

Раздел 3 ЯЗЫКОЗНАНИЕ

УДК 519.2:801.82(045)

DOI: 10.22213/2618-9763-2021-2-50-60

В. А. Баранов, доктор филологических наук, профессор

Ижевский государственный технический университет имени М. Т. Калашникова, Ижевск, Россия

О. Ф. Жолобов, доктор филологических наук, профессор

Казанский (Приволжский) федеральный университет, Казань, Россия

ЛИНГВОСТАТИСТИЧЕСКИЙ АНАЛИЗ ДРЕВНЕРУССКОЙ АНТОЛОГИИ (РНБ, Ф.п.І. 39)*

Представлен лингвостатистический анализ 29 текстов рукописи XIII века – Толстовского сборника (РНБ, Ф.п.І. 39), содержащего оригинальные произведения древнерусского автора Кирилла Туровского и переводы текстов Кирилла Иерусалимского, Иоанна Златоуста и неизвестных авторов. Целями работы являются а) верификация выводов об авторстве трех анонимных текстов сборника, приписываемых Кириллу Туровскому, сделанных в ходе их предшествующего анализа традиционными лингвистическими методами и корреляционными методами статистики, а также б) установление возможности использования статистических методов для сопоставления текстов малых объемов.

Для измерения межтекстового расстояния применен метод Дельта и его модификация косинус-Дельта, использовано различное количество наиболее частотных слов каждого из текстов – от 50 до 250. Данные извлечены из электронного издания Толстовского сборника. Вычисления и построение дендрограмм осуществлены с помощью пакета Stylo.

Анализ дендрограмм подтвердил выводы, сделанные в других работах: три анонимных текста Слово о премудрости, Поучение на Пянтюкостии и Поучение в неделю 5-ю по Пасхе не принадлежат Кириллу Туровскому. Обнаружена зависимость группировки текстов от их объема и примененных методов. Обнаружено, что косинус-Дельта, в отличие от классической Дельты, оказывается нечувствительным к различиям в объеме текстов. Кроме того, при использовании косинус-Дельта увеличение количества наиболее частотных слов позволяет обнаружить корреляцию между группировками текстов и их дискурсивной и лексико-синтаксической близостью.

Ключевые слова: Толстовский сборник XIII века; Кирилл Туровский; анонимные тексты; атрибуция, метод Дельта; стилометрия; лингвотекстология.

Введение

Определение авторства текстов, анализ их стилевой однородности, установление плагиата находятся в ведении стилометрии, одного из филологических направлений, выделившегося из стилистики и использующего в настоящее время, в первую очередь, компьютерные методики и приемы для атрибуции проблемных текстов. Традиционно объектом стилометрии являются анонимные, подложные, приписываемые, спорные тексты с точки зрения их авторства и/или времени создания. Разработаны десятки принципиально разных подходов и методов решения таких задач. С разной степенью достоверности получены ответы об авторстве или времени написании спорных текстов на многих языках (см., например, [1–3]).

Вопросы подобного рода возникают и относительно средневековых славянских тек-

стов. Всем хорошо известна проблема оригинальности древнерусского Слова о полку Игореве, закрытая блестящей работой А. А. Зализняка [4], в которой на основе анализа, в первую очередь, строевых компонентов текста ученый доказал древнерусское происхождение текста.

Есть и другие примеры использования статистических методов для анализа средневековых славянских рукописей. Так, в работах Д. М. Мироновой [5, 6] показана эффективность автоматической кластеризации средневековых списков отрывка Евангелия от Матфея (Мф14:14–34) из 525 славянских списков XI–XVI веков.

Одной из уникальных рукописей древнерусского периода является Толстовский сборник XIII века, хранящийся в Российской государственной библиотеке и содержащий оригинальные и переводные тексты нескольких

авторов – Кирилла, епископа Туровского, жившего в XII веке (далее – КТ), Кирилла, епископа Иерусалимского (далее – КИ), проповедника IV века, и Иоанна Златоуста, архиепископа Константинопольского (347–407 гг.) (далее – ИЗ), – а также тексты неизвестных авторов. В сборнике находятся три текста, которые в ряде научных изданий приписываются КТ: два анонимных поучения – Поучение в неделю 5-ю по Пасхе, Поучение на Пянтикостии (см. [7, с. 324] и Слово о премудрости (см. [8, с. 494; 9, с. 464; 10, с. 34]).

В работах [11, с. 84–86; 12; 13] в ходе лингвотекстологического анализа показано, что анонимные Поучение в Неделю 5-ю по Пасхе и Слово о премудрости не принадлежат Кириллу Туровскому, а Слово на Рождество Христово с именем Иоанна Златоуста является копией – по-видимому, славянской.

В работе [14] изложены результаты статистического анализа приписываемых КТ текстов Толстовского сборника. Три анонимных (далее – проблемных) текста были сопоставлены с шестью текстами КТ, семью текстами КИ, текстом ИЗ и тремя текстами неизвестных авторов. В качестве метода использован кластеризационный анализ, позволяющий, как известно, организовать исследуемые объекты на основе их признаков в группы (кластеры), внутри которых объекты более схожи, чем объекты из разных групп, и представить наглядно в виде дендрограммы. Показано, что все три текста с большой долей вероятности не принадлежат перу КТ¹.

Одним из важных критериев объективности результатов любого эксперимента является его верификация, например, с помощью применения других методов по отношению к тем же данным.

Целями данной работы являются установление:

а) с помощью метода Дельта статистического расстояния между тремя проблемными текстами Толстовского сборника и текстами КТ, а также другими текстами сборника;

б) возможности применения метода Дельта в отношении текстов малых объемов.

Метод Дельта

Метод *Burrows's Delta* (иначе – *Classic Delta*) предложен Дж. Ф. Бёрроузом в 2001 г. [18] и имеет несколько модификаций – *Eder's Delta*, *Hoover's Delta*, *Rotated Delta*, *Cosine Delta* и др. [19–22].

За два десятилетия были разработаны и апробированы на большом количестве материала инструментальные средства, в частности пакет *Stylo* в среде *R*, позволяющие наглядно представить кластеризацию текстов на разных языках [23–33], в том числе на современном русском языке [34–36] и на русском языке XVIII и XIX веков [37].

Считается, что метод не работает на текстах малых объемов – менее 5000 слов². Однако в литературе описано применение метода и для анализа документов небольшого объема. Так, в [39] с помощью этого метода исследован корпус текстов, «связанных с личностью Л. П. Берии и предположительно созданных им, а также набор документов, связанных с личностью его заместителя В. Меркулова», наименьший из которых содержит 400 текстовых форм [40, с. 601].

