

---

**Research Data**

---

01 Feb 2024

**Multi-dimensional scaling of web-scraping results from the A.A Zalizniak Grammatical Dictionary and the Russian National Corpus. Creating a corpus fragment of all possible word-forms of modified Russian sound verbs using web-scraping methodology. Compilation of a summary table for the present tense, past tense, future tense, imperative, imperfective and perfective gerund forms**

Irina Ivliyeva

*Missouri University of Science and Technology, ivliyeva@mst.edu*

Perry Koob

*Missouri S&T Information Technology, koobp@mst.edu*Follow this and additional works at: [https://scholarsmine.mst.edu/research\\_data](https://scholarsmine.mst.edu/research_data)Part of the [Russian Linguistics Commons](#)

---

**Recommended Citation**

Ivliyeva, Irina and Koob, Perry, "Multi-dimensional scaling of web-scraping results from the A.A Zalizniak Grammatical Dictionary and the Russian National Corpus. Creating a corpus fragment of all possible word-forms of modified Russian sound verbs using web-scraping methodology. Compilation of a summary table for the present tense, past tense, future tense, imperative, imperfective and perfective gerund forms" (2024). *Research Data*. 12.

[https://scholarsmine.mst.edu/research\\_data/12](https://scholarsmine.mst.edu/research_data/12)

This Data is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Research Data by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact [scholarsmine@mst.edu](mailto:scholarsmine@mst.edu).

---

## Multi-dimensional scaling of web-scraping results from the A.A Zalizniak Grammatical Dictionary and the Russian National Corpus. Creating a corpus fragment of all possible word-forms of modified Russian sound verbs using web-scraping methodology. Compilation of a summary table for the present tense, past tense, future tense, imperative, imperfective and perfective gerund forms

### Abstract

This project attempts not only to improve the method of web extraction in relation to the source material (lexical-semantic group of Russian sound verbs, semantically modified at the word-forming level) but also systematize the search results using the format that is convenient and effective for a linguistic researcher. The main tasks at this stage of the research were to collect, compile (combine results), analyze and present a summary index of all possible verbal forms of sound verbs, recorded in electronic versions of dictionaries of the Russian language and in the database of the [Russian National Corpus](#) (RNC).

The output takes into account not only all documented modifications of sound verbs from the [A.A Zalizniak Grammatical Dictionary](#) and the [Russian National Corpus](#), but also reveals lacunae, doublets, and indicates new (potential) units.

The results of the study may be useful for the development of various web applications for the search, collection, and visualization of linguistic material of different volumes and degrees of complexity. Possibilities of combinatorial optimization in the processing of open and closed linguistic databases can be particularly important when extracting information from various digital lexicographic sources (across a single or multiple languages), from national linguistic corpora, as well as from digital text collections.

В данном проекте предпринята попытка не только усовершенствовать методику веб-извлечения применительно к исходному материалу (лексико-семантической группе русских звуковых глаголов, модифицированных на словообразовательном уровне), но и обобщить представление результатов поиска в формате, который максимально удобен и эффективен для исследователя-лингвиста.

Основными задачами на данном этапе исследования были - собрать, скомпилировать (соединить результаты), проанализировать и представить обобщающий индекс всех возможных глагольных форм глаголов звучания, зафиксированных в электронных версиях словарей русского языка и в базе данных [Национального корпуса русского языка](#) (НКРЯ).

С учетом результатов предыдущих этапов исследований (2022, 2023) была усовершенствована четырехпозиционная система нумерации исходных форм глагола, их словоформ и модифицированных дериватов. Успешно проведено системное извлечение и фиксирование результатов поиска в виде многомерной сводной интерактивной таблицы, единицы которой снабжены отсылками к НКРЯ - представительной коллекции текстов [на русском языке](#) общим объемом [более 2 млрд слов](#), оснащенной лингвистической [разметкой](#) и инструментами поиска.

В сводных таблицах задокументированы не только все словоформы звуковых глаголов, представленные в [Грамматическом словаре А.А. Зализняка](#) и НКРЯ, но и выявлены существующие дублиеты, внутриязыковые лакуны, указаны новые (потенциальные) единицы.

