

УДК 811

UDC

DOI: 10.17223/18572685/68/18

Корпусные библиометрические аспекты славяноведческой тематики в русскоязычных научных электронных ресурсах в контексте категорий исторического познания

А.В. Бочаров

Томский государственный университет

Россия, 634050, г. Томск, пр. Ленина, 36.

E-mail: inspert2015@gmail.com

Авторское резюме

Анализируется задача измерения полноты, репрезентативности и сбалансированности при создании корпуса русскоязычных текстов по славяноведческой литературе. Расчет данных критериев формирования дисциплинарного корпуса может быть произведен с использованием библиометрических и наукометрических подходов. Разные категории исторического познания могут обозначать явления, сопоставимые с тематиками и проблематикой исследований в славяноведении. Всего проверено 76 исторических категорий (историческая аналогия, историческая память, историческая возможность историческое время, историческая личность и т.д.). Все исторические категории, сопряженные с тематикой славяноведения, либо с тематикой русинов, были ранжированы по частотности поисковой выдачи полнотекстовых ресурсов в scholar.google.ru и academia.edu. Выявлены и проанализированы следующие показатели: соотношение упоминания славяноведения совместно с разными славянскими языками; представленность славяноведения в ряду других страноведческих дисциплин; совместная встречаемость различных лингвистических дисциплин со славяноведением; соотношение и распределение категории исторического познания в текстах со славяноведческой тематикой. Изучено распределение частоты связи исторических категорий со славяноведением. Показана специфичность влияния славяноведческой тематики на популярность разных категорий и особенности функционирования исторического сознания в славяноведческом научном сообществе. Сделан вывод, что корпусная сбалансированность может рассматриваться на четырех уровнях в рамках возможности создания: 1) филологического корпуса с текстами по различным направлениям языкознания и литературоведения;

2) историографического корпуса с разнотематическими и разножанровыми историческими текстами на одном языке; 3) страноведческого корпуса с текстами по различным страноведческим дисциплинам; 4) дисциплинарного корпуса только по тематике славяноведения.

Ключевые слова: корпусная лингвистика, славяноведческий корпус, библиометрия, исторические категории, цифровые ресурсы славистики.

Corpus bibliometric aspects of Slavic studies in Russian-language scholarly electronic resources in the context of categories of historical knowledge

A.V. Bocharov

Tomsk State University

Lenin Ave, 36, Tomsk, 634050, Russia

E-mail: inspert2015@gmail.com

Abstract

The article analyses the task of measuring completeness, representativeness, and balance when creating a corpus of Russian-language texts on Slavic literature. The criteria for a disciplinary corpus can be calculated using a bibliometric and scientometric approaches. Different categories of historical knowledge can denote phenomena comparable to the themes and problems of research in Slavic studies. A total of 76 historical categories have been tested (historical analogy, historical memory, historical possibility, historical time, historical personality, etc.). All historical categories associated with Slavic studies or with Rusins have been ranked according to the frequency of search results of full-text resources in scholar.google.ru and academia.edu. The following indicators have been identified and analysed: representation of Slavic studies in a number of other regional studies disciplines; joint occurrence of various linguistic disciplines with Slavic studies; correlation and distribution of the category of historical knowledge in texts with Slavic themes. The author has studied the frequency distribution of the connection between historical categories and Slavic studies and shown the specificity of the influence of Slavic subjects on different categories and functioning of historical consciousness in the Slavic academic community. He concludes that corpus balance can be considered at four levels within the possibility of creating: 1) a philological corpus with texts in various areas of linguistics and literary criticism;

2) a historiographic corpus with historical texts of various themes and genres in one language; 3) a regional corpus with texts on various regional disciplines; 4) a disciplinary corpus only on the subject of Slavic studies.

Keywords: corpus linguistics, Slavic corpus, bibliometrics, historical categories, digital resources of Slavic studies.

Введение

Корпусная революция в филологии во взаимодействии с цифровизацией научных публикаций может и должна ставить задачу создания специализированных дисциплинарных корпусов.

Методы корпусной лингвистики в славяноведении в последнее десятилетие все активнее используются вслед за остальными языковедческими дисциплинами. В русскоязычных академических публикациях эта тенденция, в частности, продемонстрирована в ряде работ [1; 3–5; 7].

