



Том XXV 2022

АРИСТЕЙ

ARISTEAS • VOL. XXV • MMXXII • MOSQVAE

ВЕСТНИК КЛАССИЧЕСКОЙ ФИЛОЛОГИИ И АНТИЧНОЙ ИСТОРИИ

Журнал носит имя Аристея, греческого поэта VII или VI века до н. э., происходившего из города Проконнеса, который находился на одноименном острове в Мраморном море. По свидетельствам многих античных авторов, и прежде всего Геродота, Аристей первым посетил северо-восточную окраину известного тогда мира, т. е. побывал на территории современной Восточной Европы, и написал поэму под названием «Аримаспея», в которой рассказал о живущих там народах. Именно Аристей заложил историографическую традицию помещать севернее Понта Эвксинского (Черного моря) реальных и легендарных скифов и киммерийцев, затем исседонов, аримаспов, грифов и гипербореев, живущих уже на берегу Северного океана. На протяжении многих веков вплоть до Нового времени эти этногеографические представления о Восточной Европе, происходящие от Аристея, оставались доминирующими.

ARISTEAS • VOL. XXV • MMXXII



Университет Дмитрия Пожарского

25

ТОМ XXV (2022)

ВЕСТНИК КЛАССИЧЕСКОЙ ФИЛОЛОГИИ И АНТИЧНОЙ ИСТОРИИ

АРИСТЕЙ



Университет Дмитрия Пожарского
МОСКВА

⋈ VOL. XXV (MMXXII) ⋈

PHILOLOGIA CLASSICA ET HISTORIA ANTIQVA

ARISTEAS



In aedibus Academiae Pozharskianae
MOSQVAE

ISSN: 2220-9050

Журнал издается Университетом Дмитрия Пожарского

Учредитель журнала – Русский фонд содействия образованию и науке

Редакционный совет

А. Аврам (университет Ле Мана, Франция), М. фон Альбрехт (Гейдельбергский университет, Германия), Н.В. Брагинская (НИУ ВШЭ, Москва), С. Браунд (Университет Британской Колумбии, Ванкувер, Канада), А.Ю. Виноградов (НИУ ВШЭ, Москва), Е.Л. Ермолаева (СПб. государственный университет, Санкт-Петербург), А.В. Коптев (университет Тампере, Финляндия), А.Е. Кузнецов (МГУ им. М.В. Ломоносова, Москва), К.Г. Красухин (ИЯ РАН, Москва), А.В. Махлаюк (ННГУ, Нижний Новгород), В.Ф. Новодранова (МГМСУ им. А.И. Евдокимова, Москва), С.Ю. Сапрыкин (МГУ им. М.В. Ломоносова, ИВИ РАН, Москва), О.В. Сидорович (НИУ ВШЭ, Москва), А. Страмалья (Университет Бари, Италия), И.Е. Суриков (ИВИ РАН, Москва), М. Финкельберг (Тель-Авивский университет, Израиль), Г.Р. Цецхладзе (Оксфордский университет, Великобритания), Ю.А. Шичалин (ПСТГУ, Москва)

Editorial council

Alexandru Avram (Le Mans Université, France), Michael von Albrecht (Universität Heidelberg, Germany), Nina V. Braginskaya (RSUH, HSE, Moscow), Susanna Braund (University of British Columbia, Vancouver, Canada), Andrey Yu. Vinogradov (HSE, Moscow), Elena L. Ermolaeva (St. Petersburg State University, St. Petersburg), Alexander V. Koptev (University of Tampere, Finland), Alexander E. Kuznetsov (Lomonosov MSU, Moscow), Konstantin G. Krasukhin (Iling of RAS, Moscow), Aleksander V. Makhlayuk (Lobachevsky State University, Nizhny Novgorod), Valentina F. Novodranova (Evdokimov MSMSU, Moscow), Sergey Yu. Saprykin (Lomonosov MSU, IWH of RAS, Moscow), Olga V. Sidorovitch (HSE, Moscow), Antonio Stramaglia (Università degli Studi di Bari Aldo Moro, Italia), Igor E. Surikov (IWH of RAS; Moscow), Margalit Finkelberg (Tel Aviv University, Israel), Gocha R. Tsetschladze (Linacre College, University of Oxford, Great Britain), Yuri A. Shichalin (Saint Tikhon's Orthodox University of the Humanities, Moscow)

Редакционная коллегия

А.В. Подосинов (главный редактор; ИВИ РАН; МГУ им. М.В. Ломоносова), А.В. Белоусов (заместитель главного редактора; МГУ им. М.В. Ломоносова; РГГУ), С.А. Степанцов (ИМЛИ РАН), Г.С. Беликов (ответственный секретарь; МГУ им. М.В. Ломоносова); И.А. Макаров (ИВИ РАН; РГГУ), А.В. Мосолкин (МГУ им. М.В. Ломоносова), Б.М. Никольский (ИМЛИ РАН), В.В. Файер (НИУ ВШЭ), М.В. Шумилин (РАНХиГС; ИМЛИ РАН), Е.А. Щербакова (РАНХиГС)

Editorial board

Alexander V. Podossinov (Editor-in-Chief; IWH of RAS; Lomonosov MSU, Moscow), Alexey V. Belousov (assistant Editor-in-Chief; Lomonosov MSU, RSUH, Moscow), Sergey A. Stepantsov (assistant Editor-in-Chief; Institute of World Literature of RAS, Moscow), Grigory S. Belikov (executive secretary; Lomonosov MSU, Moscow), Igor A. Makarov (IWH of RAS, RSUH, Moscow), Alexey V. Mosolkin (Lomonosov MSU, Moscow), Boris V. Nikolsky (Institute of World Literature of RAS, Moscow), Vladimir V. Fayer (HSE, Moscow), Mikhail V. Shumilin (Institute of World Literature of RAS, RANEPА, Moscow), Elisaveta F. Sherbakova (RANEPА, Moscow)

На первой странице обложки – *Кифаред. Краснофигурная амфора. V в. до н.э. Аттика. The Metropolitan Museum of Art. Нью Йорк.*

Адрес издательства: 117246 Москва, Научный проезд, д. 8, с. 1, каб. 430

Сайт: aristeas.ru

Е-mail: aristeas.classics@gmail.com

Информационная поддержка: www.librarius.narod.ru

© Редколлегия журнала, 2022

© Коллектив авторов, 2022

© А.В. Белоусова, дизайн и верстка, 2022

© Оформление, Русский фонд содействия образованию и науке, 2022

СОДЕРЖАНИЕ



СОДЕРЖАНИЕ	5
CONTENTS	7
ОТ ГЛАВНОГО РЕДАКТОРА	9
ЛАТЫНЬ СЕГОДНЯ	
<i>Michaelis Shumilin</i>	
Saturarum liber imperfectus	12
СТАТЬИ	
<i>О.В. Алиева</i>	
Опыт измерения стилистической однородности методом Delta на материале Платоновского корпуса	19
<i>Н.А. Филянов</i>	
Свидетельство Диодора Сицилийского об италийском посольстве к Митридату в период Союзнической войны: достоверность и проблемы интерпретации	38
<i>А. Грюнерт</i>	
Медицинская метафора лечебницы в «Беседах» Эпиктета и «Беседах на псалмы» Оригена	51
<i>А.О. Денисов</i>	
Отображение античной мифологии на средневековых картах	71
ПУТЕШЕСТВИЯ В НЕИЗВЕСТНУЮ АНТИЧНОСТЬ	
<i>Е.В. Приходько</i>	
Пидны – эллинистическая крепость на рубежах Ксанфа. Часть 3. Вдоль крепостных стен	88
MISCELLANEA	
<i>М.В. Шумилин</i>	
Stat. Silu 4.2.7	167

ПУБЛИКАЦИИ

*Элий Аристид***К обвиняющим его в том, что он не выступает с речами****Против плясунов***(Перевод с древнегреческого С.И. Межерской)*170*Йоргос Сеферис***Стихи об античности***(Перевод с новогреческого и комментарии О.П. Цыбенко)*191

КРИТИКА И БИБЛИОГРАФИЯ

*А.В. Подосинов***Рецензия на книгу: Dueck D. Illiterate Geography in Classical Athens and Rome (Routledge Monographs in Classical Studies). Routledge; London; New York, 2021, xi+265 p., illustrations. Cased. ISBN 978-0-367-43970-5**242*В.К. Хрусталёв***Некоторые замечания к русскому переводу эпиграфического закона *de repetundis*****(«Вестник древней истории», 2007, № 4)**254

ХРОНИКА

*П.А. Евдокимов***Международная конференция «За пределами Кипра:****исследования внешних связей в Средиземноморье с эпохи поздней бронзы****до конца классического периода» (zoom.us, 8-11 декабря 2020 г.)**269

IN MEMORIAM

*Lynette Mitchell (Exeter)***Peter John Rhodes (10th August 1940 – 27th October 2021)**287