---

Результаты данного исследования могут быть полезны в разработке различных веб-приложений для поиска, сбора и визуализации лингвистического материала различных объёмов и степеней сложности. Возможности комбинаторной обработки открытых и закрытых баз данных могут быть особенно значимы при извлечении информации из цифровых лексикографических источников (на одном или нескольких языках), из национальных языковых корпусов, а также из электронных текстовых коллекций.

### **Disciplines**

Russian Linguistics

Ivliyeva, Irina V., Koob, Perry. **Multi-dimensional scaling of web-scraping results from the A.A Zalizniak Grammatical Dictionary and the Russian National Corpus. Creating a corpus fragment of all possible word-forms of modified Russian sound verbs using web-scraping methodology. Compilation of a [summary table](#) for the present tense, past tense, future tense, imperative, imperfective, and perfective gerund forms.** August 2023 – February 2024. Working theses. March 1, 2024.

This project attempts not only to improve the method of web extraction in relation to the source material (lexical-semantic group of Russian sound verbs, semantically modified at the word-forming level) but also systematize the search results using the format that is convenient and effective for a linguistic researcher.

The main task at this stage of the research was to collect, compile (combine results), analyze and present a summary index of all possible verbal forms of sound verbs, recorded in electronic versions of dictionaries of the Russian language and in the database of the [Russian National Corpus](#) (RNC).

The output takes into account not only all documented modifications of sound verbs from the [A.A Zalizniak Grammatical Dictionary](#) and the [Russian National Corpus](#), but also reveals lacunae, doublets, and indicates new (potential) units.

The results of the study may be useful for the development of various web applications for the search, collection, and visualization of linguistic material of different volumes and degrees of complexity. Possibilities of combinatorial optimization in the processing of open and closed linguistic databases can be particularly important when extracting information from various digital lexicographic sources (across a single or multiple languages), from national linguistic corpora, as well as from digital text collections.

Ивлиева, И.В., Куб, Перри. **Многомерное шкалирование результатов веб-извлечения из Грамматического словаря А.А. Зализняка и базы данных Национального корпуса русского языка (НКРЯ). Создание фрагмента корпуса словоформ модифицированных глаголов звучания русского языка методом веб-извлечения. Составление [сводной таблицы](#) форм настоящего, прошедшего и будущего времени, повелительного наклонения, деепричастий несовершенного и совершенного вида.** Август 2023 – Февраль 2024. Рабочие тезисы. 1 марта 2024.

В данном проекте предпринята попытка не только усовершенствовать методику веб-извлечения применительно к исходному материалу (лексико-семантической группе русских звуковых глаголов, модифицированных на словообразовательном уровне), но и обобщить представление результатов поиска в формате, который максимально удобен и эффективен для исследователя-лингвиста.

Основными задачами на данном этапе исследования были - собрать, скомпилировать (соединить результаты), проанализировать и представить обобщающий индекс всех возможных глагольных форм глаголов звучания,

зафиксированных в электронных версиях словарей русского языка и в базе данных [Национального корпуса русского языка](#) (НКРЯ).

С учетом результатов предыдущих этапов исследований ([2022](#), [2023](#)) была усовершенствована четырехпозиционная система нумерации исходных форм глагола, их словоформ и модифицированных дериватов. Успешно проведено системное извлечение и фиксирование результатов поиска в виде многомерной сводной интерактивной таблицы, единицы которой снабжены отсылками к НКРЯ - представительной коллекции текстов [на русском языке](#) общим объемом *более 2 млрд слов*, оснащенной лингвистической [разметкой](#) и инструментами поиска. В сводных таблицах задокументированы не только все словоформы звуковых глаголов, представленные в [Грамматическом словаре А.А. Зализняка](#) и НКРЯ, но и выявлены существующие дублиеты, внутриязыковые лакуны, указаны новые (потенциальные) единицы.

Результаты данного исследования могут быть полезны в разработке различных веб-приложений для поиска, сбора и визуализации лингвистического материала различных объёмов и степеней сложности. Возможности комбинаторной обработки открытых и закрытых баз данных могут быть особенно значимы при извлечении информации из цифровых лексикографических источников (на одном или нескольких языках), из национальных языковых корпусов, а также из электронных текстовых коллекций.