Если ставить задачу создания корпуса русскоязычных текстов по филологической, страноведческой или славяноведческой литературе, то такой корпус должен составляться с учетом критериев полноты, репрезентативности и сбалансированности. Расчет данных критериев формирования дисциплинарного корпуса может быть произведен с использованием библиометрических и наукометрических подходов. Учет библиометрической статистики также становится устойчивым трендом в самых разных научных дисциплинах. В русскоязычном славяноведении последних лет присутствуют работы с библиометрическим уклоном [6; 8; 9].

Цель исследования – провести разведочный обзор представленности в научных цифровых платформах полнотекстовых русскоязычных ресурсов по славяноведению с учетом соотношения различных тематик, которые можно учесть как критерии корпусной сбалансированности. Проблематика оценки сбалансированности относительной доли текстов в корпусе в статье дополнительно проиллюстрирована на примере сравнительной библиометрической оценки тематики русинов и русинского языка.

Сбалансированность славяноведческого корпуса по разным дисциплинарным направлениям

Очевидно, что сбалансированность для состава специализированного дисциплинарного корпуса не может быть идентичной сбалансированности национального языкового корпуса. Если в национальном корпусе речь идет о пропорциональности разных типов и жанров

текстов, то в корпусе научных (в том числе научно-образовательных) текстов понятие сбалансированности может опираться на категориально-понятийную и терминологическую специфику научных дисциплин, которым посвящен корпус. Разметка подобного корпуса должна опираться также на предметные общенаучные и узкодисциплинарные рубрикаторы.

Соотношение славяноведческой тематики на разных славянских языках в академических текстах продемонстрировано в табл. 1. Все приведенные в статье частоты актуальны на апрель 2022 г. В Google Scholar индексируются публикации, изданные с 1990-х гг. по настоящее время.

Частота текстов с упоминанием научной дисциплины на конкретном языке в первую очередь коррелирует с количеством носителей данного языка. Однако свой вклад в разницу частот на разных языках вносит также уровень развития в разных странах как научной гуманитаристики в целом, так и популярности конкретной научной дисциплины (славяноведения в нашем случае). Также дополнительный фактор в современной наукометрии – это доля англоязычных статей в национальной науке.

Таблица 1

**Работы в Google Scholar с упоминанием названия дисциплины
на разных славянских языках**

Язык	Поисковый запрос	Статистика поисковой выдачи (кол-во текстов)	Общее число говорящих
English	slavic studies slavistics	328 000	Ок. 600–700 млн (2003 г.)
Русский	славяноведение славистика	15 600	Ок. 260 млн (2014 г.)
Čeština	slavistika	5 090	Ок. 10,62 млн (2016 г.)
Hrvatski	slavistiku	3 620	Ок. 21 млн (2006 г.)
Українська	словянознавство славистика	994	Ок. 27,3 млн (2019 г.)
Polski	slawistyka	848	Ок. 40 млн (2007 г.)
Беларуский	славяназнаўства славістыка	213	Ок. 6,4 млн (2009 г.)

В сбалансированности корпуса по славяноведению необходимо также учесть соотношения текстов с упоминанием разных славянских языков совместно со славяноведением. Объем текстов с упоминанием менее изученных языков не должен превышать объем более изученных. На рис. 1. продемонстрировано соотношение количества текстов с совместным упоминанием славяноведения (славистики) с разными славянскими языками в русскоязычных публикациях, проиндексиро-

ванных в Google Scholar. Так, русинский язык является пока одним из наименее упоминаемых славянских языков в русскоязычных текстах (после верхне- и нижнелужицкого языков, которые, впрочем, имеют значительный (в три раза) разрыв с русинским языком).

Если посмотреть общую частоту поисковой выдачи в scholar.google.ru (со снятой опцией «показывать цитаты»), то для запроса (славяноведение | славистика) частота равна 15 700, а для запроса («русини» | русинский) частота в scholar.google.ru равна 4 540. Поиск в Google Scholar делался со снятой опцией «показывать цитаты», чтобы статистика указывала в большей степени на тексты, игнорируя упоминания этих же текстов в библиографических ссылок. Тематика именно русинов взята здесь как пример сравнительной доли представленности изучения отдельного языка и народа в рамках славяноведения в контексте истории.



