

К СТИЛЕМЕТРИЧЕСКОМУ ПОРТРЕТУ ТЕКСТОВ СТУДИЙСКОЙ ТРАДИЦИИ: ОПЫТ АНАЛИЗА СРЕДНЕ- И ПОЗДНЕВИЗАНТИЙСКОГО КОРПУСА МЕТОДОМ DELTA

Э. Ю. КАНАЕВА
ПСТГУ, Москва
kanaeva.elga@gmail.com

ELGA I. KANAeva

St. Tikhon's Orthodox University for Humanities

TOWARDS A STYLOMETRIC PROFILE OF THE STUDITE TEXTS: AN ANALYSIS OF A MIDDLE AND LATE
BYZANTINE CORPUS USING THE DELTA METHOD

ABSTRACT. The article presents a case study in the stylometric analysis to Middle Byzantine texts, focusing on the attribution of the "Catechetical Homilies" to Patriarch Antony III Studites (10th century). Confirmation or refutation of this hypothesis could make an important contribution to research on the intellectual history of Studite monasticism. The primary aim is to evaluate the efficacy of various Delta metric modifications (classic Burrows's, Eder's, and Cosine) for a corpus of homiletic, hagiographic, and epistolary works, and to apply a calibrated method to this contested attribution. A representative corpus of key Studite tradition authors and their contemporaries was compiled. Through a series of experiments, optimal analysis parameters—including metric choice, the number of most frequent words, and the necessity of lemmatization—were determined by calibrating the method against control samples: 14th-century texts with undisputed authorship and the corpus of St. Symeon the New Theologian. The study assesses the stylistic proximity between Antony Studites' undisputed and disputed texts. The results confirm the general applicability of the Delta method to Middle Greek material, while highlighting its parameter dependence. For longer texts (>10,000 words), classic Delta with lemmatization and a set of 500–750 most frequent words yielded the best outcomes. For shorter texts (~5,000 words), Eder's Delta or Cosine Delta applied to raw word forms with 250 most frequent words proved more effective. The method exhibited high sensitivity to textual genre. The findings support the attribution of the "Catechetical Homilies" to Antony Studites. Additionally, the analysis revealed a strong stylistic affinity between the corpora of Symeon the New Theologian and his disciple Niketas Stethatos and suggested a hypothesis of common authorship for two 10th-century hagiographical works. These results position stylometry as a viable auxiliary tool for investigating intellectual and literary networks within the Byzantine monastic tradition.

KEYWORDS: Stylometry, Anthony the Studite, Symeon the New Theologian, the Studite Monastery, Byzantine monasticism, hagiography.

* Исследование выполнено за счет гранта Российского научного фонда № 24-78-10165, «Богословские памятники византийской литературы XI-XV вв.: проблемы языка, содержания и подлинности». Организация выполнения проекта – Православный Свято-Тихоновский гуманитарный университет. This study has been funded by the Russian Science Foundation, grant no. 24-78-10165. The project was carried out at St. Tikhon's Orthodox University for the Humanities.

Введение. Стилеметрия: от классической филологии и обратно

Задача атрибуции и атетирования авторства текста является почтенной проблемой классической филологии. Исследования гомеровского вопроса, активизированные в конце XVIII века Ф. А. Вольфом, а также многократные попытки структурировать сохранившийся корпус Платона привели к возникновению стилеметрического аргумента. Поиск уникального авторского «почерка» (fingerprint в англоязычной литературе), который позволил бы отличить одного автора от другого, занимал и продолжает занимать учёных с XVIII века¹. Основы стилеметрического подхода, «закон стилистического сродства»², а также сам термин «стилеметрия» (в русскоязычных публикациях также часто встречается вариант «стилометрия») были введены в науку в 1897 г. польским филологом Винценцы Лютославским в контексте платоновских исследований³. Идея метода состоит в количественном сравнении текстов на основе наименее контролируемых автором параметров, относящихся не к содержанию, а к форме высказывания. Поскольку корректная атрибуция является важнейшим условием для реконструкции истории идей и понятий, стилеметрия становится предметом интереса не только филологов, но и историков мысли.

С развитием вычислительных методов и технологий, а также с оцифровкой языковых корпусов стало возможным экспериментально исследовать эффективность различных стилеметрических подходов на больших массивах текстов. В результате проверки множества характеристик (повторяемость цепочек символов, частотность и распределение частей речи, частотность грамматических конструкций, длина слов и предложений и др.) наилучшие результаты стали показывать методы, основанные на анализе частотности слов и словоформ⁴. В свою очередь, среди различных метрик изме-

¹ Егунов 1996, 104–128.

² Lutosławski 1897, 152.

³ Kovalev 2024, 351–356.

⁴ Орехов 2023, 146–156.

рения расстояния между текстами по этому параметру в среднем значительную эффективность продемонстрировала метрика Delta Бёрроуза⁵, которая за два десятилетия стала базовым цифровым методом решения задач атрибуции⁶. Распространению этой технологии существенно способствовало создание общедоступной программной библиотеки *stylo* для языка R с удобным интерфейсом и средствами визуализации, сделавшей метод доступным для исследователей, далёких от программирования⁷.

Стилеметрия, изначально порождённая поисками ответов на вопросы классической филологии, в своей цифровой версии утратила связь с исходным материалом. Первые и наиболее заметные успехи были достигнуты в рамках романо-германской, а затем и русской филологии⁸. Лишь в последнее время стали появляться работы, посвящённые апробации метода на древнегреческих текстах. Греческий корпус обладает рядом свойств, выделяющих его на фоне тех языков, на которых метод уже доказал свою эффективность: сравнительно большой словарь, развитая система аналитических форм, стремление авторов к подражанию стилю великих предшественников, обилие псевдоэпиграфов, а также сравнительно небольшой объём многих сохранившихся текстов. Эти свойства оказывают непосредственное влияние на работу математической модели, лежащей в основе метода Delta, поэтому его применение к древнегреческим текстам требует уточнения ряда параметров: оптимального количества наиболее частотных слов (MFW), минимального объёма анализируемого фрагмента, целесообразности использования лемм или словоформ, а также сравнительной эффективности различных модификаций метода.