## I. ЦЕЛИ И ЗАДАЧИ.

Система глагольных форм в русском языке имеет **матричную структуру**. Известно, что современные словари не регистрируют **все** возможные формы слов. Воспроизведение в бумажном издании формы слов даже одной части речи - технически неисполнимая задача. Однако с появлением электронных версий словарей впервые возникает шанс её решить. Зачем исследователям нужны таблицы глагольных форм? Во-первых, чтобы заполнить «матрицы» (таблицы) каждого глагола всеми имеющимися в русском языке актуальными словоформами, которые зафиксированы в лексикографических ресурсах. Во-вторых, чтобы выявить дублиеты, внутриязыковые лакуны, а также указать на потенциальные единицы.

## II. ХРОНОЛОГИЯ ПРОЕКТА.

В период с июня 2023 по февраль 2024 года проводились отладка программы для веб-извлечения, оптимизация представления результатов поиска, расширение диапазона и синхронизация результатов поиска **одновременно** в словаре Зализняка и НКРЯ. Подробнее об этом см.: **Ivliyeva, Irina V., Koob, Perry**. Experimental multi-dimensional scaling of web-scraping results from the A.A Zalizniak Grammatical Dictionary and the Russian National Corpus. Creating a corpus fragment of all possible word-forms of

modified Russian sound verbs using web-scraping methodology. Compilation of a summary table for the present tense, future tense, imperative, imperfective, and perfective gerund forms. April 2022 – May 2023. [Electronic resource]. Working theses. Research Data. June 5, 2023. – Access mode: [https://scholarsmine.mst.edu/research\\_data/11/](https://scholarsmine.mst.edu/research_data/11/)

11 февраля 2024 года была предпринята успешная попытка внедрить окончательно отлаженную версию программы веб-извлечения и получить систематические результаты, которые отражены в данной работе в виде обобщающей **СВОДНОЙ ТАБЛИЦЕ** словоформ глаголов звучания, модифицированных на словообразовательном уровне: формы настоящего времени, прошедшего времени, будущего времени, повелительного наклонения, и их деепричастные формы несовершенного и совершенного вида.

**Пример 1.** Страница глагола АХНУТЬ в электронном словаре А.А. Зализняка <http://morphologija.ru/>

The screenshot shows the website 'Морфология.ру' with a search bar containing the word 'ахнуть' and a red 'Искать' button. Below the search bar, the page displays the following information:

**Словоформы слова "ахнуть"**

ахнуть → **ахнуть** – инфинитив, действ. зал.

Часть речи: инфинитив - **ахнуть**

Часть речи: глагол

	Настоящее время	Прошедшее время	Будущее время	Повелительное наклонение
<b>Я</b>		ахнул; ахнула; ахнуло	ахну	
<b>Ты</b>		ахнул; ахнула; ахнуло	ахнешь	ахни
<b>Он/она</b>		ахнул; ахнула; ахнуло	ахнет	
<b>Мы</b>		ахнули	ахнем	ахнем; ахнемте
<b>Вы</b>		ахнули	ахнете	ахните
<b>Они</b>		ахнули	ахнут	

Часть речи: деепричастие

Несовершенный вид	Совершенный вид
	ахнув; ахнувши

III. **СВОДНАЯ ТАБЛИЦА.** Условные обозначения и соответствия.

1) Колонки.

Колонка	Тип данных	Title
A	Порядковый номер слова или словоформы	Number
B	Исходное слово - центр модификационного гнезда и словообразовательного гнезда (инфинитив)	Root word
C	Модифицированный дериват (инфинитив)	Word (derivative)
D	Личные местоимения	Personal Pronouns
E	Настоящее время (словоформы)	Present tense
F	Прошедшее время (словоформы)	Past tense
G	Будущее время (словоформы)	Future tense
H	Повелительное наклонение (словоформы)	Imperative
I	Деепричастие, несовершенный вид (словоформы)	Imperferctive Gerund
J	Деепричастие, совершенный вид (словоформы)	Perferctive Gerund

2) Система нумерации по позициям в колонке А.