СОКРАЩЕНИЯ290

ПРАВИЛА ОФОРМЛЕНИЯ ПРИСЫЛАЕМЫХ СТАТЕЙ291

CONTENTS



CONTENTS	7
FOREWORD OF EDITOR-IN-CHIEF	9
LATIN TODAY	
<i>Michaelis Shumilin</i>	
Saturarum liber imperfectus	12
ARTICLES	
<i>O.V. Alieva</i>	
Experience in measuring stylistic homogeneity by the Delta method on the material of the <i>Platonic Corpus</i>	19
<i>N.A. Filyanov</i>	
Testimony of Diodorus Siculus about the Italian embassy to Mithridates during the Social war: reliability and problems of interpretation	38
<i>Anna Grünert</i>	
The Medical Metaphor of the Hospital in the “Discourses” by Epictetus and the “Homilies on the Psalms” by Origen	51
<i>A.O. Denisov</i>	
Mapping ancient mythology in the Middle Ages	71
TRAVELS IN THE UNKNOWN ANTIQUITY	
<i>E.V. Prikhodko</i>	
Pydnae – the Hellenistic fortress on the borders of Xanthos. Part 3. Along the walls of the fortress	88
MISCELLANEA	
<i>M.V. Shumilin</i>	
Stat. Silu 4.2.7	167
PUBLICATIONS	
<i>Aelius Aristides</i>	
To those who criticize him because he does not declaim.	

Against those who burlesque the Mysteries (of Oratory) (Translation by S.I. Mezheritskaya)	170
<i>Giorgos Seferis</i>	
Ancient verses (Translation from modern Greek et commentary by O.P. Tsybenko)	191
REVIEWS	
<i>A.V. Podossinov</i>	
Book review: Dueck D. Illiterate Geography in Classical Athens and Rome (Routledge Monographs in Classical Studies). Routledge; London; New York, 2021, xi+265 p., illustrations. Cased. ISBN 978-0-367-43970-5	242
<i>W.K. Chrystaljow</i>	
Some critical notes on the Russian translation of the epigraphic law <i>de repetundis</i> (“Journal of Ancient History” 2007, № 4)	254
CHRONICLE	
<i>P.A. Evdokimov</i>	
International Digital Conference «Beyond Cyprus: Investigating Cypriot connectivity in the Mediterranean from the Late Bronze Age to the end of the Classical period» (zoom.us, December 8-11, 2020).	269
IN MEMORIAM	
<i>Lynette Mitchell (Exeter)</i>	
Peter John Rhodes (10th August 1940 – 27th October 2021)	287
ABBREVIATIONS	290
INFORMATION FOR AUTHORS	291



LECTORI BENEVOLO SALUTEM!

Дорогой читатель, несмотря на наступившее лихолетье, которое, я надеюсь, не продлится слишком долго, мы продолжаем работу над изданием нашего журнала, и этот том (уже 25-й!) получился, как мне кажется, столь же интересным и многогранным, как и предыдущие.

Открывается том латинскими стихотворениями, которые написал наш с Вами современник, московский филолог-классик Михаил Шумилин. Латинской прозой балуются многие, в том числе иногда и я (с моей латинской статьи начинался I том «Аристея»), но писать на древнем языке стихи в свойственной ему ритмической и стилистической манере – это уже высший класс владения языком, выше которого, как кажется, двигаться некуда. В нашей стране после Октябрьской революции таким уровнем владения языком обладало не больше пяти человек, в числе которых были великий петербургский филолог-классик Яков Боровский и наш молодой современник (тоже из Санкт-Петербурга) Денис Кейер, стихи которого мы публиковали в XV томе «Аристея». С гордостью замечу, что Шумилин – член редколлегии «Аристея», много делающий для сохранения высокого качества его публикаций.

Раздел «Статьи» содержит работы О.В. Алиевой (проводится количественное измерение стилистической неоднородности текстов на материале Платоновского корпуса), Н.А. Филянова (анализируется информация Диодора Сицилийского об италийском посольстве к Митридату VI Евпатору во время Союзнической войны 91–88 гг. до н. э.), А. Грюнерт (сопоставляются метафоры лечебницы в языческих философских текстах, в частности у Эпиктета, и Церкви как места духовного исцеления, в частности у Оригена), А.О. Денисова (исследуется античная мифология, представленная на средневековых картах). Кстати, два последних автора – аспиранты ВШЭ и МГУ соответственно.

В разделе «Путешествия в неизвестную античность» публикуется третья, завершающая, часть очерка Е.В. Приходько о Пидне – крепости на юго-западном побережье Ликии. В этот раз читатель совершит прогулку вдоль великолепно сохранившихся стен крепости, любуясь на прекрасные фотографии, сделанные самим автором. Должен отметить, что настоящий очерк, как и другие этого автора, выполнен на высоком профессиональном, исследовательском уровне.

В разделе «Miscellanea» мы снова встречаемся с Михаилом Шумилиным, который отстаивает достоверность рукописного чтения одного места в «Сильвах» римского поэта Стация перед лицом его современной критики.

Наш традиционный раздел «Публикации», в котором мы издаем новые переводы античных и антикизирующих авторов, представлен в этот раз переводом двух речей греческого автора II в. н. э. Элия Аристида, выполненным петербургским антиковедом, крупнейшим знатоком творчества Аристида С.И. Межеричкой. Большую подборку переводов стихотворений великого греческого поэта, лауреата Нобелевской премии в области поэзии Йоргоса Сефериса публикует известный переводчик античной и новогреческой литературы О.П. Цыбенко. Все стихотворения посвящены античности, которую Сеферис осмысляет и истолковывает как живую, имманентно присутствующую в нашей жизни реальность. Переводу предшествует большое Введение переводчика.

В разделе «Критика и библиография» публикуется рецензия А.В. Подосинова на книгу израильского классика-филолога Даниэлы Дуэк «Illiterate Geography in Classical Athens and Rome», в которой автор исследует, как, какими путями и в какой степени воспринимались географические сведения в необразованной части греческого и римского общества. Здесь же мы разместили критические замечания В.К. Хрусталёва на русский перевод с латинского языка эпиграфического закона о вымогательстве, выполненный Т.Г. Мякиным и опубликованный в «Вестнике древней истории» в 2007 г.

«Хроника» содержит подробный рассказ П.А. Евдокимова о конференции на Кипре, состоявшейся в 2020 г. и посвященной внешним связям Кипра в конце II – первой половине I тыс. до н. э.

В конце тома мы публикуем некролог, написанный специально для нашего журнала профессором университета в Экзетере Линетт Митчелл и посвященный памяти выдающегося британского историка античного мира Питера Джона Роудса (10 августа 1940 – 27 октября 2021 г.). Питер Роудс был большим другом русских антиковедов, несколько раз приезжал в Россию, бывал в Москве, Петербурге, Перми, читал здесь лекции, проводил летние школы и конференции совместно с российскими коллегами. Память о Питере Роудсе будет жить в сердцах многих людей, знавших этого замечательного, обаятельного человека и глубокого исследователя.

Недавно нас постигла еще одна горькая утрата. 19 февраля 2022 г. ушла из жизни Ольга Викторовна Смыка – выдающийся знаток античной и византийской литературы, талантливый переводчик греческой поэзии, великолепный преподаватель, воспитавший множество историков античного мира, обучая их древнегреческому языку на кафедре древних языков исторического факультета МГУ им. М.В. Ломоносова, на которой она проработала 50 (!) лет. Многие переводы византийской поэзии, выполненные Ольгой Викторовной, публиковались и в

нашем журнале. В предыдущем XXIV томе «Аристея» она поместила перевод первой части поэмы Иоанна Цеца, называвшейся «До Гомера», с обещанием в следующих томах опубликовать вторую часть («По Гомеру») и третью (После Гомера»). Неожиданная смерть не позволила ей закончить этот труд.

Все, кто был знаком с Ольгой Викторовной, дружил с ней или учился у нее, никогда не забудут этого удивительного человека, относившегося по-матерински к своим студентам, с большой любовью и заботой к коллегам, особенно молодым, которым нужна была помощь и поддержка, с неизменной улыбкой и радушием принимавшего всех окружающих ее людей. Я говорю об этом с полной уверенностью и правом, поскольку много лет проработал с Ольгой Викторовной на кафедре, которую я возглавляю и которая теперь осиротела без нее. Пусть земля ей будет пухом! Вечная память!



