

Ф. Н. Двинятин^a<https://orcid.org/0000-0001-7111-0133>✉ f.dvinyatin@spbu.ru**Б. В. Ковалев**^{ab}<https://orcid.org/0000-0002-1904-1844>✉ bvkovalev@yandex.ru^a Санкт-Петербургский государственный университет
(Россия, Санкт-Петербург)^b Союз писателей Санкт-Петербурга
(Россия, Санкт-Петербург)

«СИРИН НЕ УСТУПАЕТ ЛЕОНОВУ»: МЕТОД DELTA ДЛЯ СТИЛЕМЕТРИЧЕСКОГО АНАЛИЗА РУССКИХ РОМАНОВ МЕЖВОЕННОГО ПЕРИОДА

Аннотация. В статье анализируются стилистические особенности русскоязычных романов Набокова в контексте русской межвоенной прозы. Основным методом является Delta Бёрроуза — один из наиболее надежных стилеметрических инструментов, позволяющих сравнивать тексты между собой на основании распределения наиболее частотных служебных слов. Также для классификации текстов применяются методы SVM и KNN. В исследовании ставятся две цели: 1) осуществить первичную кластеризацию стилей ведущих русских прозаиков межвоенной эпохи, а также выяснить место среди них стиля Набокова (Сирин) при помощи современных алгоритмов; 2) испытать эти алгоритмы на способность информативно и адекватно решать поставленные задачи стилистического распределения и на согласованность результатов. Выясняется, что стилистически близкими к романам Набокова являются тексты Леонова, Газданова и Грина. Особенно отчетливое сходство обнаруживается между поздними романами Набокова, «Скутаревским» Леонова и «Вечером у Клэр» Газданова. Восемь русскоязычных романов Набокова делятся на две группы. Первая — «Машенька», «Защита Лужина», «Король, дама, валет», «Подвиг», «Камера обскура». Вторая — романы 1930-х годов «Отчаяние», «Приглашение на казнь», «Дар».

Ключевые слова: В. Набоков, Г. Газданов, Л. Леонов, межвоенная проза, стилеметрия, Stylo, Delta, цифровая филология

Благодарности. Работа выполнена при поддержке СПбГУ, шифр проекта 95434615.

Для цитирования: Двинятин Ф. Н., Ковалев Б. В. «Сирин не уступает Леонову»: метод Delta для стилеметрического анализа русских романов межвоенного периода // Шаги / Steps. Т. 10. № 3. 2024. С. 207–229.

Поступило 5 февраля 2024 г.; принято 4 июня 2024 г.

F. N. Dviniatin ^a

<https://orcid.org/0000-0001-7111-0133>
✉ f.dvinyatin@spbu.ru

B. V. Kovalev ^{ab}

<https://orcid.org/0000-0002-1904-1844>
✉ bukovalev@yandex.ru

^a St. Petersburg State University
(Russia, Saint Petersburg)

^b Writers' Union of St. Petersburg
(Russia, Saint Petersburg)

“SIRIN IS NOT INFERIOR TO LEONOV”: BURROWS’S DELTA FOR STYLOMETRIC ANALYSIS OF RUSSIAN NOVELS OF THE INTERWAR PERIOD

Abstract. The article analyzes the stylistic features of Nabokov’s Russian-language novels in the context of Russian interwar prose. The main method is Burrows’s Delta — one of the most reliable stylometric tools that allows comparing texts with each other based on the distribution of the most frequent function words. In addition, SVM and KNN methods are also used to classify texts. The study has two objectives: 1) to carry out a primary clustering of the styles of leading Russian prose writers of a certain era, and also, through the use of modern algorithms, to determine the place among them of the style of Nabokov (Sirin); 2) to test these algorithms for their ability to informatively and adequately solve the tasks of stylistic distribution and for the consistency of the results. It turns out that stylistically close to Nabokov’s novels are the texts of Leonov, Gazdanov and Grin. A particularly clear similarity is found between Nabokov’s late novels, Leonov’s *Skutarevsky* and Gazdanov’s *An Evening with Claire*. Nabokov’s eight Russian-language novels are divided into two groups. The first consists of *Mashenka*, *The Defense*, *King*, *Queen*, *Knave*, *Glory*, *Laughter in the Dark*. The second is made up of the novels of the 1930s: *Despair*, *Invitation to a Beheading*, *The Gift*.

Keywords: V. Nabokov, G. Gazdanov, L. Leonov, interwar prose, stylometry, Stylo, Delta, digital philology

Acknowledgements. The authors acknowledge the support of Saint Petersburg State University for research project № 95434615.

To cite this article: Dviniatin, F. N., & Kovalev, B. V. (2024). “Sirin is not inferior to Leonov”: Burrows’s Delta for stylometric analysis of Russian novels of the interwar period. *Shagi / Steps*, 10(3), 207–229. (In Russian).

Received February 5, 2024; accepted June 4, 2024

© F. N. DVINIATIN, & B. V. KOVALEV



Введение

Стилистический анализ — сложная и многомерная процедура. Необходимо выбрать методологию — одну из множества существующих или комбинацию нескольких. Следует решить, будет ли изучаться текст сплошь, как стилистически отмеченный в каждой точке, или только некоторые, привилегированные в стилистическом отношении, его компоненты (и тогда какие?). Наряду с этим нужно решить, будут ли рассматриваться выборочные фрагменты или весь массив изучаемых текстов. В идее стиля при самых разных ее пониманиях заключена идея отличия или различия (один стиль отличается от другого, стили различаются между собой — о стилях и говорят для того, чтобы констатировать эти отличия или различия), и поэтому практически обязательны сопоставления авторов друг с другом и всех их, вместе или по отдельности, — с языковой картиной эпохи. Таким образом, даже при традиционном филологическом подходе исследование стиля принадлежит пограничью между «пристальным» филологическим чтением и более обзорным подходом. Но когда речь идет о многих авторах и текстах, подключение количественных подходов даже не столько желательно, сколько насущно. Это было ясно и несколько десятилетий назад, и тем более сегодня — в свете прогресса квантитативных и компьютерных методов.

У нашего исследования две цели. Первая — осуществить первичную кластеризацию стилей ведущих русских прозаиков определенной эпохи, а также выяснить место среди них стиля Набокова (Сирина), вокруг которого в первую очередь строится исследование. Для этого используются наиболее надежные на сегодняшний день компьютерные методы, и прежде всего Delta Бёрроуза. Другая цель — испытать эти алгоритмы, во-первых, на согласованность результатов (результаты применения различных алгоритмов могут оказаться солидарными, либо противоречащими друг другу, либо взаимодополняющими), во-вторых — на способность информативно и адекватно решать поставленные задачи стилистического распределения.

Метод

Для решения исследовательских задач была избрана Delta Бёрроуза [Burrows 2002]. Этот метод получил значительное распространение в зарубежной филологии (в первую очередь в сфере цифровой текстологии) и с некоторых пор активно применяется отечественными исследователями.

Метод основан на предположении, что распределение наиболее частотных слов в текстах того или иного автора не является случайным, а служит отображением индивидуального авторского стиля. Существенно, что значительная доля самой частотной лексики приходится на служебные слова — предлоги, союзы, местоимения и пр., — которые не связаны с тематикой текста. Вычисления проводятся следующим образом. Для каждого текста в анализируемом корпусе берется некоторое количество

наиболее частотных слов и далее сравниваются их частотности, представленные как меры z-scores, т. е. стандартизированные оценки, показывающие разброс значений относительно средних. Z-score вычисляется по следующей формуле:

$$z(f_i(D_1)) = \frac{f_i(D_1) - \mu_i}{\sigma_i}. \quad \#(1)$$

где $f_i(D_1)$ — частота слова в тексте D_1 , μ_i — средняя частота слова по выборке, а σ_i — стандартное отклонение этой частоты. Соответственно, мера Delta представляет собой сумму взятых по модулю разниц между мерами z-scores у двух сравниваемых текстов, поделенную на количество слов:

$$\Delta_{Bur} = \frac{1}{n} \sum_{i=1}^n |z_i(D_1) - z_i(D_2)|. \quad \#(2)$$

где n — общее количество слов, i — конкретное слово, а D_1 и D_2 — сравниваемые тексты.

Как видим, Delta измеряет относительные частоты, что позволяет сравнивать между собой тексты разной длины. На основании подсчетов вычисляется расстояние между изучаемыми текстами: для текстов, которые написаны разными авторами, оно оказывается больше, чем для текстов, сочиненных одним автором.