В колонке А имеются четырёхступенчатые/четырёхпозиционные номера.

A	B	C	D	E	F	G	H	I	J
number	root	word	pronoun	present	past	future	imperative	Imp. gerund	Perf. gerund
1.0.0.0	ахать	<a href="#">ахать</a>	--	--	--	--	--	<a href="#">ахая</a>	--

**Позиция 1:** основное слово (исходный немодифицированный глагол, колонка В).

A	B
1.0.0.0	<a href="#">ахать</a>
2.0.0.0	<a href="#">бабахать</a>

**Позиция 2:** модифицированный дериват (модифицированный глагол, колонка С).

A	B	C
1.1.0.1	ахать	<a href="#">ахнуть</a>
1.2.0.0	ахать	<a href="#">взахать</a>
1.3.0.1	ахать	<a href="#">заахать</a>

**Позиция 3:** словоизменительная парадигма (личные местоимения, колонка D и словоформы в колонках E, F, G, H).

А	В	С	Д	Е	Ф	Г	Н
1.1.1.1	ахать	<a href="#">ахнуть</a>	Я	--	<a href="#">ахнул</a>	<a href="#">ахну</a>	--
1.1.1.2	ахать	<a href="#">ахнуть</a>	Я	--	<a href="#">ахнула</a>	<a href="#">ахну</a>	--
1.1.1.3	ахать	<a href="#">ахнуть</a>	Я	--	<a href="#">ахнуло</a>	<a href="#">ахну</a>	--
1.1.2.1	ахать	<a href="#">ахнуть</a>	Ты	--	<a href="#">ахнул</a>	<a href="#">ахнешь</a>	<a href="#">ахни</a>
1.1.2.2	ахать	<a href="#">ахнуть</a>	Ты	--	<a href="#">ахнула</a>	<a href="#">ахнешь</a>	--
1.1.2.3	ахать	<a href="#">ахнуть</a>	Ты	--	<a href="#">ахнуло</a>	<a href="#">ахнешь</a>	--
1.1.3.1	ахать	<a href="#">ахнуть</a>	Он	--	<a href="#">ахнул</a>	<a href="#">ахнет</a>	--
1.1.3.2	ахать	<a href="#">ахнуть</a>	Она	--	<a href="#">ахнула</a>	<a href="#">ахнет</a>	--
1.1.3.3	ахать	<a href="#">ахнуть</a>	Оно	--	<a href="#">ахнуло</a>	<a href="#">ахнет</a>	--
1.1.4.1	ахать	<a href="#">ахнуть</a>	Мы	--	<a href="#">ахнули</a>	<a href="#">ахнем</a>	<a href="#">ахнем</a>
1.1.4.2	ахать	<a href="#">ахнуть</a>	Мы	--	<a href="#">ахнули</a>	<a href="#">ахнем</a>	<a href="#">ахнемте</a>
1.1.5.0	ахать	<a href="#">ахнуть</a>	Вы	--	<a href="#">ахнули</a>	<a href="#">ахнете</a>	<a href="#">ахните</a>
1.1.6.0	ахать	<a href="#">ахнуть</a>	Они	--	<a href="#">ахнули</a>	<a href="#">ахнут</a>	--

**Позиция 4:**

Немодифицированный глагол, деепричастие несовершенного вида, колонка I.

А	В	С	Д	Е	Ф	Г	Н	І	Ј
1.0.0.0	ахать	<a href="#">ахать</a>	--	--	--	--	--	<a href="#">ахая</a>	--

Модифицированный глагол, деепричастия совершенного вида, колонка J).

А	В	С	Д	Е	Ф	Г	Н	І	Ј
1.3.0.1	ахать	<a href="#">заахать</a>	--		--	--	--	--	<a href="#">заахав</a>
1.3.0.2	ахать	<a href="#">заахать</a>	--		--	--	--	--	<a href="#">заахавши</a>

**3) Двойной прочерк (--)** в графе показывает, что в электронном словаре

А.А. Зализняка <http://morphologija.ru/> данное слово (или его словоформа) не обнаружены.