Несмотря на активное и результативное применение стилометрических методов, выводы некоторых авторов относительно объективности их применения для выявления авторства иногда неутешительны³.

Понятно, что одним из выходов из ситуации является перепроверка результатов с помощью нескольких статистических методов, а также с помощью традиционных филологических и лингвистических методик и приемов.

Математическое обоснование меры Дельта приведено, например, в [42] и представляет собой среднее «абсолютных значений разностей стандартизованных оценок частот слов из $\{w_i\}$ между текстами D и D' » [43, с. 4]:

$$\Delta(D, D') = \frac{1}{n} \sum_{i=1}^n |z(f_i(D)) - z(f_i(D'))|,$$

где D, D' – тексты; n – исследуемое множество слов выборки; z – оценка частоты употребления слова w_i в тексте D ; $f_i(D)$ – частота слова w_i в тексте D .

¹ Статистическому анализу Толстовского сборника посвящена также работа [15]; с помощью статистических методов исследованы и некоторые другие славянские рукописи [16; 17].

² Ср. еще более жесткие ограничения: «надежные результаты с его помощью могут быть получены только на жанрово гомогенном корпусе, а каждый составляющий этот корпус текст должен быть по объему не меньше 10 000 слов» [38, с. 76].

³ Ср. выводы о применении статистических методов для определения авторства романа «Тихий Дон» М. А. Шолохова: «В случае с определением авторства романа „Тихий Дон“ ни одно исследование не продемонстрировало даже возможность... комплексного подхода. Именно поэтому выводы оказались столь различны: одни считали авторство М. А. Шолохова бесспорным, другие его отрицали полностью, третьи пытались доказать соавторство с разными писателями, четвертые обнаруживали тематический и сюжетный плагиат, пятые отыскивали лексические и стилистические заимствования» [41, с. 71].

В свою очередь:

$$z(f_i(D)) = \frac{f_i(D) - \mu_i}{\sigma_i},$$

где μ_i – средняя частота слова в выборке; σ_i – стандартное отклонение средней частоты слова в выборке.

Метод позволяет установить расстояние между двумя текстами на основании оценки многомерных расстояний между ними.

Модификацией метода является метод косинус-Дельта, использующий измерение межтекстового расстояния с помощью формулы:

$$\Delta_C(D, D') = 1 - \frac{\vec{f}(D) \cdot \vec{f}(D')}{\|\vec{f}(D)\| \|\vec{f}(D')\|},$$

где $\vec{f}(D) = (f_1(D), f_2(D), \dots, f_n(D))$,

$$\vec{f}(D') = (f_1(D'), f_2(D'), \dots, f_n(D')) [44, с. 600].$$

Визуализация расстояний может быть осуществлена с помощью применения алгоритмов построения дендрограмм, демонстрирующих степень схожести текстов в виде иерархических деревьев, в которых «листья» соответствуют отдельному тексту, «ветви» – группам текстов.

Метод Дельта и некоторые его модификации реализованы в программном пакете *Stylo* для

среды *R*, созданном М. Эдером, Я. Рубицким и М. Кестемонтом [45]. Пакет имеет также средства визуализации измерений в виде дендрограмм.

Материал и его преобработка

Материалом исследования явились 29 текстов Толстовского сборника XIII века, созданные тремя разными авторами – КТ, КИ и ИЗ, а также неизвестными авторами (см. список источников). Три текста, как уже отмечалось, в ряде работ приписываются КТ – Поучение в неделю 5-ю по Пасхе, Слово о премудрости, Поучение на Пянтикостии.

Каждый текст был выгружен из корпуса «Манускрипт»¹, транслитерирован буквами современного алфавита; удалены диакритика, титла и пунктуационные знаки, выносные буквы внесены в строку:

ПООУЧЕНИЕ В НДЕ ЕЮ ПО посце Гси блгсви оче Азь убо о друзи и братье надеяхся на всяку неделю боле собрати въ цркви людии на послушание бжствьныхъ словесъ ныне же мне приходить нъ аще быхъ о собе глалъ то добре бысте творили не приходяще...

Поучение в неделю 5-ю по Пасхе, начало

Количество словоупотреблений в текстах колеблется от пяти с небольшим тысяч до 440 (таблица).

Количество словоупотреблений в каждом из текстов

Текст	F	Текст	F	Текст	F	Текст	F	Текст	F
Abg	1506	CJ_1	924	CJ_7	1426	CJ_13	4508	CT_Desc	2746
An_East	586	CJ_2	1566	CJ_8	601	CJ_14	2218	CT_Fath	2065
An_Pent	479	CJ_3	1627	CJ_9	1182	CJ_15	1856	CT_Paral	1872
An_Wisd	440	CJ_4	3991	CJ_10	2485	CJ_16	1213	CT_Thom	1251
Aphr	1352	CJ_5	1347	CJ_11	2176	CT_Asc	1560	Life_Bas	5257
Chrys_Nat	1698	CJ_6	2942	CJ_12	3979	CT_Blind	1726	–	–

Результаты экспериментов

При измерении межтекстового расстояния с помощью метода Дельта использовались расстояния классическая Дельта и косинус-Дельта, которые вычислялись для 50, 100, 150, 200 и 250 наиболее частотных словоформ.

Результаты измерения межтекстового расстояния были визуализированы с помощью дендрограмм, см., например, рис. 1.

¹ Тесты Толстовского сборника // Славянское письменное наследие. URL: <http://manuscripts.ru/mns/portal.main?p1=74> (дата обращения: 14.05.2021).

другу узлы, тем более похожи друг на друга перечни наиболее частотных слов.

Сопоставление группировки текстов на дендрограммах и графах позволяет говорить о следующих наиболее общих закономерностях:

- во всех случаях тексты КТ и КИ противопоставляются;

- в большинстве случаев тексты неизвестных авторов и Злат_Рожд ближе текстам КТ, чем текстам КИ;

- три проблемных текста, входя в одну подгруппу, сближаются то с текстами КИ, то с группой текстов КТ и неизвестных авторов;

- при выборках 50, 100, 150 словоформ оба метода демонстрируют устойчивую близость проблемных текстов и текстов КИ,

- при выборках 200 (классическая Дельта и косинус-Дельта) и 250 (классическая Дельта) словоформ проблемные тексты вместе с КТ, текстами неизвестных авторов и Злат_Рожд противопоставлены текстам КИ, но проблемные тексты при этом не сближаются с КТ.