Ольга Викторовна Смыка
(24 июля 1945 – 19 февраля 2022)

Москва, март 2022 г.
Главный редактор журнала
А.В. Подосинов



Опыт измерения стилистической однородности методом Delta на материале Платоновского корпуса

О.В. Алиева

oalieva@hse.ru

Школа философии и культурологии НИУ ВШЭ, Москва, Россия

Аннотация: В статье рассматривается возможность количественного измерения стилистической неоднородности текстов на материале Платоновского корпуса с применением метода измерения стилистической разницы, известного как Delta Берроуза. Автор приходит к выводу, что использование для машинной классификации ограниченного числа авторских профилей, соотносимых с заведомо известными или предполагаемыми колебаниями авторского стиля, малоинформативно. Вместо этого предлагается использование шорт-листов на отрезках разной длины, когда для каждого блока испытуемого текста определяются его ближайшие соседи. При этом в работе с блоками в 3000 и 5000 слов подтвердилась установленная другими количественными методами принадлежность «Филеба» к «поздней» группе, что подтверждает эффективность Delta. Результаты, полученные на блоках меньшей длины, более подвержены случайным колебаниям частотности, однако они могут приниматься во внимание в тех случаях, когда есть дополнительные основания допускать стилистические сдвиги в определенных отрывках.

Ключевые слова: Delta, стилометрия, Платон, частотные слова, «Филеб»

Для цитирования: Алиева О.В. Опыт измерения стилистической однородности методом Delta на материале Платоновского корпуса. *Аристей* XXV (2022): 19–37.

Measuring Stylistic Homogeneity with Burrows' Delta: An Experiment with Corpus Platonicum

O.V. Alieva

oalieva@hse.ru

HSE University, Moscow, Russia

Abstract: This paper considers the possibility of quantitative measurement of Plato's style using Burrows' Delta. The author concludes that the result is not very informative if a limited number of author's profiles, each corresponding to the known or assumed fluctuations in the author's style, is used for machine classification. Instead, minimal Delta distances calculated for the whole corpus can be used for making short-lists for each segment of a test text. Thus, with blocks of 3000 and 5000, Delta confirms the traditional assignment of Plato's *Philebus* to the late group, listing among its closest neighbors the *Laws*, the *Sophist* and the *Statesman*. For smaller samples (1000 words), dialogues of the middle group make appearance in the short-lists for some blocks. Even if these results are more likely to be affected by random noise, they can be taken into consideration if there are additional reasons to suspect stylistic shifts in the given blocks.

Keywords: Delta, stylometry, Plato, most frequent words, *Philebus*

To cite this article: Alieva O.V. Experience in measuring stylistic homogeneity by the Delta method on the material of the Platonian Corpus. *Aristeas XXV* (2022): 19–37.

1. Постановка вопроса

Количественные методы исследования традиционно привлекаются в платоноведении с двумя целями: атрибуция (установление подлинности) и относительная датировка текстов¹. В большинстве случаев вопросы атрибуции и датировки решаются применительно к диалогам целиком, что не позволяет судить о стилистической гомогенности текстов. С другой стороны, сравнение небольших отрывков (500–1000 слов)² сталкивается с той трудностью, что частотные показатели в таких отрывках больше подвержены случайным колебаниям.

Установление минимальной длины отрывка, необходимой для корректной классификации, – первоочередная задача для исследователя, которого интересуют вопросы, связанные не только с атрибуцией и датировкой текстов це-

Я сердечно благодарю Б.В. Орехова за щедрое содействие и содержательную критику первых версий этой статьи, которая помогла существенно ее доработать. За возможные недостатки, разумеется, отвечаю только я. – О. А.

¹ См. Brandwood 1990; 1992.

² Ledger 1989; Tarrant 2010.

ликом, но и с их стилистической однородностью. Сама эта однородность или неоднородность, в свою очередь, может объясняться разными причинами: от соавторства до жанровых колебаний. В этом исследовании нас интересуют не столько причины таких колебаний, сколько возможности их количественного описания.

По сути своей сравнение авторских стилей в рамках одного корпуса не отличается от сравнения стилистических расхождений между различными авторами: в некоторых экспериментах ранние «Европейцы» Генри Джеймса и его же поздние «Послы» не опознаются как сочинения одного автора³. В интересах классификации в таких случаях имеет смысл задать несколько профилей для одного автора, и именно это мы попробуем сделать для Платона с привлечением новейшего метода измерения стилистической разницы, Delta Берроуза⁴, подтвердившего свою эффективность в многочисленных исследованиях⁵.

2. Метод Delta

Прежде всего следует пояснить статистический смысл метода⁶. Суть его заключается в том, что для корпуса текстов рассчитывается частотность ряда показателей; это могут быть слова (словоформы) или так называемые n-граммы, то есть последовательности n символов подряд. Для сравнения берутся самые частотные слова⁷, среди которых будет значительная доля служебных, в наименьшей степени связанных с тематикой текста (предлоги, союзы, частицы и т. п.). Поскольку сравниваемые тексты, как правило, имеют разную длину, в стилометрических исследованиях принято брать для сравнения относительную, а не абсолютную частотность; Берроуз идет еще дальше, предлагая использовать так называемые z-scores, то есть стандартизированные оценки⁸, показывающие разброс значений относительно средних. Z-score вычисляется по формуле: $Z = \frac{x - \mu}{sd}$, где случайная величина x – это значение частотности, μ – математическое ожидание (среднее), а sd – стандартное отклонение. Иными словами, z-score показывает, на сколько стандартных отклонений x отстоит от ожидаемого. Зная z-scores для заданных слов у известных авторов/текстов, можно сравнить их с z-scores спорного текста; искомая дистанция Delta вычисляется как сумма взятых по модулю разниц между z-scores у двух сравниваемых текстов, поделенная на количество слов: $\Delta_B = \frac{1}{n} \times \sum_i^n |z_{i,A} - z_{i,B}|$, где i – конкретное слово, n – общее число слов, а A и B – сравниваемые авторы

³ Hoover 2004b: 457.

⁴ Burrows 2002.

⁵ В том числе на русскоязычном материале: Orekhov 2020.

⁶ Хорошее объяснение см. в кн.: Savoy 2020: 34–39.

⁷ Впрочем, Rybicki, Eder 2011 показывают, что это необязательно для некоторых языков.

⁸ См. об этом Kenny 1982: 51–60.

(знак | указывает, что суммируется абсолютное значение разницы). Чем больше дистанция, тем менее вероятно авторство.

В очень простом примере⁹ на атрибуцию текста Платону или Исократу (хотя Delta позволяет выдвигать и большее число кандидатов) подробно показано, как можно рассчитать эту дистанцию в R с использованием пакета Stylo¹⁰. Даже на десяти словах, выбранных для удобства объяснения, «Менон» (условно «спорный» текст в этом примере) уверенно приписывается Платону. В большинстве успешных экспериментов количество слов (далее mfw) колеблется от нескольких десятков до нескольких сотен, причем результаты классификации могут меняться со сменой mfw или их числа. В частности, Д. Хувер рекомендует отбраковывать слова, слишком характерные для одного текста (процедура, получившая название culling), или личные местоимения, частотность которых чувствительна к жанру (драма или нарратив)¹¹.

Предлагалось множество модификаций классической Delta; так, квадратическая Delta Аргамона использует не манхэттенскую метрику, как у Берроуза, а евклидово расстояние¹², но ее эффективность ставилась под сомнение¹³. Хувер предлагал учитывать только z-scores выше определенного значения или z-scores с противоположными знаками¹⁴. Высокую эффективность показала Delta с использованием косинусного расстояния¹⁵, а также так называемая Delta Эдера¹⁶. Все перечисленные методы программно реализованы в пакете Stylo. Некоторые испытания этих методов проводились и на греческом корпусе¹⁷. Дальнейшие эксперименты опираются на классическую Delta.

3. Минимальная выборка

Как показывают многие эксперименты, важнейшим условием успешной классификации является длина сравниваемых отрывков. Для англоязычной прозы назывались цифры от 5000 до 2000 слов; для латинской – 2500 слов¹⁸. Для изучения стилистической однородности диалогов Платона отрывки в 5000 слов могут быть малоинформативны¹⁹: длина одной книги в «Государстве»

⁹ Очень простой пример расчета Delta в R (всего на 10 словах для удобства) в моем аккаунте RPub: https://rpubs.com/locusclassicus/delta_very_simple_example

¹⁰ Eder, Rybicki, Kestemont 2016.

¹¹ Hoover 2004b.

¹² Argamon 2008.

¹³ Jannidis et al. 2015.

¹⁴ Hoover 2004a.

¹⁵ Smith, Aldridge 2011; Evert et al. 2017.

¹⁶ Eder 2015b.

¹⁷ См. эксперимент с греческим эпосом у Eder 2011; Koentges 2020 на платоновском материале.

¹⁸ Eder 2015a и Eder 2017.

¹⁹ Впрочем, см. Tarrant 2011.

Платона – примерно 8000–9000 слов, а весь «Филеб» насчитывает около 17000 слов. Шаг в 5000 слов применим к «Роману о Розе» Гийома де Лорриса и Жана де Мёна, в котором почти 270 000 слов²⁰, но едва ли подходит для текстов меньшего объема. Здесь успех будет во многом обусловлен контекстом: «различить Хемингуэя и Диккенса всегда будет проще, чем сестер Бронте»²¹.