Подобное исследование на базе открытого корпуса Diorisis было проведено О. В. Алиевой⁹. В выборку вошли тексты 14 авторов с V в. до н. э. до II в. н. э. (от Геродота до Плутарха), писавших в различных жанрах. Исследовательница подтвердила эффективность метода и сделала ряд важных наблюдений: для древнегреческих текстов требуемый объём анализируемых фрагментов может быть сокращён до 1000–5000 слов, что свидетельствует об их высокой чувствительности к стилеметрическому анализу. В последующих публикациях О. Алиева расширяла как объём привлекаемого корпуса (например, в одной из работ¹⁰ использовался корпус в 694 тыс. слов из

⁵ Burrows 2002.

⁶ Evert 2017, 114–116.

⁷ Eder 2016.

⁸ Мамаев 2018, 4; Орехов 2023, 153–157.

⁹ Алиева 2022.

¹⁰ Алиева 2024b.

«Perseus Digital Library's Greek Collection»¹¹, а в других¹² – корпус платоновских текстов, включая *pseudoplatonica*), так и вариативность применяемых технологий (различные метрики расстояний¹³ и различные способы обработки и визуализации матриц вычисленных расстояний¹⁴). Эти публикации, если следовать выражению самого автора, обнаруживают как «блеск», так и «нищету» стилиметрического подхода: с одной стороны, различные метрики убедительно показывают стилистическую близость текстов и тождество авторов для известных произведений (фаворит испытаний, косинусное расстояние, даёт более 96% точности при разумном подборе параметров¹⁵); с другой стороны, стилиметрия никогда не может стать решающим аргументом при решении вопроса об атрибуции или атетировании текста, поскольку сама по себе близость или удалённость стилей двух текстов не несёт информации об авторстве. Вывод неутешителен: во многих случаях «количественные методы оказываются дырявым неводом там, где улов достигается простой удочкой»¹⁶. Тем не менее, необходимо учитывать, что автор берётся за сложнейшую задачу классической филологии – упорядочивание платоновского корпуса, включающего высококачественные подражания, тексты близких по академической среде авторов и эволюцию стиля отдельных писателей. Можно ожидать, что для корпусов меньшей стилистической плотности «блеск» стилиметрии может проявиться в большей степени, чем её «нищета», хотя она всегда останется лишь вспомогательным методом анализа.

Крупное исследование применимости стилиметрии к древнегреческим текстам было проведено в 2020 году Грантом Стори и Дэвидом Мимно¹⁷. В своей публикации они решали вопрос о том, оказывает ли жанр произведения существенное влияние на его стилистические особенности, учитывая свойственный греческой литературе приём подражания. Авторы исследовали стилиметрические расстояния между 464 работами 92 авторов с VIII в. до н. э. до VI в. н. э., опубликованных в «Perseus Digital Library's Greek Collection». В рамках этого исследования был сделан ряд значимых выводов, которые мы обсудим ниже.

Упомянутые исследования рассматривают вопросы применимости стилиметрии к античному и позднеантичному литературному материалу, а в этой

¹¹ Smith 2000.

¹² Алиева 2024a; Алиева 2025.

¹³ Алиева 2024b.

¹⁴ Алиева 2025.

¹⁵ Алиева 2024b, 17, 23–24.

¹⁶ Алиева 2024a, 44

¹⁷ Storey 2020.

статье мне хотелось бы дополнить результаты апробации метода наблюдениями, сделанными на основе текстов средне- и поздневизантийского периодов (IX–XIV вв.). Данное исследование не претендует на полноту апробации численных методик (особенно на фоне упомянутых выше работ), а остаётся в рамках жанра «опытов» их применения. Надеюсь, тем не менее, что этот опыт будет полезен тем, кто рассматривает возможность применения количественного анализа к византийским текстам.

В фокусе моего внимания находится текст «Огласительных поучений», приписываемых Антонию III Студиту, патриарху Константинопольскому (973–978). Окончательная атрибуция этих текстов может дать ценный материал для изучения истории византийской мысли, поскольку Антоний является главным претендентом на роль недостающего звена в цепи традиции от прп. Феодора Студита к прп. Симеону Новому Богослову¹⁸ (анализу содержания и истории этого преемства я надеюсь посвятить отдельную статью). Целью этой статьи является, таким образом, поиск наиболее результативных параметров стилеметрического исследования для византийского материала на базе инструментов, доступных учёному, неискущённому в программировании, и проверка атрибуции «Огласительных поучений» Антонию Студиту средствами стилистического анализа. Текст был издан в 2004 году Жюльеном Леруа на основании материалов, подготовленных Оливье Делуи¹⁹.

Хотя у нас нет более подходящих кандидатур на авторство этих текстов, атрибуция их патриарху Антонию не является неоспоримой. Против этой атрибуции выдвигается ряд доводов, один из которых состоит в отсутствии сведений о том, что патриарх Антоний когда-либо был Студийским игуменом. Хотя сам по себе аргумент от недостаточности информации не является решающим, дополнительные свидетельства в её пользу могли бы повысить уровень достоверности атрибуции.

Источники

В нашем случае необходимо сравнить стиль четырёх «Огласительных поучений» со стилем текстов, чья принадлежность патриарху Антонию бесспорна. Такими текстами являются его письмо к императору Василию II, письмо к Дорофею и молитва, а также гомилия, произнесённая во Влахернах²⁰. Стилеметрический анализ чувствителен к жанру произведений, поскольку разные литературные формы требуют разных средств выразительности. Поэтому

¹⁸ Turner 1990, 137; Papadopoulos-Kerameus 1905, 807.

¹⁹ Delouis 2004.

²⁰ Delouis 2004; Darrouzès 1960; Sternbach 1975.

наилучшие результаты он даёт при сравнении текстов одного жанра. К сожалению, в нашем случае достичь этого идеала невозможно. Однако если анализ покажет единство стиля, несмотря на разницу в жанрах, это будет ещё более сильным аргументом в пользу атрибуции.

Для проведения анализа был сформирован корпус текстов, включающий:

– *Основную группу (тексты авторов, писавших в жанре огласительных поучений или близких к ним)*: «Огласительные поучения» ряда авторов: приписываемые Антонию Студиту, прп. Симеона Нового Богослова, прп. Феодора Студита, Никиты Стифата, прп. Неофита Затворника, Иоанна Мавропода; близкие по жанру тексты: поучения патриарха Иоанна Оксита, «Слово особенное к Пастырю» Иоанна Лествичника; другие тексты тех же авторов: письма и молитва Антония Студита, «Послание об исповеди» прп. Симеона Нового Богослова, чья атрибуция не вызывает сомнений после работы Холя²¹, Житие Симеона Нового Богослова Никиты Стифата²².