Обсуждение результатов

Метод Дельта, как и кластерные методы, использованные в предыдущем нашем исследовании того же материала, демонстрирует мегатекстовый характер рукописного источника: действительно, Толстовский сборник – это мегатекст-антология, структурные единицы которого связаны тематическим единством или родством, а отчасти также календарной отнесенностью [46]). Кроме того, метод Дельта вновь выявляет отчетливое разграничение двух основных частей антологии – Слов Кирилла Туровского и Оглашений Кирилла Иерусалимского, при том что остальные тексты занимают колеблющиеся по расстоянию от основных частей позиции.

Вместе с тем метод дает несколько иные результаты. Так, в этом случае, в частности, нет четкой выделенности анонимного Слова о премудрости: его статистические расстояния, судя по дендрограммам с разными характеристиками, флуктуируют по отношению к другим текстам, а само Слово входит в разные подкластеры в зависимости от количества исходных по частотности словоупотреблений.

Расходятся в ряде результатов и метод классическая Дельта и его модификация – косинус-Дельта. Так, классическая Дельта определенно нивелирует расстояния между короткими текстами, которые насчитывают от 440 (Слово о премудрости) до 601 (8-е Оглашение Кирилла Иерусалимского) словоупотребления: для уров-

ня 150 частотных слов три анонимных текста и 8-е Оглашение образуют подкластер в группе Оглашений, а для уровня 200 – в группе Слов Кирилла Туровского¹. Для уровня 250 слов подкластер трех анонимных текстов сохраняется, хотя 8-е Оглашение из него выпадает. Косинус-Дельта не фиксирует подобного сближения, поэтому можно предположить, что с помощью этого метода статистические расстояния не нивелируются в зависимости от количества словоупотреблений в тексте.

В то же время косинус-Дельта дает довольно резкое, скачкообразное изменение позиции анонимных текстов: так, для уровня 150 и 250 частотных слов они распределены по разным подкластерам в блоке Оглашений, а для уровня 200 слов образуют подкластер в блоке Слов Кирилла Туровского. Безусловно, такие колебания фиксируют семантико-синтаксическую обособленность анонимных текстов и противопоставленность их Словам Кирилла Туровского. Косинус-Дельта, судя по всему, реагирует на дискурсивную природу текстов. Об этом свидетельствует устойчивое положение компилятивного Слова на Рождество Христово в кластере Слов Кирилла Туровского. Торжественные, риторически украшенные Слова Кирилла Туровского также в определенном смысле являются компиляцией и в жанрово-стилистическом плане точно соответствуют Слову на Рождество Христово с именем Иоанна Златоуста. Устойчиво образуют подкластер Житие Василия Великого и два апокрифических текста, которые объединяет нарративный тип дискурса. Таким образом, статистические расстояния между текстами в соответствии с методом косинус-Дельта – это семантико-синтаксические и дискурсивные расстояния.

Выводы

Статистический анализ с помощью метода Дельта 29 текстов Толстовского сборника позволил подтвердить выводы о близости и контрастности этих текстов друг другу, сделанные в [47], о непринадлежности трех проблемных текстов Ан_Пасх, Ан_Прем, Ан_Пянт перу Кирилла Туровского [48, с. 84–86; 49–51].

Использование различного количества исходных данных (от 50 до 250 наиболее частотных слов каждого из текстов) и двух вариантов метода Дельта дало возможность подтвердить стабильность противопоставленности текстов КТ и КИ и одновременно обнаружить варьирующиеся характеристики, например, различную степень близости Слова о премудрости

¹ Это сближение Оглашения и трех анонимных текстов намечается уже на уровне 100 частотных слов.

и в целом всех трех проблемных текстов то текстам КИ, то текстам КТ.

Важным итогом следует считать выявление зависимости конфигурации дендрограмм от объема текстов и примененных методов. Так, косинус-Дельта, в отличие от метода «классическая Дельта», оказывается нечувствительным к различиям в объеме текстов. Кроме того, при использовании косинус-Дельта увеличение количества наиболее частотных слов позволяет обнаружить корреляцию между группировками текстов и их дискурсивной и лексико-синтаксической близостью.

Сокращения

СБТол – Сборник Слов и поучений («Толстовский сборник»), втор. пол. XIII в. (РНБ, Ф.п.1. 39), 184 л. / О. Ф. Жолобов и др.; проект «Манускрипт». URL: http://manuscripts.ru/mns/main?p_text=96362255 (дата обращения: 01.01.2020).

Авг – Легенда об Авгаре (Перенесение неруководного образа...)

Ан_Пасх – (Анонимное) поучение в неделю 5-ю по Пасхе

Ан_Прем – (Анонимное) слово о премудрости

Ан_Пянт – (Анонимное) поучение на Пянтикопии

Афрод – Сказание Афродитиана

Жит_Вас – (Житие) святителя Василия

Злат_Рожд – Слово Иоанна Златоуста на Рождество Христово

КИ1 – КИ16 – Кирилла Иерусалимского оглашения с 1-го по 16-е

КТ_Возн – Слово Кирилла Туровского на Вознесение Господне

КТ_Рассл – Слово Кирилла Туровского о расслабленном

КТ_СвОт – Слово Кирилла Туровского на сбор святых отцов...

КТ_Слеп – Слово Кирилла Туровского о слепце

КТ_Снят – Слово Кирилла Туровского о снятии тела Христова с креста...

КТ_Фом – Слово Кирилла Туровского на Фоминыю неделю (без начала)

Библиографические ссылки

1. Мартыненко Г. Я. Стилеметрия: возникновение и становление в контексте междисциплинарного взаимодействия // Структурная и прикладная лингвистика. 2014. № 10. С. 3–23.

2. Мартыненко Г. Я. Стилеметрия: возникновение и становление в контексте междисциплинарного взаимодействия. Ч. 2. Первая половина XX века // Структурная и прикладная лингвистика. 2015. № 11. С. 9–28.

3. Мартыненко Г. Я. Стилеметрия: возникновение и становление в контексте междисциплинарного взаимодействия. Ч. 3. Вторая половина XX века: ста-

новление новой стилиметрии // Структурная и прикладная лингвистика / под ред. И. С. Николаева. СПб., 2019. С. 5–33.

4. Зализняк А. А. Слово о полку Игореве. Взгляд лингвиста. М. : Языки славянской культуры, 2004. 352 с.

5. Миронова Д. М. Применение кластерного анализа в текстологии // Структурная и прикладная лингвистика. Вып. 11 : межвуз. сб. / под ред. А. С. Герда и И. С. Николаева. СПб. : Изд-во С.-Петербург. ун-та, 2015. С. 155–160.

