Вопросы лексикографии. 2023. № 28. С. 83–107
Russian Journal of Lexicography. 2023, 28, pp. 83–107

ЭЛЕКТРОННАЯ ЛЕКСИКОГРАФИЯ

ELECTRONIC LEXICOGRAPHY

Научная статья
УДК 81`373
doi: 10.17223/22274200/28/5

Лексикографические принципы создания и использования электронной фразеологической картотеки

Николай Фёдорович Алефиренко[1], Мария Михайловна Голикова[2]

[1,2] Белгородский государственный национальный исследовательский университет, Белгород, Россия
[1] alefirenko@bsu.edu.ru
[2] golikova@bsu.edu.ru

Аннотация. Детально рассматриваются принципы создания и использования фразеологической картотеки на базе платформы «1С: Предприятие». Не только показано теоретическое и практическое обоснование продуктивности подобного подхода, но и дана подробная инструкция по использованию готового приложения, а также продемонстрированы возможности его использования при изучении образных свойств фразеологизмов.

Ключевые слова: фраземика, фразеологическая единица, электронная картотека, 1С: Предприятие, язык эмигрантской прозы

Для цитирования: Алефиренко Н.Ф., Голикова М.М. Лексикографические принципы создания и использования электронной фразеологической картотеки // Вопросы лексикографии. 2023. № 28. С. 83–107. doi: 10.17223/22274200/28/5

© Н.Ф. Алефиренко, М.М. Голикова, 2023

Original article
doi: 10.17223/22274200/28/5

Lexicographic principles of creation and use of electronic phraseological card files

Nikolay F. Alefirenko[1], Mariya M. Golikova[2]

[1, 2] Belgorod National State University, Belgorod, Russian Federation
[1] alefirenko@bsu.edu.ru
[2] golikova@bsu.edu.ru

Abstract. The article considers the problem of creating an electronic phraseological card file that allows one to save and process a large amount of information, is easy to manage and access for untrained users. The relevance of the work is due to the fact that the automation of research processes will make it possible not only to simplify the work of a researcher, but also to quickly analyze large amounts of information. To do this, it is necessary to create an automated card file. The experience of working on the phraseological card file shows the effectiveness of its formation using the 1C: Enterprise platform. The article aims to present the principles of the development of an electronic system for storing and processing data extracted from literary texts and its implementation in the practical activity of a researcher-philologist. The work with the card file was carried out on the material of the works of Russian emigrant writers of the first, second, third and fourth, distinguished by individual researchers, waves. Such linguistic material has not been sufficiently studied, and its individual elements (for example, the phraseology of second-wave emigrants or the linguistic and cultural features of the phraseological image) are completely a lacuna in modern language science. The card file was formed by a continuous sampling method based on works of art of large, medium and small genres. The method of computer modeling was used to prepare the architecture of the card file, and linguistic statistics techniques were used to process the entered data. The article shows the ways to create an automated card file based on the 1C: Enterprise platform. Such work is complicated by the fact that this software is not originally intended for research activities, its resources make it possible to organize a system of interrelated elements that ensure the preservation of all fragments of the linguistic card and, most importantly, their subsequent operational statistical processing. This approach determines the scientific novelty of the work because it significantly expands the range of electronic tools applicable for linguistic research in general and lexicographic research in particular. The created card file is focused on the study of the figurative component of the phrasemics of writers of the Russian diaspora; however, the proposed approach makes it possible to apply this experience to other studies of not only phraseological, but also other linguistic units.

Keywords: phrasemics, phraseological unit, electronic card file, 1C: Enterprise, language of emigrant prose

For citation: Alefirenko, N.F. & Golikova, M.M. (2023) Lexicographic principles of creation and use of electronic phraseological card files. *Voprosy leksikografii – Russian Journal of Lexicography.* 28. pp. 83–107. (In Russian). doi: 10.17223/22274200/28/5

Введение

Любое лингвистическое исследование, как известно, основывается на анализе собранного эмпирического материала. Традиционно найденные в исследуемом материале языковые единицы фиксируются на специально созданных карточках, структура оформления которых является произвольной, но всегда содержит контекст употребления единицы и точную ссылку (выходные данные книги, страница – для художественных текстов; номер газеты, дата выпуска телепередачи – для текстов СМИ и т.п.). Создание лексической картотеки дает возможность оперативно найти расположение искомой языковой единицы в исследуемом речевом произведении и ее смыслообразующий контекст. Однако формирование подобной картотеки представляет собой непростую задачу, так как проводится исключительно вручную. Это, разумеется, не представляет особых сложностей для небольших исследований, однако серьезно усложняет обработку огромных пластов информации. В таких ситуациях более удобным и продуктивным представляется использование автоматизированных систем обработки и подсчета данных, работающих на основе электронной картотеки.

Создание картотеки является чрезвычайно важным при работе с фразеологическими единицами (ФЕ), поскольку для их глубинного дискурсивно-семантического исследования обязательным является выявление особенностей контекстуального употребления всех устойчивых словесных комплексов. Это дает возможность рассмотреть все грани фразеологической семантики, особенности сочетаемости ФЕ, а в отдельных случаях указывает и на определенные детали их историко-культурного ореола типа ФЕ *окопная правда*, появившаяся после Великой Отечественной войны: «– Мы их еще в школе проходили, это же “певцы окопной правды”» [1]. Зависимость ФЕ от контекста обусловливает и более развёрнутую структуру карточки. Необходимость фиксации точного расположения цитаты в тексте произведения

обусловливается не только каноническими предписаниями проведения лингвистического исследования, но и экономией времени ученого. При этом обработка такой картотеки очень трудоемка. Если классификация производится по какому-то одному критерию, то карточки можно систематизировать, например, складывая в отдельные стопки. Но что делать, если оснований для классификации несколько, а число карточек – более тысячи? В таком случае гораздо удобнее автоматизировать процесс систематизации и подсчета, формируя электронную картотеку.

Крупные базы данных позволяют хранить и систематизировать огромные пласты информации из разных областей знания. В отдельных научных областях также используется специализированное программное обеспечение, однако исследования в области гуманитарных дисциплин проводятся преимущественно классическими способами, практически игнорируя современные автоматизированные технологии. Во многом это обусловливается недостаточной доступностью программных решений для проведения исследований в области гуманитарных наук.

Впервые мысли о необходимости автоматизации работы с лексическим и фразеологическим материалом были высказаны академиком А.П. Ершовым еще в 1978 г. Ученый стоял у истоков Машинного фонда русского языка, ставшего впоследствии образцом для Национального корпуса русского языка. Фразеография стала важной частью фонда, но, безусловно, ее электронная реализация изначально была далека от идеальной. В 1993 г. на недостатки репрезентации в Машинном фонде русского языка фразеологического материала указывал также известный лексикограф П.Н. Денисов. Ученый отмечал, что, при всех своих преимуществах, обработка ФЕ базируется в основном на свойствах отдельных слов, автоматически перенося их на устойчивые словосочетания. В результате различие между словом и интегрированным по смыслу словосочетанием в основном устанавливается на грамматическом и коммуникативном уровне, оставляя без должного внимания их семантическое или стилистическое своеобразие [2. С. 146]. Здесь же учёный с сожалением отмечает, что «до создания лингвистически и практически ценных фразеологических словарей непосредственно на ЭВМ еще далеко» [2. С. 146].