– *Дополнительную группу*: 9 житий X–XI вв.: Димитриана Хитридского, Илии Спелеота, патр. Антония Кавлея, совместное житие Косьмы Песнописца и Иоанна Дамаскина, отдельное Иоанна Дамаскина, Илии Гелиопольского, Ирины Хрисовалантской, Илии Нового, акты Давида, Симеона и Георгия Митилинских²³. Включение данной группы преследовало цель увеличения общего объёма корпуса в целевом периоде, а также проверки влияния жанра сочинения на стилиметрический анализ.

– *Контрольную группу*: письма середины XIV в.: Варлаама Калабрийского, свт. Григория Паламы и Григория Акиндина²⁴. Включение данной группы обосновано задачей дополнительной валидации метода на текстах с бесспорной атрибуцией.

²¹ Holl 1898, 1–36.

²² Издания в порядке перечисления: Delouis 2004; Syméon 1963; Papadopoulos-Kerameus 1904; Darrouzès 1961; Sotiroudis 1998; Bollig 1882; Creazzo 2004; Migne 1864, T. 88, Col. 1165–1209; Delouis 2004; Darrouzès 1960; Sternbach 1975; Turner 2009; Hausherr 1928.

²³ Издания в порядке перечисления: Grégoire 1907; Acta 1868, sept., t. 3; Leone 1989; Papadopoulos-Kerameus 1897; Papadopoulos-Kerameus 1907; Migne 1864, T. 94, col. 429–490; Rosenqvist 1986; Taibbi 1962; Gheyn 1899.

²⁴ Schirò 1954; Meyendorff 1962; Constantinides-Hero 1983.

Таким образом, в корпус вошли 53 произведения 21 автора общим объёмом 242 600 слов²⁵.

В этом корпусе тексты Никиты Стифата имеют особое значение по следующим причинам: во-первых, они соединяют собой основную и дополнительную группу источников, поскольку Никита писал как в жанре огласительных поучений, так и в жанре житий. Во-вторых, Никита был издателем «Огласительных поучений» Симеона Нового Богослова, которые также вошли в выборку. Наличие этих текстов позволяет оценить результаты стилеметрического анализа: оставил ли Никита заметный авторский сигнал при подготовке к изданию текстов Симеона? Обладали ли Никита-житописатель и Никита-проповедник единым стилем?

Параметры экспериментов

Метрика. Исследование Стори и Мимно показало, что дивергенция Йенсена – Шеннона на греческих текстах даёт верную атрибуцию на 2,5% чаще, чем классическая Burrows' Delta. В свою очередь, классическая Delta показала лучшие результаты, чем все остальные проверявшиеся в исследовании метрики; в частности, точность её предсказаний была в среднем на 4,5% выше, чем у результатов, полученных с помощью косинусного расстояния²⁶ [Storey, p. 3–4]. Метрика Йенсена – Шеннона была точна в 92,5% случаев, тогда как метрика Бёрроуза – в 90% случаев. Расстояние Delta Эдера в исследовании не проверялось. В работе О. В. Алиевой с наибольшей точностью классифицировало тексты косинусное сходство, несколько уступили ему расстояния Танимото и Лаббе²⁷; дивергенция Йенсена – Шеннона в этой работе не исследовалась. Несмотря на выявленное отставание Delta, она остаётся довольно точным инструментом анализа. Существенным же её преимуществом является доступность для исследователя, поскольку все необходимые численные методы уже реализованы и свободно распространяются в виде упомянутого

²⁵ Тексты житий были выгружены из «Dumbarton Oaks Hagiography Database» (<https://www.doaks.org/research/byzantine/resources/hagiography>). Доступ к остальным источникам был получен через подписку Thesaurus Linguae Graecae. Полный список анализируемых источников с их параметрами приведён в табличном виде в связанном с этим исследованием репозитории, там же расположены графики, отражающие результаты экспериментов (<https://github.com/kanaevaelga-web/Stylo-metryByz26>).

²⁶ Storey 2020, 3–4.

²⁷ Алиева 2024b, 23–24.

выше пакета «stylo» для языка R. В этом пакете представлены и другие метрики, каждая из которых в исследовании Стори – Мимно показала результаты, худшие, чем Delta. Поэтому для настоящего исследования была отобрана серия модификаций метода Delta: классическая Burrows' Delta, Delta Эдера и косинусная Delta.

MFW. Для каждого текста в исследуемом корпусе были вычислены векторы наиболее частотных слов (Most Frequent Words, MFW). Греческий язык очень богат служебными частями речи. Стори и Мимно приводят в своей статье вектор из 250 наиболее частотных слов в их корпусе из 464 текстов. По этому вектору видно, что первое значимое слово (ἔφη) появляется только на 82-й позиции, следующее (πόλις) – на 89-й, а следующее (γενέσθαι) – лишь на 132-й, после чего плотность содержательных слов заметно возрастает (ср.: для английского языка слово said возникает уже на 35-м месте, а после 60-го большинство слов являются значимыми) [Storey, p. 27-30] Поэтому для греческих текстов показалось нецелесообразным брать минимальное MFW за 100, как это часто делают при исследовании европейской литературы. Кроме того, поскольку ряд источников в нашем корпусе имеет объём около 5000 слов, было малоосмысленно брать длину вектора MFW более 1000 слов, поскольку такая близость по объёму между вектором частотности и размером отрывка неминуемо должна привести к возникновению статистического шума. С учётом этих соображений для проведения экспериментов было выбрано пять размерностей вектора MFW: 150, 250, 500, 750 и 1000 слов.

Нормализация текста. Повышает ли точность анализа приведение слов к их словарным формам? Ольга Алиева на своём материале не обнаружила значимого различия между результатами, полученными с помощью анализа исходных словоформ и приведённых слов²⁸. Стори и Мимно сразу отказались от работы с лемматизированным корпусом, поскольку считают, что использование различных форм слова является важной информацией, составляющей стилистический портрет автора (впрочем, они оговариваются, что признают вероятность того, что на лемматизированном корпусе другие метрики могут оказаться даже более результативными, чем дивергенция Йенсена – Шеннона). Также авторы указывают на потенциальные проблемы при автоматической нормализации корпуса, связанные с неоднозначностью выбора форм, к которой может быть приведён ряд словоформ²⁹. В нашем исследовании стилеметрический анализ проводится как для слов, так и для словоформ,

²⁸ Алиева 2022, 27–30

²⁹ Storey 2020, 7, 17

для проверки влияния фактора нормализации текста на результаты предсказаний. Автоматическая лемматизация производилась при помощи обучаемой конвейерной модели Udrpipe для R³⁰.

