



www.volsu.ru

РАЗВИТИЕ И ФУНКЦИОНИРОВАНИЕ РУССКОГО ЯЗЫКА

DOI: <https://doi.org/10.15688/jvolsu2.2025.5.1>

UDC 81'373.21

LBC 81.053.16



Submitted: 29.05.2024

Accepted: 31.03.2025

ONOMASTIC CONTENT OF GEOGRAPHICAL NORMATIVE SCIENTIFIC TEXTS: A QUANTITATIVE APPROACH

Sergey N. Andreev

Smolensk State University, Smolensk, Russia

Nina A. Maksimchuk

Smolensk State University, Smolensk, Russia

Abstract. The article focuses on topical issues of statistical processing of educational books, which might contribute to their content peculiarities identification. The aim of this interdisciplinary research, within the framework of which this study is conducted, is to determine the role and place of proper names in the structure of universally mandatory knowledge and its individual fragments. The analysis undertaken resulted in identifying and structuring the onomastic space of geographical normative scientific macrotexts. The application of quantitative analysis revealed the composition of the onomastic core and periphery of these macrotexts, as well as the basic and variable components of onyms and their distribution across the text. The composition and frequency of toponyms in geographical normative scientific texts are noted to serve as markers of the content of the geographical excerpt of the normative scientific picture of the world (universally mandatory geographical knowledge). The findings allow for conclusions not only about the composition and structure of the onomastic content of normative scientific macrotexts but also about the degree of its stability over time. The proposed methodology for determining the onomastic (informational) load of a text can be applied to the analysis of textbooks in other subjects. The analysis applied to onomastic content of geographical normative scientific texts, which forms particular onomastic space, has significant practical relevance, related to the objectification of expert evaluations of existing textbook texts and the prospect of developing recommendations for newly created ones.

Key words: proper name, universally mandatory scientific knowledge, normative scientific text, onomastic load of a text, onomastic background, quantitative analysis.

Citation. Andreev S.N., Maksimchuk N.A. Onomastic Content of Geographical Normative Scientific Texts: A Quantitative Approach. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 2. Yazykoznanie* [Science Journal of Volgograd State University. Linguistics], 2025, vol. 24, no. 5, pp. 6-20. (in Russian). DOI: <https://doi.org/10.15688/jvolsu2.2025.5.1>

ОНОМАСТИЧЕСКОЕ НАПОЛНЕНИЕ ГЕОГРАФИЧЕСКИХ НОРМАТИВНО-НАУЧНЫХ ТЕКСТОВ: КВАНТИТАТИВНЫЙ ПОДХОД

Сергей Николаевич Андреев

Смоленский государственный университет, г. Смоленск, Россия

Нина Алексеевна Максимчук

Смоленский государственный университет, г. Смоленск, Россия

Аннотация. Исследование посвящено решению актуальных задач определения содержательных характеристик текстов учебной литературы с применением методов статистической обработки их лексического наполнения. Работа выполнена в русле междисциплинарных исследований, нацеленных на установление роли и места имен собственных в структуре общеобязательного знания и его отдельных фрагментов. В результате проведенного анализа выделено и структурировано ономастическое пространство географических нормативно-научных макротекстов. С применением квантитативного анализа выявлены состав ономастического ядра и периферии данных макротекстов, распределение онимов по тексту. Показано, что состав и частотность топонимов в географических нормативно-научных текстах служат маркерами содержания географического фрагмента нормативно-научной картины мира (общеобязательного географического знания). Полученные результаты позволяют сделать выводы не только о составе и структуре ономастического наполнения нормативно-научных макротекстов, но и о степени его устойчивости во времени. Предлагаемая методика оценки ономастической (информационной) нагруженности текста может быть использована для анализа учебников по другим предметам. Анализ ономастического наполнения географических нормативно-научных текстов, формирующего определенное ономастическое пространство, имеет практическую значимость, связанную с объективацией экспертизы текстов существующих учебников и с перспективой разработки рекомендаций для вновь создаваемой учебной литературы.

Ключевые слова: имя собственное, общеобязательное научное знание, нормативно-научный текст, ономастическая нагруженность текста, ономастический фон, квантитативный анализ.

Цитирование. Андреев С. Н., Максимчук Н. А. Ономастическое наполнение географических нормативно-научных текстов: квантитативный подход // Вестник Волгоградского государственного университета. Серия 2, Языкознание. – 2025. – Т. 24, № 5. – С. 6–20. – DOI: <https://doi.org/10.15688/jvolsu2.2025.5.1>

Введение

Представляемое исследование выполнено в рамках разработки концепции изучения языкового выражения содержания общеобязательного научного знания. Научный фрагмент общеобязательного знания человека (нормативно-научная картина мира, далее – ННКМ) складывается из научных и научно-практических сведений, усвоению которых в значительной степени посвящен процесс образования. Особенность этого фрагмента состоит в том, что он складывается не стихийно, а в результате целенаправленного сообщения отобранных и одобренных обществом научных сведений, овладение которыми регламентируется социумом и закрепляется в

нормативных документах (образовательных стандартах, программах, учебниках и т. д.). Основным средством трансляции общеобязательных научных сведений служат нормативно-научные тексты (далее – ННТ) – специально созданные и принятые обществом научные тексты, функционирующие в учебной литературе. Основной формой представления научного знания в учебных целях служат учебники как нормативно-научные макротексты. Одним из двух (наряду с термином) опорных элементов нормативно-научной картины мира, существующей в сознании носителей научного знания, является имя собственное (далее – ИС).

Предмет исследования в данной статье – ономастическое наполнение (топонимы) нор-

мативно-научных текстов учебников географии, цель – разработка методики определения ядерного и периферийного компонентов общеобязательного научного знания путем со-поставительного анализа структуры ономастического наполнения школьных учебников географии разных временных периодов.

Работа имеет междисциплинарный характер, поскольку поставленная цель объединяет несколько смежных дисциплин: ономастику (языковой материал представлен ономастической лексикой), когнитивную лингвистику (имя собственное рассматривается как носитель общеобязательного научного знания, которое влияет на формирование личности), теорию учебно-научного текста (имя собственное функционирует в нормативно-научных текстах, формируя их информационную насыщенность), прикладную лингвистику.

Теоретическое значение исследования связано с тем, что оно позволит расширить представление об оимах как носителях общеобязательного научного знания и ономастическом наполнении нормативно-научных текстов как ономастическом пространстве, его практическая значимость обусловлена возможностью использования представленных методов при анализе и экспертной оценке нормативно-научных текстов действующих учебников.

Ономастическое наполнение нормативно-научных текстов определенной тематической отнесенности (в том числе – географической) формирует своего рода ономастическое пространство. Современные философские теории пространства рассматривают фундаментальные измерения человеческого мира, выражаемые в специфических нестрогих понятиях верх – низ, центр – окраина, поверхность – глубина, удаленность – близость и др., которые фиксируются уже на уровне обыденного сознания и отражают то, в каких именно координатах и векторах описывается динамика и статика человеческого существования [Философия..., 2004]. Очевидно, что значительная часть пространственных координат и векторов в их вербальном выражении будет иметь, помимо нарицательных обозначений, собственные наименования, что позволяет применить к термину пространство уточняющее определение.