Новый виток интереса к идеям цифровизации лексикографии и фразеографии возник в начале XXI в. В 2007 г. о необходимости автоматизации словарной картотеки Института лингвистических иссле-

дований РАН высказался В.П. Захаров [3]. Среди преимуществ электронной картотеки ученый выделил, например, упорядочивание накопленной информации, а также ее унификацию. Объектами автоматизации, по мнению В.П. Захарова, должны стать работа по созданию и пополнению уже готовых карточек, а также организация системы поиска их содержимого, оптимизация уже готовых карточек для новых лексикографических задач [3. С. 203]. В работе исследователя говорится о существовании сразу трех баз данных, содержащих непосредственно сами словники, цитаты, а также библиографическую информацию. Подводя итог своим размышлениям, В.П. Захаров указывает на необходимость использования ресурсов, позволяющих задействовать средства лингвостатистической обработки информации, отмечая, что «необходимо дать пользователю программно-лингвистическую систему, функциональные возможности которой позволят эффективно оперировать данными корпуса в сочетании с картотеками и словарями» [3. С. 206]. Все это может способствовать не только работе по составлению словарей, но и проведению разнообразных исследований.

На сегодняшний день компьютерная лексикография («совокупность программных средств и методов, направленных на разработку словарей электронного вида» [4]) активно развивается в рамках компьютерной лингвистики. Современные технологии направлены не только на хранение информации, но и на отладку ее автоматического сбора, обработки и последующей трансформации в словарный вид. Так, например, картотека «Лексического атласа русских народных говоров», формируемая специалистами Института лингвистических исследований РАН, тоже представлена в электронном формате и «состоит из четырёх частей: добавление файла, экспедиции, ответы (собственно диалектные материалы) и техническое обслуживание» [5. С. 167]. Доступ к такой картотеке может осуществляться сразу с нескольких компьютеров. Более глубокий подход к разработке принципов фразеографирования, их апробация должны обеспечить системную работу над совершенствованием Национального корпуса русского языка. Однако столь многопрофильное программное обеспечение не всегда доступно для локальных исследований, проводимых отдельными учеными или малыми группами. Прежде всего, проблематичной оказывается работа с программным обеспечением, которое требует серьезной специальной подготовки. Она может стать невоз-

можной для исследователя, чья сфера научных интересов напрямую не связана с компьютерной лингвистикой и/или составлением академических словарей. Электронная картотека, конечно, может формироваться с помощью крупных систем управления базами данных (например, Microsoft Office Access). Однако работа с ней тоже требует предварительной подготовки, а сам формат представления данных не всегда применим в частных лингвистических исследованиях.

Материалы и методы

Основным материалом нашей работы выступают ФЕ, собранные методом сплошной выборки из произведений русских писателей-эмигрантов. Исследовались художественные тексты эмигрантов разных волн, что не только дало возможность выявить влияние времени и места эмиграции на образную составляющую ФЕ, но и, безусловно, сказалось на объеме и структуре картотеки. Нами были изучены произведения разных жанров, написанные И.А. Буниным, В.В. Набоковым, А.И. Куприным и другими (1-я волна), Л.Д. Ржевским, В.И. Юрасовым, Г.А. Андреевым (2-я волна), В.П. Аксеновым, В.П. Некрасовым (3-я волна), Д.И. Рубиной, А.Н. Рыбаковым, Б. Акуниным (4-я волна). При составлении картотеки основное внимание было сосредоточено на фразеологических образах, что позволило точно охарактеризовать этнокультурную сущность описываемой ФЕ в иноязычных условиях. «Лингвокреативность этнокультурной духовности человека, – отмечает Н.Ф. Алефиренко, – проявляется не только в созидании первичных языковых образов, но и в их творческой адаптации к новым дискурсивным ситуациям речепорождения» [6. С. 11]. Эта характеристика языковых образов, прежде всего, присуща и образам фразеологическим, что особенно важно для изучения идиоматики эмигрантской прозы, поскольку именно ФЕ являются чрезвычайно значимыми для выявления социокультурных особенностей языка писателей русского зарубежья. По меткому замечанию М.Я. Гловинской, именно эти единицы в большей мере подвержены влиянию чужого языка: они регулируются «не общими, а специальными индивидуальными правилами» [7. С. 446].

Значимым для создания и использования фразеологической картотеки является лингвостатистический принцип, предполагающий использование элементов лингвоквантитативного метода для совершен-

ствования механизма подсчета ФЕ. Для подготовки архитектуры картотеки применялся также метод компьютерного моделирования, который позволил рассчитать оптимальную структуру будущей картотеки, наиболее соответствующую специфике исследования.

Методологической базой исследования послужили приёмы компьютерной лексикографии и фразеографии [4, 5, 8, 9]. Теоретическую основу составили классические лексикографические исследования [10], а также работы по изучению лексической и фразеологической образности [6] и специфики языка русской эмиграции [7]. Новизна исследования заключается в том, что впервые для лингвистических исследований продуктивно используется изначально не предназначавшееся для нужд лингвистики программное обеспечение.

Результаты исследования

Учитывая, что все исследования в области лексикологии и фразеологии предполагают активное создание карточек и последующую их обработку, а существующие программные решения имеют ограниченную возможность использования (как с лингвоприкладной, так и с технической стороны), представляется уместным поиск новых способов автоматизации картотеки. Так, было принято решение разработать собственную систему хранения и обработки информации на базе платформы «1С: Предприятие 8.3». Эта платформа изначально не была ориентирована на научно-исследовательскую работу. Однако ее гибкость позволяет создать уникальную конфигурацию, удовлетворяющую потребности конкретного пользователя. Позитивный опыт использования платформы в научно-исследовательских целях (ранее была разработана конфигурация для обработки данных о нескольких тысячах медиатекстов из социальных сетей [11, 12]) позволил максимально использовать ресурсы приложения. Учебная версия программы распространяется легально и бесплатно, а ее возможностей вполне достаточно для ведения картотеки. К преимуществам такого решения можно отнести:

– возможность подготовки персонализированной конфигурации, учитывающей индивидуальные особенности каждого конкретного исследования;

– невысокие системные требования, обеспечивающие корректную работу даже на компьютерах с невысокой мощностью;

– простоту использования, понятный интерфейс;

– возможность оперативной обработки огромных массивов данных;

– широкий диапазон статистических ресурсов;

– возможность автоматического построения диаграмм, выгрузки информации в документы форматов .docx, .xlsx.

Безусловно, использование подобного программного решения может вызвать некоторые сложности. Так, наиболее значимым изъяном можно считать сложность первоначальной разработки конфигурации для последующего создания картотеки. Написание такой конфигурации требует серьёзной профессиональной подготовки и может осуществляться только программистом. Однако при составлении грамотного и подробного проекта картотеки разработка конфигурации может стать единоразовым процессом, а взаимодействие с ней способно протекать без помощи IT-специалиста.

Исходя из вышеизложенного, ясно, что подготовка плана и структуры будущей картотеки – первый и самый важный этап разработки. Производить его нужно с учетом специфики будущего исследования. Это дает возможность заранее прописать те или иные параметры (например, фиксацию наличия/отсутствия в исследуемой ФЕ того или иного компонента – соматизма, зоонима и т.п.). Безусловно, набор заданных параметров можно в любой момент пополнять, хотя этот процесс, во-первых, часто может быть достаточно трудоемким, а во-вторых, потребует, скорее всего, дополнительной обработки ранее заполненных карточек.