Функция `size.penalize` из того же пакета `Stylo` дает возможность проверить корректность классификации текстов разной длины средствами машинного обучения с учителем (доступные методы включают в себя `Delta`, метод опорных векторов, метод *k*-ближайших соседей, наивный байесовский классификатор и др.). Алгоритм извлекает из текстов случайные выборки все большей и большей длины и сравнивает их с обучающей выборкой для классификации; по умолчанию для каждой заданной длины отрывка проводится 100 итераций.

Прежде чем опробовать функцию на Платоновском корпусе, мы убедились, что высокий процент успешных классификаций невозможно получить случайно; для этого функция была опробована на трех исократых,²² созданных случайной склейкой случайных выборок из его речей.²³ Количество успешных атрибуций на отрывках разной длины нигде не превысило 39%, среднее значение 12%.

Также следовало убедиться, что высокий процент успешных классификаций невозможен, если в авторских профилях оказываются стилистически близкие тексты; поскольку априорно мы не можем быть уверены в гомогенности какого-либо текста или группы текстов, то такие исократы-близнецы были созданы искусственно путем склейки случайных выборок, но на этот раз из двух текстов – «Панафинейской речи» и «Об обмене имуществом»²⁴. Здесь максимальный успех составил 57%, среднее значение успешных атрибуций 36%.

4. Платон в профиль

Серия экспериментов²⁵ позволяет установить, что классификатор `Delta` уверенно (более 95% корректных атрибуций на отрывках всего в 1000 слов с использованием 100 `mfw`) отличает платона 1, автора «Протагора» и «Горгия»,

²⁰ Eder 2016.

²¹ Eder 2017.

²² Ниже я пишу с маленькой буквы имена, если они отсылают не к реальным авторам, а к их смоделированным профилям.

²³ По ссылке: https://rpubs.com/locusclassicus/fantom_isocrates

²⁴ По ссылке: https://rpubs.com/locusclassicus/twins_isocrates

²⁵ Все эксперименты проводились с использованием корпуса `Diorisis`: Vatri, McGillivray 2018. Хотя `Delta` допускает работу со словоформами, я пользовалась лексемами. Код для их извлечения доступен по ссылкам: https://rpubs.com/locusclassicus/all_plato_lemmata, <https://rpubs.com/locusclassicus/laws>, <https://rpubs.com/locusclassicus/republish>. Немногочисленные ошибки лемматизации, о которых см. Vatri and McGillivray 2020, по моим наблюдениям, последовательны и существенно не влияют на результат.

от платона 2, автора восьмой и девятой книг «Государства», и платона 3, автора восьмой и девятой книг «Законов» (Профиль 2). 1500 слов и 100 mfw достаточно для $\geq 98\%$ успешных атрибуций, если платон 1 – автор «Большого Гиппия» и «Евтифрона», платон 2 – автор второй и третьей книг «Государства», а платон 3 – автор шестой и одиннадцатой книг «Законов» (Профиль 1)²⁶.

Конечно, из этого не следует, что комбинации, дающие более стабильный результат, лучше представляют «типичные стили» Платона (или псевдо-Платона, учитывая спорное авторство «Гиппия Большого»²⁷ и «Евтифрона»²⁸). Строго говоря, мы не можем знать, что считать корректной классификацией, если два или три авторских профиля Платона собраны из диалогов, стилистическое сходство которых мы вынуждены постулировать либо на основании предыдущих стилеметрических исследований и собственной интуиции, либо на основании уже вычисленной заранее Delta. Последнее особенно интересно: сформировав профили на основе предварительно вычисленной дистанции Delta, мы получаем стабильно успешные результаты в тестах `size.penalize`²⁹. Это своего рода подтасовка.

5. «Филеб» и три платона

Ничто не мешает, однако, принять некоторые авторские профили в качестве моделей для сравнения. Но интерпретация машинной классификации в таком случае должна быть очень осторожной и учитывать конкретные расстояния. Мы провели эксперимент, сравнив с тремя платонами из Профилей 1–2 текст «Филеба», разделенный на отрывки в 1000 слов (выбор текста в данном случае был достаточно произвольным, поскольку нас интересовала стабильность классификации в принципе)³⁰. Результаты представлены в *табл. 1*.

Табл. 1. Сравнение отрывков «Филеба» (1000 слов) с разными профилями Платона; машинная классификация с учителем методом Delta

	P11=HiMa., Euthph. P12=R.2, 3 P13=Lg.6, 11	P11=Prt., Grg. P12=R.8, 9 P13=Lg.8, 9	P11=HiMa., Euthph. P12=R.2, 3 P13=Lg.1, 2	P11=Prt., Grg. P12=R.8, 9 P13=Lg.1, 2
Philebus_1	P12	P11	P13	P11
Philebus_2	P12	P11	P13	P13
Philebus_3	P13	P11	P13	P13
Philebus_4	P12	P11	P13	P13

²⁶ По ссылке: https://rpubs.com/locusclassicus/sample_size_1. https://rpubs.com/locusclassicus/three_platos.

²⁷ Nails 2019, 372.

²⁸ Solovyov 2020.

²⁹ По ссылке: https://rpubs.com/locusclassicus/three_platos

³⁰ По ссылке: https://rpubs.com/locusclassicus/philebus_1000

Philebus 5	P12	P12	P13	P13
Philebus 6	P11	P11	P13	P13
Philebus 7	P12	P13	P13	P13
Philebus 8	P12	P12	P13	P13
Philebus 9	P12	P12	P13	P13
Philebus 10	P12	P12	P13	P13
Philebus 11	P12	P11	P13	P11
Philebus 12	P12	P12	P13	P13
Philebus 13	P12	P12	P13	P13
Philebus 14	P12	P13	P13	P13
Philebus 15	P12	P11	P13	P13
Philebus 16	P12	P11	P13	P13
Philebus 17	P12	P11	P13	P13

Первые два сравнения (Профили 1 и 2) приводят к достаточно неожиданному результату, учитывая, что «Филеб» традиционно считается позднейшим диалогом Платона³¹. При сравнении с Профилем 1 почти все отрывки уходят платону 2, автору «Государства»; при сравнении с Профилем 2 автору «Государства» остается всего 6 блоков, а большая часть уходит платону 1, автору «Горгия» и «Протагора». На долю «позднего» платона 3, вопреки ожиданиям, не выпадает почти ничего. Однако классификация резко меняется, если мы подставляем в обоих профилях первые две книги «Законов» (Профили 3 и 4). (Изначально мы не стали этого делать, потому что в экспериментах с `size.penalize` в таких комбинациях возрастала путаница между платонами 2 и 3, особенно на отрывках в 500 слов³²).

Интуитивно результат в последних двух столбцах кажется более правдоподобным, однако это не значит, что первые две классификации ошибочны. Речь идет лишь о том, что дистанция между «Филебом» и «Государством» (или «Протагором» и «Горгием») в этом случае меньше, чем дистанция между «Филебом» и отдельными книгами «Законов». «Законы» сами по себе неоднородный текст, и некоторые исследователи считают, что вмешательства Филиппа Опунтского в этот диалог могли быть довольно значительны³³.

Все это подтверждает, что даже незначительные изменения в обучающей выборке могут приводить к существенному изменению результата³⁴. При использовании блоков длиной 5000 слов «Филеб» ведет себя похожим образом (табл. 2)³⁵. Из этого следует, что не только длина отрывка или набор `mfw` влияет на результаты сравнений, но и, прежде всего, изначально заданные авторские профили.

³¹ См. примеч. 1.

³² См., например, Additional tests 1&7 по ссылке https://rpubs.com/locusclassicus/sample_size_1.

³³ Nails, Thesleff 2003.

³⁴ Eder, Rybicki 2012.

³⁵ По ссылке: https://rpubs.com/locusclassicus/philebus_5000

Табл. 2. Сравнение отрывков «Филеба» (5000 слов) с разными профилями Платона; машинная классификация с учителем методом Delta (расстояние до победителя)

	P11=HiMa., Euthph. P12=R. 2, 3 P13=Lg. 6, 11	P11=Prt., Grg. P12=R. 8, 9 P13=Lg. 8, 9	P11=HiMa., Euthph. P12=R. 2, 3 P13=Lg. 1, 2	P11=Prt., Grg. P12=R. 8, 9 P13=Lg. 1, 2
Philebus_1	P12 1.056063 (R. 2) 1.020746 (R. 3)	P11 1.146171 (Grg.) 1.075359 (Prt.)	P13 1.0146754 (Lg. 1) 0.9209539 (Lg. 2)	P13 1.069868 (Lg. 1) 1.018049 (Lg. 2)
Philebus_2	P12 1.153919 (R. 2) 1.105528 (R. 3)	P12 1.191461 (R. 8) 1.194538 (R. 9)	P13 1.065575 (Lg. 1) 1.019719 (Lg. 2)	P13 1.170271 (Lg. 1) 1.157308 (Lg. 2)
Philebus_3	P12 1.0076794 (R. 2) 0.8882321 (R. 3)	P12 1.019441 (R. 8) 1.046299 (R. 9)	P13 0.9104088 (Lg. 1) 0.9071319 (Lg. 2)	P13 1.011570 (Lg. 1) 1.028162 (Lg. 2)

6. Закрытая игра

Если передать результаты измерения Delta функции, которая их визуализирует в виде дендрограммы (рис. 1), то можно заметить, что в Профиле 1 и 2 у всех платонов отдельный кластер, а три блока «Филеба» (5000 слов) сгруппированы отдельно (интересно, однако, что блок 1 достаточно удален от блоков 2 и 3, что видно на всех графиках). При добавлении книг 1–2 «Законов» первый блок «Филеба» (11a–26e) присоединяется к кластеру «Законов», а блоки 2 (26e–42d) и 3 (42d–59a) «Филеба» соединяются с этим общим кластером еще дальше. На всех четырех графиках они стоят особняком.