Delta с успехом применяется в рамках цифровой текстологии, решая задачи атрибуции. Метод был испытан многими западными филологами на материале разных языков: английского [Gladwin et al. 2017; Hoover 2004] немецкого [Jannidis, Lauer 2014], арабского [Al-Falahni et al. 2015], испанского [Calvo Tello 2019; Hernández-Lorenzo, Byszuk 2023].

Некоторое время назад Delta стала активно использоваться и в русскоязычном научном пространстве: атрибутировались романы Ильфа и Петрова «Двенадцать стульев» и «Золотой теленок» [Мамаев и др. 2018], письма Берии [Петров и др. 2019]. Кроме того, Delta применялась для атрибуции романа «Тихий Дон» [Великанова, Орехов 2019; Iosifyan, Vlasov 2020; Маслинский 2022], метод проверялся на корпусе древнегреческих текстов [Алиева 2022a; 2022b]; анализировались тексты и переводы В. Набокова [Орехов 2021] и различные переводы «Илиады» на русский язык [Орехов 2020].

Для нас также существенно то, что Delta используется не только в рамках цифровой текстологии, но и для анализа стилистических перемен в творчестве одного или нескольких авторов [Calvo Tello 2019; Алиева 2022b; Ковалев 2023; Skorinkin, Orekhov 2023]. Delta может показать раз-

личия не только между текстами принципиально разных писателей, но и между текстами одного автора, написанными в разный период времени.

Delta служит средством иерархической кластеризации стилей. Мы можем принять ее работу с большим или меньшим доверием, но она делает именно это. Можно сказать, что при обращении к Delta на первый план в определении стиля выходит его отношение к другим соотносимым стилям (по принципу «стиль — это то, что отличается от других стилей, и одним из них он ближе, а от других дальше»), а на задний план отходят существенные признаки стиля. В то же время — и в этом своего рода парадокс — Delta иерархически кластеризирует именно те стили, о которых мы знаем из читательских впечатлений и/или историко-литературной оценки и/или поэтологического анализа. Описывать их методами различной точности и полноты можно без всякого обращения к Delta или другому алгоритму, выполняющему те же задачи, но иерархическая кластеризация стилей осуществляется именно Delta. Примечательно, что эта кластеризация носит топологический характер (по понятным причинам скорее плоскостной, чем собственно пространственный).

Формирование корпуса

Отбор авторов и текстов задавался прежде всего двумя критериями:

1) хронологическим: тексты должны быть написаны не ранее 1920-х и не позднее конца 1940-х годов;

2) жанровым; для корректного анализа при помощи Delta необходимо соблюдать требование жанровой¹ гомогенности — в этом исследовании мы сосредоточились на крупной художественной прозе.

В результате корпус состоит из следующих текстов:

- Русскоязычные романы В. Набокова «Машенька», «Защита Лужина», «Король, дама, валет», «Камера обскура», «Подвиг», «Приглашение на казнь», «Отчаяние», «Дар»;
- «Козлиная песнь», «Труды и дни Свистонова» К. Вагинова;
- «Вечер у Клэр», «Призрак Александра Вольфа» Г. Газданова;
- «Военная тайна», «Судьба барабанщика» А. Гайдара;
- «Бегущая по волнам», «Блистающий мир» А. Грина;
- «Дом в Пасси», «Золотой узор» Б. Зайцева;
- «Голубая книга», «Сентиментальные повести» М. Зощенко;
- «Двенадцать стульев», «Золотой теленок» И. Ильфа и Е. Петрова;
- «Белеет парус одинокий», «Время, вперед» В. Катаева;
- «Барсуки», «Скутаревский» Л. Леонова;
- «Зависть», «Три толстяка» Ю. Олеши;

¹ Для Delta нет существенной разницы между большой повестью и коротким романом, однако, к примеру, драмы и романы одного и того же автора будут «восприняты» программой как тексты совершенно разные по стилю; поэтому в рамках одного замера нежелательно смешение рассказов и романов.

- «Голый год», «Красное дерево» Б. Пильняка;
- «Петр Первый», «Хождение по мукам»² А. Н. Толстого;
- «Кюхля», «Смерть Вазир-Мухтара» Ю. Тынянова;
- «Братья», «Города и годы» К. Федина.

При подборе авторов и текстов мы исходили из общей установки и дополнительных принципов ее корректировки. «Нулевая» составная часть общей установки была в том, что замеры носят неизбежно предварительный характер и могут — а отчасти и должны — корректироваться, причем как в ходе самого данного исследования, так и после его завершения, в последующих обращениях к материалу. Поэтому цель сразу составить идеальный корпус авторов и текстов не ставилась, а всяческие дополнения и новые замеры если и не приветствовались, то считались вполне законными. Все это нужно учитывать при оценке дальнейшего. Собственно, общая установка состояла в том, что в корпус должны войти произведения наиболее значительных, по данным усредненных критических оценок и вхождений в «каноны» разных типов, авторов выбранной эпохи, при этом авторы должны быть в достаточной степени разнообразны — в пределах общих черт романного письма этих лет. От каждого из авторов изначально предполагалось брать по два произведения: два, а не одно, чтобы всегда контролировать, видит ли алгоритм общность стиля одного автора; два, а не три или больше — чтобы не перегружать формируемый корпус.

При этом с самого начала предполагалось, что Набоков, центральный для исследования автор, будет представлен всеми восемью русскими романами. Общий объем произведений в корпусе изначально хотелось сохранить равным приблизительно полусотне — это означало выбор примерно двух десятков авторов; причем, чтобы оставить резерв для новых подключений, сперва можно было не выбирать полностью этот примерно пятидесятитекстовый объем. Уже на этом этапе некоторые авторы отпадали по хронологическим или стилистическим основаниям, таким как слишком резкие отличия, могущие увести алгоритм от выявления более тонких сходств и различий.

Основных причин для отказа было три: неточное совпадение поколения-эпохи, жанра, проблемы с текстологией.

Так, из авторов первого ряда были не включены те, которые в значительной степени состоялись в предыдущий период: это прежде всего Горький и Бунин (в отличие от Куприна представленные в эти годы некоторыми главными своими текстами); при этом старшее поколение диаспоры представлено Зайцевым, а также Шмелевым³. В корпусе нет Шолохова: исследование его текстов методом Delta уже проводилось [Великанова, Орехов 2019; Маслинский 2022], объем «Тихого Дона» и некоторая вну-

² «Хождение по мукам» было разбито на три единицы в соответствии с частями романа.

³ В специальном исследовании, проводившемся практически синхронно с этим на материале исторических романов, богато представлены и тексты Алданова.

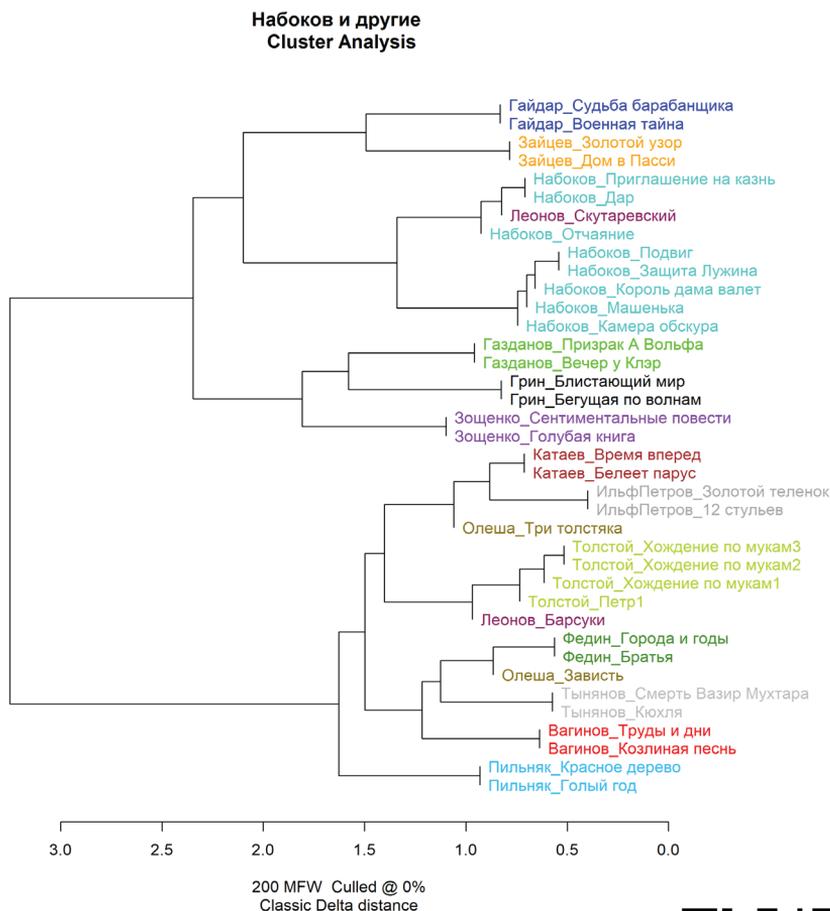
тренняя разнородность текста (эволюция стиля и различные по жанровой специфике фрагменты) могли оттянуть на себя слишком много внимания. По схожим причинам нет и Булгакова — в первую очередь из-за текстологических проблем⁴. Бабель не представлен как автор преимущественно малой прозы; в случае с Зощенко это препятствие в большей степени устранимо. Здесь необходимо повторить то, что уже сказано о возможности перепроверки результатов: добавить в корпус Бунина и/или Шолохова, Булгакова, Бабеля и др. при необходимости несложно. Проза советских авторов подбиралась так, чтобы при опоре на лучшие образцы представить разные тенденции: более традиционную и более обновленную (сказ, орнаментальность и т. д.), городскую, «интеллигентскую» и также народную; более «системную», представленную печатающимися авторами, тесно связанными с публичным литературным процессом, и более «частную», ту, что пишется «в стол», в условиях некоторой изоляции; московскую, петроградскую и одесскую и др.