Наше исследование посвящено изучению специфики ФЕ из произведений русских писателей-эмигрантов. Многоаспектное изучение ФЕ обусловило сравнительную сложность подготовленной конфигурации. Нельзя не учитывать оптимизированность ресурса под ключевую задачу. Это в полной мере соответствует подходам, свойственным прикладной лингвистике, где «под оптимизацией понимается такая модель языковой системы (или подсистемы), при которой этот объект сохраняет в результирующем представлении только те существенные свойства, которые необходимы для данной практической задачи» [8. С. 13]. При этом специфика платформы «1С: Предприятие» обусловливает потенциальную возможность корректировать созданную конфигурацию, что способствует ее универсальности, возможности использования в качестве основы для картотек иного характера. Главная страница подготовленной нами конфигурации поз-

воляет сразу перейти к списку сохраненных карточек. Слева располагается меню, обеспечивающее навигацию между основными разделами (*Общее, Фразеологизмы, Отчеты, Обработки*) (рис. 1).

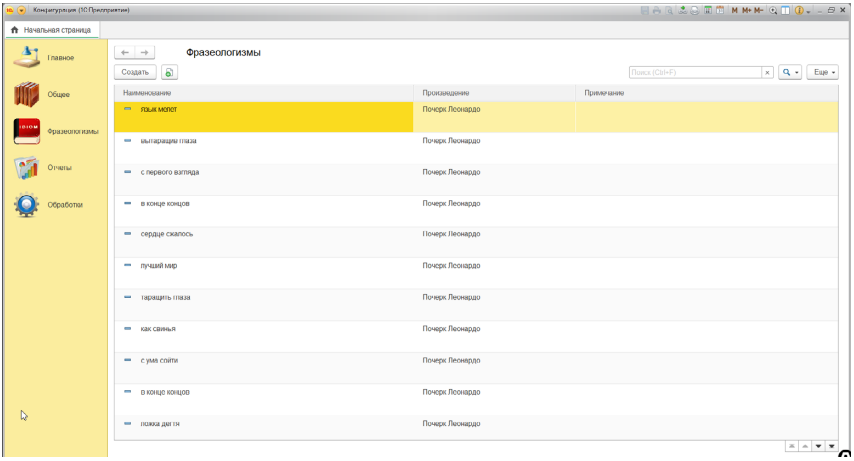

Рис. 1. Начальная страница

I. Раздел *Общее* содержит справочники, функционирующие в рамках текущей конфигурации. Под справочником в 1С понимается «прикладной объект, предназначенный для хранения данных, имеющих постоянный характер» [13]. Иными словами, такой справочник представляет собой список элементов. Если в бизнес-среде такими элементами могут быть списки сотрудников, реализуемых товаров и т.п., то в рамках исследуемой картотеки к ним можно отнести такие постоянные параметры, как название произведения, из которого берется единица, фамилия автора этого произведения, лексемы-компоненты той или иной группы и даже сами сохраняемые ФЕ. При этом справочники в пределах одной конфигурации взаимосвязаны, один может включать в свой состав элементы другого, что и позволяет делать картотеку максимально насыщенной, а также находить единицы, обладающие схожими характеристиками.

В разделе *Общее* нашей картотеки представлены справочники, информация в которых указывает на источник ФЕ: *Авторы, Книги, Произведения, Страны*. Последний справочник носит более служеб-

ный характер и необходим для корректного внесения информации об авторах исследуемых текстов.

Открывая любой справочник, пользователь видит перечень внесенных элементов. Можно ознакомиться с содержанием каждого из этого элементов, добавить новый, удалить старый и т.п. При этом добавлять элементы можно не только внутри справочника, но и непосредственно при составлении карточки, что существенно облегчает процесс работы с картотекой (рис. 2).

Рис. 2. Справочник «Авторы»

Расположение справочников в блоке *Общие* демонстрирует последовательность действий при работе с картотекой. Так, в первую очередь необходимо сохранить данные об авторе произведения-источника. Набор сведений напрямую зависит от особенностей исследования. Для нашей работы важным представлялось зафиксировать хронологические рамки эмиграции писателя, чтобы выявить возможные изменения в образной составляющей его фразеологического фонда и гипотетическую взаимосвязь таких изменений с новым местом жительства автора. Исходя из этого, карточка автора содержит элементы, представленные на рис. 3.

Отметим, что карточка автора содержит данные, необходимые для конкретного исследования. При необходимости их набор может варьироваться. Включение в структуру карточки полей с датами дает возможность автоматически классифицировать ФЕ из произведений, написанных в определенное время. Это же касается и указания стран эмиграции. Такие элементы позволяют проверить ряд исследовательских гипотез, согласно которым те или иные особенности ФЕ могут зависеть от времени или места создания произведения-источника.

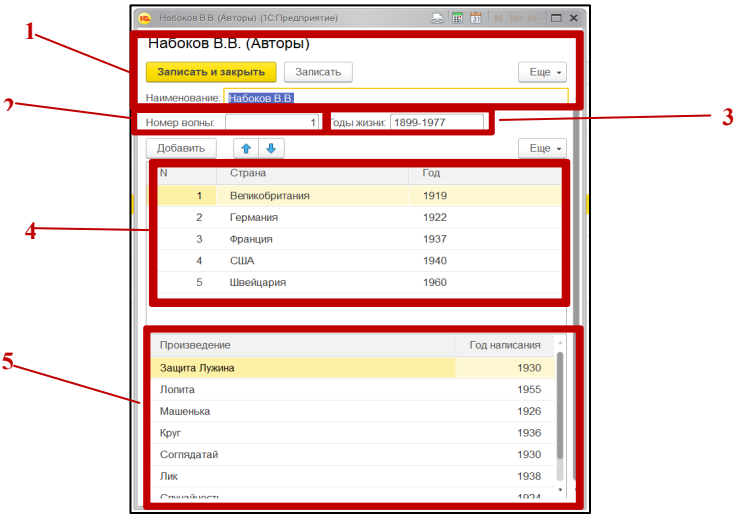

Рис. 3. Карточка автора: *1* – фамилия и инициалы авторы, кнопки для сохранения изменений в карточке; *2* – волна эмиграции; *3* – годы жизни писателя; *4* – история эмиграции (с годом переезда в страну); *5* – перечень исследуемых произведений, годы их написания

Сохранив данные об авторе, пользователь может приступить к добавлению книги в соответствующий справочник. Этот этап необходим вне зависимости от того, сколько произведений входит в состав одного издания. Карточка книги (в широком смысле слова) выглядит так, как представлена на рис. 4.

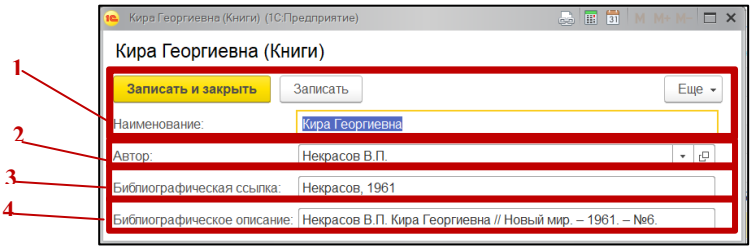

Рис. 4. Карточка автора: *1* – название книги, включаемое в библиографическое описание; *2* – автор произведения (выбирается из списка); *3* – заготовка для библиографической ссылки; *4* – заготовка для библиографического описания

Поле «Наименование» может заполняться произвольно, оно необходимо только для работы с самой картотекой. На наш взгляд, удобнее вносить туда реальное название книги-источника, чтобы впоследствии оптимизировать поиск нужного издания. Поле «Автор» содержит выпадающий список, куда автоматически попадают фамилии с уже сохраненных ранее карточек. Если нужного автора в перечне нет, то его можно добавить, нажав на соответствующую кнопку. Благодаря этому откроется поле для добавления новой карточки автора. Поле «Библиографическая ссылка» содержит основу для автоматической генерации ссылок. При добавлении нового ФЕ в его карточку вносится номер страницы, благодаря чему автоматически генерируется и точная ссылка. Это же касается библиографического описания. Такой подход существенно облегчает последующую работу с ФЕ и позволяет быстро вставлять ссылку на каждый используемый в тексте исследования пример.

