выборочного метода в эмпирическом исследовании, методов статистического моделирования. Научная новизна работы определяется концептуальным подходом к БУС — толковому словарю в бумажном формате — как лексикографическому корпусу в исследовании употребительности (частотности) фразеологических зон в словарных статьях. Параметрами репрезентативности фразеологических зон определены генеральная лингвистическая совокупность и выборочная лингвистическая совокупность как рабочий корпус. Фразеостатистические модели сгруппированы в соответствии с областью определения — совокупностью всех заголовочных слов, зафиксированной одной буквой русского алфавита. Рассмотрение фразеологических зон направлено

© Золотых Л.Г., Барышникова Е.Н., Космачёва О.Ю., 2023



This work is licensed under a Creative Commons Attribution 4.0 International License
<https://creativecommons.org/licenses/by-nc/4.0/legalcode>

на детальную проработку их количественной характеристики и процентного соотношения параметров статистической выборки. Приведена технология исследования и формула расчета показателей фразеостатистического моделирования. Результаты исследования иллюстрируются диаграммами и определяют объективность формирования исследовательского корпуса, его ядерных и периферийных пространств. Эти факторы станут определяющими для моделирования словаря пилотного корпуса электронного словаря, направленного на смысловую реализацию фразеологических единиц.

Ключевые слова: фразеологические единицы, ФЕ, БУС, выборочная совокупность, фразеостатистические модели, исследовательский корпус

История статьи:

Дата поступления: 01.07.2022

Дата приема: в печать 15.12.2022

Для цитирования:

Золотых Л.Г., Барышникова Е.Н., Космачёва О.Ю. Репрезентативность фразеологических зон «Большого универсального словаря русского языка» // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2023. Т. 14. № 2. С. 285–304. <https://doi.org/10.22363/2313-2299-2023-14-2-285-304>

Representativeness of the Phraseological Zones in “The Great Universal Dictionary of the Russian Language”

Lidia G. Zolotykh¹  , Elena N. Baryshnikova² ,
Olga Yu. Kosmacheva³ 

¹ Sichuan University,
29, Vangqianqiu str., Chengde, China, 610065

² RUDN University,
6, Miklukho-Maklaya str., Moscow, Russian Federation, 117198

³ Astrakhan State University n.a. V.N. Tatischev,
20a, Tatischev str., Astrakhan, Russian Federation, 414056

 scu1959rus@mail.ru

Abstract. The phraseological array and the universal content of the phraseological zones of dictionary entries in “The Great Universal Dictionary of the Russian Language”, wherein phraseological units are repeatedly presented depending on the grammatically supporting component according to the alphabetical arrangement of a heading word, are of interest in the aspect of a research corpus for creating a vocabulary of an online Russian phraseological dictionary (VORPD). The purpose of the study is to determine the indicators of the representativeness of phraseological zones in the structure of the explanatory dictionary and present phrase-statistical models of the frequency ratio of heading words as a structural-semantic component of phraseological units in alphabetical order. The materials of the research make 2545 phraseological zones grouped according to the letters of the Russian alphabet. Combinations of the linguo-statistical method, the sampling method in the empirical study, and the methods of statistical modelling were used as a part of the research corpus. The scientific novelty of the research is determined by the conceptual approach to the “The Great Universal Dictionary of the Russian Language” — an explanatory paper dictionary — as a lexicographic corpus in the study of the usage (frequency) of phraseological zones in dictionary entries. The parameters of the representativeness of phraseological zones define the general linguistic set and the

selective linguistic set as a working corpus. Phrase-statistical models are grouped according to the domain of definition — the set of all heading words fixed by one letter of the Russian alphabet. A consideration of phraseological zones is aimed at a detailed study of their quantitative characteristics and the percentage ratio of statistical sampling parameters. The technology of the study and the formula for calculating the indicators of phrase-statistical modelling are given. The results of the study are illustrated by diagrams and determine the objectivity of the research corpus, its nuclear and peripheral spaces. These factors will be a determinant for modelling the vocabulary of the pilot corpus of the electronic dictionary aimed at the semantic realization of phraseological units.

Keywords: phraseological units, PhU, GUDRL, selective set, phrase-statistical models, research corpus

Article history:

Received: 01.07.2022

Accepted: 15.12.2022

For citation:

Zolotykh, L.G., Baryshnikova, E.N. & Kosmacheva, O.Yu. (2023). Representativeness of the Phraseological Zones in “The Great Universal Dictionary of the Russian Language”. *RUDN Journal of Language Studies, Semiotics and Semantics*, 14(2), 285–304. <https://doi.org/10.22363/2313-2299-2023-14-2-285-304>

Введение

Осмысление лексикографических традиций в диалектическом взаимодействии «субъективности автора с объективностью представленного в лексикографическом труде общественного знания» [1. С. 16] и соотнесение тенденций лексикографии с направлениями развития современной лингвистики [2. С. 147] определили потребность создания антропоцентрических словарей, чтобы «помочь человеку, во-первых, в формировании языка как принадлежности сознания, а во-вторых, в эффективном использовании этого языка» [3. С. 132]. Результатом многолетнего фундаментального академического исследования в направлении синергетического лексикографирования стал «Большой универсальный словарь русского языка» (В.В. Морковкин, Г.Ф. Богачёва, Н.М. Луцкая, 2016 г.), многомерное пространство которого представляет интенсивное лексикографирование, репрезентируя универсальную лингвистическую разработку номинативных единиц, составляющих лексическое ядро современного русского языка [4–7]. Степенью всеохватности «Большого универсального словаря русского языка» (далее — БУС) «могут служить параметры описания слова, которых в рассматриваемом словаре насчитывается порядка семнадцати» [6. С. 112], в том числе характеристика фразеологической ценности как способности «слова входить в употребительные и коммуникативно важные фразеологические единицы» [8. С. 2].

Фразеологический параметр как интерпретатор фразеологической ценности заголовочного слова в структуре словарной статьи БУС представлен в части, выделенной традиционным для лексикографии символом (◇) и названной авторами «фразеологической зоной», которая отображает расширен-

ный подход к объекту фразеологии. Лексикографическая разработка фразеологических зон в структуре БУС показывает структурно-семантическую активность заголовочного слова в составе различных фразеологических единиц (далее — ФЕ), направляя пользователя через «многолучевые» ссылки к словарным статьям, репрезентирующим употребление заголовочного слова как фразеологического компонента.