#### IV. МНОГОМЕРНОЕ ШКАЛИРОВАНИЕ.

В период с июня 2023 года по февраль 2024 года была также успешно выполнена задача создания алгоритма поиска (выборки) каждой словоформы звуковых глаголов, зарегистрированной словаре Зализняка, и одновременного соединения результатов **В СВОДНОЙ ТАБЛИЦЕ** с результатами поиска в базе данных Национального корпуса русского языка. Также были отшлифованы принципы реализации интерактивной ссылки (многомерное шкалирование результатов веб-извлечения) для каждой словоформы в сводной таблице.

Таким образом, **В СВОДНОЙ ТАБЛИЦЕ** каждая словоформа *будущего, прошедшего и настоящего времени, повелительного наклонения*, а также *деепричастные* формы глаголов звучания сопровождаются **электронной сноской** на результаты поиска в НКРЯ. Исследователь нажимает на словоформу мышкой и попадает на соответствующую страницу поиска в [Национальном корпусе русского языка](#). Все элементы, представленные в колонках С, Е, F, D, H, I, J снабжены электронными отсылками к Русскому Национальному Корпусу <https://ruscorpora.ru/>.

**Пример 2.** Колонка G, будущее время, словоформа **ахну**.

A	B	C	D	E	F	G
Number	Root	Word	Pronoun	Present	Past	Future
1.1.1.1	ахать	<a href="#">ахнуть</a>	Я	--	<a href="#">ахнул</a>	<a href="#">ахну</a>
1.1.1.2	ахать	<a href="#">ахнуть</a>	Я	--	<a href="#">ахнула</a>	<a href="#">ахну</a>

Данная словоформа снабжена отсылкой к Национальному Корпусу русского языка <https://ruscorpora.ru/>:

**ахну** =HYPERLINK("https://ruscorpora.ru/explore?req=ахну","ахну")

При нажатии на слоформу **ахну** в таблице производится отсылка к следующим результатам поиска в НКРЯ:

The screenshot shows the main search interface of the NKRYA. At the top, the word 'ахну' is entered in the search bar. Below the search bar, there are two main sections: 'Поиск точных форм в выбранном корпусе' and 'Лексико-грамматический поиск в корпусе'. The first section shows 20 texts and 22 examples, while the second shows 1 text and 1 example. A blue arrow points to the search bar area.

Корпус *Основной* • 131 488 текстов • 374 449 975 слов  
омонимия снята, синтаксически размечен

Запрос  • 47 корпусов • 5 видов поиска

**Поиск точных форм в выбранном корпусе** ?

• 20 текстов • 22 примера  
[Показать все примеры](#)

**Лексико-грамматический поиск в корпусе** ?

• 1 текст • 1 пример  
[Показать все примеры](#)

При нажатии на «Уточнить запрос» производится отсылка к количественным и качественным результатам поиска:

The screenshot shows the 'Подкорпус' (Subcorpus) search interface. The search bar contains the word 'ахну'. Below the search bar, there are three search options: 'Поиск точных форм', 'Поиск коллокаций', and 'Лексико-грамматический поиск'. A blue arrow points to the search bar area.

НАЦИОНАЛЬНЫЙ КОРПУС русского языка

Корпус *Основной* • 131 488 текстов • 374 449 975 слов  
омонимия снята, синтаксически размечен

Подкорпус

Поиск точных форм ? ^

Слово или фраза

Поиск коллокаций ? v

Лексико-грамматический поиск ? v

В 2020—2023 годах Корпус развивается благодаря поддержке гранта МОН № 075-15-2020-793 «Компьютерно-лингвистическая платформа нового поколения для цифровой документации русского языка: инфраструктура, ресурсы, научные исследования».