С одной стороны, такая группировка дает повод заподозрить, что все наши платоны подобраны недостаточно удачно. С другой стороны, как показывают дальнейшие тесты (см. ниже табл. 6 и 7), показатели Delta стабильно >1 у текстов разной длины (т. е. при сравнении небольших отрывков с текстами целиком). Учитывая ограничения закрытой игры с тремя кандидатами, мы попробуем сыграть в открытую игру и сравнить блоки «Филеба» со всеми диалогами в Платоновском корпусе как взятыми целиком, так и разбитыми на отрывки аналогичной длины.

7. Открытая игра

Процедура заключается в следующем: текст «Филеба» делится на отрывки 5000, 3000, 1000 и 500 слов. Отрывки сравниваются со всем корпусом, разбитым на такой же длины отрезки (диалоги меньше указанной длины исключаются). Для каждого сравнения составляется шорт-лист из 5 кандидатов.

Чтобы выполнить эту задачу, нам нужно для каждого сочетания загрузить и разбить на отрывки корпус, посчитать частотность слов для каждого отрывка и рассчитать Delta. На этом этапе функция для расчёта дистанции `dist.delta` воз-

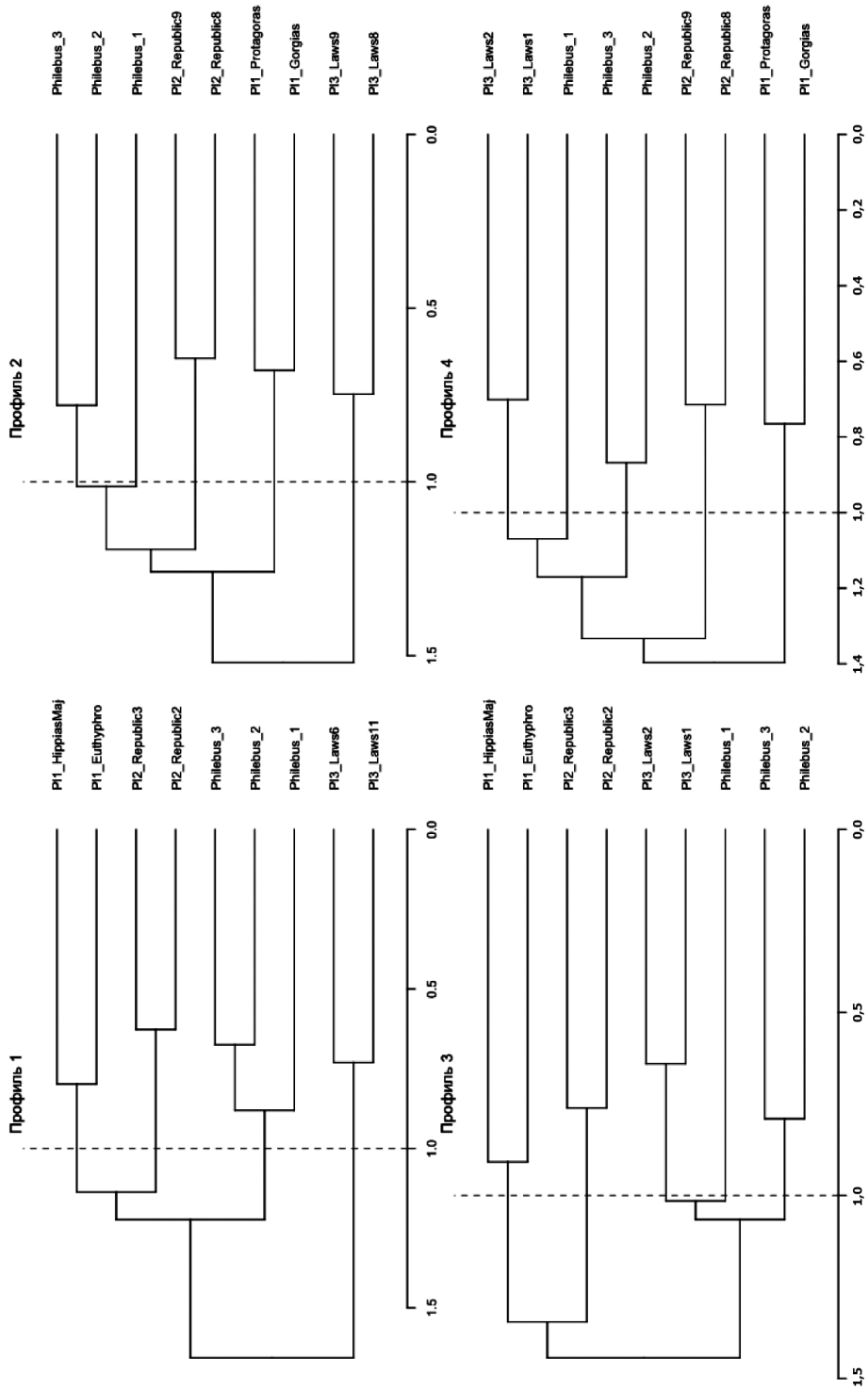


Рис. 1
Кластерный анализ на основе Delta между блоками «Филеба» (5000 слов) и авторскими профилями

вращает большую матрицу (в частности, для отрывков в 5000 слов это матрица размером 82×82 ; для отрывков в 3000 слов – 156×156 ; для отрывков в 1000 слов – 534×534). Визуализировать ее можно в виде круговой дендрограммы, но для анализа такой график может быть неудобен. Поэтому мы написали небольшую функцию, которая извлекает из матрицы столбцы, соответствующие испытуемому диалогу, для каждого столбца находит пять минимальных значений (в возрастающем порядке) и возвращает имена рядов, соответствующие этим значениям. В принципе, все то же самое можно проделать при помощи существующей функции `perform.delta` (применив к результату `$ranking`), однако ее особенность в том, что она рассматривает как «кандидата» лишь символы до знака подчеркивания, в то время как при разбиении диалогов на отрывки функция `make.samples` из пакета `Stylo` создает имена типа `Sophist_1`. Таким образом, она не подходит, если мы хотим знать не только ближайший диалог, но и номер отрывка из этого диалога, который вошел в шорт-лист.

Шорт-лист для отрывков в 5000 слов (табл. 3) в целом подтверждает традиционное отнесение «Филеба» к поздней группе. Обратим внимание на то, что из ближайших к «Филебу» книг «Законов» – книги 1–4 (на некотором удалении 7); все остальные книги, как видно на круговой дендрограмме³⁶, кластеризуются отдельно. Это подтверждает, что изначально платон 3 в Профилях 1 и 2 был выбран не очень удачно. «Софист» и «Политик» – ожидаемые соседи для диалога «поздней» группы, однако появление на близком расстоянии ($\Delta = 0,76$) одного из блоков «Теэтета» достаточно интересно. Обратим внимание на низкие показатели Δ , зачастую лишь незначительно отличающиеся для пяти кандидатов. Также заметим, что для второго блока «Филеба» Δ чуть выше.

Табл. 3. Шорт-лист для отрывков в 5000 слов (расстояние Δ);
здесь и далее серым выделены диалоги, не входящие в группу «Законы»,
«Софист» и «Политик»

	[1]	[2]	[3]	[4]	[5]
Philebus_1	Statesman_2 0.7363448	Statesman_1 0.7473091	Laws4_1 0.7647666	Theaetetus_1 0.7656733	Laws3_1 0.7799669
Philebus_2	Laws2_1 0.8063787	Sophist_2 0.8384637	Statesman_2 0.8694230	Statesman_1 0.8741261	Sophist_1 0.8820220
Philebus_3	Laws4_1 0.7366366	Laws2_1 0.7465235	Statesman_1 0.7494409	Laws3_1 0.7601559	Laws7_1 0.7606171

Шорт-лист для отрывков в 3000 слов (табл. 5)³⁷ представлен теми же диалогами: «Законы» (книги 1–4, 7 и 10), «Софист», «Политик». Вновь один блок

³⁶ По ссылке: https://rpubs.com/locusclassicus/philebus_shortlist_5000

³⁷ По ссылке: https://rpubs.com/locusclassicus/philebus_shortlist_3000

«Теэтета» представлен в списке ближайших соседей для первых 3000 слов «Филеба» (11а–20с, см. Приложение). Разница между кандидатами снова измеряется в сотых долях, что говорит об их близости, а наибольшие дистанции снова у третьего блока (он частично совпадает с блоком 2 в табл. 4).