Таким же образом происходит и работа со справочником *Произведение*. Его карточка невелика, так как ключевую информацию уже содержат в себе предыдущие справочники (рис. 5).

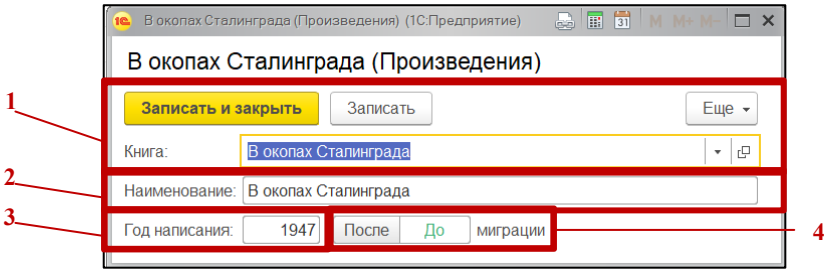

Рис. 5. Карточка произведения: *1* – название книги (выбирается из выпадающего списка); *2* – название добавляемого произведения; *3* – год создания произведения; *4* – время создания произведения по отношению ко времени эмиграции автора

Принцип сохранения произведения остается тем же. Поле «Наименование» обеспечивает создание связи между *Произведением* и *Книгой*. Установление переключателя «До/после эмиграции» позволяет быстро сортировать произведения, что может быть чрезвычайно полезно, например, при сопоставительном анализе.

II. Следующий раздел созданной картотеки – *Фразеологизмы*. Он, в свою очередь, делится на два подраздела, первый из которых включает собственно ФЕ и их поля, а второй – отдельные компоненты таких ФЕ, имеющие лингвокультурную значимость (например, соматизмы, зоонимы и т.п.). Стоит отметить, что сохранение и систематизация ФЕ в данной картотеке тоже производятся особым образом, так как она создавалась для исследования, направленного, прежде всего, на образную составляющую выражений. Возник вопрос: как систематизировать ФЕ таким образом, чтобы выражения с идентичной образной составляющей, реализованной посредством разных лексем (например, *голова* и *башка*), были объединены? Для решения подобной проблемы нами было выделено две группы ФЕ, условно названные «цари» и «холопы». Под «*царями*» понимаются головные ФЕ, а «*холопы*» представляют собой реализации таких «царей» в тексте. При этом, открыв карточку каждого «царя», можно сразу увидеть весь перечень его «холопов» (рис. 6).

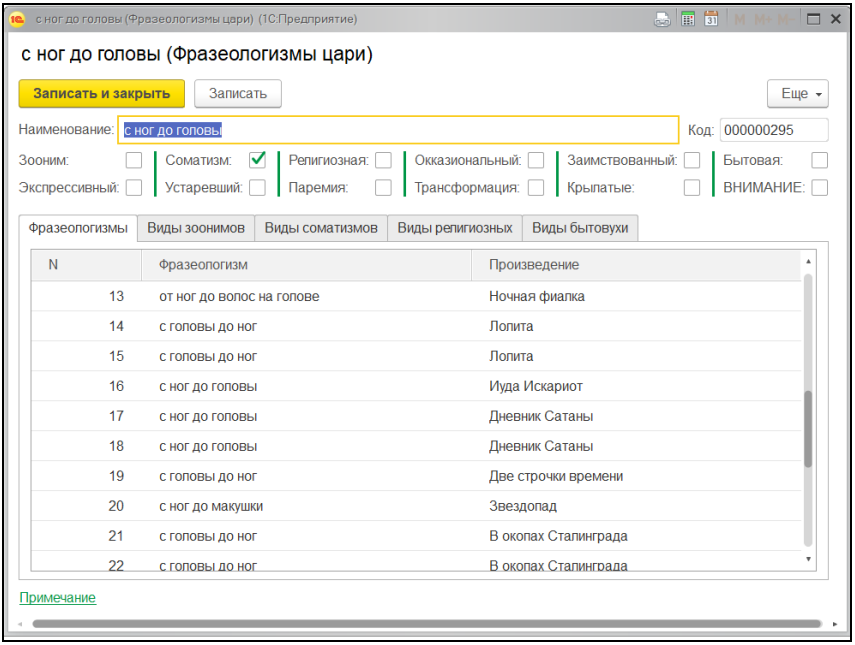

Рис. 6. Карточка «царя»

Как видим, такой подход позволяет выявить все возможные реализации той или иной ФЕ, что дает возможность показать потенциально значимые элементы фразеологического образа (например, его смысловую зависимость от порядка слов, влияние стилистической окраски и т.п.).

Статистическая обработка ФЕ также становится более продуктивной, так как автоматически могут подсчитываться как «цари», так и «холопы». С полным перечнем «царей» и «холопов» можно ознакомиться в соответствующих справочниках, используемых приложением «1С: Предприятие». Если по ошибке какой-то «царь» вводился дважды (например, дублирование произошло за счет фиксации ФЕ со стрежневыми глаголами – видовыми парами), то с помощью специально продуманного сервиса дубли можно объединить вместе со всеми содержащимися в них «холопами».

Особого внимание заслуживает непосредственно сама карточка фразеологизма. Как уже отмечалось ранее, в нашем случае она содержала в себе потенциальные компоненты-маркеры, способные пролить свет на особенности этноязыкового сознания носителя текста, в том числе с учетом возможного влияния эмиграции. Подобные компоненты могут иметь большое значение и в тех случаях, когда речь идет о других исследованиях лингвокультурной направленности. В иных случаях набор фиксируемых параметров с легкостью можно скорректировать на этапе создания конфигурации.

С учетом всего вышеизложенного, карточка каждой сохраняемой ФЕ выглядит так, как представлено на рис. 7.

Структура карточки призвана максимально отражать все особенности фиксируемой ФЕ. Так, исследователь получает возможность выделить особенность использования выражения в каждом конкретном случае (о необходимости такого процесса было сказано выше). По умолчанию поле *3* автоматически заполняется информацией из поля *2*, что облегчает ввод информации в случаях, когда значимых расхождений между «царем» и «холопом» нет. Автоматические процессы предусмотрены и для поля *4*. Так, название произведения выбирается из выпадающего списка (как и в случае с остальными подобными полями, система начинает предлагать подходящие варианты по первым буквам вводимых слов), однако по умолчанию в этом поле уже указывается то произведение, которое было в предыдущей карточке. Это помогает ускорить процесс ввода ФЕ, извлеченных из одного и того же текста.