При нажатии на «ИСКАТЬ» выдаётся следующий результат:

The screenshot shows the NCRJ search interface. At the top, it displays 'Корпус' (Corpus) with 'Основной' (Main) selected, and statistics: '131 488 текстов • 374 449 975 слов' (131,488 texts • 374,449,975 words). Below this is a search bar with 'Задать' (Set) and a 'Подкорпус' (Subcorpus) section. The search query is 'Запрос' (Query) with 'Вернуться к поиску' (Return to search) and '20 текстов • 22 примера' (20 texts • 22 examples). A button 'Добавить в сравнение' (Add to comparison) is visible. The search results section shows 'Конкорданс KWIC График Статистика Частотность 1-граммы 2-граммы 3-граммы' (Concordance KWIC Graph Statistics Frequency 1-grams 2-grams 3-grams). The first result is: '1. Николай Солодников, Тамара Эйдельман, Сергей Бунтман. Как правильно преподавать историю в школе? (2017.09)'. Below the result, there is a snippet: 'А если я не ахаю, я потом не ахну, когда увижу списки погибших каких-нибудь или чей-то подвиг или чью-то гнусность в истории.' A blue arrow points to the 'Искать' (Search) button.

Поскольку НКРЯ является закрытой системой, изменений статистических результатов поиска не предполагается. Трудностей доступа при помощи компьютерной программы (веб-извлечения) к Национальному корпусу русского языка не обнаружено.

## V. ПРИЛОЖЕНИЯ

### Appendix 1. [Приложение 1 Words not in morfologija-2024-02-11](#)

Список единицы, которые не зарегистрированы в электронном словаре А.А. Зализняка <http://morfologija.ru/>

### Appendix 2. [Приложение 2 Double pronoun dual action - 2024-02-11](#)

Совпадение форм будущего времени и императива глаголов звучания, или форм совместного действия.

### Appendix 3. [Приложение 3 Double perfective gerund- 2024-02-11](#)

Омонимия форм модификаций у глаголов близкой семантики (при отсутствии указания ударения в исходном слове в источнике поиска). Совершенный вид.

### Appendix 4. [Приложение 4 Double imperfective gerund- 2024-02-11](#)

Омонимия форм модификаций у глаголов близкой семантики (при отсутствии указания ударения в исходном слове в источнике поиска). Несовершенный вид.

**Примечание.** В приложениях нет электронных ссылок на НКРЯ. Обращайтесь к [СВОДНОЙ ТАБЛИЦЕ](#).

## VI. ПУБЛИКАЦИИ АВТОРОВ ПРОЕКТА ПО ДАННОЙ ТЕМАТИКЕ (2019-2024)

**Ivliyeva, Irina V., Koob, Perry.** Experimental multi-dimensional scaling of web-scraping results from the A.A Zalizniak Grammatical Dictionary and the Russian National Corpus. Creating a corpus fragment of all possible word-forms of modified Russian sound verbs using web-scraping methodology. Compilation of a summary table for the present tense, future tense, imperative, imperfective, and perfective gerund forms. April 2022 – May 2023. [Electronic resource]. Working theses. Research Data. June 5, 2023. – Access mode: [https://scholarsmine.mst.edu/research\\_data/11/](https://scholarsmine.mst.edu/research_data/11/)

**Ивлиева И.В.** Принципы лексикографического описания терминологии словообразовательного синтеза (на материале глаголов звучания русского языка) // Тема номера. *Интерактивная наука*. – 2023. – 6 (82). – С. 7-15. – ISSN 2414-9411. – DOI 10.21661/r-560185. [Translation: Principles of lexicographic description for word formation synthesis terminology (on the material of verbs of sound in the Russian language)]. [Electronic resource]. Feature article. *Interactive science*, 2023, 6 (82). P. 7-15. Access mode: [https://interactive-science.media/ru/article/560185/discussion\\_platform?utm\\_source=ticket&utm\\_medium=email&utm\\_campaign=request\\_onsite&utm\\_term=ru&utm\\_content=discussion\\_platform](https://interactive-science.media/ru/article/560185/discussion_platform?utm_source=ticket&utm_medium=email&utm_campaign=request_onsite&utm_term=ru&utm_content=discussion_platform)

**Ivliyeva, Irina V., Koob, Perry.** “Creating a corpus of all possible word forms of modified Russian sound verbs using web-scraping methodology. Compilation and adjustment of summary tables for the future tense, imperative, and gerund forms. Forms of “dual action”. Experimental multi-dimensional scaling of web-scraping results.” (2022). [Electronic resource]. Research Data. March 2022. – Access mode: [https://scholarsmine.mst.edu/research\\_data/9/](https://scholarsmine.mst.edu/research_data/9/)