6. Миронова Д. М. Автоматизированная классификация древних рукописей (на материале 525 списков славянского Евангелия от Матфея XI–XVI вв.) : дис. ... канд. филол. наук: 10.02.21 – Прикладная и математическая лингвистика. СПб., 2017. 315 с.

7. Сводный каталог славяно-русских рукописных книг, хранящихся в СССР. XI–XIII вв. М. : Наука, 1984. 406 с.

8. Сводный каталог славяно-русских рукописных книг, хранящихся в России, странах СНГ и Балтии. XIV век. Вып. 1 (Апокалипсис – Летопись Лаврентьевская). М. : Индрик, 2002. 768 с.

9. Зализняк А. А. Слово о полку Игореве. Взгляд лингвиста. М. : Языки славянской культуры, 2004. 352 с.

10. Словарь древнерусского языка (XI–XIV вв.): в 10 т. М. : Рус. яз. ; Азбуковник, 1988. Т. VI: (овадь – покласти). 2000. 608 с.

11. Жолобов О. Ф. О контрастирующих орфографических системах в рукописи XIII в. (к интернет-изданию Толстовского сборника) // Древняя Русь. Вопросы медиевистики. 2018. 3 (73). С. 77–89.

12. Жолобов О. Ф. Слово-притча о премудрости в списках XII–XVI вв. // Научное наследие В. А. Богородицкого и современный вектор исследований Казанской лингвистической школы : тр. и материалы Междунар. конф. (Казан. федер. ун-т, 14–17 окт. 2018 г.): в 2 т. / под общ. ред. К. Р. Галиуллина, Е. А. Горобец, Э. А. Исламовой. Казань : Изд-во Казан. ун-та, 2018. Т. 1. С. 85–90.

13. Новак М. О. Слово на Рождество Христово в Толстовском сборнике XIII в. (РНБ, Ф.п.1.39): атрибуция и композиция // Девятое Римские Кирилло-Мефодиевские чтения: материалы конф. (Рим – Салерно, 4–9 февр. 2019 г.). С. 137–142.

14. Oleg F. Zholobov, Victor A. Baranov and Maria O. Novak. Anonymous Vs. Attributed: Cluster Analysis of Tolstovskĭ Sbornik Texts and Its Interpretation in Terms of Cultural Heritage // Journal of Siberian Federal University. Humanities & Social Sciences, 2020. Vol. 13(12). Pp. 2056–2081. DOI: 10.17516/1997-1370-0704.

15. Баранов В. А., Жолобов О. Ф. Лингвостатистическое исследование частотных слов в Словах Кирилла Туровского (по рукописи РНБ, Ф.п.1.39) // Slověne. 2020. Vol. 9. No. 1. С. 29–80. DOI: 10.31168/2305-6754.2020.9.1.2. URL: http://slovene.ru/2020_1_Baranov_Zholobov.pdf (дата обращения: 20.04.2021).

16. Баранов В. А. Корреляционный анализ биграмм русских евангельских списков XI–XIV веков //

Интеллектуальные системы в производстве. 2018. Т. 16. № 4. С. 65–74.

17. Баранов В. А., Жолобов О. Ф. Древнерусские рукописи как объект статистического анализа // Интеллектуальные системы в производстве. 2019. Т. 17. № 4. С. 63–77. DOI 10.22213/2410-9304-2019-4-63-77.

18. Burrows J. 'Delta': A Measure of Stylistic Difference and a Guide to Likely Authorship // *Literary and Linguistic Computing*. 2002. Vol. 17. No. 3. Pp. 267–287.

19. Hoover D. L. Testing Burrows's Delta // *Literary and Linguistic Computing*. 2004. Vol. 19. No. 4. Pp. 453–475.

20. Argamon S. Interpreting Burrows's Delta: Geometric and Probabilistic Foundations // *Literary and Linguistic Computing*. 2008. Vol. 23. No. 2. Pp. 131–147.

21. Improving Burrows' Delta – An empirical evaluation of text distance measures / F. Jannidis, St. Pielstrom, C. Schoch, Th. Vitt // *Digital Humanities Conference*. 2015. Sydney, Australia. URL: https://www.researchgate.net/publication/280086768_Improving_Burrows%27_Delta_-_An_empirical_evaluation_of_text_distance_measures (дата обращения: 21.04.2021).

22. Understanding and explaining Delta measures for authorship attribution / St. Evert, Th. Proisl, F. Jannidis, Is. Reger, St. Pielström, Chr. Schöch, Th. Vitt // *Digital Scholarship in the Humanities*. 2017. Vol. 32. Issue 2. Pp. 114–116.

23. Hoover D. L. Testing Burrows's Delta // *Literary and Linguistic Computing*. 2004. Vol. 19. No. 4. Pp. 453–475.

24. Hoover D. Delta, Delta Prime, and Modern American Poetry: Authorship Attribution Theory and Method // *Proceedings of the 2005 ALLC/ACH Conference*. URL: <http://tomcat-stable.hcmc.uvic.ca:8080/ach/site/xhtml.xq?id=73> (дата обращения: 21.04.2021).

25. Eder M., Rybicki J. PCA, Delta, JGAAP and Polish Poetry of the 16th and the 17th Centuries: Who Wrote the Dirty Stuff? // *Digital Humanities 2009: Conference Abstracts*. MD College Park. 2009. Pp. 242–244.

26. Rybicki J., Eder M. Deeper Delta across genres and languages: Do we really need the most frequent words? // *Literary and Linguistic Computing*. 2011. Vol. 26. No 3. Pp. 315–321.

27. Abdul Razzaq A. A., Mustafa T. K. Burrows-Delta Method Fitness for Arabic Text Authorship Stylometric Detection // *International Journal of Computer Science and Mobile Computing*. 2014. Vol. 3. Issue 6. Pp. 69–78.

28. Jannidis F., Lauer G. Burrows's Delta and Its Use in German Literary History // *Distant Readings. Topologies of German Culture in the Long Nineteenth Century Studies in German Literature Linguistics and Culture* / Eds. M. Erlin, L. Tatlock. Rochester: Camden House, 2014. Pp. 29–54.

29. Towards a better understanding of Burrows's Delta in literary authorship attribution / St. Evert, Th. Proisl, F. Jannidis, St. Pielström, Chr. Schöch, Th. Vitt // *NAACL HLT Fourth Workshop on Computational Linguistics for Literature*. Denver, Colorado. 2015. Pp. 79–88. URL: https://www.researchgate.net/publication/279954745_Towards_a_better_understanding_of_Burrows's_Delta

[_in_literary_authorship_attribution](#) (дата обращения: 21.04.2021).