В лингвистике появление термина ономастическое пространство связано с именами В.Н. Топорова и А.В. Суперанской. В.Н. Топоров трактует (топ)ономастическое пространство как совокупность принадлежащих ему собственных наименований [Топоров, 2004]. А.В. Суперанская, понимая ономастическое пространство как совокупность всех имен собственных, употребляющихся в определенном социуме в определенную историческую эпоху, подчеркивает, что его объем и характер могут меняться [Суперанская, 1973].

Используя понятие ономастического пространства в целях нашего исследования, мы обозначаем этим термином совокупность ономастических единиц нормативно-научных текстов разных уровней – от ННТ одного раздела до ННТ отдельного учебника (макротекста) или всей учебной линии – предметного гипертекста [Максимчук, Голубничая, 2018]. Анализ ономастического пространства любого ННТ может быть комплексным, одновременно охватывая оими всех представленных в нем разрядов, или последовательным, проводимым по отдельным разрядам. В данной работе детально анализируются топонимы (конкретно – названия стран) как основной разряд оимов в географических текстах – носителях базового географического знания.

Поскольку наполнение ономастического пространства как составной части (фрагмента) лексической системы языка позволяет квалифицироваться в терминах теории поля [Супрун, 2000; Максимчук, 2017; и др.], то применение квантитативного анализа частот имен собственных позволяет определить состав ядра и периферии данной ономастической структуры, уровень информационной нагруженности одноименных нормативно-научных текстов и, следовательно, выявить степень сходства и различия в представлении базового географического знания в учебниках разных поколений.

В перспективе комплексный анализ ономастического наполнения нормативно-научных текстов как носителей существенного объема общеобязательного научного знания в сочетании с корпусным анализом других характеристик учебных текстов (см., например: [Куприянов, Солнышкина, Лехницкая, 2023; Монахов, Турчаненко, Чердаков, 2023;

Лапошина, Лебедева, 2024; и др.]) и методом экспертных оценок позволит выработать концепцию построения оптимального учебного текста с учетом его содержательных задач, потребностей и возможностей адресата.

Материал и методы

Материалом для проведения исследования послужили ННТ учебников географии. Выбор обусловлен несколькими причинами. География относится к традиционным предметам общего образования, что свидетельствует о понимании обществом важности формирования системы географических знаний как компонента научной картины мира для ориентации человека в окружающем его физическом (материальном) мире в целом и мире разных стран и народов в частности. В современной ситуации, характеризующейся повышенной geopolитической турбулентностью, значительными изменениями в природной географической среде, освоение молодыми людьми базового географического знания может способствовать, согласно нормативным образовательным документам, формированию целостного представления о роли России в современном мире, ценностных ориентаций личности. Кроме того, в структуре географических знаний имя собственное занимает особое место, поскольку все географические объекты, известные человеку, имеют индивидуальные наименования – топонимы, составляющие один из самых многочисленных разрядов ономастической лексики.

При определении источников материала исследования (учебников) учитывались следующие параметры. Во-первых, учебники должны быть рекомендованы к использованию высшими образовательными институтами, чем подтверждается их статус носителей общеобязательного научного знания, структура которого закреплена в нормативных документах. Во-вторых, они должны быть существенно разделены во времени, что позволяет рассматривать их содержание в динамическом аспекте, оценивая степень его стабильности и актуальности. В-третьих, источники должны принадлежать разным авторам, что обеспечивает, на наш взгляд, большую объективность выявляемых сходства и раз-

личий в структуре представляемого общеобразовательного научного знания.

Источниками послужили учебники географии для 10 класса 1991 г. (ГЕО-1) и 10–11 классов 2018 г. (ГЕО-2), отвечающие всем трем требованиям.

В работе использованы описательный, структурный, сравнительно-сопоставительный методы, квантитативный анализ (сбор, подготовка, преобразование данных, кодирование, создание базы данных, визуализация результатов, статистический анализ).

Отбор имен собственных из текстов проводился программными средствами с использованием кода на Python и двух модулей: регулярных выражений (Regular Expressions) и спейси (spaCy).

Рассматривались все разряды и структурные типы онимов, включая аббревиатуры. Не учитывались иноязычные имена собственные, приводимые на латинице (их количество в исследуемых массивах ограничивается несколькими единицами).

Изучение ономастической структуры географических ННТ проводилось при помощи квантитативного анализа частот. Были рассмотрены следующие аспекты: 1) уровень ономастической насыщенности текста и частотное ядро ономастической структуры; 2) периферия ономастической структуры; 3) характер распределения частот вводимых в учебный текст онимов от его начала к концу; 4) использование ономастической системы для описания ведущих в политико-экономическом отношении стран мира. На всех этапах исследования анализ проводится как в рамках отдельных учебников, так и в их сопоставлении друг с другом.

Результаты

Уровень ономастической насыщенности текста и частотное ядро ономастической структуры

В результате анализа были получены данные о количестве онимов и их соотношении с общим объемом текста, что позволило определить уровень ономастической насыщенности текста (далее – ОНТ). Под ономастической насыщенностью понимается количе-

ство ономастических единиц на единицу текста. В свою очередь, единицей текста считается завершенный текстовый отрезок любой протяженности, объем которого в данном случае диктуется особенностями учебно-образовательных (нормативно-научных) требований к полноте раскрытия содержания текста, обозначенного его заглавием.

Определение уровня ОНТ достигается путем установления процентного отношения проприальной лексики к общему количеству лексических единиц рассматриваемого текстового массива.

Нами были установлены следующие значения: ОНТ в ГЕО-1 составляет 6,56 %, в ГЕО-2 немногим больше – 7,82 %. Таким образом, уровень ономастической насыщенности текста в обоих источниках сходный: на каждые 15 (ГЕО-1) и 13 (ГЕО-2) слов приходится один оним.

Очевидно, что общий уровень ОНТ зависит не только от количества различных онимов, используемых в тексте, но и от их повторяемости – частотности. В таком случае для характеристики ономастического наполнения сопоставляемых текстов важным оказывается вопрос о влиянии на уровень их ономастической насыщенности конкретных ономастических единиц.

Вопросы о критериях разграничения частотной и малочастотной лексики и выявление границы между этими двумя частотными областями лингвистических единиц в настоящее время часто решается при помощи использования критерия Хирша, широко применяемого в наукометрии [Hirsch, 2005]. Этот критерий основан на поиске такой лексической единицы, частота которой совпадает с ее рангом [Popescu, 2007; Popescu, Altmann, 2006]. Использование данного критерия позволило получить значимые результаты в области изучения соотношения синсемантической и автосемантической лексики, структуры текста, при исследовании вопросов золотого сечения и др. [Andreev, Místecký, Altmann, 2018, с. 93–95; Místecký. 2018; Pan, Qiu, Liu, 2015; Zhang, Liu, 2020].