Рис. 1. Соотношение количества текстов с совместным упоминанием славяноведения (славистики) с разными славянскими языками в русскоязычных публикациях, проиндексированных в Google Scholar

Аналогичная поисковая выдача при поиске Google на сайте academia.edu почти всегда отличается в меньшую сторону, поскольку размещение текста в социальной сети происходит по инициативе авторов, а не автоматически, по факту появления изданий в электронных библиотеках или сайтах. В academia.edu для запроса славяноведение | славистика – частота равна 3 490, для запроса «русини» | русинский частота равна 951. Стоит отметить, что в поисковый запрос в данном случае не включалось слово «русин» (в единственном числе

и без кавычек), т.к. иначе в статистику выдачи вмешалось бы название журнала «Русин», не все статьи которого непосредственно посвящены русинскому языку или этносу. Если сравнивать поисковые выдачи по тому же запросу («русин» | русинский) до выхода журнала в 2005 г. и после этого, то будет заметна существенная разница: в Google Scholar до 2004 г. включительно размер поисковой выдачи равен 429 публикациям, с 2005 г. – 4040 публикациям. При этом, если с 2005 г. исключить из поиска название журнала через знак минус, т.е. – «Русин», то разница уменьшится приблизительно на одну тысячу – 2 960 публикаций. Такой результат означает, что выход журнала был не единственным, но очень важным фактором для развития русинской тематики в научной периодике.

Таким образом, упоминания тематики русинов в научных электронных библиотеках занимает приблизительно 29% от общего количества русскоязычных публикаций по теме славяноведения. В академической социальной сети academia.edu тематика русинов занимает приблизительно 27% от общего количества русскоязычных публикаций по теме славяноведения на этом же ресурсе (если считать, что упоминание русинов указывает на славяноведческую тематику, даже при условии, что сами понятия «славяноведение» или «славистика» в тексте не упоминаются).

Для организации систематизированного поиска сопряженности славяноведения с различными дисциплинами и категориями использовался авторский сервис «Конструктор комбинаторных поисковых запросов по истории и социальным наукам» (www.lib.tsu.ru/inspert/). Этот онлайн сервис позволяет генерировать поисковые запросы с учетом синонимии, омонимии и тематических тезаурусов.

В перспективе при составлении страноведческого корпуса пропорциональность представленности русскоязычных текстов по отдельным дисциплинам должна соответствовать их соотношению в ряду библиометрических частот. В ряду других страноведческих дисциплин в русскоязычных публикациях в Google Scholar наблюдаются следующие соотношения (табл. 2). Ожидаемо славяноведение (славистика) (15 600) сопоставимо с русистикой (22 900). Остальные соотношения частот можно объяснить как различной развитостью русскоязычных научных школ, так и сложностью изучения языков, а также востребованностью языков в системе международных отношений. Данные носят накопительный характер с точки зрения хронологии публикаций, т.е. это количество текстов, накопленное к текущему 2022 г. Для того чтобы сравнивать диахронические тенденции в развитии различных дисциплин или дисциплинарной терминологии, необходимо делать поисковые запросы и сравнивать результаты по разным годам. Такой

анализ не входит в задачи нашего исследования, хотя, безусловно, в корпусной сбалансированности необходимо учитывать также распределение объемов текстов по разным хронологическим периодам.

Таблица 2

**Славяноведение в ряду других страноведческих дисциплин
в русскоязычных публикациях, проиндексированных в Google Scholar**

Дисциплина	Частота (кол-во текстов)
Русистика	22 900
Славяноведение славистика	15 600
Африканистика	6 910
Тюркология	6 230
Германистика	3 090
Синология	2 990
Уралистика финно-угроведение	2 070
Арабистика	1 750
Романистика	1 700
Индология	1 510
Алтаистика	774
Индоевропеистика	691
Иудаика	671
Полонистика	454
Иранистика	363
Скандинавистика	226
Японистика	212
Семитология	120
Израилеведение	113
Албанистика	85
Санскритология	77
Балтистика	56

Распределение упоминания в корпусе специальной дисциплинарной терминологии (лингвистической в нашем случае) связано с анализом тематических тенденций в рамках сбалансированного корпуса.

