

RESEARCH

ЛИТЕРАТУРНАЯ КРИМИНАЛИСТИКА: АВТОРСТВО ИЛИ ПЛАГИАТ?

Ю.А. Тамбовцев, Л.А. Тамбовцева, Ю.Ю. Тамбовцева
Новосибирск, Россия

Summary: It is possible to prove the authorship with the help of some methods of mathematical linguistics, e.g. that of «Chi-criterion». Based on the frequency of occurrence of simple prepositions in the text it was proved that Bahtin's texts were not written by Voloshinov or Medvedev. This result may stop the endless discussions of the authorship of the texts by Bahtin.

Под плагиатом подразумевается присвоение авторства произведения, которое может иметь место на уровне идей или конкретного текста. При этом нужно отметить, что плагиат на уровне идей определить трудно. В настоящее время достаточно хорошо определяется плагиат только на уровне конкретных слов [Tambovtsev et al. 2007].

Уголовный кодекс Российской Федерации определяет плагиат как нарушение авторских и смежных прав, см. статью 146 «Нарушение авторских и смежных прав»:

- «1. Присвоение авторства (плагиат), если это деяние причинило крупный ущерб автору или иному правообладателю, – наказывается штрафом в размере до двухсот тысяч рублей или в размере заработной платы или иного дохода осужденного за период до восемнадцати месяцев, либо обязательными работами на срок от ста восьмидесяти до двухсот сорока часов, либо арестом на срок от трех до шести месяцев.
2. Незаконное использование объектов авторского права или смежных прав, а равно приобретение, хранение, перевозка контрафактных экземпляров произведений или фонограмм в целях сбыта, совершенные в крупном размере, – наказываются штрафом в размере до двухсот тысяч рублей или в размере заработной платы или иного дохода осужденного за период до восемнадцати месяцев, либо обязательными работами на срок от ста восьмидесяти до двухсот сорока часов, либо лишением свободы на срок до двух лет.
3. Деяния, предусмотренные частью второй настоящей статьи, если они совершены:
 - группой лиц по предварительному сговору или организованной группой;
 - в особо крупном размере;

– лицом с использованием своего служебного положения, – наказываются лишением свободы на срок до шести лет со штрафом в размере до пятисот тысяч рублей или в размере заработной платы или иного дохода осужденного за период до трех лет либо без такового».

Целью работы является попытка определить степень схожести текстов М.М.Бахтина, В.Н.Волошинова и П.Н.Медведева по частоте употребительности некоторых лингвистических единиц. Выполняя эту задачу, мы сможем установить точную близость друг другу некоторых текстов и частей текстов одного и того же автора с одной стороны, и разных авторов – с другой стороны. Это позволит нам решить проблему плагиата в отношении текстов М.М.Бахтина. Некоторые литературоведы (В.М.Алпатов, С.Г.Бочаров, Н.Л.Васильев, В.В.Кожин, Н.А.Паньков, И.В.Пешков и др.) считают, что вопрос авторства Волошинова и Медведева еще раз следует перепроверить, т.к. они подозреваются в плагиате текстов Бахтина [Васильев 1991; БПМ 1993; Пешков 1995]. В то же время нельзя не согласиться с В.М.Алпатовым, Н.Л.Васильевым и И.В.Пешковым, которые заявили, что все «за» и «против» плагиата были уже высказаны несколько раз и требуется или новый материал, или какой-либо новый подход к этой проблеме, чтобы полностью прояснить ситуацию. Известно, что применение математических методов дает надежный результат при установлении авторства [Фукс 1975; Morton 1978; Мартыненко 1983; Holmes 1985; Фоменко 1994; Фоменко и др. 1996; Марусенко и др. 2001; Тамбовцев и др. 2008а, 2008б; Тамбовцев 2010].

В предыдущей работе мы рассматривали проблему плагиата, связанную с текстом «Почты духов», который был опубликован под именем И.А.Крылова. Мы рассматривали возможность написания текстов «Почты духов» как И.А.Крыловым, так и другими писателями его времени. В частности, мы проверили слухи о том, что большая часть текстов «Почты духов» была написана Радищевым. Анализ распределения некоторых лингвистических единиц показал, что Радищев не приложил руку к написанию писем из «Почты духов». В то же время мы доказали, что письма гномов и других волшебных существ были однозначно написаны И.А.Крыловым [Тамбовцев и др. 2008а: 92-95].

Для установления авторства можно привлекать разные грамматические признаки. Мы решили выбрать в качестве признаков функциональные, т.е. служебные, слова, в частности предлоги. В тексте предлоги образуют определенные спектры по частоте употребления. Для подсчетов достаточно взять не все предлоги, а только их часть. Нами вычислены расстояния между текстами с точки зрения типологии распределения в них этих выбранных лингвистических единиц.