В качестве примера приведем данные по одному из разделов ГЕО-2. Ранжирование имен собственных, используемых в последней теме (главе) этого учебника, обнаруживает,

что они образуют последовательность по убыванию частот: *Земля* (ранг 1; частота 12), *США* (2; 12), *Россия* (3; 11), *Мировой океан* (4; 10), *ООН* (5; 9), *Африка* (6; 8), *Азия* (7; 7), *Европа* (8; 5) и т. д. Согласно этому критерию точка отсечения (H-point) устанавливает границу, которая проходит по ониму *Азия*, имеющему частоту 7, совпадающую с его рангом. В тех случаях, когда полного совпадения нет, предложена методика определения условной границы [Maćutek, Popescu, Altmann, 2007]. Для приведенной последовательности привлечение этого дополнительный критерий не потребовалось.

Пороговым значением для ранжированного по частоте ряда имен собственных ГЕО-1 стал оним *Европа* с частотой и рангом 26. На 26 онимов частотной части списка приходится 42 % всех словоупотреблений ИС в учебнике, причем на первые 15–32 %.

В ГЕО-2 граница проходит по слову *Лондон* с частотой и рангом, равными 30. Слова частотного списка охватывают 42 % словоупотреблений всех онимов в ГЕО-2, первые 15 слов – 36 %.

Рассмотрим этот показатель для 15 наиболее частотных онимов в обоих источниках, поскольку на эти онимы приходится около трети всех словоупотреблений ИС. Состав и частотный ранг данных единиц являются весьма важными показателями для характеристики содержательной стороны текста. Полученные данные представлены в таблице 1, в которой также приводится процентное выражение использования данного онима и его ранг по частотности.

Как видно из таблицы, в ядро высокочастотных онимов входят названия ведущих в экономическом и политическом отношении стран, а также частей света. При этом данные обоих учебников почти полностью совпадают. Единственное исключение состоит в том, что в ГЕО-1 в число самых частотных попал оним *Земля*, а в ГЕО-2 его место занял оним *Бразилия*. Однако и здесь изменения крайне незначительные: оним *Земля* в ГЕО-2 стал 17 по рангу, а оним *Бразилия* с 16 ранга в ГЕО-1 поднялся на 12 ранг в ГЕО-2.

Совпадение обнаруживается и по рангам: 4 из 15 наиболее частотных онимов – *США*, *СССР / Россия*, *Индия*, *Азия* – имеют полное

Таблица 1. Ядро ономастической структуры в сопоставляемых учебниках**Table 1. The onomastic structure core in compared textbooks**

ГЕО-1			ГЕО-2		
Ранг	Оним	%	Ранг	Оним	%
1	<i>США</i>	5,14	1	<i>США</i>	5,73
2	<i>СССР</i>	3,21	2	<i>Россия</i>	4,30
3	<i>Япония</i>	3,01	3	<i>КНР</i>	3,40
4	<i>Африка</i>	2,42	4	<i>Европа</i>	2,78
5	<i>Западная Европа</i>	2,28	5	<i>Япония</i>	2,61
6	<i>Индия</i>	1,93	6	<i>Индия</i>	2,39
7	<i>Канада</i>	1,84	7	<i>Африка</i>	2,17
8	<i>Азия</i>	1,69	8	<i>Азия</i>	2,17
9	<i>Австралия</i>	1,64	9	<i>Канада</i>	1,83
10	<i>ФРГ</i>	1,64	10	<i>Латинская Америка</i>	1,55
11	<i>Земля</i>	1,62	11	<i>ФРГ</i>	1,46
12	<i>Франция</i>	1,52	12	<i>Бразилия</i>	1,43
13	<i>Латинская Америка</i>	1,52	13	<i>Австралия</i>	1,40
14	<i>Великобритания</i>	1,30	14	<i>Франция</i>	1,39
15	<i>КНР</i>	1,15	15	<i>Великобритания</i>	1,27

ранговое соответствие, а ФРГ и Великобритания занимают смежные позиции в сопоставляемых текстах. Если опустить это единственное несовпадение по лексическому составу, то для остальных единиц обоих списков коэффициент ранговой корреляции Спирмена показывает большое сходство $R = 0,71$ (статистически значим для $p < 0,05$). Причем в четырех случаях (*США, СССР / Россия, Индия, Азия*) имеет место и сходство рангов. Таким образом, на уровне частотной лексики ономастическое наполнение географических ННТ в высшей степени стабильно.

В текстах обоих рассматриваемых источников лидируют онимы *США – Россия / СССР*, однако, как будет показано ниже, ономастическая структура представления этих стран значительно различается, что связано уже с другой характеристикой текста. Кроме них в частотный список онимов (выше точки Хирша) входят все страны – основатели БРИКС и все страны «Большой семерки».

Состав, а также количество употребляемых в географических текстах онимов и их частотность в определенной степени являются отражением объективного места объекта ономастической номинации на географической и экономической картах мира (и, следовательно, в структуре географического знания). Полученные данные о частотах, употребленных в географических текстах имен собственных, можно рассматривать как предварительные ориентиры в содержании географического

фрагмента ННКМ (общеобязательного географического знания).

Характеризуя совокупность наиболее частотных онимов анализируемых географических макротекстов не только как ономастическое ядро их лексического наполнения, но и как центральную часть общеобязательного географического знания, обратим внимание на другие компоненты данной полевой структуры и, прежде всего, на ее периферию.

Периферия ономастической структуры

При изучении частотных показателей значение имеют не только высокочастотные, но и малочастотные единицы – гапаксы (см. об этом: [Popescu, Altmann, 2006]). Мы обнаружили имена собственные, используемые в учебнике один раз (Нарах legomena): *Баотоу, Брисбен, Корба, Науру, Того, Эссен* и др. (ГЕО-1); *Анхой, Бурунди, Лхасой, Теночтитлан, Цзянсу* и др. (ГЕО-2), и имена собственные, которые используются в учебнике два раза (Нарах dislegomena): *Бокаро, Лагос, Оман* и др. (ГЕО-1), *Гуандун, Луанда, Свазиленд* и др. (ГЕО-2).

В таблице 2 приводятся данные о количестве таких единиц в процентах от всех ИС и указано количество словоупотреблений этих малочастотных слов относительно всех словоупотреблений ИС.

По этим показателям оба учебника снова представляются достаточно сходными, хотя в случае с гапаксами речь идет не о сход-

*Таблица 2. Процентный состав гапаксов в текстах, %**Table 2. Percentage composition of hapaxes in texts, %*

Тип ИС	Лексические единицы		Словоупотребления	
	ГЕО-1	ГЕО-2	ГЕО-1	ГЕО-2
Нарах legomena	47,47	55,83	8,74	9,76
Нарах dislegomena	15,43	14,27	5,68	4,99

стве списков лексических единиц, а о том месте, которое они занимают в текстах. Согласно данным таблицы 2, в обоих учебниках около половины всех онимов употребляется только один раз: в ГЕО-1 – несколько меньше половины, в ГЕО-2 – несколько больше. Если к этим данным добавить данные об ИС с частотой использования, равной двум, то остальная проприальная лексика составит менее 40 % в ГЕО-1 и менее 30 % в ГЕО-2 от общего количества всех онимов. Однако частотные онимы охватывают около 85 % всех словоупотреблений ИС в обоих учебниках.

Таким образом, содержательная (познавательная) значимость частотных ИС в формировании фрагмента географического знания подтверждается объективными (количественными) показателями, а сами онимы из приведенного списка можно квалифицировать как ключевые имена (слова) данных географических макротекстов.