Интерес к образной семантике номинативных единиц, сохраняющих многообразные речемыслительные практики лингвокультуры и способных демонстрировать своими узуальными и окказиональными смыслами ценностные отношения в их современном понимании, расширил границы фразеологизма как объекта лексикографического описания. Для исследователей образного слова достоверность традиционных фразеологических словарей является постоянным научным базисом. Однако следует отметить интеграцию научной фундаментальности и доступного изложения материала, присущую современным фразеологическим словарям¹, что определяет значимость их популяризаторской функции не только в научном, но и в обыденном познании.

Представление разносторонних характеристик ФЕ в разных лексикографических трудах определяется научными направлениями авторов-составителей, что отражается в структурных особенностях словарных статей, в частности в широкой проекции фразеографирования как одной из универсальных особенностей БУС. Изучение фразеологической модели как типовой схемы когнитивно-дискурсивного отношения компонентного состава фразеологического знака к его свободносинтаксическому прототипу выявляет первично-денотативные объекты этноязыковой картины мира и позволяет сделать общие выводы о типичном предметно-чувственном опыте. Привлечение корпусного анализа для исследования изменений фразеологической семантики выявляет эволюцию этого предметно-чувственного опыта для определенного лингвокультурного сообщества [9; 10]. В этом отношении БУС интересен своими фразеологическими зонами, представляющими многократно ФЕ в за-

¹ *Алефиренко Н.Ф., Золотых Л.Г.* Фразеологический словарь. Культурно-познавательное пространство русской идиоматики. М.: ЭЛПИС, 2008. 469 с. *Баско Н.В., Зимин В.И.* Фразеологический словарь русского языка. 5–11 классы: 1000 фразеологизмов: значение, употребление, история происхождения. М.: АСТ-Пресс, 2010. 412 с. (Серия «Настольные словари школьника»). Программа «Словари XXI века»). *Большой фразеологический словарь русского языка: значение, употребление, культурологический комментарий; отв. ред. В.Н. Телия.* М.: Ридерз Дайджест, 2012. 781 с. (Фундаментальные словари). *Жуков А.В., Жукова М.Е.* Словарь современной русской фразеологии: свыше 2100 фразеологических единиц, значение, происхождение, употребление, примеры из современной прессы; Ин-т русского яз. им. В.В. Виноградова РАН. М.: АСТ-ПРЕСС, 2015. 411 с. (Настоящие словари русского языка). *Зимин В.И.* Словарь-тезаурус русских пословиц, поговорок и метких выражений: более 22000 пословиц, поговорок, молвушек, присловий, приговорков, присказок, загадок, примет, дразнилок, считалок. М.: АСТ-Пресс, 2016. 729 с. (Программа «Словари XXI века»).

висимости от грамматически опорного компонента соответственно алфавитному расположению заголовочного слова.

Рассмотрение потенциала фразеологических зон БУС в перспективе формирования словника лингвистической базы данных для проекта электронного фразеологического словаря (АСПОТ) обусловлено функциональными способностями заголовочного слова в структуре ФЕ, «отображение которых в печатных словарях вызывает много вопросов и споров» [11. С. 47]. Возможность онлайн словарей включать большое количество лексических моделей и давать расширенный комментарий в словарной статье [11; 12] определяет перспективу объединения вокруг заголовочного слова как ключевого всех более или менее устойчивых выражений — от речевых формул до паремий и крылатых выражений. Изучение репрезентативности фразеологических зон БУС — авторитетного фразеологического источника — направлено на создание базы данных электронного фразеологического словаря нового типа. Значимость моделирования такого словаря определяется инновационной поисковой системой, позволяющей посредством гиперссылок без ограничений пространства и линейной организации обрабатывать слова и выражения в списке заголовочных слов, чтобы пользователь мог найти ФЕ, соответствующую целевому контексту [13]. Также отмечается проблема недостаточного культурного контекста во фразеологических словарях, что определяет актуальность количественного и качественного расширения смысловой реализации ФЕ [14]. Проект электронного фразеологического словаря предлагает модель представления широкому кругу пользователей примеров презентации как узальных, так и модифицированных ФЕ в когнитивно-смысловом пространстве текстов разных жанров и стилей.

Наряду с Oxford Advanced Learner's Dictionary (OALD), который на протяжении многих лет практически безраздельно доминирует в продаже словарей ESL [15], универсальность «Большого универсального словаря русского языка» как одноязычного толкового словаря определяет его широкую предназначенность «от лиц, прекрасно знающих, преподающих и профессионально описывающих русский язык, до лиц изучающих его как родной или иностранный» [8. С. 3]. Этот фактор, а также фразеологический массив и особенность его подачи пользователям Словаря актуализировали наш интерес к фразеологическим зонам БУС как объекту исследования в направлении выявления их употребительности (частотности) в структуре словарных статей в аспекте создания словника электронного фразеологического словаря.

Цель исследования — определить показатели репрезентативности фразеологических зон в структуре «Большого универсального словаря русского языка» и представить фразеостатистические модели частотного соотношения заголовочных слов как структурно-семантического компонента ФЕ в алфавитном ряду.

В представленном исследовании фразеологические зоны в структуре словарных статей БУС рассматриваются как часть своеобразного лингвистического корпуса русского языка, многоаспектно репрезентирующего свод ФЕ, что аргументируется коммуникативной задачей антропоцентрической лексикографии, реализуемой инновационными направлениями современной фразеологии.

Факторы выбора БУС для создания исследовательского корпуса

В электронных словарях, как и в традиционных (в бумажном варианте), словник является обязательным компонентом, отражающим языковую реальность и определяющим сферу охвата словарного материала. Проектирование электронного словаря смысловой реализации ФЕ как продукта компьютерной лингвистики невозможно без обращения к современным корпусным технологиям. Однако нельзя отрицать значимость исследовательского корпуса, который создается «не *post factum* — после проведения какого-либо исследования, а до его проведения», — корпуса данных [16. С. 27] для изучения разных функциональных векторов ФЕ и реализации широкого спектра фразеологических задач.