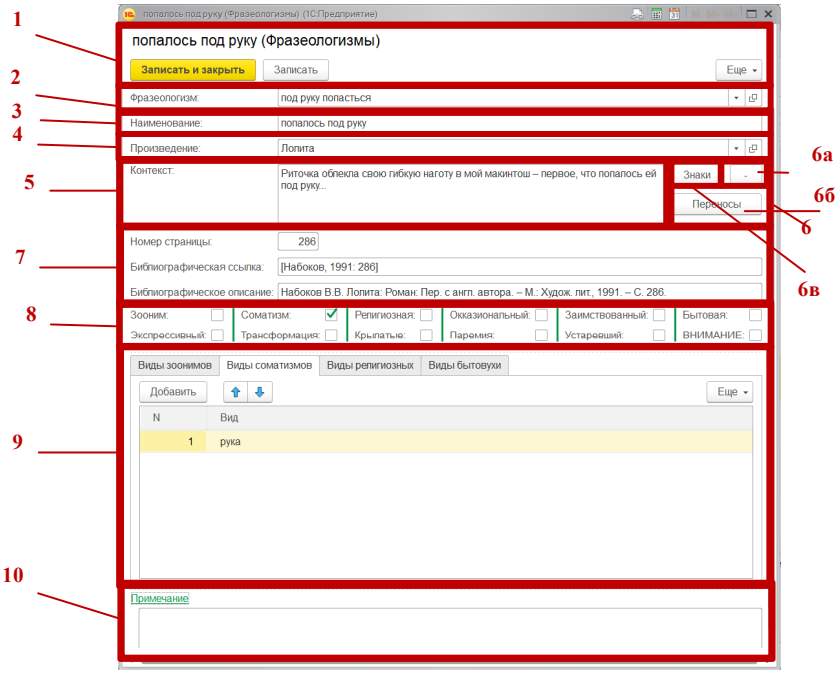

Рис. 7. Карточка ФЕ: *1* – «шапка» карточки с указанием ФЕ и кнопки для сохранения изменений; *2* – ФЕ («царь»); *3* – конкретная реализация ФЕ в отдельно взятом случае («холоп»); *4* – произведение-источник; *5* – поле для ввода контекста, в котором употребляется ФЕ; *6* – служебные кнопки для форматирования; *7* – библиографические сведения; *8* – блоки для маркирования значимых элементов ФЕ; *9* – перечень значимых компонентов в составе ФЕ; *10* – поле для ввода примечаний

Отдельное поле предусмотрено для ввода контекста – одного-двух предложений, позволяющих увидеть, как ФЕ реализуется внутри текста, какие дополнительные смыслы у нее появляются и т.п. Контекст может не вводиться вручную, а копироваться. Это происходит в тех случаях, когда в качестве источника используется электронный ресурс, а также помогает существенно сэкономить время на ввод. Однако нередко копируемый текст может переноситься с некорректным форматированием: лишними пробелами, разрывами строк, дефисами вместо тире. Такие недочеты можно устранить и вручную, однако гораздо проще воспользоваться специальными кнопками:

– *6а* – изменение дефиса на обычное тире;

– *6б* – удаление разрывов строк;

– *6в* – удаление лишних пробелов, корректировка неверных символов.

Отдельного внимания заслуживает библиографический блок. Пользователю достаточно указать лишь номер страницы, где находится найденное выражение. Все остальные данные (библиографическое описание, библиографическая ссылка) генерируются автоматически. Это ускоряет и упрощает введение таких ссылок в канву научных работ, так как избавляет исследователя от необходимости составлять их самостоятельно.

Примечательным становится восьмой блок в карточке. Здесь собран набор маркеров, представляющихся важными в рамках проводимого исследования. Это не только традиционно выделяемые лингвокультурологами маркеры (в том числе с учетом особой специфики настоящего исследования и его компьютерной реализации [14]), но и ряд дополнительных отметок, представлявшихся необходимыми. Например, отдельно были отмечены ФЕ, содержащие в своём составе зоонимы и соматизмы. Выбор подобных единиц обусловливается их спецификой. Так, Е.С. Яковлева отмечает, что зоонимические ФЕ «возникли в результате переосмысления фенотипических признаков и повадок домашних и диких животных» [15. С. 3]. Этот механизм обусловливает, с одной стороны, универсальность определенных зоонимических образов, а с другой – культурно обусловленные различия в их смысловой интерпретации [16]. Особого внимания заслуживают зооморфизмы, так как они способны стать важным инструментом для характеристики этноязыкового сознания [17]. Схожая ситуация наблюдается и применительно к компонентам-соматизмам. М. Горды пишет, что «вопрос человеческого тела – это вопрос сущности/природы самого человека» [18. С. 43]. Подобные механизмы самопознания присущи представителям различных лингвокультур, сами лексемы могут носить характер универсалий (части тела людей не меняются в зависимости от национальности), однако на понятийном уровне этнокультурные различия могут быть довольно значимыми, что обусловливается различными доминантами у того или иного этноса (на первый план могут выходить разные признаки) [19. С. 190]. Кроме того, изучение соматических ФЕ «позволяет не только приоткрыть “завесу” какой-либо этнической лингвокультуры, но и выявить

особенности ее дискурсивно-когнитивного субстрата» [20. С. 16]. Все это обусловливает высокую значимость зоонимических и соматических компонентов при изучении фразеологических образов.

Фиксация подобных компонентов производилась совместно с заполнением девятого поля, где представлены сами выделенные компоненты. Так как в рамках проводимого нами исследования на первый план выдвигалась образная составляющая выражений, родственные лексемы фиксировались как общая единица (например, *кот, котик* и *кошачий* отмечались как *кот*). Исключение делалось для тех случаев, когда лексемы различались по стилистическим характеристикам, условиям употребления или же обладали иными значимыми различиями (например, *кот, кошка* и *котенок* фиксировались раздельно, так как разница в возрастных или гендерных параметрах, просматривающаяся при сопоставлении семантики этих лексем, может играть важную роль в смысловом аспекте). Отдельно фиксировались также лексемы, способные иллюстрировать значимые детали быта носителей языка (*корыто, калоша, цент* и т.п.). Сюда же могли быть отнесены компоненты, не связанные напрямую с бытом человека, но способные охарактеризовать его деятельность (*карты, шарманка, пилюля*) или же реалии окружающей среды (*камень, елка*), что может иметь большее значение в лингвокультурном аспекте и даже принимать на себя роль символов [21]. Кроме того, маркер «бытовые» применялся и к ФЕ, не содержащим в своем составе специфичных компонентов, но репрезентирующих те или иные культурные или исторические факты (*изменник Родины, железный занавес*). Двойную функцию также выполнял маркер «религиозные».

Маркирование ФЕ производится не только на основе наличия в их составе тех или иных компонентов. Так, отдельно выделяются ФЕ высокой экспрессивности, устаревшие выражения, ФЕ, подвергшиеся той или иной трансформации, крылатые слова и паремии (широкое понимание фразеологии дает нам основание причислять две последние отмеченные группы к ФЕ). Отдельно отмечены заимствованные ФЕ, исследование которых представляется особенно важным при работе с текстами писателей-эмигрантов. Особая служебная отметка «Внимание» позволяет маркировать нестандартные выражения, требующие более детального изучения и/или помогающие получить интересные наблюдения.

Отдельное поле предназначено для записи примечаний. Оно позволяет делать заметки, полезные для последующей работы с ФЕ. Здесь могут фиксироваться сведения о происхождении ФЕ (например, некоем литературном или историческом источнике), отмечаться особенности трансформации, а также любые другие детали, которые становятся очевидными при чтении текста произведения-источника. Это дает возможность сохранять важные наблюдения и задействовать их при подготовке научных работ.