Принципиальная правомерность отнесения имен собственных к разряду ключевых слов географического НИТ, помимо признания имени собственного одним из основных носителей общеобязательного знания, поддерживается положением о том, что наиболее значимыми в тексте являются те слова, которые позволяют представить текст в свернутом виде [Новиков, 2007, с. 56]. Логично предположить, что указанная последовательность отношений должна быть справедлива и в обратном направлении: признанное в современной ономастике рассмотрение имени собственного как свернутого текста это подтверждает.

Ключевые слова в значительной степени характеризуют понятие информативности текста, под которой понимается системный признак (категория) текста, «связанный со свойством текста фиксировать знания о мире, отражающие авторское мировосприятие, выраженное в конкретной речевой форме» [Жеребило, 2011]. Однако для имен собственных считаем недостаточным связывать информа-

тивность текста только с ключевыми (частотными) онимами, тем более что в абсолютном выражении в географических НИТ преобладают онимы средней и малой частотности.

В структуре ономастического наполнения географических текстов эти онимы составляют ближнюю и дальнюю периферию, располагающиеся в границах между ядром и зоной гапаксов и составляющие зону информационного (ономастического) расширения. В нее входят ИС с показателями частотности 3–4, тематически (содержательно) соотносящиеся с одной из единиц ядра. Единицы зоны ономастического расширения находятся с ключевым ономом в отношениях включения: страны – моря – реки – горы – равнины – города и т. д.

Зона гапаксов, объединяющая Нарах legomena (с учетом или без Нарах dislegomena), составляет так называемый ономастический фон, участвующий в формировании географического знания, основным манифестантом которого является тот или иной ключевой оним текста (определение степени участия конкретных онимов зоны информационного расширения и зоны гапаксов в формировании предметного знания относится к перспективам исследования). В отличие от единиц зоны расширения, гапаксы могут иметь в разной степени отдаленные, но прямые связи с ключевыми онимами или не иметь таких связей. Однако в любом случае количество онимов-гапаксов в составе ономастического фона НИТ в значительной степени и определяет уровень ономастической (информационной) нагруженности текста.

Распределения частот онимов от начала текста к его концу

Особенность оценки ОНТ состоит в том, что она может устанавливаться в статическом и динамическом аспектах. С позиций статики под ономастической (информационной) нагруженностью текста будем понимать ко-

личество разных онимов в процентном соотношении с общим количеством лексических единиц в тексте (фрагменте, разделе и т. д.). Информационная нагруженность в динамическом ракурсе обусловливается тем, как быстро возрастает количество онимов в текстах от начала учебника к концу. Для установления этого показателя можно использовать методику определения сумм с нарастающим итогом, в ряде исследований позволившую получить существенные результаты [Naumann, Popescu, Altmann, 2012].

На рисунке 1 по оси x отложены Введение и 7 тем учебника ГЕО-1, по оси y –

количество использования онимов. От темы к теме количество словоупотреблений онимов возрастает, так как каждый раз к словоупотреблениям в данной теме прибавляются словоупотребления онимов, использованных раньше.

Аналогичные данные по росту числа словоупотреблений онимов во Введении и 12 главах ГЕО-2 приводятся на рисунке 2.

Чтобы выявить характер нарастания количества онимов и того, есть ли какая-нибудь упорядоченность в этом нарастании, можно использовать степенную функцию [Naumann, Popescu, Altmann, 2012; Andreev 2016]:

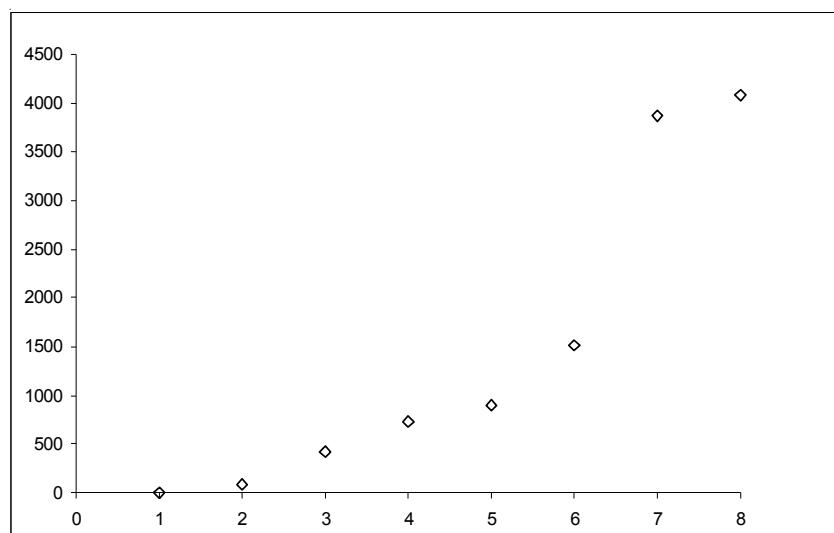


Рис. 1. Суммы словоупотреблений ИС с нарастающим итогом в ГЕО-1

Fig. 1. Cumulative sums of onyms frequencies in GEO-1

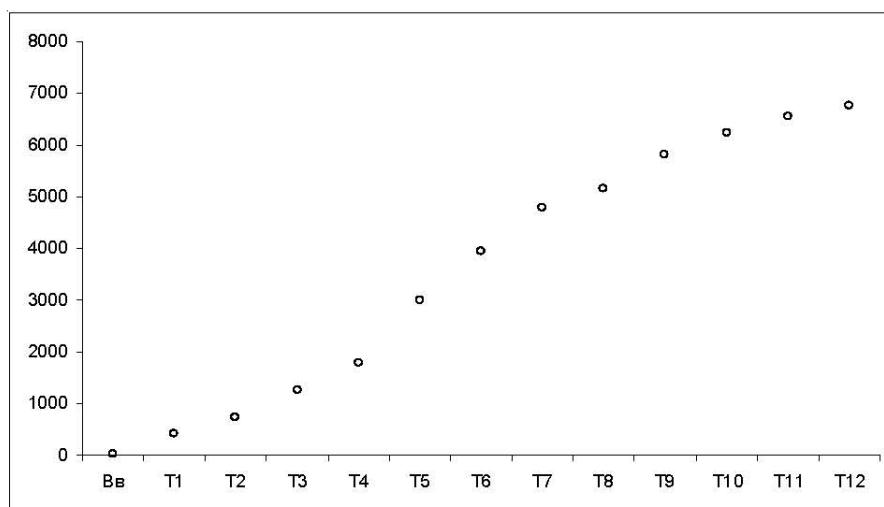


Рис. 2. Суммы словоупотреблений ИС с нарастающим итогом в ГЕО-2

Fig. 2. Cumulative sums of onyms frequencies in GEO-2

$$y = ax^b,$$

где a и b – параметры.

Ее применение показано на рисунках 3 и 4.