В табл. 3 показаны частоты в совместной встречаемости лингвистических дисциплин со славяноведением (славистикой) в Google Scholar. Использовался список языковедческих дисциплин, представленный в одной из поисковых форм сервиса InSpert – «Клиограмме истории идей». Если характеризовать состав оцифрованных текстов, то подавляющее большинство публикаций в результатах комбинаторного поиска – это научные статьи или разделы монографий, учебных пособий или диссертаций, в которых искомые ключевые слова встречаются в одном смысловом контексте. Иногда в небольшом проценте случаев в поисковые результаты попадают сборники статей в общем файле, в которых ключевые слова встречаются под одной обложкой, но в

разных статьях. Другой особый случай, когда на одной pdf-странице перед началом одной статьи идет конец и список литературы другой статьи, в котором встречаются искомые ключевые слова. Такие ситуации не указывают на прямую связь славяноведения с искомым понятием. Однако из-за редкости подобных случаев (несколько процентов от результата поиска) они не искажают общую статистическую картину состава текстов для составления корпуса по славяноведению.

Таблица 3

Совместная встречаемость лингвистических дисциплин со славяноведением (славистикой) в Google Scholar

Лингвистическая дисциплина	Частота в Google Scholar	Лингвистическая дисциплина	Частота в Google Scholar
синтактика	18 600	лингводидактика	825
семантика	14 900	идиоматика	755
лингвистика	8 900	контрастивная лингвистика	658
поэтика	8 230	топонимика	614
грамматика	5 800	корпусная лингвистика	390
лексикология	4 990	прагмалингвистика	375
лексикография	4 800	акцентология	301
этнолингвистика	4 160	идеография	235
этимология	4 100	морфонология	225
когнитивная лингвистика	3 060	нарратология	141
стилистика	2 640	терминоведение	117
фонетика	2 620	ностратика	110
социоллингвистика	2 390	семиология	86
фразеология	2 370	фоносемантика	58
семиотика	2 270	этносемантика	52
фонология	2 060	лингвоперсонология	44
риторика	1 780	лексикостатистика	42
диалектология	1 700	семитология	40
текстология	1 630	дериватология	35
морфология	1 620	глоссематика	27
психоллингвистика	1 460	грамматология	26
компьютерная лингвистика	1 410	силлогистика	25
фольклористика	1 410	психопоэтика	24
ономасиология	1 330	меметика	5
орфография	1 270	стилометрия	4
лингвокультурология	1 180		

Частота связи лингвистических субдисциплин со славяноведением указывает:

- во-первых, на степень изученности и актуальности различных

языковедческих аспектов в славяноведении. Поэтому, например, высокочастотны лексикология или этимология, но малочастотно совместное упоминание со славяноведением таких дисциплин, как фоносемантика или лингвоперсонология;

– во-вторых, на универсальность лингвистических явлений, изучаемых соответствующими дисциплинами. Поэтому, например, синтактика и семантика самые частотные;

– в-третьих, на специфичность и закреплённость терминологии в именовании дисциплин. Поэтому, например, лингвостатистика или психопозитика – малочастотные.

Категории исторического познания как критерии тематической сбалансированности славяноведческого корпуса

Славяноведение, наряду с другими страноведческими дисциплинами, является комплексной дисциплиной, изучающей языки, литературу, фольклор, историю, материальную и духовную культуру. Для частотной аналитики структурно-содержательных аспектов корпуса текстов по таким комплексным дисциплинам уместной и наиболее обобщающей будет проверка сопряженности с категориями исторического познания.

Историческая категория – это абстрактный конструкт для наиболее обобщающих интерпретаций явлений исторического процесса и наиболее общепринятых группировок понятий в историческом познании.

Разные исторические категории могут обозначать явления, вполне сопоставимые с тематиками и проблематикой исследований в славяноведении. Всего проверено 76 исторических категорий (историческая аналогия, историческая память, историческое время, историческая личность и т.д.). Далее, в целях экономии, первое слово «исторический» в названии категорий не будет повторяться каждый раз. Весь список категорий представлен в одной из поисковых форм сервиса InSpert – «Матрице исторического познания». Исторические категории, наряду с языковедческими понятиями, могут стать критериями для разметки ключевых слов в дисциплинарном корпусе. Пропорциональная представленность разных исторических категорий в корпусе также должна отвечать критериям сбалансированности.