Табл. 4. Шорт-лист для отрывков в 3000 слов (расстояние Δ)

	[1]	[2]	[3]	[4]	[5]
Philebus_1	Laws7_4 0.7211966	Sophist_4 0.7508173	Statesman_3 0.7570131	Statesman_1 0.7837819	Theaetetus_3 0.7913422
Philebus_2	Laws2_2 0.6721729	Laws1_2 0.7159967	Laws3_2 0.7172953	Laws4_1 0.7248398	Statesman_1 0.7255297
Philebus_3	Statesman_1 0.7819919	Laws10_2 0.7935629	Laws2_1 0.8015238	Sophist_2 0.8083551	Sophist_1 0.8387032
Philebus_4	Sophist_4 0.7032408	Laws3_2 0.7408127	Laws10_2 0.7410721	Statesman_1 0.7542996	Laws7_1 0.7592156
Philebus_5	Sophist_2 0.6696491	Statesman_3 0.7192088	Laws2_2 0.7270360	Statesman_4 0.7340047	Statesman_1 0.7486723

При сравнении отрывков в 1000 слов картина складывается гораздо более разнообразная (табл. 5)³⁸, однако в этом случае сложно сказать, насколько появление в шорт-листе таких неожиданных лидеров, как «Письма» или «Соперники» является следствием статистического шума. В целом, «поздняя группа» в лице «Законов», «Софиста» и «Политика» сохраняет преимущество, однако заметно присутствие неплатоновского материала («Письма», «Послезаконие», «Соперники»), а также диалогов, традиционно относимых к «средней» («Государство», «Кратил») или «ранней поздней» («Теэтет») группам.

Табл. 5. Шорт-лист для отрывков в 1000 слов (расстояние Δ);
серым выделены диалоги, не входящие в группу «Законы», «Софист»
и «Политик»

	[1]	[2]	[3]	[4]	[5]
Philebus_1	Sophist_11 0.6228654	Statesman_2 0.7177375	Laws7_10 0.7288833	Sophist_10 0.7315933	Laws9_3 0.7323748
Philebus_2	Epistles_4 0.7622403	Laws7_8 0.7628634	Laws7_11 0.7638279	Epinomis_5 0.7677371	Laws3_8 0.7730543
Philebus_3	Sophist_1 0.7711457	Sophist_11 0.7784318	Laws6_1 0.7889345	Laws1_7 0.7921384	Epistles_12 0.7976978
Philebus_4	Laws2_4 0.6311966	Sophist_11 0.6585850	Statesman_2 0.7072646	Laws7_11 0.7295352	Sophist_6 0.7346180
Philebus_5	Laws7_11 0.6174606	Laws4_2 0.6843917	Laws7_2 0.6850423	Laws7_8 0.6912151	Statesman_2 0.6949880

³⁸ По ссылке: https://rpubs.com/locusclassicus/philebus_shortlist_1000

Philebus_6	Laws4_2 0.7746729	Laws7_8 0.8040109	Sophist_9 0.8066185	Sophist_6 0.8124312	Statesman_1 0.8246621
Philebus_7	Laws3_5 0.6825926	Statesman_7 0.7041804	Statesman_8 0.7182453	Epinomis_5 0.7264019	Cratylus_14 0.7511094
Philebus_8	Sophist_6 0.6949492	Statesman_3 0.7753968	Sophist_2 0.7851942	Statesman_1 0.7892776	Statesman_16 0.8116827
Philebus_9	Laws7_11 0.8628637	Laws2_2 0.8689965	Laws3_7 0.8770885	Sophist_10 0.8929529	Laws2_1 0.9071359
Philebus_10	Laws2_7 0.8021291	Sophist_11 0.8130075	Statesman_16 0.8136990	Statesman_3 0.8358486	Laws1_7 0.8429030
Philebus_11	Sophist_9 0.7420126	Laws10_6 0.7602156	Laws2_2 0.7809632	Laws7_11 0.7883295	Laws9_3 0.7915022
Philebus_12	Laws2_5 0.7198274	Laws8_2 0.7212377	Laws8_3 0.7213431	Republic8_7 0.7253949	Statesman_8 0.7293098
Philebus_13	Sophist_6 0.6838848	Laws3_7 0.6931290	Statesman_8 0.6938999	Theaetetus_1 0.7107045	Republic1_1 0.7160107
Philebus_14	Sophist_6 0.7147558	Laws1_7 0.7395857	Statesman_8 0.7542160	Sophist_11 0.7655062	Statesman_16 0.7721228
Philebus_15	Lovers_2 0.6982947	Statesman_7 0.7023505	Laws7_8 0.7103996	Theaetetus_1 0.7117087	Laws4_2 0.7124543
Philebus_16	Statesman_1 0.7954937	Laws2_5 0.8101638	Laws2_1 0.8114642	Laws3_4 0.8188293	Laws3_3 0.8221680
Philebus_17	Laws3_7 0.6723772	Laws1_8 0.6772724	Laws7_11 0.6957418	Laws6_10 0.7166983	Laws6_1 0.7179309

Дальнейшее снижение длины отрывка до 500 слов³⁹ влечет за собой появление в шорт-листе новых диалогов из числа подложных («Алкивиад 2», «Феаг»). «Соперники» исчезают, однако блок 30 на том же месте оказывается близок «Письмам» и «Феагу». «Письма» оказываются первыми для блоков 24 и 30, а «Послезаконие» – для блока 3. Среди диалогов, считающихся подлинными, расширяют свое присутствие «Государство» (№ 1 для блоков 2, 12, 26) и «Кратил» (№ 1 и № 2 для блока 1), к которым добавляются «Федр» (№1 для блока 8), «Пир» (№1 для блока 13) и некоторые другие. Разумеется, влияние случайного шума на таких небольших выборках будет выше, так что относиться к этим результатам надо с большой осторожностью.

8. Важные оговорки

Поскольку Delta вычисляется с учетом среднего и стандартного отклонения, а оба этих показателя чувствительны к любым изменениям в корпусе, дистанция между двумя текстами может меняться в зависимости от того, в каком окружении они сравниваются. Статистически может иметь смысл уравновесить наш исходный корпус так, чтобы большие тексты вроде «Государства» и

³⁹ Мы не приводим таблицу в силу ее объема, однако результат доступен по ссылке: https://rpubs.com/locusclassicus/philbus_shortlist_500

«Законов» были представлены лишь частью книг, а особенно длинные диалоги (например, «Горгий») – лишь частью отрывков. Однако при этом мы будем вынуждены сделать некоторые допущения, которые, в свою очередь, тоже могут повлиять на результат. Поэтому мы решили оставить корпус в неизменном виде.

Точно так же, если мы решим сравнить *блоки* испытуемого диалога с диалогами *целиком*, результат может выглядеть иначе, поскольку в этом случае изменится вклад каждого блока в значение среднего и стандартного отклонения и потому что стилистические колебания в рамках одного текста потеряют значение. Так, весь корпус без испытуемого диалога составляет 55 элементов (считая отдельно книги «Государства» и «Законов»), а при разделении на блоки в 1000 слов – 534 элемента. Разделив «Филеб» на блоки в 1000 слов, мы получили 17 элементов. При расчете среднего и стандартного отклонения в комбинации 55+17 или 534+17 разница может оказаться существенной, а это, в свою очередь, влияет на z-scores. Впрочем, у нас нет оснований заведомо полагать, что влияет в худшую сторону, и такие вычисления можно провести хотя бы в порядке эксперимента.

С другой стороны, эта проблема решается путем расчета среднего и стандартного отклонения *без учета испытуемого текста* (при измерении Delta возможны оба варианта, и они предусмотрены создателями Stylo). Таким образом, мы имеем еще две возможности для сравнения: (1) «Филеб» vs. тексты корпуса целиком, z-scores с учетом всего; (2) «Филеб» vs. тексты целиком, z-scores рассчитываются только по корпусу. Эти возможности будут испытаны на отрывках в 1000 слов, поскольку именно на этом уровне классификация становится уже достаточно вариативной, но еще не такой «шумной», как с отрывками в 500 слов.

9. Отрывки 1000 слов vs. тексты целиком

Результаты, достигнутые при сравнении блоков «Филеба» (1000 слов) с диалогами корпуса целиком (z-scores с учетом всего), представлены в таблице 6.⁴⁰ *Dubia* и *spuria* полностью исчезают, «Тезтет» появляется в числе лауреатов для блоков 3 и 4 и неожиданно выходит на первое место в блоке 9 (что достаточно любопытно, т. к. в литературе предлагалось читать это место «Филеба» в свете «Тезтета»⁴¹). Сразу три места забирают диалоги ранней и средней (условно) группы при сравнении с блоком 11. Кое-где сохраняет позиции «Кратил». В основном же диалог остается уверенно «поздним».