Как отмечалось ранее, выставление всех маркеров в карточке позволяет генерировать отчеты. Типы таких отчетов программируются при создании конфигурации, однако в любой момент можно добавить новые формы, а также настраивать существующие, самостоятельно указывая данные для подсчета. Так, например, одним из важных этапов работы становится общий подсчет введенных ФЕ, как целиком, так и в рамках отдельно обозначенных групп (рис. 8).

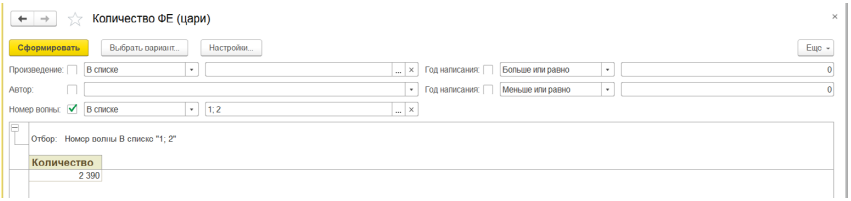

Рис. 8. Отчет (количество «царей»)

Как видно на рис. 8, возможен подсчет как всех ФЕ в картотеке, так и тех, что ограничены определенными рамками, например, волнами эмиграции, автором, временем написания и т.п. Это делает форму отчёта максимально универсальной и позволяет быстро получать любые необходимые сведения. Более сложные типы отчетов позволяют получать данные сразу по нескольким параметрам, а также подсчитывать процентное соотношение тех или иных единиц (рис. 9).

Приведенный фрагмент отчета показывает, как может демонстрироваться количество ФЕ, соответствующих заданным параметрам, а также позволяет увидеть автоматический подсчет процентов таких ФЕ по отношению к общему количеству выражений. Разбивка результатов по отдельным писателям и произведениям позволяет получать подробную информацию о фразеологической палитре творчества

каждого из исследуемых авторов. При этом для получения данных по каждому из авторов в целом предусмотрен отдельный отчет (рис. 10).

Процентное соотношение ФЕ

Сформировать | Выбрать вариант... | Настройки... | Еще

Выбранные поля: Всего, Зоонимы, Количество зоонимов1, Соматизмы, Количество соматизмов1, Р...
Группировка: Номер волны, Автор, Произведение
Отбор:
Произведение.После миграции: Да
Сортировка: Номер волны, Автор, Произведение
Произведение.Книга.Автор.Номер волны: 1

Номер волны / Автор / Произведение	Всего	Зоонимы К-во	Зоонимы %	Соматизмы К-во	Соматизмы %	Религиозные К-во	Религиозные %	Окказиональные К-во	Окказиональные %	Заимствованны К-во	Заимствованны %
1	5 262	101	1,37	1 069	21,42	562	12,48	26	0,30	52	0,
Андреев Л.Н.	644	6	0,88	119	21,10	75	9,37			10	1
Дневник Сатаны	158			26	16,46	23	14,56			10	6
Жили-были	33			4	12,12	2	6,06				
Иуда Искариот	57	1	1,75	26	45,61	3	5,26				
Красный смех	72	1	1,39	11	15,28	6	8,33				
Сашка Жегулев	324	4	1,23	52	16,05	41	12,65				
Бунин И.А.	1 234	23	0,69	227	16,64	112	10,32	11	0,33		
Господин из Сан-Франциско	26			4	15,38						
Грамматика любви	14			2	14,29	1	7,14				
Деревня	364	18	4,95	63	17,31	39	10,71	9	2,47		
Жизнь Арсеньева	496	3	0,60	90	18,15	35	7,06	1	0,20		
Легкое дыхание	8			1	12,50						
Митина любовь	120			18	15,00	7	5,83				
Солнечный удар	18			5	27,78	2	11,11				
Суходол	149	2	1,34	40	26,85	19	12,75	1	0,67		
Темные аллеи	14			1	7,14	4	28,57				
Чистый понедельник	25			3	12,00	5	20,00				
Куприн А.И.	1 454	40	2,26	321	24,86	164	11,10	8	0,42	11	0
Гранатовый браслет	112	7	6,25	24	21,43	10	8,93	5	4,46	1	0
Дочь великого Барнума	27	1	3,70	10	37,04	1	3,70				
Жанета	93	3	3,23	20	21,51	10	10,75			4	4
Инна	9			4	44,44	1	11,11				
Колесо времени	95	2	2,11	15	15,79	15	15,79			2	2
Куст сирени	18			4	22,22	2	11,11				
Ленинка	21			5	23,81	2	9,52				

Рис. 9. Отчет (маркированные ФЕ)

Статистика по авторам

	Аксенов В.П.		
	Количество (отн. эмиграции)		Всего
	До	После	
ФЕ с зоонимами	21	37	58
ФЕ с соматизмами	118	152	270
ФЕ с бытовыми	20	22	42
ФЕ с религиозными	67	30	97
Крылатые	17	28	45
Заимствованные	14	30	44
Паремии	5	9	14
Окказиональные	2	1	3
Экспрессивные	178	173	351
Трансформация	55	63	118
Устаревшие	3	15	18

Всего ФЕ у автора: 1 533

Рис. 10. Отчет (отдельный писатель)

В приведенном отчете даются общие сведения о количестве ФЕ того или иного типа, встречающихся у автора в произведениях, написанных до и после эмиграции.

Отдельные формы отчетов предусмотрены для экспликации отношений одной группы ФЕ к другой (или ко всем ФЕ в целом). При этом пользователь может изучать как все выражения из картотеки, так и отдельные (например, взятые из произведений конкретного автора) (рис. 11).

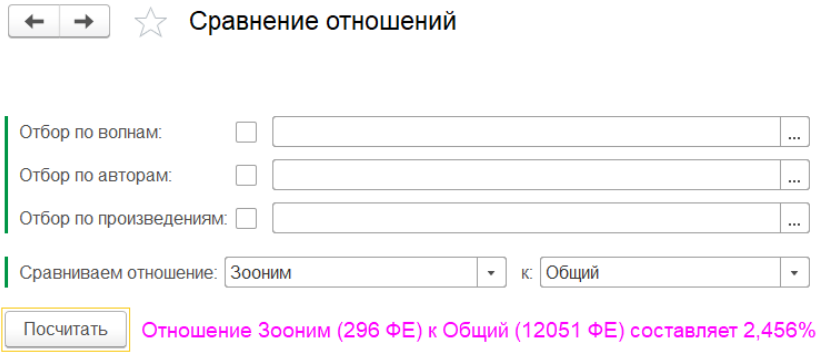

Рис. 11. Отчет (пропорции)

В рамках научных работ зачастую используется представление данных в виде диаграмм. Этот формат является более наглядным, помогает ярче продемонстрировать соотношение тех или иных категорий. Представленная нами картотека позволяет автоматически генерировать диаграммы по заданным параметрам. Пример такой диаграммы, показывающей количество ФЕ с соматизмами, демонстрирует рис. 12.

Пользователь получает возможность варьировать перечень источников исследуемых ФЕ, выбирать единицы, которые необходимо отразить в диаграмме, а также настраивать вид представления данных. Все это помогает получить изображение, которое можно вставить в текст исследования без дополнительных корректировок.

Создание автоматизированной картотеки на базе платформы «1С: Предприятие» открывает перед исследователем широкий диапазон возможностей и позволяет существенно оптимизировать процесс работы с большим количеством языковых единиц. Электронные карточки компактны, вмещают большое количество информации, подлежат корректировке и дополнению.