Методы обработки результатов. Пакет *stylo* предлагает несколько методов финальной обработки и визуализации результата. Мной были выбраны метод кластерного анализа – как самый наглядный и удобный для выявления степени близости между каждой парой авторов – и консенсусное дерево (Bootstrap consensus tree) – как метод, который позволяет сгладить статистические выбросы и усреднить результаты по всем векторам MFW. Таким образом, для кластерного анализа эмпирическую базу исследования составили результаты 30 независимых экспериментов (3 модификации метода × 5 размерностей MFW × 2 варианта нормализации текста – отсутствующая и на основе Udrpipe), а для консенсусного дерева – 6 экспериментов (3 метода × 2 варианта нормализации).

Вот несколько наиболее репрезентативных графиков:

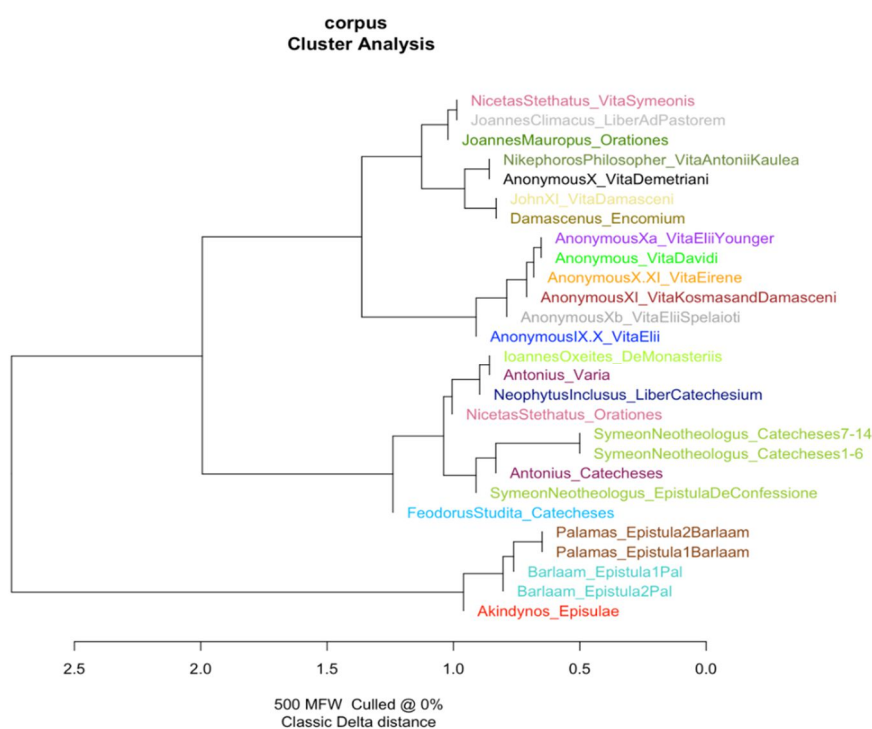


Рис. 1. Кластерный анализ: Burrows's Delta, 500 MFW, лемматизированный корпус

³⁰ URL: <https://CRAN.R-project.org/package=udrpipe> (дата обращения 12.01.2026)