Постановка эксперимента

Эксперименты были реализованы при помощи пакета Stylo [Eder et al. 2016] в программной среде R. На первом этапе мы провели серию опытов с частотным словарем объемом 100, 200, 300, 400, 500 слов (англ. most frequent words, далее — mfw) — стандарт для такого рода исследований. Использовалась стандартная процедура токенизации и нормализации токенов в пакете Stylo, лемматизация не проводилась. В современной стилеметрии вкуче с Delta используются и алгоритмы кластеризации, позволяющие представить результаты в виде дендрограммы — что выгодно отличается от традиционных таблиц наглядностью изображения. В нашем исследовании кластеры формировались по методу Варда: стилистически похожие тексты располагаются на одной «ветви» (ил. 1).

Для корректного прочтения дендрограмм мы предлагаем ввести два разграничения. С одной стороны, чтение предлагаемых схем «справа» или «слева» (в том виде, в котором они тут приведены), т. е. от самой дробной группировки к самой общей и, соответственно, наоборот, от самой общей к самой дробной; с другой стороны, осмысление итогов кластеризации без обращения к историко-литературной интерпретации (как простого понимания того, «что хочет сказать алгоритм») — и осмысление уже с подключением того, что известно за пределами данного эксперимента, т. е. закономерностей литературного процесса и представлений о стилистической эволюции, сформированных десятилетиями классических филологических наблюдений.

⁴ «На грани приемлемого» для данного исследования был и роман «Хождение по мукам» — из-за многочисленных и разновременных авторских переделок текста, а также долгой работы над полным текстом. Однако окончательная авторская версия существует, а хронологический разброс нейтрализован отдельным введением в исследование трех частей романа.



Ил. 1. Иерархическая кластеризация, визуализирующая расстояние Delta для анализируемых текстов. 200 mfw

Fig. 1. Hierarchical clustering visualizing Delta distance for analyzed texts 200 mfw

Итак, что дает чтение справа этой дендрограммы (38 текстов, 15 авторов — где Ильф и Петров, конечно, считаются одним автором)? Главный результат состоит в том, что тексты 12 авторов из 15 оказались на одной большой ветви; кроме того, с прибавлением одного текста другого автора на одной ветви оказались все тексты Набокова. Два исключения — Леонов и Олеша. Таким образом, можно сказать, что в явно преобладающем количестве случаев алгоритм считает авторский сигнал — условимся здесь и далее называть так совокупность индивидуальных стилистических характеристик, выявляемую алгоритмом. Поскольку принадлежность текстов одному автору является единственным фактическим и объективным параметром, по которому можно проверить адекватность кластери-

зации, это положение дел важно: на этом этапе у нас есть 13 случаев, когда алгоритм, не знающий внешних историко-литературных обстоятельств, предлагает кластеризацию, соответствующую им, и два случая парадоксального отклонения.

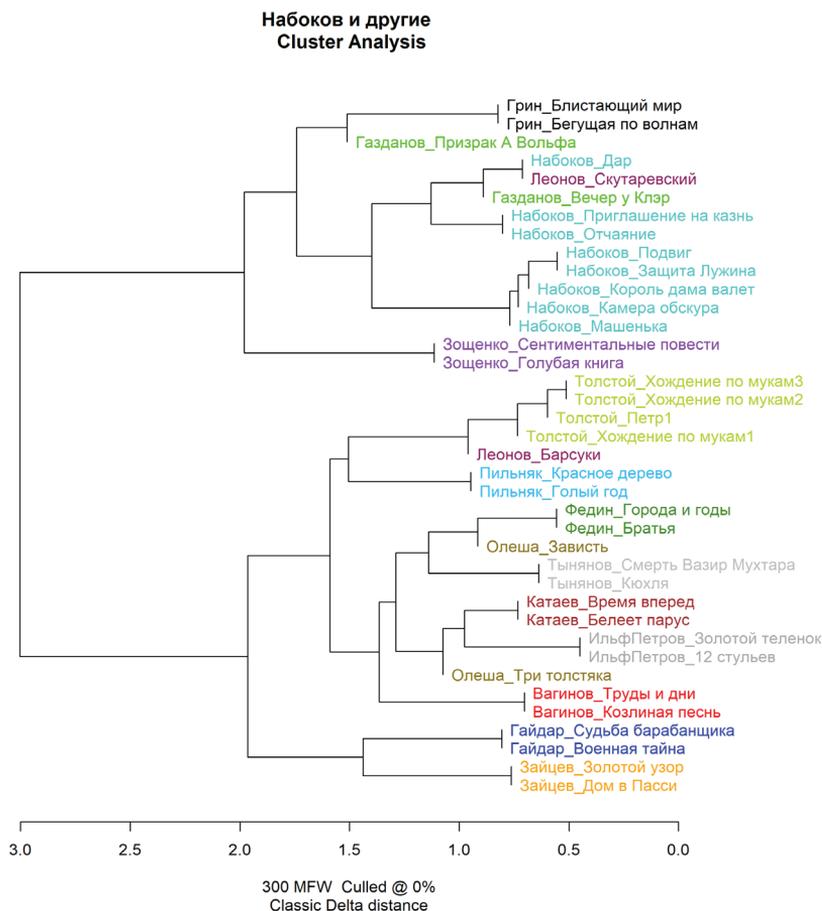
На следующем этапе (все еще чтение справа) можно проследить, как алгоритм группирует авторов и тексты по два: кому и чему он находит ближайших соседей. Здесь можно несколько уточнить констатацию из предыдущего абзаца, сказав, что 33 из 38 текстов алгоритм сближает прежде всего с текстами тех же авторов. Как же дело обстоит с самими авторами? Первичные сближения здесь таковы: для Катаева ближайшие — Ильф и Петров, для Ильфа и Петрова — Катаев; для Газданова Грин, для Грина Газданов; для Гайдара Зайцев и для Зайцева Гайдар. Это все случаи, когда один автор ближайшим образом соединен с другим, и они словно символизируют результаты разной степени «интуитивности»: более чем ожидаемое (Катаев — Ильф и Петров), совершенно неожиданное (Гайдар — Зайцев) и среднее (Грин — Газданов).

Для дальнейших констатаций нужны элементы чтения слева или полностью такое чтение слева. Мы обнаруживаем еще относительно компактную группу: к Газданову и Грину прибавляется Зошенко (простая группировка трех авторов); к Катаеву и Ильфу — Петрову примыкают «Три толстяка» Олеси (сложность здесь в том, что Олеша — «разъятый» автор, в том смысле, что алгоритм распределяет его тексты по разным ветвям); Толстой объединяется с «Барсуками» Леонова (Леонов — тоже «разъятый» автор); Федин — с «Завистью» Олеси, а Набоков — со «Скутаревским» Леонова. При этом если сближенные с двумя романами Олеси по отдельности Федин и Катаев — Ильф и Петров отдалены друг от друга в средней мере, то сближенные по отдельности с двумя романами Леонова Набоков и Толстой разведены алгоритмом достаточно решительно — как, разумеется, и сами романы Леонова. Федину и «Зависти» ближе всего Тынянов, им и Тынянову — Вагинов. Совершенно особняком стоит в этой схеме Пильняк.