⁴⁰ По ссылке: https://rpubs.com/locusclassicus/shortlist_1000_plus

⁴¹ Mooradian 1996.

Табл. 6. Шорт-лист для блоков «Филеба» (1000 слов)
при сравнении с диалогами целиком; z-scores с учетом всего
(серым выделены диалоги, не входящие в группу «Законы»,
«Софист» и «Политик»)

	[1]	[2]	[3]	[4]	[5]
Philebus_1	Sophist 0.8243880	Statesman 0.9461000	Laws7 0.9811575	Laws4 1.0102986	Protagoras 1.0130875
Philebus_2	Laws10 1.0002920	Phaedrus 1.0084294	Sophist 1.0155985	Statesman 1.0294057	Laws7 1.0341863
Philebus_3	Sophist 1.0240379	Statesman 1.0397440	Laws3 1.0993220	Theaetetus 1.1245411	Laws4 1.1360292
Philebus_4	Sophist 0.8629180	Laws2 0.9334550	Theaetetus 0.9529578	Laws7 0.9696327	Statesman 0.9743726
Philebus_5	Laws1 0.9138170	Laws4 0.9158632	Laws7 0.9166366	Statesman 0.9468616	Laws3 0.9492765
Philebus_6	Statesman 1.0699696	Laws1 1.0736744	Laws7 1.1046025	Laws2 1.1064011	Laws3 1.1230333
Philebus_7	Statesman 0.9363344	Laws2 0.9729947	Laws7 0.9917472	Laws3 0.9930119	Laws1 1.0053274
Philebus_8	Sophist 0.9905226	Statesman 1.0803036	Laws3 1.1102350	Laws2 1.1357096	Laws4 1.1979431
Philebus_9	Theaetetus 1.1151917	Sophist 1.1238661	Laws2 1.1418482	Laws1 1.1781119	Republic1 1.2011876
Philebus_10	Statesman 1.0673672	Laws3 1.1204329	Laws2 1.1521679	Sophist 1.1528030	Laws1 1.1681888
Philebus_11	Laws1 1.0585023	Gorgias 1.0605358	Laches 1.0712905	Republic6 1.0830221	Laws2 1.0843613
Philebus_12	Statesman 0.9047089	Laws7 0.9307059	Laws3 0.9850718	Laws4 0.9912397	Laws10 0.9955936
Philebus_13	Sophist 0.9146190	Statesman 0.9256685	Laws3 0.9499466	Laws2 0.9501581	Laws7 0.9939371
Philebus_14	Sophist 0.9225994	Statesman 0.9646321	Laws2 1.0253777	Laws3 1.0403143	Phaedrus 1.0483780
Philebus_15	Laws7 0.9130220	Statesman 0.9303398	Cratylus 0.9312650	Laws3 0.9319110	Sophist 0.9344285
Philebus_16	Laws2 1.0258327	Statesman 1.0716923	Laws1 1.0914840	Laws3 1.1055128	Cratylus 1.1317762
Philebus_17	Laws1 0.8216427	Laws3 0.8580213	Laws2 0.8655289	Laws7 0.8919713	Statesman 0.9251122

Вычисление z-scores без учета test set, то есть 17 блоков «Филеба», влечет за собой лишь незначительные изменения (табл. 7). «Протагор», «Федр» и (дважды) «Теэтет» остаются в числе ближайших к блокам 1–4 текстов соответственно. Блок 9 снова сближается с «Теэтетом» и «Государством». В числе ближайших к блоку 11 текстов остаются «Горгий» и «Лахет». В блоке 14 «Федра» сменяет «Теэтет» на последнем месте, он же добавляется в блоке 15 к «Кратилу». В блоке 16 «Кратил» остается на последнем месте. Из нового

можно заметить лишь появление «Послезакония» в шорт-листе для блока 7 и «Государства» в шорт-листе для блока 12.

Табл. 7. Шорт-лист для блоков «Филеба» (1000 слов) при сравнении с диалогами целиком; z-scores без учета «Филеба» (серым выделены диалоги, не входящие в группу «Законы», «Софист» и «Политик»)

	[1]	[2]	[3]	[4]	[5]
Philebus_1	Sophist 0.9676037	Statesman 1.122057	Laws7 1.170879	Laws4 1.195967	Protagoras 1.207110
Philebus_2	Sophist 1.2065061	Laws10 1.207241	Statesman 1.225184	Phaedrus 1.233250	Laws3 1.238732
Philebus_3	Sophist 1.2549495	Statesman 1.273373	Laws3 1.311981	Laws4 1.367779	Theaetetus 1.399472
Philebus_4	Sophist 1.0413330	Laws2 1.136596	Theaetetus 1.154685	Statesman 1.182364	Laws7 1.203936
Philebus_5	Laws4 1.1044829	Laws1 1.118566	Laws7 1.134280	Statesman 1.142611	Laws3 1.164133
Philebus_6	Laws1 1.3016101	Statesman 1.310024	Laws2 1.332382	Laws7 1.358906	Laws3 1.384327
Philebus_7	Statesman 1.1754862	Laws2 1.193079	Laws7 1.211349	Epinomis 1.238194	Laws3 1.239209
Philebus_8	Sophist 1.2443948	Statesman 1.373784	Laws3 1.384514	Laws2 1.418846	Laws4 1.497310
Philebus_9	Sophist 1.4273551	Theaetetus 1.431466	Laws2 1.434651	Laws1 1.495979	Republic1 1.522404
Philebus_10	Statesman 1.3215110	Laws3 1.406722	Sophist 1.439911	Laws2 1.441624	Laws1 1.456667
Philebus_11	Laws1 1.2624594	Laws2 1.285467	Gorgias 1.302802	Sophist 1.312611	Laches 1.316706
Philebus_12	Statesman 1.1075352	Laws7 1.140141	Republic9 1.189376	Laws10 1.197329	Laws4 1.203037
Philebus_13	Sophist 1.0875262	Statesman 1.106343	Laws3 1.124710	Laws2 1.136040	Laws4 1.181927
Philebus_14	Sophist 1.1322134	Statesman 1.180965	Laws2 1.254489	Laws3 1.266012	Theaetetus 1.280901
Philebus_15	Laws7 1.1119978	Cratylus 1.113894	Laws3 1.123937	Sophist 1.124668	Theaetetus 1.131898
Philebus_16	Laws2 1.3174663	Statesman 1.371384	Laws1 1.388362	Laws3 1.413595	Cratylus 1.434500
Philebus_17	Laws1 1.0080851	Laws3 1.052048	Laws2 1.056533	Laws7 1.109706	Statesman 1.12774

10. Делить или не делить?

В двух примерах выше (табл. 6 и 7) результаты существенно отличаются от тех, что были получены при делении корпуса на отрывки (табл. 5). Совсем неожиданным это назвать нельзя. В самом деле, при сохранении того же количества блоков (17) и того же количества лауреатов ($5 \times 17 = 85$) количество

претендентов на одно призовое место вырастает в разы: как уже говорилось, при делении всех диалогов на блоки в 1000 слов мы получаем 534 участника вместо 55, так что большая часть диалогов не проходит во второй тур. Кроме того, длинные диалоги вроде «Законов» оказываются представлены большим числом отрывков, так что у них чисто статистически больше шансов сформировать «парламентское большинство».

Здесь нам видятся два возможных решения. Во-первых, можно расширить число кандидатов в шорт-листе, включив туда отрывки со значением Δ ниже определенного (например, 0.8). Во-вторых, можно изменить саму систему представительства, представив каждый текст равным числом случайных выборок. Оба решения по-своему несовершенны: в первом случае релевантная дистанция назначается достаточно произвольно, во втором случае разные выборки могут приводить к разным результатам. Впрочем, на отрывках 5000⁴² (3 случайные выборки на текст) результаты во многом соответствуют полученным ранее на нормальных выборках (табл. 8; ср. табл. 3), что говорит о надежности результата.

Табл. 8. Шорт-лист для блоков «Филеба» (5000 слов)
при сравнении со случайными выборками того же объема (дистанция Δ)

	[1]	[2]	[3]	[4]	[5]
Philebus_1	Statesman 0.7416121	Statesman 0.7466936	Sophist 0.7585117	Sophist 0.7983998	Laws7 0.8002310
Philebus_2	Statesman 0.7935715	Laws1 0.8579932	Sophist 0.8606261	Sophist 0.8708460	Sophist 0.8819386
Philebus_3	Statesman 0.7384443	Statesman 0.7492319	Sophist 0.7628503	Laws1 0.7642040	Laws1 0.7678577

На отрывках в 3000 слов в числе кандидатов устойчиво присутствуют «Законы» 1–3 и 7, «Софист», «Политик» (ср. табл. 4 выше); а вот появление «Кратила», «Теэтета» и «Федра» носит скорее мерцающий характер; показатели дистанции сопоставимы с представленными в табл. 8⁴³. Вариативность на отрывках в 1000 слов становится больше⁴⁴, но уверенности в надежности таких классификаций намного меньше. Очевидно, что необходимы дальнейшие тесты на отрывках такой длины с привлечением дополнительного материала, в том числе за пределами Платоновского корпуса; этот вопрос будет рассмотрен отдельно.