Основным фактором выбора БУС для создания исследовательского корпуса и в результате отбора ФЕ для формирования словника электронного фразеологического словаря стал принцип отбора слов с учетом их коммуникативной актуальности — ценностной характеристики заголовочного слова, «которая определяется, с одной стороны, его смысловой важностью и употребительностью в рамках заданного перечня наиболее важных словарных тем и ситуаций (идеографический критерий), а с другой — его соответствием наиболее существенным интенциям (функциональный критерий)» [8. С. 8]. Для формирования лингвистической базы данных и структуризации материала пилотного корпуса смысловой реализации ФЕ важно «применение принципов и алгоритмов составления лексикографических ресурсов с использованием корпусного подхода» [17. С. 61]. На первом этапе проектирования электронного фразеологического словаря корпусный подход позволит получить словник, соответствующий критерию коммуникативной актуальности, и «сопоставить корпусу опись данных, в которой единицы хранения характеризуются по тем параметрам, которые могут оказаться важными для пользователя» [16. С. 28].

Выделение исследователями корпусной лингвистики как методологического вектора лингвистической парадигмы актуализирует применение математических методов для получения статистических данных о языковых и речевых единицах, позволяющих «проследить изменение частот и контекстов в различные периоды времени, получить данные о сочетаемостных свойствах и особенностях лексических единиц» [18. С. 25]. Анализ трудов, посвященных корпусным базам данных в направлении когнитивных, лингвистических

и педагогических исследований [19], направленных на обоснование метода корпусного анализа для изучения семантической интерпретации образных единиц в контекстах разных языков [9; 10; 18–20], позволяет аргументировать выбор БУС, где представлены 3 тысячи ФЕ², структурными компонентами которых выступают заголовочные слова, для формирования исследовательского корпуса проекта электронного фразеологического словаря.

Технология исследования фразеологических зон БУС как репрезентативного источника данных

Всеохватность «Большого универсального словаря русского языка» — это, несомненно, результат труда лексикографов с текстовыми единицами лингвистического корпуса / лингвистических корпусов. Репрезентативность в представлении реального функционирования номинативных и коммуникативных единиц стала основным критерием при выделении наиболее употребительных слов в качестве заголовочных единиц словарных статей БУС. Исследование употребительности (частотности) фразеологических зон в словарных статьях связано с концептуальным подходом к БУС в качестве лексикографического корпуса, который «рассматривается как некоторый объект, служащий моделью некоторой внешней по отношению к нему реальности», и аргументацией, что «именно репрезентативность корпуса определяет достоверность полученных на его материале результатов» [21. С. 163]. Сложность формирования информационной базы исследования — использование традиционных форм для фиксирования фразеологических зон в корпусе «Большого универсального словаря русского языка», представленного только в бумажном формате.

Для определения объема исследовательского корпуса и выявления критерия представительной употребительности (частотности) «применение статистического подсчета относительных показателей различных лингвистических единиц является отправной точкой в создании естественно-научного фундамента» [22. С. 774]. Рассмотрение фразеологических зон БУС в качестве репрезентативного источника данных направлено на детальную проработку их количественной характеристики и процентного соотношения параметров статистической выборки. Количественные параметры включения фразеологических зон в структуру словарных статей определяют объективность моделирования исследовательского корпуса и, следовательно, формирования словника пилотного корпуса электронного словаря, выделения его ядерных и периферийных пространств.

² Ср.: в *Большом фразеологическом словаре русского языка: значение, употребление, культурологический комментарий* (отв. ред. В.Н. Телия. М.: Ридерз Дайджест, 2012. 781 с.) представлено толкование 1 500 ФЕ.

Совокупную репрезентативность заголовочных слов как компонентов ФЕ представляет генеральная лингвистическая совокупность (ГЛС). В исследовании под ГЛС выступает общее количество фразеологических зон, выбранных из всех словарных статей методом сплошной выборки.

Выборочная лингвистическая совокупность (ВЛС) — это «множество» фразеологических зон словарных статей, где заголовочные единицы выступают структурными компонентами ФЕ. ВЛС является в рамках данного исследования рабочим корпусом, по первичным измерениям которого формируются фразеостатистические модели, которые сгруппированы в соответствии с областью определения — совокупностью всех заголовочных слов, зафиксированной буквой русского алфавита.

Лингвостатистические аспекты изучения репрезентативности фразеологических зон определили следующие этапы фразеостатистического моделирования:

- Сбор для частного исследования из корпуса «Большого универсального словаря русского языка» статистических данных о фразеологических зонах в качестве ГЛС.
- Фиксирование фразеологических зон (ВЛС) в соответствии с областью определения, представляющую совокупную употребительность соотношения заголовочных слов как структурно-семантического компонента ФЕ в алфавитном ряду.
- Выделение ядерных и периферийных зон употребительности фразеологических зон в алфавитном ряду и распределение по группам, согласно полученным статистическим данным.
- Представление конечности области определения каждой группы таблицей / диаграммой соответствия в алфавитном ряду.
- Определение «букв-лидеров» и «букв-аутсайдеров» по суммарному количеству заголовочных слов как репрезентантов фразеологических зон.
- Представление фразеостатистической модели каждой группы круговой диаграммой, где дана оценка репрезентативности выборки в процентном выражении.

Определение показателей употребительности (частотности) фразеологических зон по алфавитному параметру поможет выявить масштабность репрезентативности ФЕ — основы исследовательского корпуса для пилотного проекта электронного фразеологического словаря.

Материалы и методы исследования

«Большой универсальный словарь русского языка» (В.В. Морковкин, Г.Ф. Богачёва, Н.М. Луцкая, 2016) не имеет электронного варианта, поэтому процесс по отбору материала проходил с корпусом толкового словаря в бумажном формате методом сплошной выборки. Результатом сплошной выборки

ки из словарных статей БУС стали 2545 фразеологических зон, выделенных как эмпирический материал исследования, которые были сгруппированы в соответствии с буквами русского алфавита как объекты выборочной лингвистической совокупности. В рамках корпусного подхода к исследованию репрезентативности фразеологических зон использовалась комбинация лингвостатистического метода, выборочного метода в эмпирическом исследовании, методов статистического моделирования.