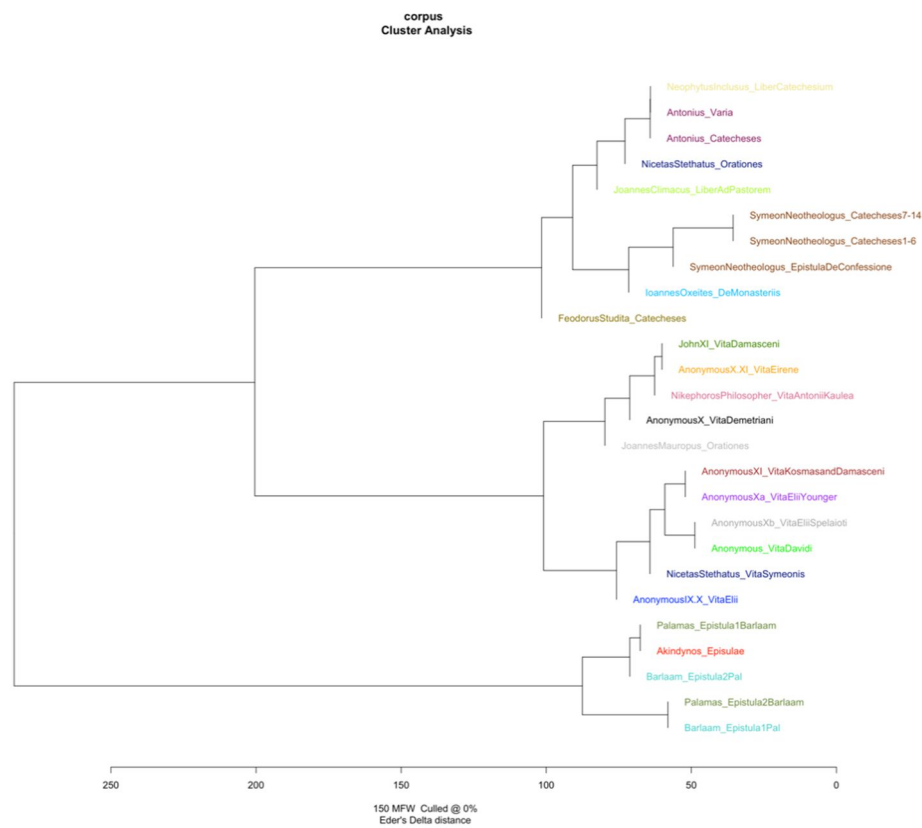


Рис. 2. Кластерный анализ: Eder's Delta, 150 MFW, корпус без нормализации

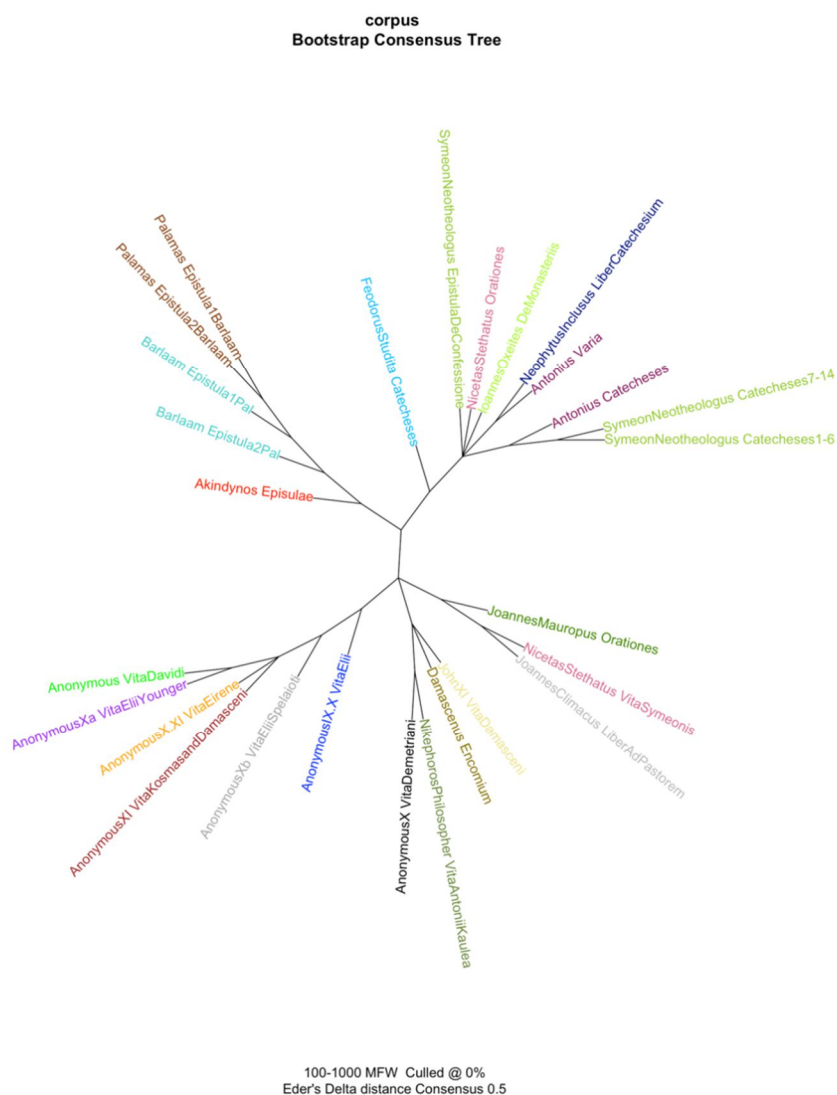


Рис. 3. Консенсусное дерево: Delta Эдера, лемматизированный корпус

Интерпретация результатов

Исследование проводилось в два этапа. Во-первых, калибровка метода: поиск оптимальных параметров (метрика, размер MFW, нормализация), по контрольным точкам. Наиболее результативными признавались такие комбинации параметров, которые: а) уверенно отделяют группу писем XIV века от текстов IX-XI вв.; б) уверенно разделяют тексты по жанрам внутри X-XI вв.; в) стабильно демонстрируют стилистическую близость для пары текстов Си-

меона (огласительные поучения и послание об исповеди, чья принадлежность Симеону уже не подвергается сомнению³¹); г) стабильно демонстрируют стилистическую близость для пары писем Варлаама; д) стабильно демонстрируют стилистическую близость для пары писем свт. Григория Паламы. Во-вторых, поиск ответа на содержательные вопросы по корпусу: устойчивые паттерны легли в основу экспериментов, направленных на выявление его стилистической структуры.

Общая валидность метода и влияние жанра произведения. В своём исследовании Стори и Мимно обнаружили, что стилистическое расстояние между авторами, пишущими в одном жанре, но существенно удалёнными друг от друга по времени, часто оказывается значительно меньше, чем то же расстояние между современниками, но писавшими в разных жанрах. Например, (IV в. н. э.) оказывается ближе к Иосифу Флавию (I в. н. э.) и Дионисию Геликарнасскому (I в. до н. э.), чем Исократ к Лисию или Платон к Ксенофону (все – V–IV вв. до н. э.)³². Авторы также проверили по этому же параметру корпусы текстов на английском и исландском языках (поскольку из языков, сохранивших многовековую письменную традицию, исландский считается в лингвистике одним из наиболее стабильных, а английский, наоборот, одним из наиболее подвижных) и пришли к выводу, что такая жанровая стилистическая устойчивость есть яркая особенность именно греческого корпуса. Из всех стилистически близких пар авторов в греческом корпусе 23% пар писали на расстоянии не менее четырёх веков, в исландском таких пар оказалось уже 2,6%, тогда как в английском – ни одной (фактически там не оказалось ни одной пары, разделённой более чем столетием).

Наши наблюдения согласуются с результатами американских исследователей. Все проверяемые комбинации параметров успешно отделили тексты XIV в. (письма) от основного корпуса, а внутри X–XI вв. – надёжно сегментировали агиографические тексты от текстов прочих жанров. Также во всех экспериментах текст «Жития прп. Симеона Нового Богослова» авторства Никиты Стифата вошёл в кластер житий и ни разу не оказывался вблизи «Огласительных поучений» того же автора. Это свидетельствует о значительном влиянии жанра древнегреческого текста на стилеметрический анализ. Из этого можно также сделать вывод, что если тексты различных жанров попадают в близкие кластеры, то их атрибуция одному автору становится весьма вероятной.

³¹ Storey 2020, 6–14.

³² Storey 2020, 11.

Письма Антония Студита не попадали в один кластер с источниками богословской переписки XIV века, но это легко объяснимо тем, что фактически письма Антония принадлежат другому жанру – текстам официального дипломатического общения. Более неожиданным может показаться то, что и «Послание об исповеди» прп. Симеона ни разу не объединялось с письмами Варлаама, Паламы и Акиндина, но это тоже, как нам кажется, объяснимо жанровыми различиями: письмо Симеона, как отмечает его издатель, представляет собой скорее образец гомилетического искусства, чем строгий трактат на богословскую тему, а жанр письма усвоен этому тексту искусственно³³.

Методологические выводы

От метода нет смысла ожидать предельной точности. Он выявляет стилистическую близость между двумя текстами, но крайне редко относит их к одному кластеру (свидетельствует о тождестве стиля). Например, на нашем корпусе только треть экспериментов дала полное тождество стиля для двух писем Варлаама и две трети экспериментов – для двух писем Паламы, хотя их общее авторство известно наверняка. Точность, на которую можно претендовать с вероятностью 90% и выше, – это расстояние в один узел для текстов одного жанра и в два узла для текстов разного жанра. Если требовать от метода большей строгости, то его точность становится практически равной случайному выбору: 60% в среднем по всем комбинациям параметров, не более 70% для отдельных наиболее удачных характеристик. Для анализа последующих результатов возьмём за истину предположение, что если для текстов одного жанра расстояние тождественности составляет ≤ 1 узла, то для текстов разных жанров разница в ≤ 2 узлов также свидетельствует в пользу общности стиля и, следовательно, повышает вероятность общего авторства.