Все исторические категории, сопряженные либо с тематикой славяноведения, либо с тематикой русинов, были ранжированы по частотности поисковой выдачи полнотекстовых ресурсов. Частотность в свою очередь была ранжирована по разрядам значности чисел (пятизначные, четырехзначные, трехзначные, двузначные, однозначные).

Сопряженность исторических категорий со славяноведением, как правило, на один разряд превышает сопряженность этих же категорий с тематикой русинов, по очевидным причинам большей обобщенности. А частотность совместной встречаемости в scholar.google.ru для большинства категорий на один разряд выше, чем на сайте academia.edu.

В табл. 4 показаны наиболее частотные категории (пятизначные числа для славяноведения и четырехзначные для тематики русинов). Конкретные примеры смысловой сопряженности славяноведения с какой-либо исторической категорией можно легко проверить в текстовых фрагментах, дающихся в поисковой выдаче гиперссылок на тексты.

Таблица 4

Наиболее частотные категории исторического познания в текстах со славяноведческой тематикой

Исторические категории	(славяноведение славистика)		(«русины» русинский)	
	Частота в scholar.google.ru	Частота в academia.edu	Частота в scholar.google.ru	Частота в academia.edu
Развитие	16 300	5 200	3 230	637
Возможность	15 300	6 650	2 720	681
Период	15 000	1 750	2 860	598
Эпоха	14 700	4 780	2 000	571
Ситуация	14 000	4 420	553	599
Причина	13 200	2 170	942	659
Факт	12 600	4 550	2360	595
Память	11 900	4 860	1780	570
Пространство	10 400	3 820	1780	608

Тематика именно русинов взята как пример сравнительной доли представленности изучения отдельного языка и народа в рамках славяноведения в контексте истории.

В табл. 5 показаны вторые по частотности исторические категории (четырёхзначные числа для славяноведения и трёхзначные для тематики русинов). Это наиболее обобщающие по смыслу категории, необходимые для функционирования исторического сознания в целом. Славяноведческая тематика не дает здесь какой-то особенной специфики в предпочтениях категорий.

Все точные значения частот приводятся в таблицах лишь для иллюстрации достоверности и проверяемости данных. Для анализа данных здесь важны не столько точные значения частот, сколько их относительные значения. Состав и объем проиндексированных поисковиками текстов может меняться со временем. Однако эти изменения будут незначительны по комбинациям совместной встречаемости тематик – за один год может прибавиться несколько статей или книг, но общая разрядность числа публикаций останется неизменной.

Таблица 5

**Вторые по частотности категории исторического познания
в текстах со славяноведческой тематикой**

Историческая категория	(славяноведение славистика)		«русини» русинский)	
	частота в scholar.google.ru	частота в academia.edu	частота в scholar.google.ru	частота в academia.edu
Личность	9 580	3 650	1 300	519
Феномен	7 240	3 050	1 080	436
Условия	7 170	1 120	1 510	431
Связь	6 210	1 370	364	408
Кризис	6 060	2 400	1 080	526
Категория	5 520	875	699	350
Истоки	5 390	1 280	1 010	459
Специфика	5 030	781	643	332
Выбор	4 330	1 110	634	493
Прогресс	4 270	2 160	629	313
Влияние	3 450	1 380	632	394
Интерпретация	3 340	948	287	254
Наследие	3 210	860	494	318
Памятник	3 120	1 090	436	345
Известность	2 820	997	462	465
Достижение	2 500	767	422	302
Объективность	2 360	764	322	263
Случайность	2 290	1 530	314	181
Предшественник	2 160	289	333	72
Шанс	2 130	944	510	168
Аналогия	2 110	947	297	263
Преимственность	2 000	655	286	239
Типичность	1 860	832	126	166
Уникальность	1 520	454	219	167
Забвение	1 510	522	235	83
Альтернативность	1 500	1520	300	220
Источник	1 460	840	174	160
Веха	1 340	384	174	155
Опыт	1 250	360	242	152
Контекст	1 220	454	169	138
Закономерность	1 040	475	131	125

Все исторические категории, сопряженные с тематикой славяноведения либо с тематикой русинов, были отсортированы по частотности поисковой выдачи полнотекстовых ресурсов. Частотность, в свою очередь, была ранжирована по разрядам значности чисел (пятизначные, четырехзначные, трехзначные, двузначные, однозначные).