Кружками обозначены наблюдаемые суммы, линия показывает теоретически ожидаемые частоты согласно этой функции. Как видно на рисунках, кружки расположены близко к линии,

а иногда и накладываются на нее. Это говорит о том, что функция хорошо отражает распределение частот. На это же указывает и коэффициент детерминации r^2 , который может иметь значения в диапазоне от 0 до 1. Для ГЕО-1 $r^2 = 0,941$ (рис. 3), для ГЕО-2 $r^2 = 0,967$ (рис. 4). В обоих случаях коэффициент детерминации близок к верхнему пределу, что свидетельствует об успешности применения данной функции.

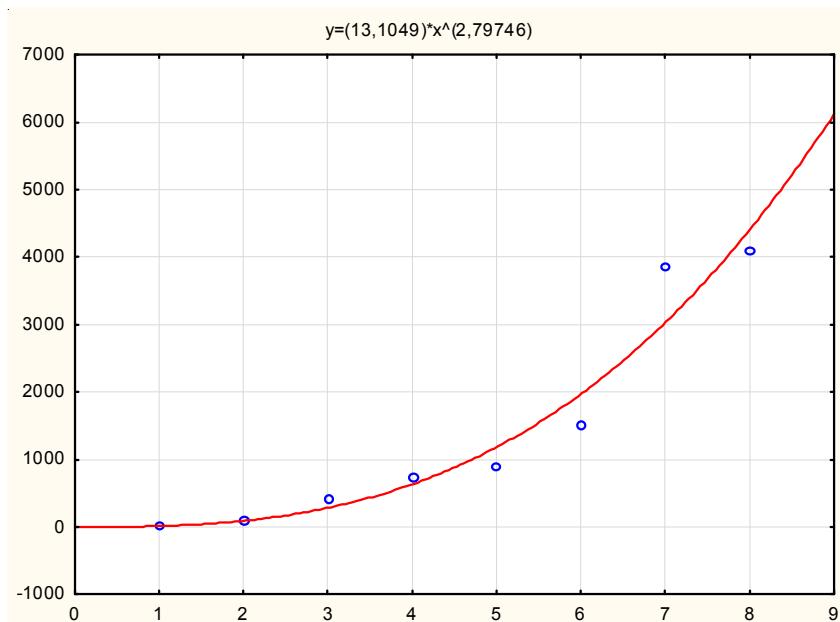


Рис. 3. Аппроксимация распределения частот онимов с нарастающим итогом в ГЕО-1

Fig. 3. Fitting the cumulative frequency distributions of onyms in GEO-1

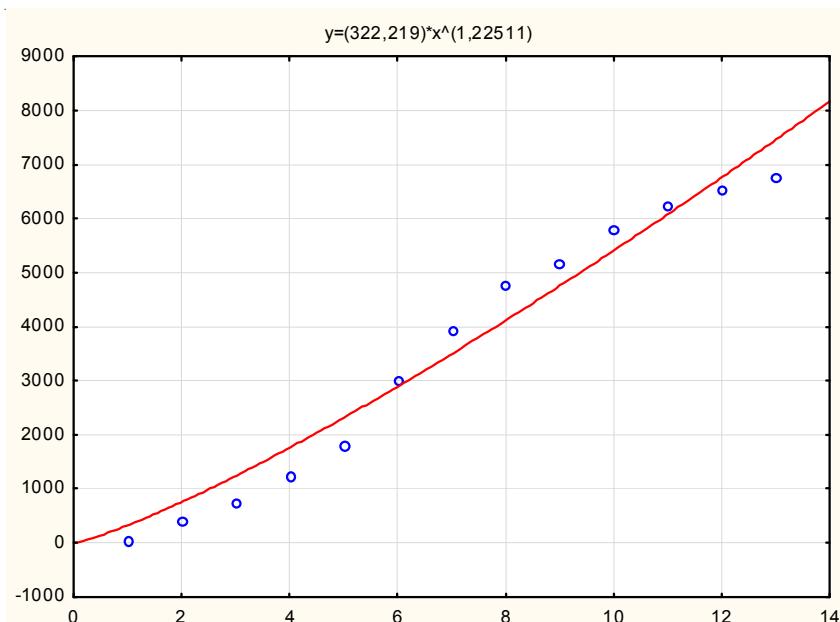


Рис. 4. Аппроксимация распределения частот онимов с нарастающим итогом в ГЕО-2

Fig. 4. Fitting the cumulative frequency distributions of onyms in GEO-2

Несмотря на то, что в обоих случаях степенная функция успешно отражает распределение накопленных частот словоупотреблений, в характере накопления ИС от начала к концу в двух учебниках наблюдаются значительные различия. В ГЕО-1 имеется резкий рост использования ИС только к концу учебника, в то время как в ГЕО-2 нарастание имеет более плавный характер.

Ономастическая система в описании двенадцати ведущих стран мира

На основе анализа имен собственных в географических ННТ представляется возможным рассмотреть ономастическую структуру описания ведущих стран мира в обоих учебниках. В число этих стран вошли страны – основатели БРИКС и члены «Большой семерки» (G7). Отметим, что в ГЕО-1 отсутствует самостоятельный раздел о России, поскольку учебник посвящен политической и экономической географии зарубежных стран, в то время как ГЕО-2, также в основном направленный на описание зарубежных стран, такой раздел включает. Кроме того, в ГЕО-1 практически не представлена Южно-Африканская Республика. Целесообразно начать анализ с тематически более объемного учебника, то есть с ГЕО-2. При подсчете онимов от-

бирались только те из них, которые относятся к рассматриваемой стране, причем анализу в этом случае подвергался не весь текст учебника, а только соответствующий параграф или раздел.

Для анализа привлекаются два параметра: количество словоупотреблений онимов (*tokens*) и количество разных онимов (*types*) – уникальных лексем. При подсчете словоупотреблений (*tokens*) каждый оним учитывался столько раз, сколько он был употреблен в тексте. При подсчете уникальных лексем (*types*) каждый оним учтен только один раз вне зависимости от того, сколько раз он встречается в тексте.

Полученные в результате подсчетов данные представлены в виде диаграммы рассеяния (рис. 5).

На этой диаграмме по оси *x* отложены частоты словоупотреблений ИС, по оси *y* – количество уникальных лексем. Как видно на диаграмме, в нижних регистрах выстраивается достаточно плотная группа, которая от левого нижнего угла вытягивается в направлении к верхнему правому. Абсолютный минимум по обоим параметрам у ЮАР, абсолютный максимум – у США. В описании США находим максимум словоупотреблений ИС и максимум различных уникальных лексических единиц ИС. Примерно в середине диаго-