Результаты исследования

Векторная проекция посредством использования «многолучевых» ссылок на то множество словарных статей, сколько компонентов имеет ФЕ, указывает на более широкий выбор фразеологических зон и, соответственно, большее количество ФЕ, чем отмечено авторами Словаря. Например, фразеологическая зона в структуре словарной статьи «КРОВЬ» содержит 40 ФЕ. Из них содержание 29 ФЕ рассматривается в данной словарной статье, ср. **Голубая кровь** *часто ирон.* <...>. **Искусственная кровь** <...>. **Кровь за кровь** <...>. **Кровь и плоть** *чья-л.; кровь от крови* *кого-л., чьей-л.* *высок.* <...>. [**Хоть**] **кровь из носу** / **из нóсу** / **из нóса** *разг.* <...>. **Кровь с молоком** <...>. **Зов крови** *высок.* <...>. **Жажда крови** *книжн., часто ирон.* <...>. **Лить (проливать) [свою] кровь** *за кого-что-л.* *высок.* <...>. **Лить (проливать) кровь** *чью-л.* <...>. **Обагрять руки кровью (в крови)** *книжн.* <...>. **Разгонять** *обычно несов.* **кровь** <...>. **В кровь; до крови** <...>. **Малой кровью** <...>. **Налиться кровью** <...>. **Смыть кровью** <...>. **С кровью** <...>. **В крови** *что-л. у кого-л.* <...>. **Купаться (топить кого-что-л.) в крови** <...>. **Руки по локоть в крови** *у кого-л.* <...>. **На крови** <...>. **Кровь чья-л.** *на ком-л.* <...>. **Кровь бросилась (ударилась, кинулась) в голову кому-л.** <...>. **Кровь бросилась в лицо** <...>. **Кровь говорит** *в ком-л.* <...>. **Кровь играет** *зд. несов.* <...>. **Кровь кипит** *несов. (бродит несов.)* *в ком-л.* <...>. **Льётся** *несов.* **кровь** <...>. **Кровь стынет (леденеет) [в жилах]** <...>.

Семантизация 11 ФЕ происходит по ключевому компоненту в других словарных статьях, ср. **Голос крови** — *см. голос.* **Жажда крови** — *см. жажда.* **До последней капли крови** — *см. капля.* **Пить кровь** *чью-л., кого-л.* — *см. пить.* **Портить себе кровь** — *см. портить.* **Войти в плоть и кровь** — *см. входить.* **Потом и кровью** — *см. пот.* **Кровью заплатит за кого-что-л.** — *см. платить.* **Кровью пахнет** — *см. пахнуть.* **Сердце кровью обливается** *у кого-л.* — *см. сердце.* **Кровь загорелась** — *см. загореться.*

Статистическая мера репрезентативности определяется параметрами выборки, определяющими объем рабочего корпуса ВЛС, соответственно параметрам ГЛС, количественный объем которых зависит от структурной многоплановости заголовочного слова как компонента ФЕ. Результат отно-

Ядро составляют фразеологические зоны, индекс которых представлен более чем в 150 словарных статьях. Общее количество фразеологических зон **группы-I** — 1193. Репрезентантами группы стали заголовочные слова, соответствующие следующим буквам русского алфавита: **В — К — Н — П — С**, которые определены как буквы-лидеры. Количественное соотношение частотной репрезентативности фразеологических зон **группы-I** представлено на рис. 2. Расположение соответствует заголовочным словам в алфавитном порядке.

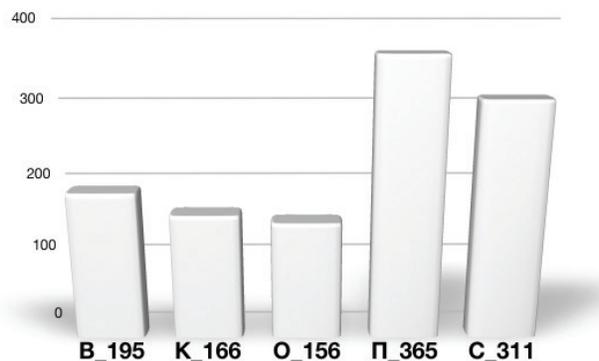


Рис. 2. Представление конечности области определения **группы-I** в алфавитном ряду
Figure 2. Representation of the finiteness of the **group-I** domain of definition in the alphabetical order

Фразеостатистическая модель каждой выделенной группы представлена на круговой диаграмме, где зафиксировано суммарное процентное соотношение (λ_i) статистической выборки фразеологических зон (Index) от общего числа всех фразеологических зон (PhU). Оценка репрезентативности выборки в процентном выражении показана расположением букв соответственно λ_i — от большего выражения к меньшему выражению по часовой стрелке.

Общая λ **группы-I** составляет 46,87 %, что иллюстрируется фразеостатистической моделью на рис. 3.

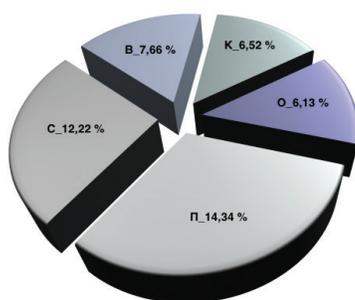


Рис. 3. Фразеостатистическая модель процентного соотношения **группы-I**
Figure 3. Phrase-statistical model of the percentage ration of the **group-I**

Количественное соотношение фразеологических зон **группы-II** представлено индексом от 100 до 150 словарных статей. Общее количество фразеологических зон **группы-II** — 727. Репрезентантами группы стали заголовочные слова, соответствующие следующим буквам русского алфавита: **Д — З — М — Н — Р — Т**. Количественное соотношение частотной репрезентативности фразеологических зон **группы-II** представлено на рис. 4. Расположение соответствует заголовочным словам в алфавитном порядке.

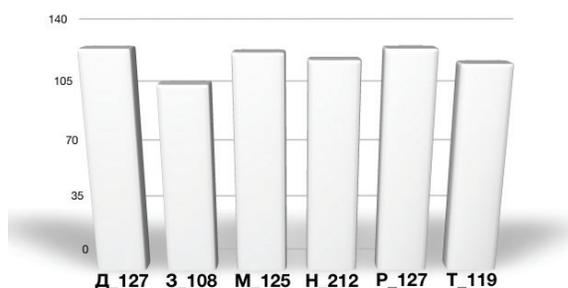


Рис. 4. Представление конечности области определения **группы-II** в алфавитном ряду
Figure 4. Representation of the finiteness of the **group-II** domain of definition in the alphabetical order

Общая λ **группы-II** составляет 28,55% что иллюстрируется фразеостатистической моделью на рис. 5.