На текстах XIV века, обладавших сравнительно большим размером (порядка 10 000 слов) в нашей серии экспериментов значительно лучше других методов показала себя классическая Delta (95% успеха против 83% для косинусного и эдеровского расстояний). Существенно выиграли эксперименты на лемматизированной версии корпуса по сравнению с нелемматизированной (93% против 70%). Лучшей размерностью вектора MFW стали 500 и 750 слов (92–100% успеха). Дальнейшее увеличение числа слов, по всей видимости, приводит только к увеличению статистического шума. Меньшее число слов не даёт методу стабилизироваться. Пример, иллюстрирующий работу метода на нормализованном корпусе, приведен на рис 1.

³³ Turner 2009, 14, 20.

В тестах, в которых оценивалась стилистическая близость текста небольшого размера (послание об исповеди Симеона, 4200 слов) огласительным словам того же автора, были получены иные результаты. Лучшие прогнозы давали косинусное и эдеровское расстояния (90-100% против 70% для классического). Уменьшение размерности вектора MFW приводило к повышению точности предсказаний: 100% исследований были точны для векторов в 150, 250, 500 слов против 67–71% для векторов большей размерности. Также работа с оригинальным корпусом приводила к заметно лучшим результатам по сравнению с нормализованной версией: 93% успехов против 80% (см. пример на рис. 2).

На небольшом тексте метод дал достаточно уверенную оценку стилистического родства, что согласуется с выводами Алиевой, а также Стори и Мимно на этот счёт.

Хотя были выявлены наиболее результативные параметры в нашей серии экспериментов, простое объединение их в «тройки» не всегда даёт наилучший результат. Например, то, что для больших текстов лучше всего показали себя параметры «классический метод», «750 MFW» и «лемматизированный корпус», к сожалению, не означает, что эксперимент, объединяющий эти три характеристики, будет наиболее результативным. Анализ результатов проведённых экспериментов приводит меня к заключению, что для получения наиболее убедительной и репрезентативной картины всё равно в каждом исследовании необходимо проводить серию экспериментов с разными параметрами.

Построение консенсусного дерева с использованием Delta Эдера на лемматизированном корпусе и других метрик на оригинальном корпусе даёт возможность быстро увидеть достаточно достоверную картину (точность 92-100%), не углубляясь в сложности подбора размерности вектора MFW. Можно рекомендовать этот инструмент для формирования предварительных гипотез, которые, в свою очередь, можно проверять, используя более тонкую настройку инструмента. Следует отметить, что при построении консенсусного дерева на лемматизированном корпусе классическая и косинусная Delta дали смазанный результат, который пришлось игнорировать при дальнейшем анализе. См. пример работы метода на Рис.3.

Содержательные выводы на основном корпусе (X–XI вв.)

Оценим стилистическое расстояние между бесспорными текстами патр. Антония Студита (чуть больше 6000 слов) и приписываемыми ему «Огласительными поучениями» (6500 слов). При использовании выявленных на предыдущем шаге параметров для малых текстов стилеметрический анализ

свидетельствует в пользу гипотезы о принадлежности «Огласительных поучений» патриарху Антонию Студиту: в 89% (8 из 9) экспериментов на оригинальном корпусе кластеры, содержащие эти тексты были ближайшими соседями, и в 100% случаев расстояние между ними не превышало ≤ 2 узлов, что позволяет говорить о стилистическом тождестве с учётом различия в жанрах. Более того, если из серии экспериментов исключить размерность MFW 500 слов, то из 6 оставшихся экспериментов (для 150 и 250 наиболее частотных слов) 100% считало проблемные тексты ближайшими соседями по стилю.

Обнаружена заметная стилистическая близость между огласительными поучениями Никиты Стифата и текстами того же жанра прп. Симеона Нового Богослова. Если исключить из рассмотрения эксперименты с размерностью 150 MFW, которой, видимо, оказалось недостаточно, 100% экспериментов показало близость стиля (расстояние ≤ 3 узлов) между этими авторами. При этом метод ни разу не предложил их отождествить или сократить расстояние до 1 узла. Вопрос о том, что произошло в этом случае – оказал ли Симеон столь значительное влияние на своего ученика или, наоборот, Никита при редактировании текстов своего учителя оставил настолько заметный личный авторский сигнал, – предстоит изучить более внимательно другими научными методами. Также выявлена некоторая стилистическая близость между «Огласительными поучениями» Антония Студита и «Посланием об исповеди» Симеона (в 21 эксперименте (70%) расстояние ≤ 3 узлов).

Очень часто к группе текстов, описанных в предыдущих пунктах, примыкают тексты Феодора Студита. Частично это объясняется сходством жанров: в корпус взяты огласительные поучения всех четырёх авторов. Однако близость с другими текстами этих авторов, как представляется, может свидетельствовать об общем интеллектуальном контексте всех четырёх авторов, связанных традицией Студийской обители в Константинополе.

Попутно исследование позволило отметить стилистическое тождество агиографических текстов – похвалы патриарху Антонию Кавлею авторства Никифора Философа и анонимного жития Димитриана Хитридского (25 из 30 экспериментов). Причём для трёх размерностей MFW отождествление происходило в 100% случаев: для 500, 750, 1000 слов. Эта степень стилистической близости превышает даже ту, что была выявлена для текстов свт. Григория Паламы. Гипотезу о возможном общем авторстве для этих текстов теперь необходимо проверить историческими и классическими методами. Отчасти она подтверждается наблюдениями Анри Грегура, издателя жития Димитриана: он указывает, что анонимный автор имел философское образование и

был близким современником своего героя, то есть писал в начале X века³⁴, что делает авторство Никифора Философа вероятным. С ним спорит Хрисанф Лопарёв³⁵, однако русский учёный не знал о существовании рукописи жития XI века, поэтому его критика не имеет достаточных оснований. С другой стороны, оба автора сходятся в том, что житие Димитриана написано человеком, близко знакомым с кипрскими реалиями и, вероятно, киприотом, а относительно Никифора Философа у нас таких сведений нет. Однако сохранилось письмо Фотия, адресованное «Никифору философу монашествующему», в котором отправитель хвалит получателя за мастерское владение азианским стилем (насыщенным сложными периодами и словесными украшениями по сравнению с аттическим стилем)³⁶, а Грегуар пишет о своём анонимном авторе, что тот вполне владеет искусством «составлять помпезные периоды и расцвеченные парафразы»³⁷. Все эти соображения подкрепляют гипотезу о тождестве адресата Фотия, автора жития Димитриана и похвалы Антонию Кавлею, но она нуждается в дополнительной проверке учёными-филологами и историками.