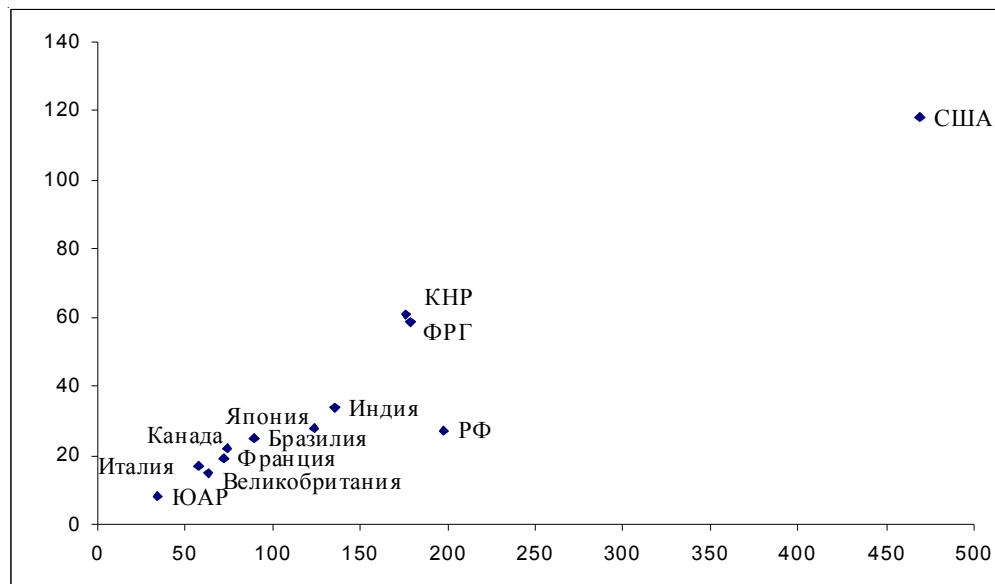


Рис. 5. Диаграмма рассеяния стран по количественным параметрам ономастической структуры их описания в ГЕО-2

Fig. 5. Scatterplot of countries by quantitative parameters of their description onomastic structure in GEO-2

нального кластера ЮАР – США располагаются Китай и ФРГ, которые образуют отдельную подгруппу. Она несколько отклоняется от оси вверх, что отражает большее разнообразие ономастического пространства данных разделов по сравнению с общей тенденцией при описании стран в этом учебнике.

По оси абсцисс к этой подгруппе близка Россия, однако ее положение относительно оси ординат значительно отличается от Китая и ФРГ (а также Индии из общего кластера) по количеству используемых уникальных ИС. Так, при описании РФ на оним *Россия* (с вариантом *РФ*) приходится 158 словоупотреблений, то есть 80 % от использования всех онимов, которые связаны с Россией. Для сравнения, в тексте о США на оним *США* с вариантом *Соединенные Штаты* приходится только 35 % словоупотреблений, а остальные онимы отражают различные географические объекты в США. При описании Германии название страны (*ФРГ / Германия*) составляет 34 % от всех онимов в тексте, Китая (*Китай / КНР*) – 48 %.

На рисунке 6 приводятся данные относительно описания стран в учебнике ГЕО-1.

В ГЕО-1 наблюдается в целом та же картина, что и в ГЕО-2. США находится в правом верхнем углу с максимальными по обоим показателям значениями. КНР образует отдель-

ный кластер, однако в него входит Индия, а не ФРГ. Остальные страны образуют единый кластер, вытянутый по диагонали.

Обсуждение результатов

Изучение ономастического наполнения географических нормативно-научных текстов показывает, что состав и количество употребляемых в них онимов, их частотность являются маркерами содержания географического фрагмента нормативно-научной картины мира (общеобязательного географического знания).

Для единиц ономастического ядра, номинирующих ведущие объекты на географической и экономической картах мира (и, следовательно, в структуре географического знания), зафиксирована суммарная частота употребления, равная приблизительно 40 % при интенсивности использования онимов в тексте с частотой 1 оним на 14 слов.

Уровень отсечения частотного списка, формирующего ядро содержания географического фрагмента нормативно-научной картины мира, от периферии совокупного ономастического пространства находится в диапазоне 26–30 рангов в ранжированном по нисходящей ряду ИС.

Малочастотная лексика, в обоих источниках выполняющая функцию информацион-

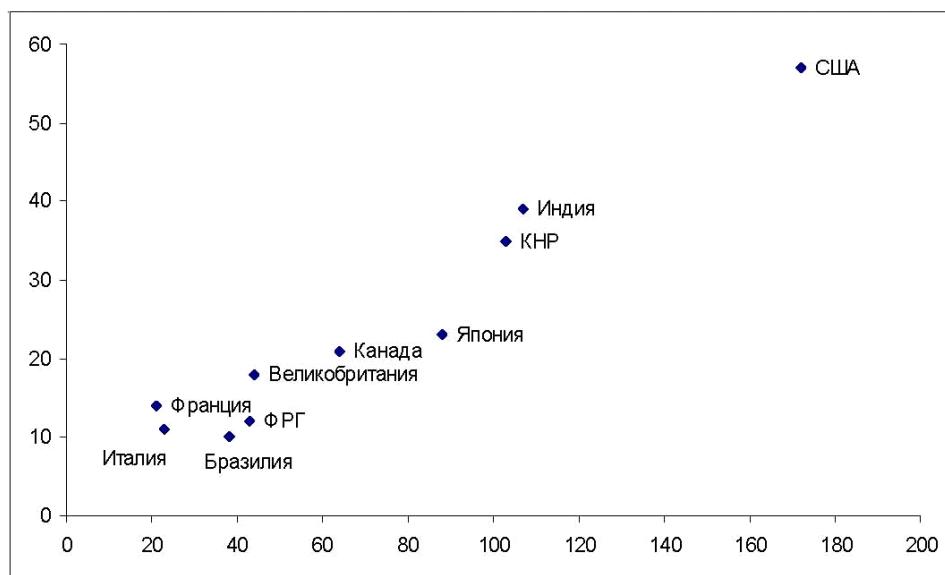


Рис. 6. Диаграмма рассеяния стран по количественным параметрам ономастической структуры их описания в ГЕО-1

Fig. 6. Scatterplot of countries by quantitative parameters of their description onomastic structure in GEO-1

ного расширения и/или создания ономастического фона, оказывает серьезное влияние на полноту раскрытия содержания ключевых имен. Охватывая менее 10 % всех словоупотреблений ИС, онимы гапакс легомена составляют около половины всех ИС в учебных текстах.

Рост использования онимов от начала учебника к концу имеет тенденцию, отражающую степенной функцией.

Несмотря на достаточно большой период времени, разделяющий выпуск анализируемых учебников, и произошедшие за это время значительные изменения в политико-экономической ситуации в стране и в мире, сравниваемые ННТ характеризуются весьма высоким уровнем сходства их ономастического пространства. Это отражается в целом ряде аспектов: в области ономастической насыщенности макротекста; идентичности ономастического состава ядра; соотношении гапаксов с частотной лексикой и, как следствие, формировании сходного ономастического фона; специфике ономастического описания стран.

Установление с помощью кванитативных методов функции нарастания ономастической (информационной) напряженности текста служит существенным основанием как для анализа концентрации ИС с применением метода, предложенного в наших предыдущих работах и основанного на суммировании частот для аппроксимации распределения признака [Андреев, 2023], так и для разработки рекомендаций по созданию новых текстов.

Выявленная близость основных характеристик ономастического пространства двух источников ННТ, находящихся на противоположных точках почти тридцатилетней временной оси, может рассматриваться с позиции соотношения статики и динамики в характеристике географического фрагмента общеобязательного научного знания. Эта близость свидетельствует в пользу объективности представляемой в учебниках информации, ее достоверности и подчеркивает определенную (и необходимую) устойчивость общеобязательного знания. В связи с полученными результатами возникает вопрос о соотношении в структуре общеобязательного знания стабильности и актуальности, что, в свою очередь, требует более четкого определения са-

мого понятия актуальности по отношению к базовому знанию. Очевидно, что в решение этого вопроса, имеющего не только образовательное, но и социокультурное значение, данные квантитативного анализа нормативно-научных текстов разной временной принадлежности могут внести необходимую объективность.