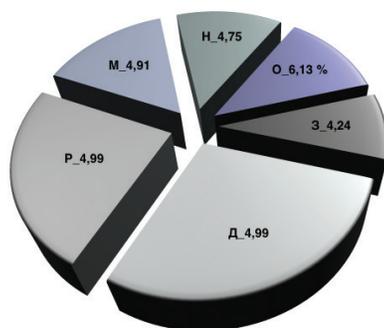


Рис. 5. Фразеостатистическая модель процентного соотношения **группы-II**
Figure 5. Phrase-statistical model of the percentage ration of the **group-II**

Количественное соотношение фразеологических зон **группы-III** представлено индексом от 50 до 100 словарных статей. Общее количество фразеологических зон **группы-III** — 363. Репрезентантами группы ста-

ли заголовочные слова, соответствующие следующим буквам русского алфавита: **Б — Г — И — Л — У**. Количественное соотношение частотной репрезентативности фразеологических зон **группы-III** представлено на рис. 6. Расположение соответствует заголовочным словам в алфавитном порядке.

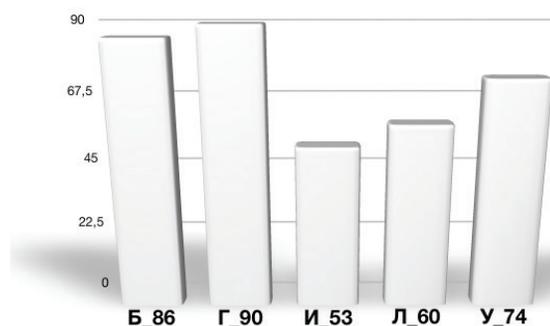


Рис. 6. Представление конечности области определения **группы-III** в алфавитном ряду
Figure 6. Representation of the finiteness of the **group-III** domain of definition in the alphabetical order

Общая λ **группы-III** составляет 14,25 %, что иллюстрируется фразеостатистической моделью на рис. 7.

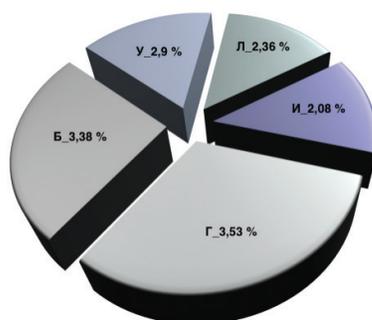


Рис. 7. Фразеостатистическая модель процентного соотношения **группы-III**
Figure 7. Phrase-statistical model of the percentage ratio of the **group-III**

Количественное соотношение фразеологических зон **группы-IV** представлено индексом от 20 до 50 словарных статей. Общее количество фразеологических зон **группы-IV** — 176. Репрезентантами группы стали заголовочные слова, соответствующие следующим буквам русского алфавита: **А — Ж — Х — Ч — Ш**. Количественное соотношение частотной репрезентативности фразеологических зон **группы-IV** представлено на рис. 8. Расположение соответствует заголовочным словам в алфавитном порядке.

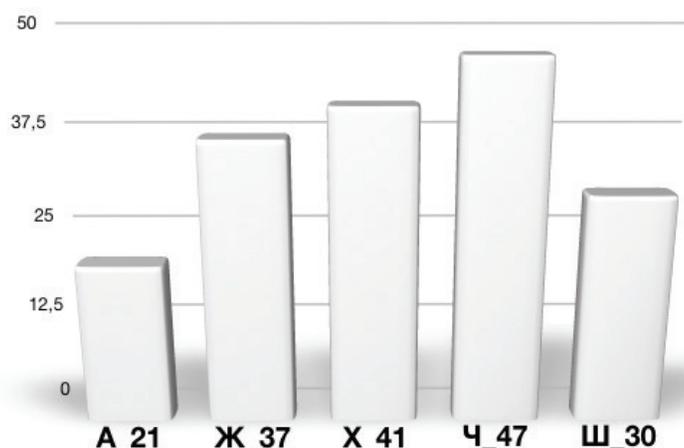


Рис. 8. Представление конечности области определения **группы-IV** в алфавитном ряду
Figure 8. Representation of the finiteness of the **group-IV** domain of definition in the alphabetical order

Общая λ **группы-IV** составляет 6,92 %, что иллюстрируется фразеостатистической моделью на рис. 9.

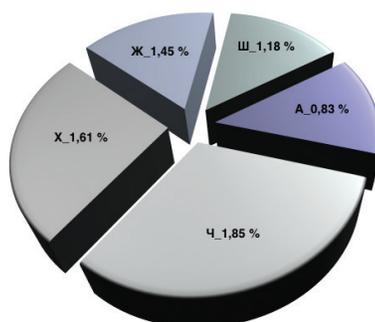


Рис. 9. Фразеостатистическая модель процентного соотношения **группы-IV**
Figure 9. Phrase-statistical model of the percentage ratio of the **group-IV**

На периферии находятся фразеологические зоны **группы-V**, индекс которых менее 20 словарных статей. Общее количество фразеологических зон **группы-V** — 86. Репрезентантами группы стали заголовочные слова, соответствующие следующим буквам русского алфавита: **Е, Ё — Ф — Ц — Щ — Э — Ю — Я**, которые определены как буквы-аутсайдеры. Количественное соотношение частотной репрезентативности фразеологических зон **группы-V** представлено на рис. 10. Расположение соответствует заголовочным словам в алфавитном порядке.

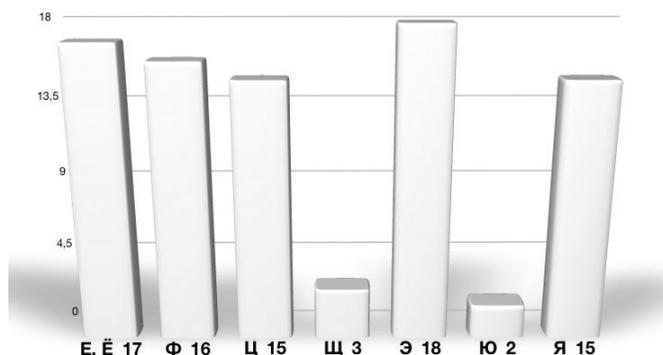


Рис. 10. Представление конечности области определения **группы-V** в алфавитном ряду
Figure 10. Representation of the finiteness of the **group-V** domain of definition in the alphabetical order

Общая λ **группы-V** составляет 3,38 %, что иллюстрируется фразеостатистической моделью на рис. 12.