Количество научных публикаций по узкой тематике может подняться с нескольких десятков до нескольких сотен или тысяч, только за период смены целого поколения исследователей, т. е. за 10–15 лет, а по

некоторым тематикам малоизученность или малоактуальность остается неизменными многие десятилетия. Эти аспекты также должны учитываться при анализе сбалансированности дисциплинарного корпуса.

В табл. 6 показаны третьи по частотности исторические категории (трехзначные числа для славяноведения и двухзначные для тематики русинов).

Таблица 6

**Третьи по частотности категории исторического познания
в текстах со славяноведческой тематикой**

Историческая категория	(славяноведение славистика)		(«русины» русинский)
	Частота в scholar.google.ru	Частота в academia.edu	Частота в scholar.google.ru Частота в academia.edu
Достоверность	881	76	156 60
Значение	811	851	118 285
Воображение	698	266	113 113
Роль	674	1090	104 453
Время	627	184	82 69
Материал	570	146	80 47
Неизбежность	563	172	96 89
Сознание	540	148	121 69
Судьба	534	261	233 81
Миф	531	123	80 50
Прецедент	488	266	130 99
Аутентичность	427	202	67 61
Необратимость	412	73	84 52
Повторяемость	410	143	27 43
Тенденция	369	259	75 86
Событие	353	166	62 52
Реальность	303	108	48 45
Фон	285	93	33 58
Травма	274	133	46 75
Перспектива	260	74	65 28
Истина	207	67	63 23
Свидетельство	161	67	49 41
Обратимость	117	64	10 21
Смысл	109	44	9 19
Фальсификация	105	196	29 30

Сопряженность исторических категорий со славяноведением, как правило, на один разряд превышает сопряженность этих же категорий с тематикой русинов, по очевидным причинам большей обобщенности. Частотность совместной встречаемости в Google Scholar для большинства категорий на один разряд выше, чем на сайте academia.edu. Отдельные несовпадения соотношений в рейтинге частот для тематики русинов с общей славяноведческой тематикой демонстрируют тот факт, что для истории разных языков формируется собственная

структура исторического сознания, влияющая на приоритеты в выборе тематик и аргументов в работах специалистов-языковедов.

Из основных тенденций в использовании исторических категорий здесь можно отметить, во-первых, влияние на все исторические процессы и феномены истории славянских народов нахождения их в составе империй (Российской и Австро-Венгерской) и Советского союза; во-вторых, важность тематики борьбы славянских народов за свою самостоятельность и самобытность; в-третьих, значительную роль в историографии, связанной со славянскими народами, личностей филологов-языковедов, фольклористов и литературоведов, по сравнению с историками.

В табл. 7 показаны наименее частотные категории исторического познания. Следует отметить, что для исторической науки в целом данные малочастотные исторические категории также не самые общеупотребительные и почти всегда дискуссионные. Кроме того, они часто относятся к пограничным дискурсивным практикам, поскольку характерны в большей степени для публицистического жанра, чем для академических текстов.

Предположительно, в качестве одного из факторов разница в частотах различных категорий может быть объяснена так называемым эффектом Матфея (Matthew effect), впервые предложенным американским социологом науки Робертом Мертоном [2: 56–63].