Выводы

Проведенное исследование позволило получить следующие результаты и сформулировать некоторые выводы.

В ходе квантитативного анализа ономастического наполнения нормативно-научных макротекстов двух школьных учебников географии были выделены и охарактеризованы ядро и периферия ономастического пространства сопоставляемых географических нормативно-научных макротекстов, определен состав единиц, выполняющих роль информационного расширения и формирующих ономастический фон.

Результаты анализа частот проприальной лексики показали высокую интенсивность использования онимов в текстах. Их распределение отражается степенной функцией, параметры которой позволяют оценить интенсивность введения имен собственных в текст и рост их количества от начала учебника к его концу.

Ономастическая структура характеризуется сильным противопоставлением частотного ядра и периферии. Школьные учебники, разработанные и изданные с большой разницей во времени, демонстрируют выраженную инерцию структуры ядра. Оно характеризуется высокой степенью устойчивости во времени как по составу лексических единиц, их частотности, так и по соотношению параметров ядра между собой (частотные ранги, точка отсечения от периферии, количество типов онимов).

Перспективой работы является применение описанной методики для исследования ономастического пространства других предметных фрагментов ННКМ с целью определения сходств и различий ономастической структуры ННТ в зависимости от предметной сферы, на основе чего предполагается разработка рекомендаций по ономастическому наполнению предметных ННТ для составителей новых учебников.

СПИСОК ЛИТЕРАТУРЫ

- Андреев С. Н., 2023. Распределение атрибутов в поэмах А.С. Пушкина и гипотеза Скinnera // Вестник Томского государственного университета. Филология. № 82. С. 5–22. DOI: 10.17223/19986645/82/1
- Жеребило Т. В., 2011. Термины и понятия: Методы исследования и анализа текста : слов.-справ. Назрань : Пилигрим. 108 с.
- Куприянов Р. В., Солнышкина М. И., Лехницкая П. А., 2023. Параметрическая таксономия учебных текстов // Вестник Волгоградского государственного университета. Серия 2, Языкоzнание. Т. 22, № 6. С. 80–94. DOI: <https://doi.org/10.15688/jvolsu2.2023.6.6>
- Лапошина А. Н., Лебедева М. Ю., 2024. Сравнительный анализ текстов учебников русского языка для начальной школы с использованием количественных методов // Русский язык в школе. Т. 85, № 5. С. 7–26. DOI: 10.30515/0131-6141-2024-85-5-7-26
- Максимчук Н. А., 2017. Терминологическое поле ономастического пространства // Известия Смоленского государственного университета. № 4 (40). С. 155–167.
- Максимчук Н. А., Голубничая А. В., 2018. Ономастическое пространство учебного макротекста: константы и переменные // Ученые записки ВГУ им. П.М. Машерова. Т. 25. С. 148–154.
- Монахов С. И., Турчаненко В. В., Чердаков Д. Н. 2023. Школьный учебный текст в аспекте терминоупотребления: корпусный анализ // Научный результат. Вопросы теоретической и прикладной лингвистики. Т. 9, № 1. С. 27–49. DOI: 10.18413/2313-8912-2023-9-1-0-3
- Новиков А. И., 2007. Текст и его смысловые доминанты / под ред. Н. В. Васильевой, Н. М. Нестеровой, Н. П. Пешковой. М. : Ин-т языкоzнания РАН. 224 с.
- Суперанская А. В., 1973. Общая теория имени собственного. М. : Наука. 246 с.
- Супрун В. И., 2000. Ономастическое поле русского языка и его художественно-эстетический потенциал. Волгоград : Перемена. 172 с.
- Топоров В. Н., 2004. Имя как фактор культуры // Топоров В. Н. Исследования по этимологии и семантике: в 3 т. Т. 1. М. : Яз. слав. культуры. С. 380–383.
- Философия : энцикл. слов., 2004 / под ред. А. А. Ивица. М. : Гардарики. 840 с.
- Andreev S., 2016. Verbal vs. Adjectival Style in Long Poems by A.S. Pushkin // Glottometrics. Vol. 33. P. 25–31.
- Andreev S., Místecký M., Altmann G., 2018. Sonnets: Quantitative Inquiries. Lüdenscheid : RAM-Verlag. 128 p. (Studies in Quantitative Linguistics ; vol. 29).
- Hirsch J. E., 2005. An Index to Quantify an Individual's Scientific Research Output // Proceedings of the National Academy of Sciences of the USA. Vol. 102, iss. 46. P. 16569–16572. DOI: 10.1073/pnas.0507655102
- Maćutek J., Popescu I.-I., Altmann G., 2007. Confidence Intervals and Test for H-Point and Related Text Characteristics // Glottometrics. Vol. 15. P. 45–52.
- Místecký M., 2018. Counting Stylistic Properties of Sonnets: A Case of Machar's Letni Sonety // Glottometrics. Vol. 41. P. 1–12.
- Naumann S., Popescu I.-I., Altmann G., 2012. Aspects of Nominal Style // Glottometrics. Vol. 23. P. 23–55.
- Pan X., Qiu H., Liu H., 2015. Golden Section in Chinese Contemporary Poetry // Glottometrics. Vol. 32. P. 55–62.
- Popescu I.-I., 2007. Text Ranking by the Weight of Highly Frequent Words // Exact Methods in the Study of Language and Text. Berlin : N. Y. : Mouton de Gruyter. P. 555–565.
- Popescu I.-I., Altmann G., 2006. Some Aspects of Word Frequencies // Glottometrics. Vol. 13. P. 23–46.
- Popescu I.-I., Altmann G., 2008. Hapax Legomena and Language Typology // Journal of Quantitative Linguistics. Vol. 15, № 4. P. 370–378. DOI: 10.1080/09296170802326699
- Zhang X., Liu H., 2020. Function Words in Male and Female Authors: A Diachronic Investigation of Modern Chinese Prose // Glottometrics. Vol. 48. P. 67–87.

ИСТОЧНИКИ

ГЕО-1 – Гладкий Ю. Н., Лавров С. Н. Экономическая и социальная география мира : проб. учеб. для 10 класса сред. шк. М. : Просвещение, 1991. 272 с.

ГЕО-2 – Максаковский В. П. География. 10–11 классы : учеб. для общеобразоват. организаций. Баз. уровень. 27-е изд. М. : Просвещение, 2018. 416 с.