При чтении слева первое ветвление выделяет, естественно, две группы. Назовем одну очень условно «набоковской» по нашему основному автору, другую — «одесско-фединской». «Набоковская» ветвится на уже описанные группы Газданов — Грин — Зошенко и Гайдар — Зайцев — Набоков + «Скутаревский». Во второй группе сразу отделяется Пильняк, ему противопоставлены все остальные. Эти «остальные» делятся на кластеры Федин — Тынянов — Вагинов — «Зависть» и Катаев — Ильф и Петров — «Три толстяка» — Толстой — «Барсуки».

Казус Леонова и Газданова

Примечательно, что в экспериментах на 300–500 mfw роман Газданова «Вечер у Клэр» атрибутируется позднему Набокову (ил. 2).



Ил. 2. Иерархическая кластеризация, визуализирующая расстояние Delta для анализируемых текстов. 300 mfw

Fig. 2. Hierarchical clustering visualizing Delta distance for analyzed texts. 300 mfw



Как видим, к Набокову, помимо «Скутаревского», прибавился «Вечер у Клэр». Зощенко теперь отдельный автор «набоковской» ветви, более тесная его связь именно с Грином и Газдановым преодолевается. Пильняк, напротив, уже не читается как отдельный в рамках второй ветви автор, а «переезжает» к Толстому и «Барсукам». Гайдар и Зайцев перемещаются из «набоковской» ветви, сильно поредевшей, в другую, но в ней отделяются от всех остальных первыми.

На этом этапе мы решили воспользоваться функцией `size.penalize`. Обычно она используется для проверки эффективности метода на отрывках разной длины при работе с различными классификаторами — в том числе и с Delta. Механизм ее работы таков: из текста извлекаются слу-

чайные выборки всё возрастающей длины, а затем они сравниваются с обучающей выборкой для классификации с применением заданного количества частотных слов (в нашем случае — 100, 200 и 500 *mfw*). Для каждого отрывка проводится сто итераций. Однако для нас важна не только оценка точности метода: функция также предоставляет матрицы смешения, позволяющие понять, между какими авторами происходит путаница и, следовательно, стилистические признаки каких писателей настолько близки, что представляются проблемными для однозначной атрибуции.

Анализ матриц смешения для романов «Скутаревский» и «Барсуки» показал, что Delta «путает» «Скутаревского» с текстами Набокова. При частотном словаре объемом 500 *mfw* роман атрибутируется Набокову на любом отрезке от 5000 до 10 000 слов в ста случаях из ста; при 200 *mfw* показатель колеблется от 73 (5000 слов) до 93 (10 000 слов). В этой связи нельзя не вспомнить слова Ходасевича, который при сопоставлении эмигрантской и советской литератур заявил, что «по объему дарования Сирин не уступает Леонову или Федину» (цит. по: [Мельников, Коростелев 2000: 98]. В какой-то степени Delta подтверждает справедливость этого сравнения — используя лишь распределение служебных частей речи. «Барсуки» же атрибутируются Толстому: при 100 *mfw* показатель колеблется от 64 до 78; при 200 *mfw* — от 82 до 92; при 500 *mfw* — от 94 до 99.

«Выделяемость» текстов Набокова и Леонова мотивируется и на языковом уровне. Оба в целом чужды той в самом широком смысле южной речевой основе, которая, возможно, сказывается в словесных предпочтениях многих крупнейших прозаиков этой (межвоенной) и предыдущей (рубеж XIX и XX вв., начало XX в.) эпохи (Чехова, Бунина, Булгакова, Платонова, Катаева, Олеси, Шолохова), хотя это и не абсолютно: преобладание *глядеть* над *смотреть*, типичное у «южных» авторов (детство и юность которых прошли в таких регионах и городах, как Приазовье, Черноземье, Дон, Киев, Одесса), характерно и для Леонова.

В употреблении наиболее частых предлогов и союзов обоим авторам свойственна нейтральность: например, для Булгакова характерна повышенная частотность предлога *в* (3,4% в общем словаре по данным Национального корпуса русского языка (НКРЯ) против 2,8–2,9% у Набокова и Леонова и 3% в среднем по корпусу за десятилетие), для Андрея Белого — «с» (1,5% против 1,3%), Шолохова — *на* (2,1% против среднего 1,6–1,7%); с другой стороны, у Белого мало *на* (1,3%), у Шолохова — *в* (2,4%) и у Платонова — *с* (0,8%), а Набоков и Леонов нейтральны по всем трем предлогам. У Булгакова и Платонова много союзов *и* (а у Шолохова мало), у Шолохова и Платонова много *а*, Набоков и Леонов вновь нейтральны.

Разумеется, различия, биографические и внутритекстовые, очевидны: участие Леонова в событиях Гражданской войны, несомненная включенность в процессы развития советской литературы и идейная и стилевая ориентация на Достоевского — это все то, чего у Набокова нет.

Однако показатели матриц смешения «Вечера у Клэр» еще радикальнее (табл. 1):

Таблица 1 / Table 1

**Показатели матриц смешения романа «Вечер у Клэр»
Confusion matrices of *An Evening with Claire***

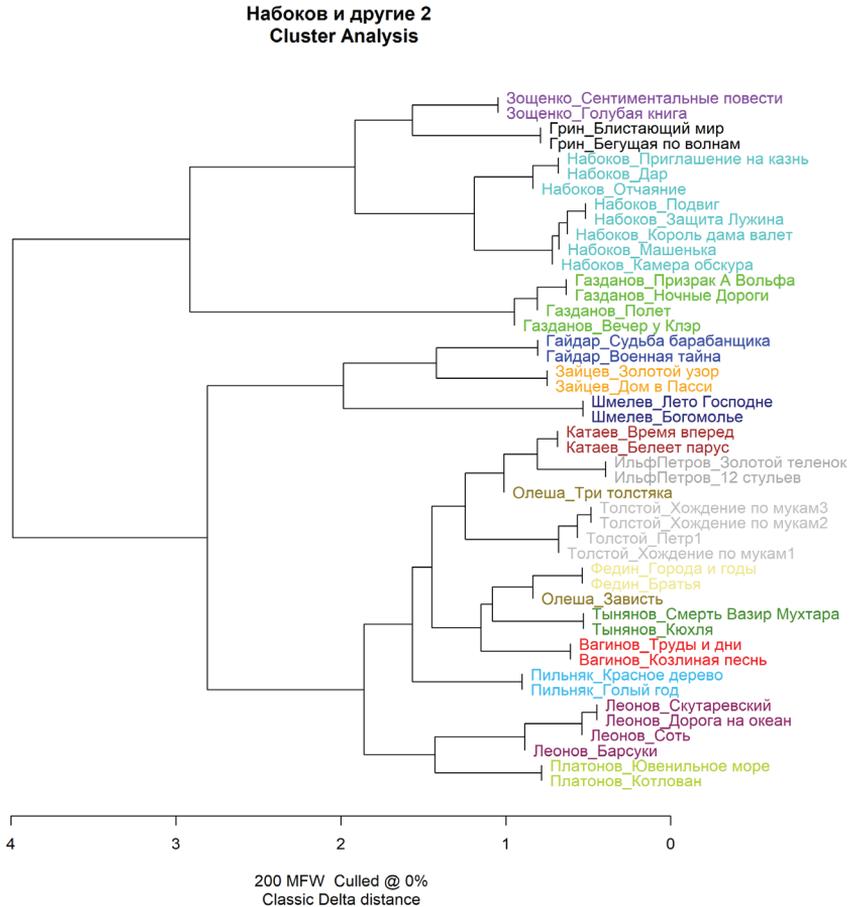
	100 mfw			200 mfw			500 mfw		
	5000	7000	10000	5000	7000	10000	5000	7000	10000
Набоков	86	88	94	95	97	98	99	100	100
Газданов	4	5	1	5	3	2	1	0	0
Леонов	10	7	5	0	0	0	0	0	0

Сближение Набокова с Газдановым, как кажется, требует объяснений в наименьшей степени. Их часто сопоставляют из-за общих черт [Шульман 2000; Кибальник 2002; 2012], таких как принадлежность к определенному поколению, к диаспоре, артистизм, преобладающий жанр короткого, отточенного, стилистически сдержанного романа, собранного вокруг центрального героя и неразветвленной фабулы. На фоне сходства заметны различия, но это не повод пренебрегать сходствами. На книжной полке героя «Тяжелого дыма» соседствуют «Вечер у Клэр» и «Защита Лужина» (а также «Двенадцать стульев»):

Тут был и случайный хлам (больше всего), и учебники по политической экономии (я хотел совсем другое, но отец настоял на своем); были и любимые, в разное время потрафившие душе, книги, «Шатер» и «Сестра моя жизнь», «Вечер у Клэр» <...>, «Защита Лужина» и «Двенадцать стульев»... (цит. по НКРЯ).