⁴² По ссылке: <https://rpubs.com/locusclassicus/random5000>

⁴³ По ссылке: <https://rpubs.com/locusclassicus/random3000>

⁴⁴ По ссылке: <https://rpubs.com/locusclassicus/random1000>

11. Заключение

Если этот небольшой эксперимент позволяет сделать некоторые выводы, то первый из них заключается в том, что создание авторских профилей, соотносимых с заведомо известными и предполагаемыми колебаниями авторского стиля, при использовании Delta в случаях, подобных нашему, не имеет большого смысла. Результаты таких сравнений не будут совсем ошибочны, однако в ситуации «закрытой игры» они могут дать ложное представление о действительном расстоянии между текстами.

Гораздо более информативно использование шорт-листов на отрезках разной длины. При этом в работе с блоками в 3000 и 5000 слов (с использованием как нормальной, так и случайной выборки) подтвердилась установленная другими количественными методами принадлежность «Филеба» к «поздней» группе, что служит дополнительным доводом в пользу эффективности метода Delta. Результаты, полученные на блоках меньшей длины, более подвержены случайным колебаниям частотности, однако они могут приниматься во внимание в тех случаях, когда есть дополнительные основания допускать стилистические сдвиги в определенных отрывках.

В рамках данной работы мы не планируем касаться вопроса о таких основаниях применительно к «Филебу», поскольку наша задача заключалась лишь в том, чтобы проверить применимость Delta к исследованиям стилистической однородности на платоновском материале. Хотя предложенную процедуру можно еще дорабатывать, она, как представляется, вполне работоспособна.

Литература/References

- Argamon Sh. 2008: Interpreting Burrows's Delta: Geometric and Probabilistic Foundations. *Literary and Linguistic Computing* 23/2: 131–147.
- Brandwood L. 1990: *The Chronology of Plato's Dialogues*. Cambridge.
- Brandwood L. 1992: Stylometry and Chronology. In: Kraut R. (ed.). *The Cambridge Companion to Plato*. Cambridge. 90–120.
- Burrows J. 2002: Delta: A Measure of Stylistic Difference and a Guide to Likely Authorship. *Literary and Linguistic Computing* 17/3: 267–287.
- Eder M. 2011: Style-Markers in Authorship Attribution: A Cross-Language Study of the Authorial Fingerprint. *Studies in Polish Linguistics* 6/1: 99–114.
- Eder M. 2015a: Does Size Matter? Authorship Attribution, Small Samples, Big Problem. *Digital Scholarship in the Humanities* 30/2: 167–182.
- Eder M. 2015b: Taking Stylometry to the Limits: Benchmark Study on 5281 Texts from Patrologia Latina. In: *Digital Humanities 2015*. Sydney.
- Eder M. 2016: Rolling Stylometry. *Digital Scholarship in the Humanities* 31/3: 457–469.

Eder M. 2017: Short Samples in Authorship Attribution: A New Approach. In: *Digital Humanities 2017*. Montreal. <https://dh2017.adho.org/abstracts/341/341.pdf>.

Eder M., Rybicki J. 2012: Do Birds of a Feather Really Flock Together, or How to Choose Training Samples for Authorship Attribution. *Literary and Linguistic Computing* 28/2: 229–36.

Eder M., Rybicki J., Kestemont M. 2016: Stylometry with R: A Package for Computational Text Analysis. *The R Journal* 8/1: 107–121.

Evert S., Proisl Th., Jannidis F., Reger I., Pielström S., Schöch Ch., Vitt Th. 2017: Understanding and Explaining Delta Measures for Authorship Attribution. *Digital Scholarship in the Humanities* 32 (Suppl. 2): ii4–ii16.

Hoover D.L. 2004a: Delta Prime? *Literary and Linguistic Computing* 19/4: 477–495.

Hoover D.L. 2004b: Testing Burrows's Delta. *Literary and Linguistic Computing* 19/4: 453–475.

Jannidis F., Pielström S., Schöch Ch., Vitt Th. 2015: Improving Burrows' Delta. An Empirical Evaluation of Text Distance Measures. In: *Digital Humanities 2015*. Sydney.

Kenny A. 1982: *The Computation of Style: An Introduction to Statistics for Students of Literature and Humanities*. Oxford.

Koentges Th. 2020: The Un-Platonic Menexenus: A Stylometric Analysis with More Data. *Greek, Roman, and Byzantine Studies* 60/2: 211–241.

Ledger G.R. 1989: *Re-counting Plato: A Computer Analysis of Plato's Style*. Oxford.

Mooradian N. 1996: Converting Protarchus: Relativism and False Pleasures of Anticipation in Plato's Philebus. *Ancient Philosophy* 16/1: 93–112.

Nails D., Thesleff H. 2003. Early Academic Editing: Plato's Laws. In: Scolnicov S., Brisson L. (eds.). *Plato's Laws: From Theory into Practice*. Sankt Augustin. 14–29.

Orekhov B.V. 2020: «Илиада» Е.И. Кострова и «Илиада» А.И. Любжина: стилиметрический аспект [Iliad by Kostrov and Iliad by Lyubzhin: the Stylometry Case]. *Aristeas* 21: 282–296.
Орехов Б.В. «Илиада» Е.И. Кострова и «Илиада» А.И. Любжина: стилиметрический аспект. *Аристей* 21: 282–296.

Rybicki J., Eder M. 2011: Deeper Delta across Genres and Languages: Do We Really Need the Most Frequent Words? *Literary and Linguistic Computing* 26/3: 315–321.

Savoy J. 2020: *Machine Learning Methods for Stylometry: Authorship Attribution and Author Profiling*. Cham.

Smith P.W.H., Aldridge W. 2011: Improving Authorship Attribution: Optimizing Burrows' Delta Method. *Journal of Quantitative Linguistics* 18/1: 63–88.

Tarrant H. 2010: Some Support from Computational Stylistics. *Hermathena* 189: 93–101.

Tarrant H. 2011: A Six-Book Version of Plato's Republic: Same Text Divided Differently, or Early Version? In: Mackay A. (ed.). *ASCS 32 Selected Proceedings*. <http://ascs.org.au/news/ascs32/Tarrant.pdf>

Vatri A., McGillivray B. 2018: The Diorisis Ancient Greek Corpus: Linguistics and Literature. *Research Data Journal for the Humanities and Social Sciences* 3/1: 55–65.

Vatri A., McGillivray B. 2020: Lemmatization for Ancient Greek: An Experimental Assessment of the State of the Art. *Journal of Greek Linguistics* 20/2: 179–196.

Nails D. 2019: *Lyudi Platona: prosopografiya Platona i drugikh sokratikov [The People of Plato: A Prosopography of Plato and Other Socratics]*. Moscow.

Solovyov R.S. 2020: *Zakony Platona kak relevantnyy kontekst shkol'nogo dialoga Evtifron [Plato's Laws as a Relevant Context of the School Dialogue Euthyphro]*. *Bogoslovskiy vestnik [Theological Herald]* 3/38: 341–351.

Приложение

Соответствие блоков «Филеба» в 1000 слов пагинации по Стефану (для блоков в 3000 и 5000 слов пагинация может быть вычислена путем объединения блоков в 1000 слов).

Philebus_1 = 11a1–14c1

Philebus_2 = 14c2–17c7

Philebus_3 = 17c7–20c8

Philebus_4 = 20c8–23c11

Philebus_5 = 23c11–26e5

Philebus_6 = 26e5–29e8

Philebus_7 = 29e8–33a6

Philebus_8 = 33a7–36b8

Philebus_9 = 36b9–39c3

Philebus_10 = 39c3–42d4

Philebus_11 = 42d5–45d7

Philebus_12 = 45d7–49a1

Philebus_13 = 49a1–52b5

Philebus_14 = 52b5–55c8

Philebus_15 = 55c8–59a1

Philebus_16 = 59a1–62b3

Philebus_17 = 62b3–65c6

Информация об авторе

Ольга Валерьевна Алиева
кандидат филологических наук
доцент Школы философии
и культурологии НИУ ВШЭ
105066 Россия, Москва, Ст. Басманная
21/4
E-mail: oalieva@hse.ru

Information about the author

Olga V. Alieva
PhD (Cand. Sci. – Philology)
associate professor, School of Philosophy
and Cultural Studies, HSE University
105066 Russia, Moscow, St. Basmannaya
21/4
E-mail: oalieva@hse.ru



СОКРАЩЕНИЯ



ВДИ	Вестник древней истории. М.
НЛО	Новое литературное обозрение. М.
OLD	Oxford Latin Dictionary. Oxford, 1968–1982.