REFERENCES

- Andreev S.N., 2023. Raspredelenie atributov v poemah A.S. Pushkina i gipoteza Skinnera [The Distribution of Attributes in Poems by Alexander Pushkin and Skinner's Hypothesis]. *Vestnik Tomskogo gosudarstvennogo universiteta. Filologiya* [Tomsk State University

- Journal of Philology], no. 82, pp. 5-22. DOI: 10.17223/19986645/82/1
- Zherebilo T.V., 2011. *Terminy i ponyatiya: Metody issledovaniya i analiza teksta: slov.-sprav.* [Terms and Concepts: Methods of Research and Text Analysis. Dictionary-Reference Book]. Nazran, Piligrim Publ. 108 p.
- Kupriyanov R.V., Solnyshkina M.I., Lehnickaja P.A., 2023. Parametricheskaya taksonomiya uchebnym tekstov [Parametric Taxonomy of Educational Texts]. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 2. Yazykoznanie* [Science Journal of Volgograd State University. Linguistics], vol. 22, no. 6, pp. 80-94. DOI: <https://doi.org/10.15688/jvolsu2.2023.6.6>
- Laposhina A.N., Lebedeva M.Yu., 2024. Sravnitelnyy analiz tekstov uchebnikov russkogo yazyka dlya nachalnoy shkoly s ispolzovaniem kolichestvennykh metodov [Comparative Analysis of Texts of Russian Language Textbooks for Primary School Using Quantitative Methods]. *Russkiy yazyk v shkole* [Russian Language at School], vol. 85, no. 5, pp. 7-26. DOI: 10.30515/0131-6141-2024-85-5-7-26
- Maksimchuk N.A., 2017. Terminologicheskoe pole onomasticheskogo prostranstva [Terminological Field of Onomastic Space]. *Izvestiya Smolenskogo gosudarstvennogo universiteta* [Bulletin of Smolensk State University], no. 4(40), pp. 155-167.
- Maksimchuk N.A., Golubnichaja A.V., 2018. Onomasticheskoe prostranstvo uchebnogo makroteksta: konstanty i peremennyye [Onomastic Space of Educational Macrotext: Constants and Variables]. *Uchenye zapiski VGU im. P.M. Masherova* [Scientific Notes of VSU Named After P.M. Masherov], vol. 25, pp. 148-154.
- Monahov S.I., Turchanenko V.V., Cherdakov D.N., 2023. Shkolnyy uchebnyy tekst v aspekte terminoupotrebleniya: korpusnyy analiz [Terminology Use in School Textbooks: Corpus Analysis]. *Nauchnyy rezul'tat. Voprosy teoretičeskoy i prikladnoy lingvistiki* [Research Result. Theoretical and Applied Linguistics], vol. 9, no. 1, pp. 27-49. DOI: 10.18413/2313-8912-2023-9-1-0-3
- Novikov A.I., 2007. *Tekst i ego smyslovye dominanty* [Text and Its Semantic Dominants]. Moscow, In-t yazykoznanija RAN. 224 p.
- Superanskaya A.V., 1973. *Obshhaya teoriya imeni sobstvennogo* [General Theory of Proper Names]. Moscow, Nauka Publ. 246 p.
- Suprun V.I., 2002. *Onomasticheskoe pole russkogo yazyka i ego khudozhestvenno-esteticheskiy potentsial* [Onomastic Field of the Russian Language and Its Artistic and Aesthetic Potential]. Volgograd, Peremena Publ. 172 p.
- Toporov V.N., 2004. Imya kak faktor kultury [Name as a Factor of Culture]. Toporov V.N. *Issledovaniya po etimologii i semantike: v 3 t. T. 1* [Toporov V.N. Research on Etymology and Semantics. In 3 Vols. Vol. 1], Moscow, Yaz. slav. kultury Publ., pp. 380-383.
- Ivin A.A., ed., 2004. *Filosofiya: entsikl. slov.* [Philosophy. Encyclopedic Dictionary]. Moscow, Gardariki Publ. 840 p.
- Andreev S., 2016. Verbal vs. Adjectival Style in Long Poems by A.S. Pushkin. *Glottometrics*, vol. 33, pp. 25-31.
- Andreev S., Místecký M., Altmann G., 2018. *Sonnets: Quantitative Inquiries*. Lüdenscheid, RAM-Verlag. 128 p. (Studies in Quantitative Linguistic; vol. 29).
- Hirsch J.E., 2005. An Index to Quantify an Individual's Scientific Research Output. *Proceedings of the National Academy of Sciences of the USA*, vol. 102, iss. 46, pp. 16569-16572. DOI: 10.1073/pnas.0507655102
- Mačutek J., Popescu I.-I., Altmann G., 2007. Confidence Intervals and Tests for the H-Point and Related Text Characteristics. *Glottometrics*, vol. 15, pp. 45-52.
- Místecký M., 2018. Counting Stylistic Properties of Sonnets: A Case of Machar's Letni Sonety. *Glottometrics*, vol. 41, pp. 1-12.
- Naumann S., Popescu I.-I., Altmann G., 2012. Aspects of Nominal Style. *Glottometrics*, vol. 23, pp. 23-55.
- Pan X., Qiu H., Liu H., 2015. Golden Section in Chinese Contemporary Poetry. *Glottometrics*, vol. 32, pp. 55-62.
- Popescu I.-I., 2007. Text Ranking by the Weight of Highly Frequent Words. *Exact Methods in the Study of Language and Text*. Berlin, New York, Mouton de Gruyter, pp. 555-565.
- Popescu I.-I., Altmann G., 2006. Some Aspects of Word Frequencies. *Glottometrics*, vol. 13, pp. 23-46.
- Popescu I.-I., Altmann G., 2008. Hapax Legomena and Language Typology. *Journal of Quantitative Linguistics*, vol. 15, no. 4, pp. 370-378. DOI: 10.1080/09296170802326699
- Zhang X., Liu H., 2020. Function Words in Male and Female Authors: A Diachronic Investigation of Modern Chinese Prose. *Glottometrics*, vol. 48, pp. 67-87.

SOURCES

- Gladkiy Yu.N., Lavrov S.N. *Ekonomicheskaya i sotsialnaya geografiya mira: prob. ucheb. dlya 10 klassa sred. shk.* [Economic and Social Geography of the World. Pilot Textbook for 10th Grade High School]. Moscow, Prosveshhenie Publ., 1991. 272 p.

Maksakovskiy V.P. *Geografiya. 10–11 klassy: ucheb. dlya obshcheyeobrazovat. organizatsiy. Baz. uroven* [Geography. 10–11 Grades. Textbook for

General Education Organizations. A Basic Level]. Moscow, Prosveshchenie Publ., 2018. 416 p.

Information About the Authors

Sergey N. Andreev, Doctor of Sciences (Philology), Professor, Department of English Language, Smolensk State University, Smolensk, Russia, smol.an@mail.ru, <https://orcid.org/0000-0001-7076-1241>

Nina A. Maksimchuk, Doctor of Sciences (Philology), Professor, Department of Russian Language, Smolensk State University, Smolensk, Russia, nmaximchuk@yandex.ru, <https://orcid.org/0000-0003-2792-1355>

Информация об авторах

Сергей Николаевич Андреев, доктор филологических наук, профессор кафедры английского языка, Смоленский государственный университет, г. Смоленск, Россия, smol.an@mail.ru, <https://orcid.org/0000-0001-7076-1241>

Нина Алексеевна Максимчук, доктор филологических наук, профессор кафедры русского языка, Смоленский государственный университет, г. Смоленск, Россия, nmaximchuk@yandex.ru, <https://orcid.org/0000-0003-2792-1355>