При всей очевидности сближения его нельзя игнорировать. Поэтому далее мы решили провести дополнительный эксперимент, добавив в корпус тексты Л. Леонова «Соть» и «Дорога на океан», а вместе с тем иные тексты Г. Газданова: «Полет», «Ночные дороги». Цель — проверить, повлияет ли увеличение количества текстов авторов, один из текстов которого атрибутировался «не по адресу», к воссоединению текстов этих авторов на одной ветви. Также в качестве сопоставительного материала были включены тексты А. Платонова («Котлован», «Ювенильное море») и И. Шмелева («Лето Господне» и «Богомолье»). Результаты представлены на ил. 3.

Тексты Леонова и Газданова «воссоединились». Теперь в корпусе 46 единиц, это почти тот предел, который был определен в начале исследования. На этот раз алгоритм считывает авторский сигнал в 16 из 17 случаев, и только «Зависть» и «Трех толстяков» продолжает рассматривать как фактически принадлежащих разным авторам, что, в целом, объяснимо в свете жанровых различий текстов. Введенный в корпус на этом этапе Шмелев примыкает к паре Зайцева (что само по



Ил. 3. Иерархическая кластеризация, визуализирующая расстояние Delta для анализируемых текстов. 200 mfw

Fig. 3. Hierarchical clustering visualizing Delta distance for analyzed texts. 200 mfw



себе ожидаемо) и Гайдара (по-прежнему интригует близость последнего с Зайцевым, а теперь и со Шмелевым). Пять (четыре плюс одна, ранее всего отделяющаяся от наиболее нагруженной) групп таковы: Набоков — Грин — Зощенко; Газданов; Шмелев — Зайцев — Гайдар; Леонов — Платонов; прочие (Пильняк внутри этой группы стоит особняком; к «одесситам» примыкает Толстой; кроме того, Федин — Тынянов — Вагинов — «Зависть»).

При помощи функции `size.penalize` мы вычислили средний показатель точности атрибуции «Скутаревского» и «Барсуков» — 1. Выравнивается и показатель «Вечера у Клэр», хотя и не столь радикально (табл. 2).

Таблица 2 / Table 2

Средние показатели успешной атрибуции («Вечер у Клэр»)
Accuracy scores (*An Evening with Claire*)

	5000	6000	7000	8000	9000	10000
100 mfw	0,80	0,72	0,88	0,91	0,90	0,96
200 mfw	0,57	0,50	0,63	0,62	0,58	0,63
500 mfw	0,45	0,52	0,63	0,57	0,56	0,67

Изменились и показатели матриц смещения (см. табл. 3).

Таблица 3 / Table 3

Показатели матриц смещения романа «Вечер у Клэр»
(расширенный корпус)
Confusion matrices of *An Evening with Claire*
(extended corpus)

	100 mfw			200 mfw			500 mfw		
	5000	7000	10000	5000	7000	10000	5000	7000	10000
Набоков	13	11	2	41	36	37	55	37	33
Газданов	80	88	96	57	63	63	45	63	67
Леонов	7	1	2	2	1	0	0	0	0

Средний показатель успешной атрибуции других романов Газданова колеблется между 0,99 и 1.

В целом, можно говорить о том, что роман «Вечер у Клэр» является наиболее стилистически схожим с русскоязычными романами Набокова из всех ранних романов Газданова. Даже при включении в выборку большего количества текстов Газданова Delta выявляет «набоковский след» в «Вечере у Клэр»⁵. Это контрастирует с тем, что мы видим в случае Леоно-

⁵ Примечательно, что С. А. Кибальник говорит об акцентированной связи «Вечера у Клэр» с «Подлинной жизнью Себастьяна Найта»: «Газданов одним из первых в русской эмиграции начал осмысление своего внутреннего транскультурного опыта и создал особую транскультурную поэтику. На этой поэтике в значительной степени построена, как и «Вечер у Клэр», газдановская «История одного путешествия», которая буквально подготовила «Подлинную жизнь Себастьяна Найта» <...> автор в этом романе как бы раздваивается на двух главных героев: Володю и его старшего родного брата Николая (имена взяты из «автобиографической трилогии» Толстого, которая является одним из претекстов обоих романов)» [Кибальник 2011: 37].

ва: после добавления его текстов в корпус показатели набоковского следа в «Скутаревском» снизились до нуля.

Итак, поскольку здесь представлены данные нескольких замеров, необходимо обобщить, какие итоги кластеризации являются неизменными, а какие меняются в зависимости от сдвига условий эксперимента. Из 17 авторов, включая двух, добавленных на заключительном этапе, 13 группировались только компактно — их тексты не смешивались с чужими и не относились в разные части схемы; исключения — Набоков, Леонов, Газданов, Олеша (последний, и только он, — во всех замерах). Всегда в одной из двух главных ветвей были Грин — Газданов — Зошенко — Набоков, всегда в другой главной ветви — Катаев — Ильф и Петров — Олеша — Толстой — Федин — Тынянов — Вагинов — Пильняк; перемещались в другую из двух главных ветвей Гайдар — Зайцев; присутствовал в обеих ветвях Леонов. Всегда рядом были Катаев — Ильф и Петров — «Три толстяка», Федин — Тынянов — «Зависть» и Гайдар — Зайцев (но Гайдар и Зайцев, оставшись вместе, переместились в другую главную ветвь). Внутри второй ветви меняли ближайшее соседство Толстой (вначале с «одесситами», потом с Пильняком и «Барсуками», потом опять с «одесситами»), Вагинов и Пильняк плавно колеблются между отделенностью и отдаленными примыканиями во второй ветви. Леонов, разделенный вначале между Набоковым и Толстым, на втором этапе объединяется с Платоновым. Выделяются асимметрии между рассеиванием авторов, малых групп и больших групп: Олеша разделен, но «Зависть» всегда рядом со стабильной парой Федин — Тынянов, а «Три толстяка» с такой же связкой Катаева, Ильфа и Петрова; Гайдар и Зайцев вместе, но меняют соседство (стоит напомнить, во второй ветви они всегда отделяются в первую очередь, так что это фактически отдельная группа).

Возможно, эти различия и асимметрии в замерах позволяют говорить о неких «посредничествах». Леонов оказывается своеобразным «посредником» между Набоковым, Толстым и Платоновым (к Набокову притягивается «интеллигентский» «Скутаревский»), Олеша — между группой Катаева, Ильфа и Петрова и связкой Федин — Тынянов. Через Зайцева и Гайдара протягивается отделенная связка Набокова, Газданова, Грина, Зошенко со второй ветвью⁶. Все это, как кажется, не контринтуитивно. Олеша смыкает традицию «южной школы» и линию «Серрапионов», близких опоязовцам. Леонов, начинавший с прозы, пронизанной сказом, орнаментальностью и почвенным натурализмом, в «Скутаревском» присоединяется к традиции артистического, психологически насыщенного романа о герое из интеллигентной среды, используя при этом язык богатый

⁶ Между Набоковым и Грином обнаруживаются схождения в фабулистике и композиции новелл, включая схемы и мотивы кажимости, заблуждения, обмана, безумия, общее внимание к показательным претекстам, таким как рассказ Амброза Бирса «Случай на мосту через Совиный ручей», использование таких вымышленных топонимов для дистопических городов, как *Лисс* и *Фиальта*. Такие рассказы, как «Мат в три хода» и «Тerra incognita» расширяют область сближения, отнюдь ее не исчерпывая. О связи Набокова и Грина см.: [Толстая 2021].

и образный, но в целом сдержанный — точки сближения с Набоковым очевидны. Связь автора еще чеховско-бунинской плеяды как с Набоковым и Газдановым, так и с авторами вроде Толстого, Федина и «бунинца» Катаева тоже сама по себе не удивительна.

Два периода русскоязычного Набокова

При проведении экспериментов мы обратили внимание на одну константу — регулярную дифференциацию романов Набокова на две группы. Первая состоит из ранних текстов «Машенька», «Защита Лужина», «Король, дама, валет», «Подвиг», «Камера обскура». Вторая — из романов 1930-х годов «Отчаяние», «Приглашение на казнь», «Дар».

Аналогичные данные фигурируют и в статье Орехова [2021], однако мы решили уточнить результаты — и приняли решение обратиться к ряду классификаторов, которыми располагает пакет Stylo: Delta, KNN и SVM.

Метод k-ближайших соседей (KNN; англ. k-nearest neighbors) основывается на оценивании сходства объектов и относит атрибутируемый объект к тому классу, к которому принадлежит и большинство его ближайших соседей. Метод опорных векторов (SVM; англ. support vector machine), разработанный В. Н. Вапником [1979], зиждется на том, что алгоритм создает гиперплоскость, которая оптимальным способом разделяет данные на классы. Векторы, расположенные с одной «стороны» гиперплоскости, относятся к одному классу, а расположенные с другой — к другому.