Заключение

Исследование работоспособности метода Delta на средне- и поздневизантийском материале показывает, что он применим для анализа греческих текстов, однако требует тщательного подбора оптимальных параметров и большого числа экспериментов. При выборе параметров наблюдаются значительные вариации эффективности: большие документы (около 10 000 слов) лучше классифицировались классической Delta с лемматизацией и вектором частотности в 500–750 слов, малые (около 5000 слов) – оказалось предпочтительнее исследовать с помощью Eder's Delta или косинусной Delta с MWF 250 слов. Тексты среднего размера примерно одинаково хорошо (но в среднем хуже, чем более крупные документы) классифицировались всеми тремя метриками, при размерах вектора частотности в 250–750 слов как на нормализованном корпусе, так и без нормализации. Жанр произведения оказывал существенное влияние на его положение в классификации текстов.

Полученные данные позволяют подкрепить гипотезу о принадлежности «Огласительных поучений» перу Антония Студита: подавляющее большинство экспериментов демонстрирует значительное стилистическое сходство с

³⁴ Grégoire 1907, 213.

³⁵ Лопарёв 1914, 369.

³⁶ Norden 1915, 371–373, no. 3.

³⁷ Grégoire 1907, 213.

его признанными произведениями, несмотря на различие жанров. Обнаружены признаки тесной взаимосвязи между стилями Симеона и Никиты Стифата, а также их умеренной близости со стилем Антония, что позволяет ставить вопрос о влияниях и заимствованиях в интеллектуальной среде византийского монашества IX–XI вв.

Дальнейшие исследования требуют расширения базы текстов, включения материалов иных жанров и исторических периодов, а также интеграции стилеметрических данных с традиционной филологией и историей. Отдельного внимания заслуживает проверка гипотезы об общем авторстве житий патр. Антония Кавлея и Димитриана Хитридского, а также более глубокое изучение механизмов сочетания жанрового и индивидуального стиля в византийской письменности.

БИБЛИОГРАФИЯ/ REFERENCES

- Алиева, О.В. (2022) "Delta Берроуза для древнегреческих текстов," *ΣΧΟΛΗ (Schole)* 16.2, 693–705.
- Алиева, О.В. (2024a) "Pseudoplatonica: Блеск и нищета стилометрии," *Философия. Журнал Высшей школы экономики* 8.3, 35–56.
- Алиева, О.В. (2024b) "Меры расстояния для определения авторства древнегреческих текстов," *Цифровые гуманитарные исследования* 1, 8–33.
- Алиева, О.В. (2025) "Новая количественная модель Платоновского корпуса 2. Филогенетические методы в стилометрии," *Вестник ПСТГУ. Серия III: Филология* 84, 54–85.
- Егунов, А.Н. (1996) "Атрибуция и атетеза в классической филологии," *Древний мир и мы* 1, 89–139.
- Лопарев, Хр.М. (1914) *Греческие жития святых VIII и IX веков: Опыт научной классификации памятников агиографии с обзором их с точки зрения исторической и историко-литературной. Диссертация*. Петроград.
- Мамаев, Н.К., Марусенко М.А., Пиотровская К.Р., Ронжин А.Л. (2018) "Метод Дельты Бёрроуза для определения авторства анонимных и псевдонимных литературных произведений на русском языке," *Proceedings of the R. Piotrowski's Readings in Language Engineering and Applied Linguistics*. Санкт-Петербург, 1–14.
- Орехов, Б.В. (2023) "Глава 6. Компьютерный анализ текста," *Цифровые гуманитарные исследования: монография*. ред. А.Б. Антопольский, А.А. Бонч-Осмоловская, Л.И. Бородкин и др. Красноярск, 120–157.
- Acta Sanctorum (1863–1940) *quotquot in orbe coluntur vel a Catholicis scriptoribus celebrantur, quae ex latinis et graecis, aliarumque gentium antiquis monumentis*. 71 vols. Paris.

- Bollig, I., Lagarde, P. de, eds. (1882) *Iohannis Euchaitorum Metropolitae quae in Codice Vaticano Graeco 676 supersunt*. Göttingen.
- Burrows, J. (2002) "Delta': a measure of stylistic difference and a guide to likely authorship," *Literary and Linguistic Computing* 17.3, 267–287.
- Constantinides-Hero, A., ed. (1983) *Letters of Gregory Akindynos, Corpus Fontium Historiae Byzantinae. Series Washingtonensis*, vol. 21. Washington.
- Creazzo, T. (2004) "Joannis Oxeitae oratio de monasteriis laicis non tradendis," *Quaderni della Rivista di Bizantinistica* 8, 61–83.
- Darrouzès, J. (1960) "Épistoliers byzantins du Xe siècle," *Archives de l'Orient Chrétien* 6, 344–346.
- Darrouzès, J., ed. (1961) *Nicétas Stéthatos, Opuscules et Lettres, vol. 81 of Sources chrétiennes*. Paris.
- Delouis, O., Leroy J. (2004) "Quelques inédits attribués à Antoine III Stoudite," *Revue des études byzantines* 62, 33–69.
- Eder, M., Rybicki, J., Kestemont, M. (2016) "Stylometry with R: a package for computational text analysis," *R Journal* 8.1, 107–121.
- Evert, S., et al. (2017) "Understanding and explaining delta measures for authorship attribution," *Digital Scholarship in the Humanities* 32.2, ii4–iii6.
- Grégoire, H. (1907) "Saint Demetrianos, évêque de Chytri (île de Chypre)," *Byzantinische Zeitschrift* 16, 204–240.
- Hausherr, I., ed. (1928) *Un grand mystique byzantin. Vie de Syméon le Nouveau Théologien par Nicétas Stéthatos*, vol. 12 of *Orientalia Christiana*. Rome: Pontificum Institutum Orientalium Studiorum.
- Holl, K. (1898) *Enthusiasmus und Bussgewalt beim Griechischen Mönchtum: Eine Studie zu Symeon dem Neuen Theologen*. Leipzig.
- Kovalev, B. V. (2024) "From classics to digital philology: on the origin and growth of stylometry," *Philologia Classica*, 19.2, 347–360.
- Leone, P.L.M. (1989) "L'Encomium in patriarcham Antonium II Cauleam del filosofo e restore Niceforo," *Orpheus: rivista di umanità classica e Cristiana* 10, 404–429.
- Lutosławski, W. (1897) *The Origin and Growth of Plato's Logic. With an Account of Plato's Style and Chronology of His Writings*. London.
- Meyendorff, J. (1962) "Επιστολαὶ πρὸς Ἀκίνδυνον καὶ Βαρλαάμ," in: P.K. Chrestou (ed.) *Γρηγορίου τοῦ Παλαμᾶ συγγράμματα* 1. Θεσσαλονίκη, 202–312.
- Migne, J.-P., ed. (1857–1866) *Patrologiae cursus completus: series graeca: posterior pars continens patres post-Nicenos usque ad Nicephorum*. 161 vols. Paris.
- Norden, E. (1915) *Die antike Kunstprosa: vom VI. Jahrhundert v. Chr. bis in die Zeit der Renaissance. Bd. 1*. Leipzig.
- Papadopoulos-Kerameus, A., ed. (1897) *Analekta hierosolymitikes stachyologias*. St. Petersburg.
- Papadopoulos-Kerameus, A. (1905) "Antonios ho Stoudites [transl. Anthony the Studite]," *Néa Siôn* 2, 807–815 (in Greek).