30. Improving Burrows' Delta – An empirical evaluation of text distance measures / F. Jannidis, St. Pielstrom, C. Schoch, Th. Vitt. *Digital Humanities Conference*, 2015, Sydney, Australia. URL: https://www.researchgate.net/publication/280086768_Improving_Burrows%27_Delta_-_An_empirical_evaluation_of_text_distance_measures (accessed 21.04.2021).

31. Savoy J. Estimating the probability of an authorship attribution // *Journal of the Association for Information Science and Technology*. 2015. Vol. 67 (6). Pp. 1–11.

32. José Calvo Tello. What does Delta see inside the Author? // *Evaluating Stylometric Clusters with Literary Metadata*. 2018. Pp. 153–161.

33. Oakes Michael P. Computer Stylometry of C.S. Lewis's *The Dark Tower* and Related Texts // *Digital Scholarship in the Humanities*. 2018. Vol. 33. Issue 3. Pp. 637–650. DOI: <https://doi.org/10.1093/lc/fqx043>

34. Метод Дельты Бёрроуза для определения авторства анонимных и псевдонимных литературных произведений на русском языке / Н. К. Мамаев, М. А. Марусенко, К. Р. Пиотровская, А. Л. Ронжин // *Proceedings of the R. Piotrowski's Readings in Language Engineering and Applied Linguistics*. СПб. : RWTH Aachen University, 2018. С. 1–14. URL: <https://clck.ru/UQKxh> (дата обращения: 20.04.2021).

35. Об авторстве «писем Берии из заточения» / В. В. Петров, М. А. Марусенко, К. Р. Пиотровская, И. Н. Маньяс, Н. К. Мамаев // *Вестник Санкт-Петербургского университета. Право*. 2019. Т. 10 (3). С. 586–605. URL: <https://lawjournal.spbu.ru/article/view/5929> (дата обращения: 21.04.2021).

36. Великанова Н. П., Орехов Б. В. Цифровая текстология: атрибуция текста на примере романа М. А. Шолохова «Тихий Дон» // *Мир Шолохова*. 2019. 1(11). С. 70–82.

37. Орехов Б. В. «Илиада» Е. И. Кострова и «Илиада» А. И. Любжина: стилиметрический аспект // *Аристей*. 2020. Т. 21. С. 282–296. URL: http://aristees.ru/21_10 (дата обращения: 20.04.2021).

38. Великанова Н. П., Орехов Б. В. Цифровая текстология: атрибуция текста на примере романа М. А. Шолохова «Тихий Дон» // *Мир Шолохова*. 2019. 1(11). С. 70–82.

39. Об авторстве «писем Берии из заточения» / В. В. Петров, М. А. Марусенко, К. Р. Пиотровская, И. Н. Маньяс, Н. К. Мамаев // *Вестник Санкт-Петербургского университета. Право*. 2019. Т. 10 (3). С. 586–605. URL: <https://lawjournal.spbu.ru/article/view/5929> (дата обращения: 21.04.2021).

40. Там же.

41. Великанова Н. П., Орехов Б. В. Указ соч.

42. Мамаев Н. К. [и др.] Указ соч.

43. Там же.

44. [Петров В. В. и др.] Указ соч.

45. Eder M., Rybicki J., Kestemont M. Stylometry with R: A package for computational text analysis // *The R Journal*. 2016, Vol. 8, No. 1. Pp. 107–121. URL:

<https://journal.r-project.org/archive/2016/RJ-2016-007/index.html> (дата обращения: 20.04.2021).

46. Жолобов О. Ф. Толстовский сборник XIII в. как мегатекст // Лингвокультурологические исследования развития русского языка в условиях полиэтнической среды: опыт и перспективы (1–4 октября 2018 г.): тр. и матер.: в 2 т. / под общ. ред. Е. А. Горобец, О. Ф. Жолобова, М. О. Новак. Казань: Изд-во Казан. ун-та, 2018. Т. 2. С. 73–77.

47. Oleg F. Zholobov, Victor A. Baranov and Maria O. Novak. Anonymous Vs. Attributed: Cluster Analysis of Tolstovskĭ Sbornik Texts and Its Interpretation in Terms of Cultural Heritage // Journal of Siberian Federal University. Humanities & Social Sciences. 2020. Vol. 13(12). Pp. 2056–2081. DOI: 10.17516/1997-1370-0704.

48. Жолобов О. Ф. О контрастирующих орфографических системах в рукописи XIII в. (к интернет-изданию Толстовского сборника) // Древняя Русь. Вопросы медиевистики. 2018. 3 (73). С. 77–89.

49. Жолобов О. Ф. Слово-притча о премудрости в списках XII–XVI вв. // Научное наследие В. А. Богородицкого и современный вектор исследований Казанской лингвистической школы: тр. и матер. междунар. конф. (Казан. федер. ун-т, 14–17 окт. 2018 г.): в 2 т. / под общ. ред. К. Р. Галиуллина, Е. А. Горобец, Э. А. Исламовой. Казань: Изд-во Казан. ун-та, 2018. Т. 1. С. 85–90.

50. Новак М. О. Указ. соч.

51. Oleg F. Zholobov, Victor A. Baranov and Maria O. Novak. Anonymous Vs. Attributed: Cluster Analysis of Tolstovskĭ Sbornik Texts and Its Interpretation in Terms of Cultural Heritage // Journal of Siberian Federal University. Humanities & Social Sciences. 2020. Vol. 13(12). Pp. 2056–2081. DOI: 10.17516/1997-1370-0704.