В обучающий корпус мы поместили ранние романы «Машенька», «Защита Лужина», «Король, дама, валет» и поздние «Дар» и «Приглашение на казнь». В тестовый корпус были включены тексты «Подвиг» и «Отчаяние» как атрибутируемые и «Камера обскура» как сопоставительный. Эксперименты проводились с частотным словарем объемом 100, 200, 300 mfw.

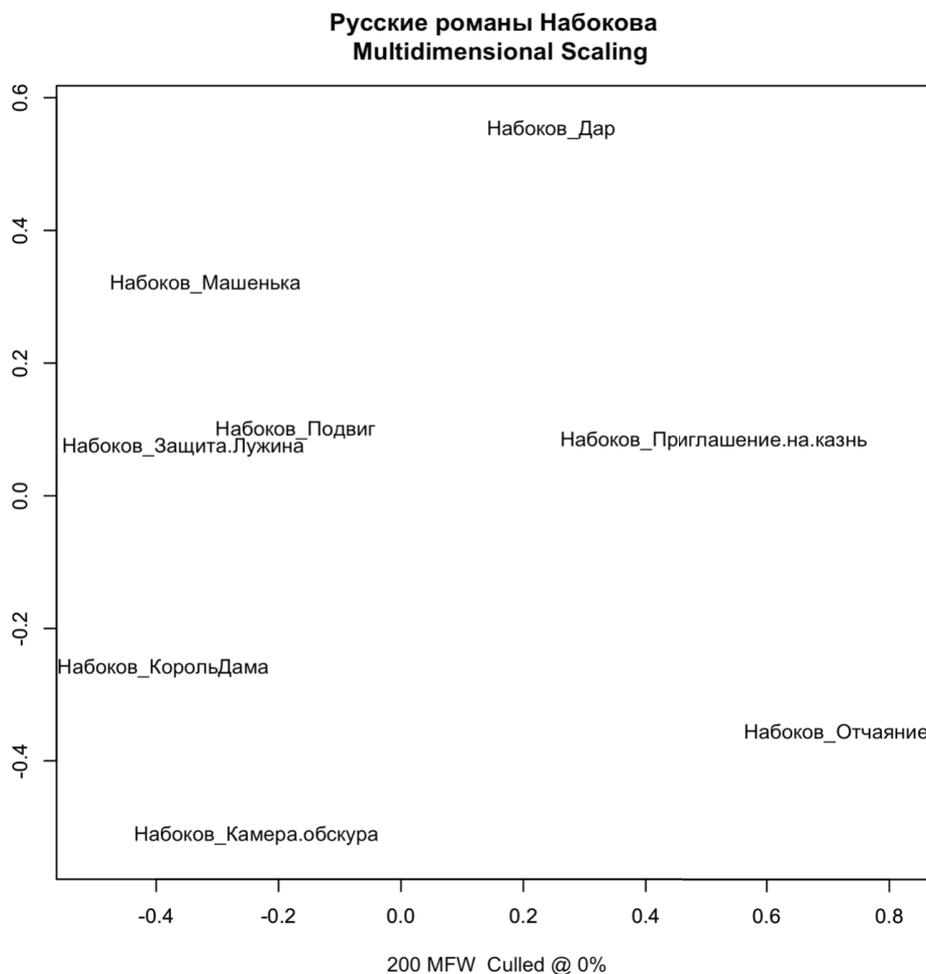
В результате были получены идентичные результаты: каждый классификатор при каждом объеме слов отнес «Подвиг» к первой группе, а «Отчаяние» — ко второй.

Аналогичные результаты были получены и при атрибуции текстов «Приглашение на казнь» (вторая группа) и «Король, дама, валет» (первая).

Особенно наглядно разделение романов Набокова на две группы видно при многомерном шкалировании (ил. 4).

Деление восьми русских романов Набокова на эти группы сохраняется при использовании целого ряда классификаторов. Мы полагаем, что этому можно отыскать мотивировки и на уровне традиционного литературоведческого анализа.

Хотя в пяти романах первой группы различаются и среда действия (три романа из «русского цикла» и два немецких), и контур фабулы (адюльтер и преступление как фабульная основа немецких, использующих топику жанровой литературы и /или раннего кинематографа, чего почти совершенно нет в «Машеньке», «Защите Лужина» и «Подвиге»), общая простота и некоторая стертость языковой ткани и микропоэтики играет свою роль.



Ил. 4. Многомерное шкалирование русскоязычных романов Набокова
200 mfw

Fig. 4. Multidimensional scaling of Nabokov's Russian novels. 200 mfw



В трех других романах существенно возрастают количество и функциональная нагруженность звуковых повторов и каламбуров, парадоксальных описаний и усложненных метафор, не прямой, иронической речи персонажей и повествователя, игры с претекстами, языковой и нарративной многослойности. «Умышленное», полное двусмысленностей и пародийности повествование сближает их с поздними русскими рассказами Набокова и с двумя шедеврами его драматургии, «Событием» и «Изобретением Вальса». Все перечисленное не примыкает непосредственно к тем лексическими предпочтениям, которые отслеживает использованный алго-

ритм, однако было бы естественно предполагать, что общая перестройка слога сказалась и в сдвиге словоупотребления. Можно с некоторой осторожностью заключить, что по сравнению с «Отчаянием», «Приглашением на казнь» и «Даром» (главным воплощением поздней манеры русского Набокова) пять первых романов выглядят образцами более прямого и прозрачного, однонаправленного и «реалистического» повествования (реалистического, конечно, только при сопоставлении с тремя другими романами).

Заключение

Мы можем сгруппировать основные итоги и выводы по четырем направлениям.

1. Осуществлена кластеризация авторов и стилей. В предложенном материале выделяются несколько центров притяжения. Во-первых, это Набоков, Газданов, Грин, рядом с ними Зошенко. Во-вторых, это Федин и Тынянов, далее Катаев, Ильф и Петров, и между этими двумя группами Олеша. В-третьих, это Гайдар и Зайцев, к которым присоединяется Шмелев. Какие черты словоупотребления и стиля Гайдара сближают его со Шмелевым и Зайцевым, двумя православными авторами, сформировавшимися задолго до революции, позднее творчество которых приходится на годы жизни в диаспоре, сказать пока сложно; в любом случае от всех советских авторов Гайдар здесь далек. При последнем, самом надежном замере сгруппировались — и это, пожалуй, в-четвертых — Леонов и Платонов. Три последние группы различаются между собой меньше, чем все они в совокупности с первой (иначе: из четырех групп трем другим резче всего противопоставлена первая). К той же большой ветви прибавляются Толстой, Тынянов и Пильняк, не имеющие устойчивых и надежных близких соседств. Разрушатся ли или укрепятся эти линии («иностранная», «одесско-серапионовская», «советско-почвенная» и «православно-гайдаровская») при прибавлении нового материала, могут выяснить новые исследования.

2. Еще раз проверен алгоритм. Он определил 16 из 17 авторских сигналов и выдал решения большей и меньшей степени понятности по дальнейшей кластеризации авторов, что позволяет говорить о достаточно высокой степени его надежности. Стоит учитывать, что речь идет о словоупотреблении и о том, что может быть отражено в словоупотреблении, хотя непосредственно к нему не относится (синтаксис в наборе наиболее частых союзов, коммуникативная перспектива в относительной частотности местоимений и т. д.). Другие характеристики текста могут изучаться другими способами, в том числе автоматическими; стоит скорее удивляться тому, как многое видит этот инструмент, основанный всего лишь на простом количественном сопоставлении словоформ.

3. Значимые как на методологическом, так и на собственном стилистическом уровне результаты получены при определенном ранжировании стилистически схожих текстов. Анализируя триаду Леонов — Газданов —

Набоков, мы обнаружили, что на малом корпусе «Скутаревский» квалифицируется как стилистически схожий с поздними русскими романами Набокова; на расширенном же корпусе Delta перестает констатировать их фундаментальное сходство. Этот результат вписывается в общую тенденцию подобного рода исследований: авторский сигнал оказывается сильнее иных факторов, это всегда показатель «первого порядка». Иное — в случае с «Вечером у Клэр» Газданова. Даже при эксперименте на расширенном корпусе Delta отмечает сходство этого романа с текстами Набокова, что, как мы предполагаем, подтверждает при помощи статистических методов идею о том, что «Вечер у Клэр» и в целом творческий корпус Газданова стилистически ближе к романам Сирина, нежели «Скутаревский» и другие тексты Леонова.