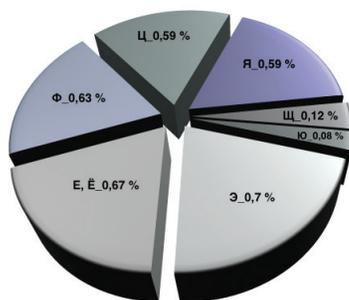


Рис. 11. Фразеостатистическая модель процентного соотношения В
Figure 11. Phrase-statistical model of the percentage ratio of the **group-V**

Словарных статей с заголовочными словами на букву «Ю» всего 7 [8. С. 1435–1436]. Фразеологические зоны представлены в структуре только 2-х словарных статей: «ЮМОР» Юмор висельника *ирон.* — шутки, острооты человека, к-рый находится в безвыходном положении. **Не понял юмора** *разг.* — употр. для выражения неодобрения по поводу чьих-л. действий или высказываний [8. С. 1436]. Более репрезентативной является фразеологическая зона словарной статьи «ЮБКА» Нижняя юбка — предмет женского белья, юбка, которая надевается под платье или обычную юбку, с тем чтобы одежда лучше сидела. Шотландская юбка — юбка в клетку, которая является частью традиционного костюма шотландских мужчин, килт. **Ни одной юбки не пропустит** кто-л. — *см. пропустить.* В юбке кто-л. — а) *шутл.* о женщине, которая проявляет качества, обладает характером какого-л. известного лица, персонажа мужского пола <...>; б) *шутл. и ирон.* о женщине, которая проявляет какие-л. качества, черты, свойственные, согласно общепринятым стереотипам, представителям какой-л., главным образом мужской, профессии <...>;

в) *шутл.* о женщине, профессия, занятие которой, по мнению большинства, является скорее мужской, чем женской <...>. **Держаться за чью-л. юбку; быть при юбке** чьей-л. — быть в полном подчинении у женщины (обычно у матери, жены), проявлять полное отсутствие самостоятельности, отдав право действовать и принимать решения женщине (обычно близкой). **Лезть под юбку** кому-л. *разг., неодобр.* — осуществлять (по отношению к женщине) какие-л. действия, имеющие характер сексуального домогательства [8. С. 1435].

Словарных статей с заголовочными словами на букву «Щ» — всего 5 [8. С. 1420–1421] и общим объемом соответствуют одной странице Словаря. Фразеологические зоны есть в структуре 3-х словарных статей, и наиболее репрезентативной является фразеологическая зона словарной статьи «**ЩИ**» **Кислые щи** — щи из квашеной капусты. **Зелёные щи** — щи из щавеля, шпината и т.п. **Суточные щи** — щи, которые едят на следующий день после из приготовления. **Пустые щи** — щи без мяса, жиров. **Щи да каша** — *пища наша погов.* — говорится в шутку, когда подают к столу или едят эти два блюда или одно из них. **Лаптем щи хлебать** *несов., разг.* — а) жить в нищете, отсталости, невежестве; б) быть совершенно необразованным, некультурным. **Тех же щей, да пожиге влей** *разг.* — о слегка видоизменённом повторении одного и того же. **Попасть (угодить ...)** *как кур во щи разг.* — глупо, неожиданно и против воли оказаться в неприятном положении. **Профессор кислых щей** — *см. профессор* [8. С. 1421].

Многозначная ФЕ **профессор кислых щей** *прост.* — а) о человеке, который изображает из себя знатока, а на самом деле плохо разбирается в чём-л. <...>; б) *пренебр.* о профессоре, который, по мнению говорящего, неизвестно чем занимается или занимается всякой ерундой <...> [8. С. 949] — является единственной косвенно-производной единицей, в структуре которой употребляется заголовочное слово **ПРОФЕССОР**. Такое радиальное перемещение ФЕ из одной словарной статьи в другую, мотивированное участием заголовочного слова как ключевого компонента структурирования ФЕ и фиксирующее единичную репрезентацию ФЕ во фразеологической зоне, перспективно как база данных в аспекте изучения фразеологических структурных моделей.

Фразеостатистические модели процентного соотношения отдельной группы определяют объективность формирования исследовательского корпуса в перспективе моделирования словника пилотного корпуса электронного словаря смысловой реализации фразеологических единиц, выделения его ядерных и периферийных пространств.

Заключение

Концептуальный подход к «Большому универсальному словарю русского языка» — толковому словарю в бумажном формате — как лексикографическому корпусу в исследовании употребительности (частотности)

фразеологических зон в словарных статьях определил направление исследования в аспекте формирования исследовательского корпуса для моделирования словника пилотного корпуса электронного словаря (АСПОТ) смысловой реализации фразеологических единиц. Основным фактором выбора БУС для создания исследовательского корпуса стал принцип отбора слов с учетом их коммуникативной актуальности — ценностной характеристики заголовочного слова как структурно-семантического компонента фразеологических единиц.

Параметрами технологии исследования фразеологических зон БУС как репрезентативного источника данных определены (1) генеральная лингвистическая совокупность — общее количество фразеологических зон, выбранных из всех словарных статей; (2) выборочная лингвистическая совокупность как рабочий корпус, по первичным измерениям которого сформированы фразеостатистические модели, сгруппированные в соответствии с областью определения — совокупностью всех заголовочных слов, зафиксированной одной буквой русского алфавита.

Выделенные группы фразеостатистических моделей представляют количественное соотношение частотной репрезентативности фразеологических зон в алфавитном ряду и оценку репрезентативности выборки в процентном выражении. Детальная проработка количественной характеристики и процентного соотношения параметров статистической выборки фразеологических зон БУС определили объективность моделирования исследовательского корпуса и, следовательно, формирования словника пилотного корпуса электронного фразеологического словаря, выделения его ядерных и периферийных пространств.