References

- Martynenko G. Ja. [Stylemetry: emergence and formation in the context of interdisciplinary interaction] *Strukturnaja i prikladnaja lingvistika*, 2014, no. 10, pp. 3–23. (in Russ).
- Martynenko G. Ja. [Stylemetry: emergence and formation in the context of interdisciplinary interaction]. Part 2. The first half of the XX century]. *Strukturnaja i prikladnaja lingvistika*, 2015, no. 11, pp. 9–8 (in Russ).
- Martynenko G. Ja. [Stylemetry: emergence and formation in the context of interdisciplinary interaction]. Part 3. The second half of the XX century: the formation of a new style. *Strukturnaja i prikladnaja lingvistika / pod redakciej I. S. Nikolaeva*. Sankt-Peterburg, 2019, pp. 5–33. (in Russ).
- Zaliznjak A. A. *Slovo o polku Igoreve. Vzgljad lingvista* [A word about Igor's regiment. Linguist's view. Languages of Slavic Culture]. Moscow, Languages of Slavic Culture Publ., 2004, 352 p. (in Russ).
- Mironova D. M. *Primenenie klasternogo analiza v tekstologii*. *Strukturnaja i prikladnaja lingvistika. Vypusk 11, mezhvuzovskij sbornik, pod redakciej A. S. Gerda i I. S. Nikolaeva* [Application of cluster analysis in textology // Structural and Applied Linguistics. Issue 11, interuniversity collection, edited by A. S. Gerd and I. S. Nikolaev] St. Petersburg, St. Petersburg University Publ., 2015, pp. 155–160. (in Russ).
- Mironova D. M. *Avtomatizirovannaja klassifikacija drevnih rukopisej (na materiale 525 spiskov slavjanskogo Evangelija ot Matfeja XI–XVI vv.) : dis. ... kand. filol. nauk: 10.02.21 – Prikladnaja i matematičeskaja lingvistika* [Automated classification of ancient manuscripts (based on 525 lists of the Slavic Gospel of Matthew XI–XVI centuries): dissertation of the candidate of philological sciences: 10.02.21 - Applied and mathematical linguistics]. St. Petersburg, 2017, 315 p. (in Russ).
- Svodnyj katalog slavjano-russkikh rukopisnyh knig, hranjashhihsja v SSSR. XI–VIII veka* [Consolidated catalog of Slavic-Russian manuscript books stored in the USSR. XI – XIII centuries] Moscow, Nauka Publ., 1984. 406 p. (in Russ).
- Svodnyj katalog slavjano-russkikh rukopisnyh knig, hranjashhihsja v Rossii, stranah SNG i Baltii. XIV vek. Vyp. 1 (Apokalipsis – Letopis' Lavrent'evskaja)* [Consolidated catalog of Slavic-Russian manuscript books stored in Russia, the CIS countries and the Baltic states. XIV century. Issue 1 (Apocalypse - Laurentian Chronicle)]. Moscow, Indrik Publ., 2002, 768 p. (in Russ).
- Zaliznjak A. A. *Slovo o polku Igoreve. Vzgljad lingvista* [A word about Igor's regiment. Linguist's view]. Moscow, Languages of Slavic Culture, 2004, 352 p. (in Russ).
- Slovar' drevnerusskogo jazyka (XI–XIV vv.): v 10 tomah* [Dictionary of the Old Russian language (XI–XIV centuries): in 10 volumes. Vol. VI: (ovad - poklasty)]. Moscow, Russian language Publ., Azbukovnik Publ., 1988, 2000, 608 p. (in Russ).
- Zholobov O. F. [On the contrasting spelling systems in the 13th century manuscript (to the Internet edition of the Tolstoy collection)]. *Drevnjaja Rus'. Voprosy medievistiki*, 2018, 3 (73), pp. 77–89. (in Russ).
- Zholobov O. F. *Slovo-pritcha o premudrosti v spiskah XII–XVI vv. Nauchnoe nasledie V. A. Bogorodickogo i sovremennij vektor issledovanij Kazanskoj lingvisticheskoj shkoly: trudy i materialy Mezhnarodnoj konferencii: v 2 tomah, pod obshhej redakciej K. R. Galiullina, E. A. Gorobec, Je. A. Islamovoj* [Wordparable about wisdom in the lists of the XII–XVI centuries // Scientific heritage of V. A. Bogoroditsky and the modern vector of research of the Kazan linguistic school, proceedings and materials of the International conference, in 2 volumes / edited by K. R. Galiullin, E. A. Gorobets, E. A. Islamova]. Kazan, Kazan Publ. University, 2018, vol. 1, pp. 85–90. (in Russ).
- Novak M. O. *Slovo na Rozhdestvo Hristovo v Tolstovskom sbornike XIII veka (RNB, F.p.I.39): atribucija i kompozicija. Devjatye Rimskie Kirillo-Mefodievskie chtenija: materialy konferencii (Rim – Salerno, 4–9 fevralja 2019 goda)* [Word for Christmas in the Tolstoy collection of the 13th century (RNL, F.p.I.39): attribution and composition // Ninth Roman Cyril and Methodius Readings, conference proceedings (Rome - Salerno, February 4–9, 2019 year)], Rome, Salerno, 2019, pp. 137–142. (in Russ).