4. Наконец, не стоит забывать, что обычно Delta используется для атрибуции и близких ей задач. Здесь она применена в рамках сопоставительно-стилевого исследования. Для относительно новых задач, может быть, потребуются новые условия использования алгоритма: возможно, новые требования к количеству сопоставляемых авторов, текстов одного автора, к жанровому и хронологическому разбросу, к самому количеству координируемых замеров. Предлагаемая работа может рассматриваться как один из первых, все еще экспериментальных примеров собственно стилистического применения метода Delta.

Литература

- Алиева 2022a — *Алиева О. В.* Delta Берроуза для древнегреческих авторов: опыт применения // СХОАН. Vol. 16. No. 2. 2022. С. 693–705. <https://doi.org/10.25205/1995-4328-2022-16-2-693-705>.
- Алиева 2022b — *Алиева О. В.* Опыт измерения стилистической однородности методом Delta на материале Платоновского корпуса // Аристей: Вестник классической филологии и античной истории. Т. 25. 2022. С. 19–37. https://doi.org/10.53084/22209050_2022_25_19.
- Вапник 1979 — *Вапник В. Н.* Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979.
- Великанова, Орехов 2019 — *Великанова Н. П., Орехов Б. В.* Цифровая текстология: атрибуция текста на примере романа М. А. Шолохова «Тихий Дон» // Мир Шолохова. 2019. № 1. С. 70–82.
- Кибальник 2002 — *Кибальник С. А.* Газданов и Набоков // Русская литература. 2002. № 3. С. 22–41.
- Кибальник 2011 — *Кибальник С. А.* Набоков и Газданов. (О романе Набокова «Полная жизнь Себастьяна Найта») // Новый филологический вестник. 2011. № 1(16). С. 36–46.
- Кибальник 2012 — *Кибальник С. А.* Псевдоединство младоэмигрантов: Гайто Газданов и Владимир Набоков // Вопросы литературы. 2012. № 1. С. 276–287
- Ковалев 2023 — *Ковалев Б. В.* Delta и «бум»: романы Марио Варгаса Льюиса через призму стилиметрии // Литература двух Америк. № 15. 2023. С. 116–141. <https://doi.org/10.22455/2541-7894-2023-15-116-141>.
- Мамаев и др. 2018 — *Мамаев Н. К., Марусенко М. А., Пиотровская К. Р., Ронжин А. Л.* Метод Дельты Берроуза для определения авторства анонимных и псевдоаноним-

- ных литературных произведений на русском языке // Proceedings of the R. Piotrowski's Readings in Language Engineering and Applied Linguistics. Saint Petersburg, Russia, November 27, 2017 / Ed. by A. Ronzhin, T. Noskova, A. Karpov. [Б. м.]: [б. и.], 2018. С. 107–119 [в электрон. версии паг. отдельная]. (CEUR Workshop Proceedings; Vol. 2233). URL: https://ceur-ws.org/Vol-2233/Paper_9.pdf.
- Маслинский 2022 — *Маслинский К. А.* Уточненная цифровая текстология: еще раз к вопросу об авторстве романа «Тихий Дон» // Русская литература. 2022. № 1. С. 247–254. <https://doi.org/10.31860/0131-6095-2022-1-247-254>.
- Мельников, Коростелев 2000 — Классик без ретуши. Литературный мир о творчестве Владимира Набокова: Критические отзывы, эссе, пародии / Под общ. ред. Н. Г. Мельникова; Сост., подгот. текста Н. Г. Мельникова, О. А. Коростелева. М.: Нов. лит. обозрение, 2000.
- Орехов 2020 — *Орехов Б. В.* «Илиада» Е. И. Кострова и «Илиада» А. И. Любжина: стилеметрический аспект // Аристей: Вестник классической филологии и античной истории. Т. 21. 2020. С. 282–296.
- Орехов 2021 — *Орехов Б. В.* Текст и перевод Владимира Набокова через призму стилеметрии // Новый филологический вестник. 2021. № 3(58). С. 200–213. https://doi.org/10.54770/20729316_2021_3_200.
- Петров и др. 2019 — *Петров В. В., Марусенко М. А., Пиотровская К. Р., Маньяс И. Н., Мамаев Н. К.* Об авторстве «писем Берии из заточения» // Вестник Санкт-Петербургского университета. Право. Т. 10. № 3. 2019. С. 586–605. <https://doi.org/10.21638/spbu14.2019.313>.
- Толстая 2021 — *Толстая Е. Д.* Атака на невыразимое: Набоков и Грин // Набоков и современники: Историко-литературный альманах. № 1 / Отв. ред. Т. О. Пономарева, О. Ю. Сконечная. СПб.: Симпозиум, 2021. С. 288–309.
- Шульман 2000 — *Шульман М. Ю.* Газданов и Набоков // Возвращение Гайто Газданова: Науч. конф., посвящ. 95-летию со дня рождения / Сост. М. А. Васильева. М.: Рус. путь, 2000. С. 15–24.
- Al-Falahni et al. 2015 — *Al-Falahni A., Bellafkin M., Romdani M., Al-Serem M.* Authorship attribution in Arabic poetry // 2015 10th International Conference on Intelligent Systems: Theories and Applications (SITA). October 20–21, 2015, Rabat, Morocco. ENSIAS, Rabat, Morocco. [N. p.]: [n. e.], 2015. URL: <https://ieeexplore.ieee.org/document/7358411>. <https://doi.org/10.1109/SITA.2015.7358411>.
- Burrows 2002 — *Burrows J.* ‘Delta’: A measure of stylistic difference and a guide to likely authorship // Literary and Linguistic Computing. Vol. 17. No. 3. 2002. P. 267–287. <https://doi.org/10.1093/lc/17.3.267>.
- Calvo Tello 2019 — *Calvo Tello J.* Delta inside Valle-Inclán: stylometric classification of periods and groups of his novels // Romanische Studien. 2019. Beihefte 6. P. 151–163.
- Eder et al. 2016 — *Eder M., Rybicki J., Kestemont M.* Stylometry with R: A package for computational text analysis // The R Journal. Vol. 8. No. 1. 2016. P. 107–121.
- Gladwin et al. 2017 — *Gladwin A. A. G., Lavin M. J., Look D. M.* Stylometry and collaborative authorship: Eddy, Lovecraft, and ‘The Loved Dead’ // Digital Scholarship in the Humanities. Vol. 32. No. 1. 2017. P. 123–140. <https://doi.org/10.1093/lc/fqv026>.
- Hernández-Lorenzo, Byszuk 2023 — *Hernández-Lorenzo L., Byszuk J.* Challenging stylometry: The authorship of the baroque play *La Segunda Celestina* // Digital Scholarship in the Humanities. Vol. 38. No. 2. 2023. P. 544–558. <https://doi.org/10.1093/lc/fqac063>.
- Hoover 2004 — *Hoover D. L.* Testing Burrows’s Delta // Literary and Linguistic Computing. Vol. 19. No. 4. 2004. P. 453–475.
- Iosifyan, Vlasov 2020 — *Iosifyan M., Vlasov I.* And Quiet Flows the Don: The Sholokhov–Kryukov authorship debate // Digital Scholarship in the Humanities. Vol. 35. No. 2. 2020. P. 307–318. <https://doi.org/10.1093/lc/fqz017>.

Jannidis, Lauer 2014 — 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 / Ed. by M. Erlin, L. Tatlock. Rochester: Camden House, 2014. P. 29–54.

Skorinkin, Orekhov 2023 — Skorinkin D., Orekhov B. Hacking stylometry with multiple voices: Imaginary writers can override authorial signal in Delta // Digital Scholarship in the Humanities. Vol. 38. No. 3. 2023. P. 1246–1266. <https://doi.org/10.1093/llc/fqad012>.