Библиографический список

1. *Дубичинский В.В.* Теоретическая и практическая лексикография. Вена, Харьков, 1998.
2. *Сидоренко К.П.* Русское лексикографическое пространство в алфавитном порядке. Рецензия на книгу: Козырев В.А., Черняк В.Д. Кто есть кто в русской лексикографии: справочная книга // Вопросы лексикографии. 2017. № 11. С. 143–148. <https://doi.org/10.17223/22274200/11/10>
3. *Морковкин В.В.* Антропоцентрический versus лингвоцентрический подход к лексикографированию // Национальная специфика языка и её отражение в нормативном словаре : сб. ст. / АН СССР, Науч. совет по лексикологии и лексикографии, Ин-т рус. яз.; отв. ред. Ю.Н. Караулов. М.: Наука, 1988. С. 131–136.
4. *Богачёва Г.Ф., Морковкин В.В.* О всеохватном лексикографическом представлении лексического ядра русского языка // Вестник Томского государственного университета. Филология. 2011. № 3(15). С. 6–11.
5. *Козырев В.А., Черняк В.Д.* Кто есть кто в русской лексикографии: справочная книга. СПб.: Своё издательство, 2016.
6. *Ольховская А.И.* Рецензия на книгу: Морковкин В.В., Богачёва Г.Ф., Луцкая Н.М. Большой универсальный словарь русского языка // Вопросы лексикографии. 2016. № 2(10). С. 110–121. <https://doi.org/10.17223/22274200/10/7>

7. Карнова О.М. Фундаментальный лексикографический труд (Рец. на кн.: Морковкин В.В. Большой универсальный словарь русского языка / В.В. Морковкин, Г.Ф. Богачёва, Н.М. Луцкая; под ред. проф. В.В. Морковкина. М.: Словари XXI века, 2016. 1456 с.) // Вестник Воронежского государственного университета. Серия: Лингвистика и межкультурная коммуникация. 2017. № 4. С. 178–180.
8. Морковкин В.В., Богачёва Г.Ф., Луцкая Н.М. Большой универсальный словарь русского языка: Государственный институт русского языка им. А.С. Пушкина; под ред. В.В. Морковкина. М.: АСТ-ПРЕСС КНИГА; Словари XXI века, 2016.
9. Изотова А.А. Идиоматическая фразеология и корпус // Вестник университета. 2014. № 21. С. 258–260.
10. Пивоварова Е.В. Корпусный анализ как инструмент для выявления семантических изменений у фразеологизмов — «ложных друзей» переводчика // Филологические науки. Вопросы теории и практики. 2020. № 13(9). С. 312–320. <https://doi.org/10.30853/filnauki.2020.9.57>
11. Фесенко О.П., Лаухина С.С. Электронные словари как продукт современной лексикографии // Омский научный вестник. 2015. № 4(141). С. 46–48.
12. Gouws R.H. Extended article comments in online dictionaries // *Lexicography*. 2022. № 9(2). P. 111–138. <https://doi.org/10.1558/Lexi.24289>
13. McAlpine J., Myles J. Capturing phraseology in an online dictionary for advanced users of English as a second language a response to user needs // *System*. 2003. № 31(1). P. 71–84. [https://doi.org/10.1016/S0346-251X\(02\)00074-X](https://doi.org/10.1016/S0346-251X(02)00074-X)
14. Szerszunowicz J. Exemplification in a phraseological dictionary an extended illustration model // *Lexicography*. 2018. № 5(1). P. 35–54. <https://doi.org/10.1007/s40607-018-0045-y>
15. MacFarquhar P.D., Richards J.C. On Dictionaries and Definitions // *RELC Journal*. 1983. № 14(1). P. 111–124. <https://doi.org/10.1177/003368828301400112>
16. Баранов А.Н. Введение в прикладную лингвистику. М.: Эдиториал УРСС, 2001.
17. Девяткова О.И., Кантышева Н.Г. Проектирование терминологического учебного словаря таможенной сферы корпусный подход // Вопросы лексикографии. 2022. № 25. С. 55–80. <https://doi.org/10.17223/22274200/25/3>
18. Долгих З.Б. Обзор ряда корпусных возможностей в сфере лингвистических исследований (на примере анализа средств градуирования в португальском языке) // Вестник Московского государственного лингвистического университета. Гуманитарные науки. 2018. Вып. 5(795). С. 21–32.
19. Павловская И.Ю., Горина О.Г. Корпусно-когнитивные методы изучения лексической стороны речи на иностранном языке // Вестник Череповецкого государственного университета. 2017. № 1. С. 132–138. <https://doi.org/10.23859/1994-0637-2017-1-76-19>
20. Павельева Т.Ю. Изучение коллокаций на основе лингвистических корпусов текстов // Вестник Тамбовского университета. Серия Гуманитарные науки. 2016. Т. 21. № 3–4 (155–156). С. 56–61. [https://doi.org/10.20310/1810-0201-2016-21-3/4\(155/156\)-56-61](https://doi.org/10.20310/1810-0201-2016-21-3/4(155/156)-56-61)
21. Богоявленская Ю.В. Репрезентативность лингвистического корпуса метод верификации достоверности полученных данных // Политическая лингвистика. 2016. № 4(58). С. 163–166.
22. Гречанин В.А. Статистические методы в исследовании текстов // Вестник Башкирского университета. 2018. № 23(3). С. 772–775.

References

1. Dubichinsky, V.V. (1998). *Theoretical and practical lexicography*. Vienna, Kharkov. (In Russ.).
2. Sidorenko, K.P. (2017). Russian lexicographic space in alphabetical order. Book review Kozyrev, V.A. & Chernjak, V.D. Who is who in Russian lexicography: a reference book. *Russian Journal of Lexicography*, 11, 143–148. <https://doi.org/10.17223/22274200/11/10> (In Russ.).