14. Oleg F. Zholobov, Victor A. Baranov and Maria O. Novak. Anonymous Vs. Attributed: Cluster Analysis of Tolstovskĭ Sbornik Texts and Its Interpretation in Terms of Cultural Heritage // *Journal of Siberian Federal University. Humanities & Social Sciences*, 2020, vol. 13(12), pp. 2056–2081. DOI: 10.17516/1997-1370-0704.
15. Baranov V. A., Zholobov O. F. [Lingvostatisticheskoe study of frequency words in the Words of Kirill Turovsky (according to the manuscript of the National Library of Russia, F.p.I.39)]. *Slověne*, 2020, vol. 9, no. 1, pp. 29–80]. (in Russ).
16. Baranov V. A. [Correlation analysis of bigrams of Russian evangelical copies of the XI-XIV centuries] *Intellektual'nye sistemy v proizvodstve*, 2018, vol. 16, no. 4, pp. 65–74. (in Russ).
17. Baranov V. A., Zholobov O. F. [Old Russian manuscripts as an object of statistical analysis] *Intellektu-al'nye sistemy v proizvodstve*, 2019, vol. 17, no. 4, pp. 63–77. DOI 10.22213/2410-9304-2019-4-63-77. (in Russ).
18. Burrows J. 'Delta': A Measure of Stylistic Difference and a Guide to Likely Authorship. *Literary and Linguistic Computing*, 2002, vol. 17, no. 3, pp. 267–87.
19. Hoover D. L. Testing Burrows's Delta. *Literary and Linguistic Computing*, 2004, vol. 19, no. 4, pp. 453–475.
20. Argamon S. Interpreting Burrows's Delta: Geometric and Probabilistic Foundations. *Literary and Linguistic Computing*, 2008, vol. 23, no. 2, pp. 131–147.
21. Jannidis F., Pielstrom St., Schoch C., Vitt Th. Improving Burrows' Delta – An empirical evaluation of text distance measures. *Digital Humanities Conference*, 2015, Sydney, Australia. Available at: https://www.researchgate.net/publication/280086768_Improving_Burrows%27_Delta_-_An_empirical_evaluation_of_text_distance_measures (accessed 21.04.2021).
22. Evert St., Proisl Th., Jannidis F., Reger Is., Pielström St., Schöch Chr., Vitt Th. Understanding and explaining Delta measures for authorship attribution. *Digital Scholarship in the Humanities*, 2017, vol. 32, issue 2, pp. 114–116.
23. Hoover D. L. Testing Burrows's Delta. *Literary and Linguistic Computing*, 2004, vol. 19, no. 4, pp. 453–475.
24. Hoover D. Delta, Delta Prime, and Modern American Poetry: Authorship Attribution Theory and Method. *Proceedings of the 2005 ALLC/ACH Conference*. Available at: <http://tomcat-stable.hcmc.uvic.ca:8080/ach/site/xhtml.xq?id=73> (accessed 21.04.2021).
25. Eder M., Rybicki J. PCA, Delta, JGAAP and Polish Poetry of the 16th and the 17th Centuries: Who Wrote the Dirty Stuff? *Digital Humanities 2009: Conference Abstracts*. MD College Park, 2009, pp. 242–244.
26. Rybicki J., Eder M. Deeper Delta across genres and languages: Do we really need the most frequent words? *Literary and Linguistic Computing*, 2011, vol. 26, no. 3, pp. 315–321.
27. Abdul Razzaq A. A., Mustafa T. K. Burrows-Delta Method Fitness for Arabic Text Authorship Stylometric Detection. *International Journal of Computer Science and Mobile Computing*, 2014, vol. 3, issue 6, pp. 69–78.
28. Jannidis F., Lauer G. Burrows's Delta and Its Use in German Literary History. *Distant Readings. Topologies of German Culture in the Long Nineteenth Century Studies in German Literature Linguistics and Culture*. Eds. M. Erlin, L. Tatlock. Rochester: Camden House, 2014, pp. 29–54.
29. Evert St., Proisl Th., Jannidis F., Pielström St., Schöch Chr., Vitt Th. Towards a better understanding of Burrows's Delta in literary authorship attribution. *NAACL HLT Fourth Workshop on Computational Linguistics for Literature*. Denver, Colorado, 2015, pp. 79–88. Available at: https://www.researchgate.net/publication/279954745_Towards_a_better_understanding_of_Burrows's_Delta_in_literary_authorship_attribution (accessed 21.04.2021).
30. Jannidis F., Pielstrom St., Schoch C., Vitt Th. Improving Burrows' Delta – An empirical evaluation of text distance measures. *Digital Humanities Conference*, 2015, Sydney, Australia. Available at: https://www.researchgate.net/publication/280086768_Improving_Burrows%27_Delta_-_An_empirical_evaluation_of_text_distance_measures (accessed 21.04.2021).
31. Savoy J. Estimating the probability of an authorship attribution. *Journal of the Association for Information Science and Technology*, 2015, vol. 67 (6), pp. 1–11.
32. José Calvo Tello. What does Delta see inside the Author? *Evaluating Stylometric Clusters with Literary Metadata*, 2018, pp. 153–161.
33. Oakes Michael P. Computer Stylometry of C.S. Lewis's *The Dark Tower* and Related Texts. *Digital Scholarship in the Humanities*, 2018, vol. 33, issue 3, pp. 637–650. DOI: <https://doi.org/10.1093/lhc/fqx043>
34. Mamaev N. K., Marusenko M. A., Piotrovskaja K. R., Ronzhin A. L. [Burroughs Delta Method for Authorship of Anonymous and Pseudonymous Literary Works in Russian]. *Proceedings of the R. Piotrowski's Readings in Language Engineering and Applied Linguistics*. St. Petersburg, RWTH Aachen University, 2018, pp. 1–14. Available at: <https://clck.ru/UQKXh> (accessed 20.04.2021). (in Russ).
35. Petrov V. V., Marusenko M. A., Piotrovskaja K. R., Man'jas I. N., Mamaev N. K. [About the authorship of "Beria's letters from captivity"]. *Vestnik Sankt-Peterburgskogo universiteta. Pravo*, 2019, vol. 10 (3), pp. 586–605. Available at: <https://lawjournal.spbu.ru/article/view/5929> (accessed 21.04.2021). (in Russ).
36. Velikanova N. P., Orehov B. V. [Digital textual criticism: attribution of text on the example of the novel "Quiet Don" by M. A. Sholokhov]. *Mir Sholohova*, 2019, 1 (11), pp. 70–82. (in Russ).
37. Orehov B. V. [Kostrov and "Iliad" by A. I. Lyubzhin: a stylometric aspect]. *Aristey*, 2020, vol. 21, pp. 282–296. Available at: http://aristeas.ru/21_10 (accessed 20.04.2021). (in Russ).
38. Velikanova N. P., Orehov B. V. [Digital textual criticism: attribution of text on the example of the novel by M. A. Sholokhov "Quiet Don"]. *Mir Sholohova*, 2019, 1 (11), pp. 70–82. (in Russ).