References

- Al-Falahni, A., Bellafkin, M., Romdani M., & Al-Serem, M. (2015). Authorship attribution in Arabic poetry. In *2015 10th International Conference on Intelligent Systems: Theories and Applications (SITA). October 20–21, 2015, Rabat, Morocco. ENSIAS, Rabat, Morocco* (n. e.). <https://ieeexplore.ieee.org/document/7358411>. <https://doi.org/10.1109/SITA.2015.7358411>.
- Alieva, O. V. (2022a). Del'ta Berrouza dlia drevnegrecheskikh avtorov: opyt primeneniia [Testing Burrows' Delta on Ancient Greek authors]. *Shole*, 16(2), 693–705. <https://doi.org/10.25205/1995-4328-2022-16-2-693-705>. (In Russian).
- Alieva, O. V. (2022b). Opyt izmereniia stilisticheskoi odnorodnosti metodom Delta na materiale Platonovskogo korpusa [Measuring stylistic homogeneity with Burrows' Delta: An experiment with Corpus Platonicum]. *Aristei: Vestnik klassicheskoi filologii i antichnoi istorii*, 25, 19–37. https://doi.org/10.53084/22209050_2022_25_19. (In Russian).
- Burrows, J. (2002). 'Delta': A measure of stylistic difference and a guide to likely authorship. *Literary and Linguistic Computing*, 17(3), 267–287. <https://doi.org/10.1093/llc/17.3.267>.
- Calvo Tello, J. (2019). Delta inside Valle-Inclán: stylometric classification of periods and groups of his novels. *Romanische Studien*, 2019(Supplement 6), 151–163.
- Eder, M., Rybicki, J., & Kestemont, M. (2016). Stylometry with R: A package for computational text analysis. *The R Journal*, 8(1), 107–121.
- Gladwin, A. A. G., Lavin, M. J., & Look, D. M. (2017). Stylometry and collaborative authorship: Eddy, Lovecraft, and 'The Loved Dead'. *Digital Scholarship in the Humanities*, 32(1), 123–140. <https://doi.org/10.1093/llc/fqv026>.
- Hernández-Lorenzo, L., & Byszuk, J. (2023). Challenging stylometry: The authorship of the baroque play *La Segunda Celestina*. *Digital Scholarship in the Humanities*, 38(2), 544–558. <https://doi.org/10.1093/llc/fqac063>.
- Hoover, D. L. (2004). Testing Burrows's Delta. *Literary and Linguistic Computing*, 19(4), 453–475.
- Iosifyan, M., & Vlasov, I. (2020). And Quiet Flows the Don: The Sholokhov–Kryukov authorship debate. *Digital Scholarship in the Humanities*, 35(2), 307–318. <https://doi.org/10.1093/llc/fqz017>.
- Jannidis, F., & Lauer, G. (2014). Burrows's Delta and its use in German literary history. In M. Erlin, & L. Tatlock (Eds.). *Distant readings: Topologies of German culture in the long nineteenth century* (pp. 29–54). Camden House.
- Kibal'nik, S. A. (2002). Gazdanov i Nabokov [Gazdanov and Nabokov]. *Russkaia literatura*, 2002(3), 22–41. (In Russian).
- Kibal'nik, S. A. (2011). Nabokov i Gazdanov. (O romane Nabokova "Podlinnaia zhizn' Sebast'iana Naita") [Vladimir Nabokov and Gaito Gazdanov (on V. Nabokov's novel *The Real Life of Sebastian Knight*)]. *Novyi filologicheskii vestnik*, 2011(1, no. 16), 36–46. (In Russian).
- Kibal'nik, S. A. (2012). Psevdoedinstvo mladoemigrantov: Gaito Gazdanov i Vladimir Nabokov [Pseudo-unity of the Young Émigrés: Vladimir Nabokov and Gaito Gazdanov]. *Voprosy literary*, 2012(1), 276–287 (In Russian).

- Kovalev, B. V. (2023) Delta i “bum”: romany Mario Vargasa L’osy cherez prizmu stilemetrii. [Delta and “boom”: Novels of Mario Vargas Llosa through the prism of stylometry]. *Literatura dvukh Amerik*, 15, 116–141. <https://doi.org/10.22455/2541-7894-2023-15-116-141>. (In Russian).
- Mamaev, N., Marusenko, M., Piotrowska, X., Ronzhin, A. (2018). Metod Del’ty Berrouza dlia opredeleniia avtorstva anonimnykh i psevdonimnykh literaturnykh proizvedenii na russkom iazyke [Burrows’s Delta for authorship attribution of Russian literary texts]. In A. Ronzhin, T. Noskova, & A. Karpov (Eds.). *Proceedings of the R. Piotrowski’s Readings in Language Engineering and Applied Linguistics. Saint Petersburg, Russia, November 27, 2017* (pp. 107–119, separate pagination in Web paper) (n. p.). https://ceur-ws.org/Vol-2233/Paper_9.pdf. (In Russian).
- Maslinsky, K. A. (2022). Utochnennaia tsifrovaia tekstologiiia: eshche raz k voprosu ob avtorstve romana “Tikhii Don” [Refined digital textual criticism: More on the question of authorship of the novel *Quiet Flows The Don*]. *Russkaia literatura*, 2022(1), 247–254. <https://doi.org/10.31860/0131-6095-2022-1-247-254>. (In Russian).
- Mel’nikov, N. G., & Korostelev, O. A. (Eds.) (2000). *Klassik bez retushi. Literaturnyy mir o tvorchestve Vladimira Nabokova: Kriticheskie otzvyvy, esse, parodii* [Classic without re-touching. Literary world about the work of Vladimir Nabokov: Critical reviews, essays, parodies]. Novoe literaturnoe obozrenie. (In Russian).
- Orekhov, B. V. (2020). “Iliada” E. I. Kostrova i “Iliada” A. I. Liubzhina: stilemetricheskii aspekt [Iliad by Kostrov and Iliad by Lyubzhin: the stylometry case]. *Aristei: Vestnik klassicheskoi filologii i antichnoi istorii*, 21, 282–296. (In Russian).
- Orekhov, B. V. (2021). Tekst i perevod Vladimira Nabokova cherez prizmu stilemetrii [Text and translation by Vladimir Nabokov through the prism of Stylemetry]. *Novyi filologicheskii vestnik*, 2021(3, no. 58), 200–213. https://doi.org/10.54770/20729316_2021_3_200. (In Russian).
- Petrov, V. V., Marusenko, M. A., Piotrovskaya, K. R., Man’ias, I. N., & Mamaev, N. K. (2019). Ob avtorstve “pisem Berii iz zatocheniia” [On the authorship of ‘Beria’s Letters from imprisonment’]. *Vestnik Sankt-Peterburgskogo universiteta. Pravo*, 10(3), 586–605. <https://doi.org/10.21638/spbu14.2019.313>. (In Russian).
- Shul’man, M. Iu. Gazdanov i Nabokov [Gazdanov and Nabokov]. In M. A. Vasil’eva (Ed.). *Vozvrashchenie Gaito Gazdanova: Nauchnaia konferentsiia, posviashchennaia 95-letiiu so dnia rozhdeniia* (pp. 15–24). Russkii put’. (In Russian).
- Skorinkin, D., & Orekhov, B. (2023). Hacking stylometry with multiple voices: Imaginary writers can override authorial signal in Delta. *Digital Scholarship in the Humanities*, 38(3), 1246–1266. <https://doi.org/10.1093/lc/fqad012>.
- Tolstoy, H. D. (2021). Ataka na nevyrazimoe: Nabokov i Grin [An attack at the unexpressible: Grin and Nabokov]. In T. O. Ponomareva, & O. Iu. Skonechnaia (Eds.). *Nabokov i sovremenniki. Istoriko-literaturnyj al’manah* (Vol. 1, pp. 288–309). (In Russian).
- Vapnik, V. N. (1979). *Vosstanovlenie zavisimostei po empiricheskim dannym* [Recovery of dependencies based on empirical data]. Nauka. (In Russian).
- Velikanova, N. P., & Orekhov, B. V. (2019). Tsifrovaia tekstologiiia: atributsiia teksta na primere romana M. A. Sholokhova “Tikhii Don” [Digital textual criticism: text attribution on the example of the novel by M. A. Sholokhov *And Quiet Flows the Don*]. *Mir Sholokhova*, 2019(1), 70–82. (In Russian).

* * *

Информация об авторах Information about the authors

Федор Никитич Двинятин

*кандидат филологических наук
доцент, кафедра русского
языка, Санкт-Петербургский
государственный университет
Россия, 199034, Санкт-Петербург,
Университетская наб., д. 7–9
✉ f.dvinyatin@spbu.ru*

Борис Вадимович Ковалев

*ассистент, кафедра романской
филологии, Санкт-Петербургский
государственный университет
Россия, 199034, Санкт-Петербург,
Университетская наб., д. 7–9
член Союза писателей Санкт-
Петербурга
✉ bvkovalev@yandex.ru*

Fyodor N. Dviniatin

*Cand. Sci. (Philology)
Associate Professor, Department
of Russian Language, St. Petersburg
State University
Russia, 199034, St. Petersburg,
Universitetskaya Emb., 7–9
✉ f.dvinyatin@spbu.ru*

Boris V. Kovalev

*Assistant, Department
of Romance Philology
Russia, 199034, St. Petersburg,
Universitetskaya Emb., 7–9
Member of the Writers' Union
of St. Petersburg
✉ bvkovalev@yandex.ru*