Таблица 7

**Наименее частотные категории исторического познания
в текстах со славяноведческой тематикой**

Историческая категория	(славяноведение славистика)		(«русинь» русинский)	
	Частота в scholar.google.ru	Частота в academia.edu	Частота в scholar.google.ru	Частота в academia.edu
Образ	93	30	29	23
Оценка	80	10	6	2
Дистанция	75	9	2	5
Справедливость	66	15	30	10
Идентичность	52	23	29	19
Понимание	49	23	6	3
Урок	46	13	14	9
Ответственность	40	9	11	4
Ошибка	26	5	4	4
Цикличность	7	3	1	2
Прогноз	2	1	0	0

Речь идет о феномене неравномерного распределения преимуществ, в котором сторона, уже ими обладающая, продолжает их накапливать и приумножать, в то время как другая, изначально

ограниченная, оказывается обделена ещё сильнее и, следовательно, имеет меньшие шансы на дальнейший успех. В наукометрии эффект Матфея обуславливает повышение цитируемости и популярности уже популярных направлений исследования и публикаций, и пропорциональное отсутствие повышения цитируемости слаборазвитых направлений и редкоцитируемых публикаций.

Заключение

Подводя итоги перспектив использования библиометрических частотных данных, заметим, что каждый ряд цифр в приведенных таблицах можно рассматривать как указание на материал для отдельного исследования на пересечении славяноведения с историографией или культурологией, а также поводом для обоснования научной актуальности в выборе научных направлений и подходов.

Из анализа всего комплекса статистических данных можно сделать три основных вывода:

Упоминание дисциплин в библиометрической частотности – это средство проектирования сбалансированности дисциплинарных корпусов. Терминологический аспект библиометрической частотности – одна из возможных целей корпусного лингвистического моделирования и проверки статистических гипотез о тематических тенденциях.

Распределение частоты сопряженности исторических категорий со славяноведением зависит от двух факторов, которые не обязательно коррелируют:

1) общая частота категории, отражающая её важность и популярность в историческом сознании в целом.

2) специфичность влияния славяноведческой тематики на важность и популярность разных категорий, обуславливающая особенности функционирования исторического сознания в славяноведческом научном сообществе.

Славяноведение пересекается по своей дисциплинарной принадлежности с языкознанием, страноведением и с историей. Соответственно, корпусная сбалансированность может рассматриваться на четырех уровнях в рамках возможности создания:

– филологического корпуса с текстами по различным направлениям языкознания и литературоведения;

– историографического корпуса с разнотематическими и разножанровыми историческими текстами на одном языке (или с параллельными переводами);

– страноведческого корпуса с текстами по различным страноведческим дисциплинам;

– дисциплинарного корпуса только по тематике славяноведения.

ЛИТЕРАТУРА

1. Баранов В. Исторический корпус как цель и инструмент корпусной палеославистики // *Scripta & e-Scripta*. 2015. Vol. 14/15. P. 39–62.
2. Merton R.K. The Matthew effect in science: The reward and communication systems of science are considered // *Science*. 1968. Vol. 159, № 3810. P. 56–63. DOI: 10.1126/science.159.3810.56
3. Shimko E. Корпусная лингвистика как способ изучения переводческого наследия Кирилла и Мефодия // *Konštantínove listy*. 2021. Vol. 14, № 1. P. 184–193. DOI: 10.17846/CL.2021.14.1.184-193
4. Изотов А.И. Новые направления славянского языкознания: корпусная лингвистика // *Язык, сознание, коммуникация: сб. ст. / ред. В.В. Красных, А.И. Изотов. М.: МАКС Пресс, 2015. Вып. 51. С. 82–93. ISBN: 978-5-317-05157-0*
5. Молдован А.М. Проблемы и задачи корпусного изучения славянской письменности // *Славянский альманах*. 2014. Т. 2013. С. 25–34.
6. Николенко А.В. П.Н. Берков и Информационно-библиографическая комиссия Международного комитета славистов (к 60-летию основания) // *Современные проблемы книжной культуры: основные тенденции и перспективы развития: материалы IX Междунар. науч. семинара: в 2 ч. Ч. 1 / сост. Л.А. Авгуль, Д.Н. Бакун. М., 2018. С. 98–106. ISBN: 978-5-9908169-2-3*
7. Узенёва Е.С. Аксиологические исследования славянских языков и культур // *Славянский мир в третьем тысячелетии*. 2020. Т. 15, № 3/4. С. 217–225. DOI: 10.31168/2412-6446.2020.15.3-4.15
8. Шапиро Н. Коллекции и фонды русистики и славистики в библиотеках США // *Библиография. Научный журнал по библиографоведению, книговедению и библиотековедению*. 2015. № 6. С. 103–127.
9. Шатъко Е.В. Славистические литературоведческие журналы Сербии после 1991 г. // *Литературно-критическая периодика в странах Центральной и Юго-Восточной Европы: структура, типология, социокультурный контекст: [сб. тр.] / под ред. Л.Ф. Широковой. М.: Ин-т славяноведения РАН, 2020. С. 203–215. DOI: 10.31168/2618-8554.2020.12*