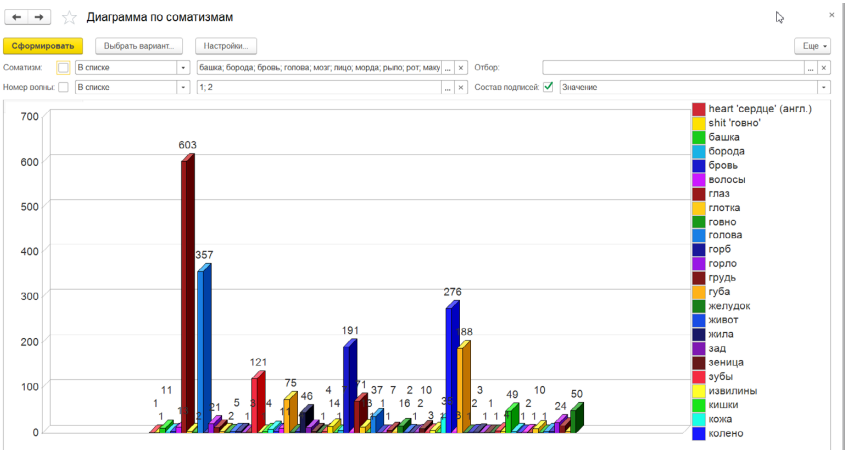

Рис. 12. Отчет (диаграмма)

Заключение

Однако, несмотря на отлаженное функционирование, картотека, на наш взгляд, подлежит дальнейшей доработке. Так, уместным представляется расширения числа задействованных маркеров, генерируемых отчетов и иных аспектов рассмотрения ФЕ. Выявленные лексикографические принципы создания фразеологической картотеки помогут устранить имеющиеся технические упущения при совершенствовании компьютерной методологии, оптимизирующей работу над электронными словарями XXI в.

Список источников

1. *Аксенов В.П.* Бумажный пейзаж. Анн-Арбор : Ardis, 1983.

2. *Денисов П.Н.* [Рец. на:] В.Н. Телия (отв. ред.). Фразеография в Машинном фонде русского языка // Вопросы языкознания. 1993. № 5. С. 145–146.

3. *Захаров В.П.* Словарная картотека Института лингвистических исследований РАН как объект автоматизации // Компьютерная лингвистика и интеллектуальные технологии : труды междунар. конф. «Диалог 2007» (Бекасово, 30 мая – 3 июня 2007 г.) / под ред. Л.Л. Иомдина, Н.И. Лауфер, А.С. Нариньяни, В.П. Селегея. М. : Изд-во РГГУ, 2007. С. 202–206.

4. *Сарыгул К.* Компьютерная (электронная) лексикография // Актуальные проблемы гуманитарных и естественных наук. 2016. № 4-2. URL: https://cyberleninka.ru/article/n/kompyuternaya-elektronnaya-leksikografiya (дата обращения: 29.10.2022).

5. *Глебова О.В., Чихачев К.Б.* Электронная картотека ЛАРНГ. Современное состояние и перспективы // Лексический атлас русских народных говоров (Материалы и исследования) 2020 / отв. ред. С.А. Мызников. СПб. : ИЛИ РАН, 2020. 1000 с.

6. *Алефиренко Н.Ф.* Языковые образы как единицы этнокультурного сознания // Гуманитарные исследования. 2021. № 4 (80). С. 10–18.

7. *Язык* русского зарубежья: Общие процессы и речевые портреты / отв. ред. Е.А. Земская. Москва ; Вена : Языки славянской культуры: Венский славистический альманах, 2001. 496 с.

8. *Большакова Е.И., Клышинский Э.С., Ландэ Д.В., Носков А.А., Пескова О.В., Ягунова Е.В.* Автоматическая обработка текстов на естественном языке и компьютерная лингвистика : учеб. пособие. М. : МИЭМ, 2011. 272 с.

9. *Краснобаева-Чёрная Ж.В.* Терминологический банк данных «Классификационные параметры фразеологических единиц» как электронный терминографический продукт: опыт проектирования // Вопросы лексикографии. 2020. № 18. С. 117–132.

10. *Никитина Т.Г., Рогалёва Е.И.* Современная учебная фразеография: инновационные параметры бумажного словаря // Вопросы лексикографии. 2021. № 20. С. 67–90.

11. *Голикова М.М.* Событийная повестка в корпоративной коммуникации: анализ социальных медиа вуза. Белгород, 2020. 166 с.

12. *Алефиренко Н.Ф., Голикова М.М.* Междисциплинарный подход как основа современной проектной деятельности (на примере создания электронного фразеологического словаря) // Информационные и коммуникативные технологии. Проектная деятельность в образовательном и информационно-коммуникативном процессе: опыт и перспективы : сб. науч. ст. по материалам III Всерос. науч.-практ. конф. с междунар. участием. Симферополь, 2020. С. 184–188.

13. *Глоссарий* разработчика 1С. URL: https://its.1c.ru/db/v8devgloss#content:22:hdoc (дата обращения: 29.10.2022).

14. *Булыгина Е.Ю., Трипольская Т.А.* Национально-культурный компонент в семантике слова и способы его представления в базе данных прагматически маркированной лексики // Вопросы лексикографии. 2017. № 11. С. 5–19.

15. *Яковлева Е.С.* Национально-культурная специфика компонентов-зоонимов, репрезентирующих домашних, диких и мифологических животных во фразеологическом фонде китайского и английского языков : дис. … канд. филол. наук. Белгород, 2019. 24 с.

16. *Куражова И.В.* Имена животных как отражение ценностной картины мира в английской лингвокультуре : дис. ... канд. филол. наук. Иваново, 2007. 201 с.

17. *Гридина Т.А., Коновалова Н.И.* Зооморфизмы как основа моделирования фразеологической семантики: русско-польские соответствия // Русин. 2019. Т. 56. С. 198–212.

18. *Горды М.* Соматическая фразеология современных русского и польского языков. Щецин, 2010. 365 с.

19. *Кошарная С.А.* Фразеологизация концепта в контексте этнокультуры (на материале ФЕ, вербализующих концепт «Рука») // Фразеология, познание и куль-

тура : сб. докл. 2-й Междунар. науч. конф. Т. 1: Фразеология и познание. Белгород, 2010. С. 189–193.

20. *Абдельхамид С.А.* Фраземы с соматическим компонентом в автохтонной языковой картине мира // Вестник Адыгейского государственного университета. Сер. 2: Филология и искусствоведение. 2015. С. 15–21.

21. *Гнездилова Н.С.* Библейская символика камня в русской фразеологии // Вестник ТГПУ. 2020. № 2 (208). С. 68–75.

References

1. Aksenov, V.P. (1983) *Bumazhnyy peyzazh* [Paper landscape]. Ann-Arbor: Ardis.

2. Denisov, P.N. (1993) [Rets. na:] V.N. Teliya (otv. red.). Frazeografiya v Mashinnom fonde russkogo yazyka [[Book Review] Teliya, V.N. (ed.) Phraseography in the Machine Fund of the Russian Language]. *Voprosy yazykoznaniya.* 5. pp. 145–146.