3. Morkovkin, V.V. (1988). Anthropocentric versus linguocentric approach to lexicography. In *National specificity of language and its reflection in the normative dictionary*. Moscow: Nauka. pp. 131–136. (In Russ.).
4. Bogacheva, G.F. & Morkovkin, V.V. (2011). On all-embracing lexicographic presentation of the lexical core of the Russian language. *Tomsk State University Journal of Philology*, 3(15), 6–11. (In Russ.).
5. Kozyrev, V.A. & Chernjak, V.D. (2016). *Who is who in Russian lexicography a reference book*. Saint Petersburg: Svojo publ. (In Russ.).
6. Olkhovskaya, A.I. (2016). Book review Morkovkin, V.V., Bogacheva, G.F. & Lutsкая, N.M. The big universal dictionary of the Russian language. *Russian Journal of Lexicography*, 2(10), 110–121. <https://doi.org/10.17223/22274200/10/7> (In Russ.).
7. Karpova, O.M. (2017). Book review Morkovkin, V.V., Bogacheva, G.F. & Lutsкая, N.M. The academic dictionary of Russian phraseology. *Voronezh State University. Linguistics and Intercultural Communication*, 4, 178–180. (In Russ.).
8. Morkovkin, V.V., Bogacheva, G.F. & Lutsкая, N.M. (2016). *The Great Universal Dictionary of the Russian Language*. Moscow. (In Russ.).
9. Izotova, A.A. (2014). Idiomatic phraseology and corpus. *Vestnik Universiteta*, 21, 258–260. (In Russ.).
10. Pivovarova, E.V. (2020). Corpus Analysis as a Tool for Identifying Semantic Changes of Phraseological Units — Translator’s “False Friends”. *Philology. Theory & Practice*, 13(9), 312–320. <https://doi.org/10.30853/filnauki.2020.9.57> (In Russ.).
11. Fesenko, O.P. & Laukhina, S.S. (2015). Electronic dictionaries as a product of modern lexicography. *Omsk Scientific Bulletin*, 4(141), 46–48. (In Russ.).
12. Gouws, R.H. (2022). Extended article comments in online dictionaries. *Lexicography*, 9(2), 111–138. <https://doi.org/10.1558/lexi.24289>
13. McAlpine, J. & Myles, J. (2003). Capturing phraseology in an online dictionary for advanced users of English as a second language a response to user needs. *System*, 31(1), 71–84. [https://doi.org/10.1016/S0346-251X\(02\)00074-X](https://doi.org/10.1016/S0346-251X(02)00074-X)
14. Szerszunowicz, J. (2018). Exemplification in a phraseological dictionary an extended illustration model. *Lexicography*, 5(1), 35–54. <https://doi.org/10.1007/s40607-018-0045-y>
15. MacFarquhar, P.D. & Richards, J.C. (1983). On Dictionaries and Definitions. *RELC Journal*, 14(1), 111–124. <https://doi.org/10.1177/003368828301400112>
16. Baranov, A.N. (2001). *Applied Linguistics*. Moscow: URSS publ. (In Russ.).
17. Devyatkova, O.I. & Kantysheva, N.G. (2022). Design of a terminological learner’s dictionary in the customs sphere A corpus approach. *Russian Journal of Lexicography*, 25, 55–80. <https://doi.org/10.17223/22274200/25/3> (In Russ.).
18. Dolguikh, Z.B. (2018). An Overview of some basic possibilities of corpus in the field of linguistic research (on the example of the analysis of graduation means in the Portuguese language). *Vestnik of Moscow State Linguistic University. Humanities*, 5(795), 21–32. (In Russ.).
19. Pavlovskaya, I.Yu. & Gorina, O.G. (2017). Corpus-based cognitive oriented methods of lexical analysis in foreign language studies. *Cherapovets State University Bulletin*, 1, 74–80. <https://doi.org/10.23859/1994-0637-2017-1-76-19> (In Russ.).
20. Paveleva, T.Yu. (2016). Collocations study based on the text corpora. *Tambov University Review. Series Humanities*, 21 (3–4 (155–156)), 56–61. [http://doi.org/10.20310/1810-0201-2016-21-3/4\(155/156\)-56-61](http://doi.org/10.20310/1810-0201-2016-21-3/4(155/156)-56-61) (In Russ.).
21. Bogoyavlenskaya, Yu.V. (2016). Representativeness of text corpus method of verification of data reliability. *Political Linguistics*, 4(58), 163–166. (In Russ.).
22. Grechachin, V.A. (2018). Statistical methods in the study of texts. *Bulletin of Bashkir University*, 23(3), 772–775. (In Russ.).

Сведения об авторах

Золотых Лидия Глебовна, доктор филологических наук, профессор, профессор факультета русского языка, Институт иностранных языков и культуры Сычуаньского университета; *сфера научных интересов*: лингвокультурология, когнитивная лингвистика, исследование фразеологической семантики; *e-mail*: scu1959rus@mail.ru
ORCID: 0000-0002-1425-5405; SPIN-код 7721–2060

Барышникова Елена Николаевна, кандидат филологических наук, доцент, доцент кафедры русского языка и методики его преподавания филологического факультета, Российский университет дружбы народов; *сфера научных интересов*: языковые контакты в лексикологии русского языка; лингвокультурологический анализ заимствованной лексики в русском языке; психолингвистика; билингвизм; методика обучения русскому языку как иностранному; лингвистические, лингвокультурологические и методические основы создания национально-ориентированных дидактических материалов по русскому языку как иностранному; *e-mail*: baryshnikova-en@rudn.ru
ORCID: 0000-0003-4228-1763; elibrary SPIN-код 9044–9815; ResearcherID AAM-8931-2020

Космачёва Ольга Юрьевна, кандидат филологических наук, доцент кафедры факультета филологии и журналистики, Астраханский государственный университет им. В.Н. Татищева; *сфера научных интересов*: исследование фразеологической семантики, лингвокультурология, когнитивная лингвистика; *e-mail*: olga_ast@list.ru
ORCID: 0000-0002-3722-3131; SPIN-код 5314-9461

Information about the authors:

Lidia G. Zolotykh, Dr.Sc. (Philology), Professor, Professor of the Russian Language Faculty, Institute of Foreign Languages and Culture of the Sichuan University; *Research interests*: linguoculturology, cognitive linguistics, study of phraseological semantics; *e-mail*: scu1959rus@mail.ru
ORCID: 0000-0002-1425-5405; SPIN code 7721–2060

Elena N. Baryshnikova, PhD in Philology, Associate Professor, Associate Professor of the Department of Russian Language and Methods of its Teaching at the Faculty of Philology, RUDN University, *Research interests*: Language Contacts in the lexico-semantic Russian language skills; linguocultural analysis of borrowed vocabulary in Russian; psycholinguistics; bilingualism; methods of teaching Russian as a foreign language; linguistic, linguocultural and methodical foundations of nation-making educational materials on Russian as a foreign language; *e-mail*: baryshnikova-en@rudn.ru
ORCID: 0000-0003-4228-1763; elibrary SPIN code 9044–9815; ResearcherID AAM-8931-2020

Olga Yu. Kosmacheva, PhD in Philology, Associate Professor of the Department of Philology and Journalism, Astrakhan State University n.a. V.N. Tatischev. *Research interests*: study of phraseological semantics, linguocultural science, cognitive linguistics; *e-mail*: olga_ast@list.ru
ORCID: 0000-0002-3722-3131; SPIN code 5314-9461