REFERENCES

1. Baranov, V. (2015) *Istoricheskiy korpus kak tsel' i instrument korpusnoy paleoslavistiki* [Diachronic OCS corpus as an object and instrument of corpus palaeoslavitic]. *Scripta & e-Scripta*. 14/15. pp. 39–62.
2. Merton, R.K. (1968) The Matthew effect in science: The reward and communication systems of science are considered. *Science*. 159(3810). pp. 56–63. DOI: 10.1126/science.159.3810.56
3. Shimko, E. (2021) *Korpusnaya lingvistika kak sposob izucheniya perevodcheskogo naslediya Kirilla i Mefodiya* [Corpus Linguistics as a research

method of Cyril's and Methodius's cultural heritage]. *Konstantínove listy*. 14(1). pp. 184–193. DOI: 10.17846/CL.2021.14.1.184-193

4. Izotov, A.I. (2015) *Novye napravleniya slavyanskogo yazykoznaniya: korpusnaya lingvistika* [New directions of Slavic linguistics: corpus linguistics]. In: Krasnykh, V.V. & Izotov, A.I. (eds) *Yazyk, soznanie, kommunikatsiya* [Language, Consciousness, Communication]. Vol. 51. Moscow: MAKS Press. pp. 82–93.

5. Moldovan, A.M. (2013) *Problemy i zadachi korpusnogo izucheniya slavyanskoy pis'mennosti* [Problems and tasks of the corpus linguistics studies of the Slavic-script heritage]. In: Nikiforov, K.V. et al. (eds) *Slavyanskiy al'manakh* [Slavic Almanac]. Moscow: Indrik. pp. 25–34.

6. Nikolenko, A.V. (2018) *P.N. Berkov i Informatsionno-bibliograficheskaya komissiya Mezhdunarodnogo komiteta slavistov (k 60-letiyu osnovaniya)* [P.N. Berkov and the Information and Bibliographic Commission of the International Committee of Slavists (on the 60th anniversary of its foundation)]. In: Avgul, L.A. & Bakun, D.N. (eds) *Sovremennye problemy knizhnoy kul'tury: osnovnye tendentsii i perspektivy razvitiya* [Modern problems of book culture: main trends and development prospects]. Vol. 1. Moscow: Nauka. pp. 98–106.

7. Uzeneva, E.S. (2020) *Axiological studies of Slavic languages and cultures. Slavyanskiy mir v tret'em tysyacheletii – Slavic World in the Third Millennium*. 15(3/4). pp. 217–225 (in Russian). DOI: 10.31168/2412-6446.2020.15.3-4.15

8. Shapiro, N.G. (2015) *Kollektsii i fondy rusistiki i slavistiki v bibliotekakh SShA* [Russian and Slavic collections and funds at the libraries of USA]. *Bibliografiya. Nauchnyy zhurnal po bibliografovedeniyu, knigovedeniyu i bibliotekovedeniyu*. (6). pp. 103–127.

9. Shatko, E.V. (2020) *Slavisticheskie literaturovedcheskie zhurnaly Serbii posle 1991 g.* [Slavic literary journals of Serbia after 1991]. In: Shirokova, L.F. (ed.) *Literaturno-kriticheskaya periodika v stranakh Tsentral'noy i Yugo-Vostochnoy Evropy: struktura, tipologiya, sotsiokul'turnyy kontekst* [Literary-critical periodicals in the countries of Central and South-Eastern Europe: structure, typology, socio-cultural context]. Moscow: Institute of Slavic Studies RAS. pp. 203–215. DOI: 10.31168/2618-8554.2020.12

Бочаров Алексей Владимирович – кандидат исторических наук, доцент факультета исторических и политических наук Томского государственного университета (Россия).

Alexey V. Bocharov – Tomsk State University (Russia).

E-mail: inspert2015@gmail.com