**Ивлиева И.В.** Развитие терминологии словообразовательного синтеза. Особенности терминосистемы (на материале глаголов звучания русского языка)// *Интерактивная наука*. – 2022. – № 6 (71). – С. 86 -94. – ISSN 2414-9411. DOI 10.21661/r-557017. [Electronic resource]. - Access mode: [https://interactive-science.media/ru/article/557017/discussion\\_platform](https://interactive-science.media/ru/article/557017/discussion_platform)

**Ивлиева И.В.** Потенциальные единицы глагольного синтеза: свойства и возможности (на материале русских глаголов звучания) // *Интерактивная наука*. – 2021. – № 5 (60). – С. 11-19. – ISSN 2414-9411. [Electronic resource]. - Access mode: [https://interactive-science.media/ru/article/554295/discussion\\_platform](https://interactive-science.media/ru/article/554295/discussion_platform)

**Ivliyeva, I., Koob, Perry.** “Method of web-extraction (web scraping) of Russian verb paradigms from electronic dictionaries and databases. Matrix organization of lacunae,

their codification and classification (on the material of the verbs of sound)” (2021). [Electronic resource]. Research Data. 7. – Access mode: [https://scholarsmine.mst.edu/research\\_data/7](https://scholarsmine.mst.edu/research_data/7)

**Ivliyeva, Irina V., Koob, Perry.** “Lacunae Matrices, Web Scraping. Theses.” (2021). Russian Linguistics Research. 1. [Electronic resource]. – Access mode: [https://scholarsmine.mst.edu/russian\\_linguistics\\_research/1](https://scholarsmine.mst.edu/russian_linguistics_research/1)

**Koob, Perry B., and Ivliyeva, Irina V.** "Russian National Corpus Web Scraping Project 2019-2020." [Electronic resource]. February 1, 2020. – Access mode: [https://scholarsmine.mst.edu/artlan\\_phil\\_facwork/158/](https://scholarsmine.mst.edu/artlan_phil_facwork/158/)

**Ивлиева И.В.** Особенности семантики лакун-модификаций. К идее создания словаря лакун (на материале глаголов звучания) // Интерактивная наука. – 2020. – № 6 (52). – С. 57-70. – ISSN 2414-9411. [Electronic resource]. - Access mode: [https://interactive-science.media/ru/article/551671/discussion\\_platform](https://interactive-science.media/ru/article/551671/discussion_platform)

**Ивлиева И.В.** Роль имплицитных X-компонентов значения в синтезе русских глагольных модификаций (на примере глаголов механической фонации) // Интерактивная наука. 2019. – № 6 (40). С. 106-113. – ISSN 2414-9411. Access mode: [https://interactive-science.media/en/article/497213/discussion\\_platform](https://interactive-science.media/en/article/497213/discussion_platform)

## **VII. ЭЛЕКТРОННЫЕ ЛЕКСИКОГРАФИЧЕСКИЕ РЕСУРСЫ**

**Грамматический словарь А.А. Зализняка.** The grammatical dictionary of the Russian language by A.A. Zalyzniak, <https://morfologija.ru/>. Accessed February 11, 2024.

**Национальный корпус русского языка.** The Russian National Corpus, <https://ruscorpora.ru/>. Accessed February 11, 2024.

## **VIII. ИССЛЕДОВАТЕЛЬСКАЯ ГРУППА, КОНТАКТЫ**

За подробностями о компьютерной программе (web-scraping algorithm) и по вопросам методологического характера просьба обращаться непосредственно к научно-исследовательской группе:

**Ирина В. Ивлиева.** **Ivliyeva, Irina V.** [ivliyeva@mst.edu](mailto:ivliyeva@mst.edu), профессор кафедры искусств, иностранных языков и философии, Missouri University of Science and Technology.

**Перри Куб.** **Koob, Perry** [koobp@mst.edu](mailto:koobp@mst.edu), эксперт-администратор баз данных, Missouri University of Science and Technology.