39. Petrov V. V., Marusenko M. A., Piotrovskaja K. R., Man'jas I. N., Mamaev N. K. [About the authorship of "Beria's letters from captivity"]. *Vestnik Sankt-Peterburgskogo universiteta. Pravo*, 2019, vol. 10 (3), pp. 586-605. Available at: <https://lawjournal.spbu.ru/article/view/5929> (accessed 21.04.2021).
40. Petrov V. V., Marusenko M. A., Piotrovskaja K. R., Man'jas I. N., Mamaev N. K. [About the authorship of "Beria's letters from captivity"]. *Vestnik Sankt-Peterburgskogo universiteta. Pravo*, 2019, vol. 10 (3), pp. 586-605. Available at: <https://lawjournal.spbu.ru/article/view/5929> (accessed 21.04.2021).
41. Velikanova N. P., Orehov B. V. [Digital textual criticism: attribution of text on the example of the novel by M. A. Sholokhov "Quiet Don"]. *Mir Sholohova*, 2019, 1 (11), pp. 70–82. (in Russ).
42. Mamaev N. K., Marusenko M. A., Piotrovskaja K. R., Ronzhin A. L. [Burroughs Delta Method for Authorship of Anonymous and Pseudonymous Literary Works in Russian]. *Proceedings of the R. Piotrowski's Readings in Language Engineering and Applied Linguistics*. St. Petersburg, RWTH Aachen University, 2018, pp. 1–14. Available at: <https://clck.ru/UQKxh> (accessed 20.04.2021). (in Russ).
43. Mamaev N. K., Marusenko M. A., Piotrovskaja K. R., Ronzhin A. L. [Burroughs Delta Method for Authorship of Anonymous and Pseudonymous Literary Works in Russian]. *Proceedings of the R. Piotrowski's Readings in Language Engineering and Applied Linguistics*. St. Petersburg, RWTH Aachen University, 2018, pp. 1–14. Available at: <https://clck.ru/UQKxh> (accessed 20.04.2021). (in Russ).
44. Petrov V. V., Marusenko M. A., Piotrovskaja K. R., Man'jas I. N., Mamaev N. K. [About the authorship of "Beria's letters from captivity"]. *Vestnik Sankt-Peterburgskogo universiteta. Pravo*, 2019, vol. 10 (3), pp. 586-605. Available at: <https://lawjournal.spbu.ru/article/view/5929> (accessed 21.04.2021). (in Russ).
45. Eder M., Rybicki J., Kestemont M. Stylometry with R: A package for computational text analysis. *The R Journal*. 2016, vol. 8, no. 1, pp. 107–121. Available at: <https://journal.r-project.org/archive/2016/RJ-2016-007/index.html> (accessed 20.04.2021).
46. Zholobov O. F. *Tolstovskij sbornik XIII vek kak megatekst. Lingvokul'turologicheskie issledovanija razvitiija russkogo jazyka v uslovijah polijeticheskoj sredy: opyt i perspektivy (1–4 oktjabrja 2018 g.): trudy i mater: v 2 tomah / pod obshhej redakciej E. A. Gorobec*, O. F. Zholobova, M. O. Novak [Tolstoy's collection of the 13th century as a megatext. Linguistic and cultural studies of the development of the Russian language in a multi-ethnic environment: experience and prospects (October 1-4, 2018): works and materials. In 2 vol. Edited by E. A. Gorobets, O. F. Zholobov, M. O. Novak]. Kazan, Kazan University Publishing House, 2018, vol. 2, p. 73-77. (in Russ).
47. Oleg F. Zholobov, Victor A. Baranov and Maria O. Novak. Anonymous Vs. Attributed: Cluster Analysis of Tolstovskij Sbornik Texts and Its Interpretation in Terms of Cultural Heritage. *Journal of Siberian Federal University. Humanities & Social Sciences*, 2020, vol. 13(12), pp. 2056-2081. DOI: 10.17516/1997-1370-0704.
48. Zholobov O. F. [On the contrasting spelling systems in the 13th century manuscript (to the Internet edition of the Tolstoy collection)] *Drevnjaja Rus'. Voprosy medievistiki*, 2018, 3 (73), pp. 77-89. (in Russ).
49. *Slovo-pritcha o premudrosti v spiskah XII–XVI vekov. Nauchnoe nasledie V. A. Bogorodickogo i sovremennij vektor issledovanij Kazanskoj lingvisticheskoj shkoly: trudy i materialy Mezhdunarodnoj konferencii. v 2 tomah. Pod obshhej redakciej K. R. Galiullina, E. A. Gorobec, Je. A. Islamovoj* [Word-parable about wisdom in the lists of the KhII – XVI centuries // Scientific heritage of V. A. Bogoroditsky and the modern vector of research of the Kazan linguistic school: proceedings and materials of the International conference. in 2 volumes. Edited by K. R. Galiullin, E. A. Gorobets, E. A. Islamova]. Kazan, Kazan University Publishing House, 2018, volume 1, p. 85-90. (in Russ).
50. Novak M. O. *Slovo na Rozhdestvo Hristovo v Tolstovskom sbornike XIII veka (RNL, F.p.I.39): atribucija i kompozicija. Devjatyje Rimskie Kirillo-Mefodievskie chtenija: materialy konferencii (Rim – Salerno, 4–9 fevralja 2019 goda)* [Word for Christmas in the Tolstoy collection of the 13th century (RNL, F.p.I.39): attribution and composition // Ninth Roman Cyril and Methodius Readings, conference proceedings (Rome - Salerno, February 4-9, 2019 year)], Rome, Salerno, 2019, pp. 137-142. (in Russ).
51. Oleg F. Zholobov, Victor A. Baranov and Maria O. Novak. Anonymous Vs. Attributed: Cluster Analysis of Tolstovskij Sbornik Texts and Its Interpretation in Terms of Cultural Heritage. *Journal of Siberian Federal University. Humanities & Social Sciences*, 2020, vol. 13 (12), pp. 2056-2081. DOI: 10.17516/1997-1370-0704.

V. A. Baranov, Doctor of Philology, Professor
Kalashnikov Izhevsk State Technical University, Izhevsk, Russia
O. F. Zholobov, Doctor of Philology, Professor
Kazan (Volga Region) Federal University, Kazan, Russia

LINGVO-STATISTICAL ANALYSIS OF THE OLD RUSSIAN ANTHOLOGY (RNL, F.p.I. 39)

A linguistic analysis of 29 texts of the 13th century manuscript (the Tolstoy miscellany, RNL, F.p.I. 39), containing the original works of the ancient Russian author Kirill Turovsky and translations of the texts of Cyril of Jerusalem, John Chrysostom and unknown authors is presented. The objectives of the work are a) verification of the conclusions about the authorship of three anonymous texts of the manuscript attributed to Kirill Turovsky, made in the course of

their previous analysis using traditional linguistic methods and correlation methods of statistics, and also b) establishing the possibility of using statistical methods to compare small texts.

To measure the distance between the texts, the Delta method and its modification the Cosine-Delta were used, a different number of the most frequent words (from 50 to 250) of each of the texts were used. The data were extracted from the electronic edition of the Tolstoy miscellany. Calculations and construction of dendrograms were carried out using the Stylo package.

Dendrogram analysis confirmed the findings of other studies: the three anonymous texts – Parable of Wisdom, Sermon on Pentecost and Sermon on the 5th Sunday after Easter – do not belong to Kirill Turovsky. The dependence of the grouping of texts on the size of the texts and applied methods was found. It was found that the cosine-Delta method, in contrast to the classical Delta, turns out to be insensitive to differences in the size of texts. In addition, an increase in the number of the most frequent words when using cosine-delta allows one to find a correlation between groupings of texts and their discursive and lexico-syntactic affinity.

Keywords: Tolstoy miscellany of the XIII century; Kirill Turovsky; anonymous texts; attribution; Delta method; stylometry; linguistic textology.

Получено: 11.05.2021

Образец цитирования

Баранов В. А., Жолобов О. Ф. Лингвостатистический анализ древнерусской антологии (РНБ, Ф.п.И. 39) // Социально-экономическое управление: теория и практика. 2021. № 2 (45). С. 50–60. DOI: 10.22213/2618-9763-2021-2-50-60

For Citation

Baranov V. A., Zholobov O. F. [Lingvo-Statistical Analysis of the Old Russian Anthology (RNL, F.p.I. 39)]. *Social'no-jekonomicheskoe upravlenie: teorija i praktika*, 2021, no. 2 (45), pp. 50-60 (in Russ.). DOI: 10.22213/2618-9763-2021-2-50-60