3. Zakharov, V.P. (2007) Slovarnaya kartoteka Instituta lingvisticheskikh issledovaniy RAN kak ob"ekt avtomatizatsii [Vocabulary card file of the Institute for Linguistic Studies of the Russian Academy of Sciences as an object of automation]. In: Iomdin, L.L. et al. (eds) *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: Trudy mezhdunarodnoy konferentsii "Dialog 2007"* [Computer Linguistics and Intelligent Technologies: Proceedings of the International Conference "Dialogue 2007"]. Bekasovo. 30 May – 3 June 2007). Moscow: RSUH. pp. 202–206.

4. Sarygul, K. (2016) Komp'yuternaya (elektronnaya) leksikografiya [Computer (electronic) lexicography]. *Aktual'nye problemy gumanitarnykh i estestvennykh nauk.* 4-2. pp. 191–196.

5. Glebova, O.V. & Chikhachev, K.B. (2020) Elektronnaya kartoteka LARNG. Sovremennoe sostoyanie i perspektivy [Electronic file cabinet LARNG. Current state and prospects]. In: Myznikov, S.A. (ed.) *Leksicheskiy atlas russkikh narodnykh govorov (Materialy i issledovaniya) 2020* [Lexical atlas of Russian folk dialects (Materials and research) 2020]. St. Petersburg: ILS RAS.

6. Alefirenko, N.F. (2021) Yazykovye obrazy kak edinitsy etnokul'turnogo soznaniya [Language images as units of ethnocultural consciousness]. *Gumanitarnye issledovaniya.* 4 (80). pp. 10–18.

7. Zemskaya, E.A. (ed.) (2001) *Yazyk russkogo zarubezh'ya: Obshchie protsessy i rechevye portrety* [The language of the Russian diaspora: General processes and speech portraits]. Moscow; Vienna: Yazyki slavyanskoy kul'tury: Venskiy slavisticheskiy al'manakh.

8. Bol'shakova, E.I. et al. (2011) *Avtomaticheskaya obrabotka tekstov na estestvennom yazyke i komp'yuternaya lingvistika: ucheb. posobie* [Automatic text processing in natural language and computational linguistics: textbook]. Moscow: MIEM.

9. Krasnobaeva-Chernaya, Zh.V. (2020) The Terminological Data Bank "Classification Parameters of Phraseological Units" as an Electronic Terminographic Product: Design Experience. *Voprosy leksikografii – Russian Journal of Lexicography*. 18. pp. 117–132. (In Russian). doi: 10.17223/22274200/18/6

10. Nikitina, T.G. & Rogaleva, E.I. (2021) Modern Educational Phraseography: Innovative Parameters of a Paper Dictionary. *Voprosy leksikografii – Russian Journal of Lexicography*. 20. pp. 67–90. (In Russian). doi: 10.17223/22274200/20/4

11. Golikova, M.M. (2020) *Sobytiynaya povestka v korporativnoy kommunikatsii: analiz sotsial'nykh media vuza* [Event agenda in corporate communication: analysis of university social media]. Belgorod: [s.n.].

12. Alefirenko, N.F. & Golikova, M.M. (2020) [Interdisciplinary approach as the basis of modern project activities (on the example of creating an electronic phraseological dictionary)]. *Informatsionnye i kommunikativnye tekhnologii, proektnaya deyatel'nost' v obrazovatel'nom i informatsionno-kommunikativnom protsesse: opyt i perspektivy* [Information and communication technologies, project activities in the educational and information-communicative process: experience and prospects]. Conference Proceedings. Simferopol: [s.n.]. pp. 184–188. (In Russian).

13. 1C. (2022) *Glossariy razrabotchika 1S* [1C Developer Glossary]. [Online] Avaiblable from: https://its.1c.ru/db/v8devgloss#content:22:hdoc (Accessed: 29.10.2022).

14. Bulygina, E.Yu. & Tripol'skaya, T.A. (2017) The National-Cultural Component in the Semantics of Words and Ways of Presenting It in a Database of Pragmatically Marked Vocabulary. *Voprosy leksikografii – Russian Journal of Lexicography*. 11. pp. 5–19. (In Russian). doi: 10.17223/22274200/11/1

15. Yakovleva, E.S. (2019) *Natsional'no-kul'turnaya spetsifika komponentov-zoonimov, reprezentiruyushchikh domashnikh, dikikh i mifologicheskikh zhivotnykh vo frazeologicheskom fonde kitayskogo i angliyskogo yazykov* [National and cultural specificity of zoonym components representing domestic, wild and mythological animals in the phraseological fund of Chinese and English]. Philology Cand. Diss. Belgorod.

16. Kurazhova, I.V. (2007) *Imena zhivotnykh kak otrazhenie tsennostnoy kartiny mira v angliyskoy lingvokul'ture* [Animal names as a reflection of the value picture of the world in English linguaculture]. Philology Cand. Diss. Ivanovo.

17. Gridina, T.A. & Konovalova, N.I. (2019) Zoomorphisms as the Basis for Modelling Phraseological Semantics: Russian-Polish Equivalents. *Rusin*. 56. pp. 198–212. (In Russian). doi: 10.17223/18572685/56/12

18. Gordy, M. (2010) *Somaticheskaya frazeologiya sovremennykh russkogo i pol'skogo yazykov* [Somatic phraseology of modern Russian and Polish languages]. Szczecin: [s.n.].

19. Kosharnaya, S.A. (2010) Frazeologizatsiya kontsepta v kontekste etnokul'tury (na materiale FE, verbalizuyushchikh kontsept "Ruka") [Phraseologisation of the concept in the context of ethnoculture (on the basis of phraseological units verbalizing the concept "Ruka")]. *Frazeologiya, poznanie i kul'tura* [Phraseology, cognition and culture]. Proceedings of the 2nd International Conference. Vol. 1. Belgorod: [s.n.], pp. 189–193. (In Russian).

20. Abdel'khamid, S.A. (2015) Frazemy s somaticheskim komponentom v avtokhtonnoy yazykovoy kartine mira [Phrasemes with a somatic component in the autochthonous language picture of the world]. *Vestnik Adygeyskogo gosudarstvennogo universiteta. Seriya 2: Filologiya i iskusstvovedenie*, pp. 15–21.

21. Gnezdilova, N.S. (2020) Biblical Symbolism of Stone in Figurative Means of Russian Language. *Vestnik TGPU – Tomsk State Pedagogical University Bulletin.* 2 (208). pp. 68–75. (In Russian). doi: 0.23951/1609-624X-2020-2-68-75

Информация об авторах:
Алефиренко Николай Фёдорович – доктор филологических наук, профессор кафедры русского языка и русской литературы Белгородского государственного национального исследовательского университета (Белгород). E-mail: alefirenko@bsu.edu.ru
Голикова Мария Михайловна – аспирант, ассистент кафедры русского языка и русской литературы Белгородского государственного национального исследовательского университета (Белгород). E-mail: golikova@bsu.edu.ru

Авторы заявляют об отсутствии конфликта интересов.

***Information about the authors*:**
Nikolay F. Alefirenko, Dr. Sci. (Philology), professor, Belgorod National State University (Belgorod, Russian Federation). E-mail: alefirenko@bsu.edu.ru
Mariya M. Golikova, postgraduate student, Belgorod National State University (Belgorod, Russian Federation). E-mail: golikova@bsu.edu.ru

The authors declare no conflicts of interests.

Статья поступила в редакцию 15.09.2022;
одобрена после рецензирования 07.02.2023; принята к публикации 26.04.2023.

The article was submitted 15.09.2022;
approved after reviewing 07.02.2023; accepted for publication 26.04.2023.