- Papadopoulos-Kerameus, A., ed. (1904) *Theodoros Studites, Μεγάλη κατήχησις*. St. Petersburg.
- Papadopoulos-Kerameus, A. (1907) "Sylloge Palaistines kai Syriakes Hagiologias XIX," *Πравославный палестинский сборник* 57, 42–59.
- Rosenqvist, O., ed. (1986) *The Life of St. Irene Abbess of Chrysobalanton*. Uppsala.
- Schirò, G., ed. (1954) *Barlaam Calabro, Epistole Greche*. Palermo.
- Smith, D.A., Rydberg-Cox J.A., Crane G.R. (2000) "The Perseus Project: A digital library for the humanities," *Literary and Linguistic Computing* 15.1, 15–25.
- Sotiroidis, P. (1998) "Βίβλος τῶν κατηχήσεων," in: D.G. Tsames, N. Zacharopoulos, C. Oikonomou, I. Karabidopoulos, eds. *Ἀγίου Νεοφύτου τοῦ Ἐγκλείστου Συγγράμματα*, 2. Πάφος, 189–431.
- Sternbach, L. (1975) "Λόγος ἀναγνωσθεὶς ἐν Βλαχέρναις παρὰ τοῦ ταπεινοῦ Στουδίτου μοναχοῦ Ἀντωνίου τοῦ Τριψύχου," in: F. Makk, ed. *Traduction et commentaire de l'homélie écrite probablement par Théodore le Syncelle sur le siège de Constantinople en 626*. Opuscula Byzantina 3. Szeged, Hungary, 113–117.
- Storey, G., Mimno, D. (2020) "Like two pis in a pod: author similarity across time in the ancient Greek corpus," *Journal of Cultural Analytics*, 5.2, 1–35.
- Krivochéine, B., ed., Paramelle, J., transl. (1963) *Syméon le Nouveau Théologien. Catéchèses*. Paris.
- Taibbi, G.R., ed. (1962) *Vita di nt' Elia il Giovane*. Palermo.
- Turner, H. J. M. (2009) *The Epistles of St Symeon the New Theologian*. Oxford.
- Turner, H.J.M. (1990) *St. Symeon the New Theologian and Spiritual Fatherhood*. Leiden.
- Van den Gheyn, J. (1899) "Acta graeca ss. Davidis, Symeonis et Georgii Mitylenae in insula Lesbo," *Analecta Bollandiana* 18, 209–259.

References in Russian:

- Aliyeva, O. V. (2022) "Delta Berrouza dlya drevnegrecheskikh tekstov [transl. Testing Delta Burrow's Method for Ancient Greek Texts]," *ΣΧΟΛΗ (Scholē)* 16. 2, 693–705.
- Aliyeva, O. V. (2024) "Pseudoplatonica: blesk i nishcheta stilometrii [transl. Pseudoplatonica: Splendors and Miseries of Stylometry]," *Filosofia. Zhurnal Vysshei shkoly ekonomiki* 8.3, 35–56.
- Aliyeva, O. V. (2024) "Mery rasstoyaniia dlia opredeleniia avtorsva drevnegrecheskikh tekstov [transl. Distance Measures for Determination of Authorship of Ancient Greek Texts]," *Tsifrovye gumanitarnye issledovaniia* 1, 8–33.
- Aliyeva, O. V. (2025) "Novaia kolichestvennaia model' Platonovskogo korpusa 2. Filogeneticheskie metody v stilometrii [transl. A New Quantitative Model of the Platonic Corpus, Part 2. Phylogenetic Methods in Stylometry]," *Vestnik PSTGU. Seriya III: Filologiya* 84, 54–85.
- Egunov, A. N. (1996) "Atributsia i atezeza v klassicheskoi filologii [transl. Attribution and Denial of Authorship in Classical Philology]," *Drevnii mir i my* 1, 1996, 89–139.
- Loparev, Kh. M. (1914) *Grecheskie zhitiia sviatykh VIII i IX vekov* [transl. *Greek Lives of Saints from the 8th and 9th Centuries: Scientific Classification Experience of Hagiographic Monuments Reviewed Historically and Historico-Literarily*]. Petrograd.
- Mamaev, N. K., Marusenko M. A., Piotrovskaya K. R., Ronzhin A. L. (2018) "Metod Del'ty Berrouza dlia opredeleniia avtorstva anonimnykh i psevdonimnykh literaturnykh proizvedenii na

russkom iazyke [transl. The Burrow's Delta Method for Identifying Authorship of Anonymous and Pseudonymous Russian Literary Works],” *Proceedings of the R. Piotrowski's Readings in Language Engineering and Applied Linguistics*. Petersburg, 1–14.

Orekhov, B. V. (2023) “Glava 6. Komp'iuternii analiz teksta [transl. Chapter 6. Computerized Text Analysis],” *Tsifrovye gumanitarnye issledovaniia: monografiya*, eds. A. B. Antopolsky, A. A. Bonch-Osmolovskaia, L. I. Borodkin [et al.]. Krasnoyarsk, 120–157.