В частности, на первом этапе нами задействована частота встречаемости некоторых простых предлогов как представителей служебных слов. При распознавании авторских текстов, на наш взгляд, следует анализировать, прежде всего, те слова, которые пишущий человек употребляет без активного участия сознания. Такими словами и являются служебные слова. На наш взгляд, если кто-либо попытается подделать стиль какого-либо писателя, то он станет, прежде всего, подделывать слова, выполняющие

главную номинативную функцию, т.е. существительные, прилагательные, глаголы, причастия или наречия. Служебные слова обычно ускользают от подделок в связи с тем, что находятся на периферии сознания и служат только лишь для грамматической связи слов. В то же время стиль писателя в большой степени характеризуют служебные слова, т.е. такие слова, которые служат для выражения отношений между знаменательными словами. Таким образом, служебные слова в значительной степени отражают мыслительные особенности того или иного писателя, так как они показывают индивидуальные и присущие только ему способы выражения отношений.

Необходимость исследования вызвана тем, что авторство некоторых произведений Волошинова и Медведева подвергается сомнению. К сожалению, до сих пор их анализ методами математической статистики не проводился, хотя именно такой анализ требуется, чтобы подтвердить или опровергнуть авторство Бахтина против авторства Волошинова или Медведева.

Таким образом, на наш взгляд, необходимо, наконец, поставить точку в этом затянувшемся споре и точно установить с помощью статистических критериев, писал ли эти тексты Бахтин, Волошинов или Медведев. Это можно сделать при помощи некоторых методов установления авторства с использованием математической статистики, как мы это делали при установлении авторства Крылова и Радищева [Тамбовцев и др. 2008а; 2008б].

Научная новизна работы заключается в том, что до сих пор лингвисты вели споры о плагиате текстов указанных выше авторов без применения статистических методов для установления авторства. В данной работе обсуждаются результаты компьютерной обработки текстов, которые находятся под подозрением в плагиате. С помощью компьютера были получены частоты встречаемости лингвистических единиц, после чего был применен критерий «хи-квадрат». В данной работе на основании встречаемости в тексте предлогов была впервые предпринята попытка установить типологическую схожесть текстов не только Бахтина, Волошинова, Медведева, но и других литературоведов и писателей.

Материал исследования

В качестве материала исследования были взяты многие тексты современных писателей и писателей-классиков на русском языке. Среди этих писателей такие, как М.Горький, С.Довлатов, Л.Кассиль, И.А.Крылов, М.Ю.Лермонтов, А.С.Пушкин, Л.Н.Толстой, А.И.Солженицын [Тамбовцев и др. 2008а; 2008б]. По характеру распределения лингвистических единиц этот материал полностью не изучался. Именно поэтому мы привлекаем обширный филологический материал. Это позволит более полно сравнить статистические характеристики распределения предлогов в текстах М.М.Бахтина и других литературоведов и писателей.

Состав выборки:

- Бахтин М.М. Творчество Франсуа Рабле и народная культура средневековья и Ренессанса. М., 1965 (1990). Объем выборки – 147.346 слов.

- Бахтин М.М. Формы времени и хронотопа в романе. Очерки по исторической поэтике // Бахтин М.М. Вопросы литературы и эстетики. М., 1975, с. 234-407. Объем выборки – 49.295 слов.
- Бахтин М.М. Проблемы поэтики Достоевского. М., 1963. Объем выборки – 61.565 слов.
- Волошинов В.Н. Марксизм и философия языка. Л., 1930. Объем выборки – 23.518 слов.
- Медведев П.Н. Формальный метод в литературоведении. М., 1928. Объем выборки – 10.142 слова.

Кроме художественной прозы, была привлечена политическая, историческая и научная проза.

Метод исследования

Нами проводились подсчеты частоты 25 предлогов по всем текстам. В связи с тем что такие служебные слова, как предлоги, встречаются в тексте достаточно часто, можно применить один из методов математической статистики, который называется «хи-квадрат». Этот критерий обычно применяется в лингвистических исследованиях для того, чтобы измерить однородность или схожесть двух и более текстов по каким-либо выбранным признакам [Тамбовцев 2003: 14-16]. Схожесть двух (или более) объектов вычисляется с помощью критерия «хи-квадрат» и тогда, когда мы не знаем статистическую форму распределения признаков, так как этот критерий может применяться не только к нормальному распределению, но и к другим видам распределений. Таким образом, при применении этого критерия можно не обращать внимания на вид распределения [Herdan 1966: 36-40]. Этот критерий довольно прост в применении и даёт надёжный результат. Среди прочего этот критерий можно применить в тех случаях, когда мы хотим установить, кто является автором текста.

Мы выбираем уровень значимости в 5%, который даёт 95% надёжности. Более детальное описание применения этого критерия можно найти в любой книге по применению статистических методов в лингвистике и других науках [Brownlee 1949: 50; Herdan 1966: 36-39; Brainerd 1974: 164-170; Butler 1985: 112-114; Тамбовцев 2003: 26-33].

Необходимо отметить, что мы несколько преобразовывали величину полученного «хи-квадрата», которую использовали в качестве расстояния между авторами. Вычисленную величину «хи-квадрата» мы делили на 3,841, чтобы найти величину коэффициента ТМВ, т.е. коэффициента Тамбовцева, при одной степени свободы. Коэффициент ТМВ – это отношение эмпирического значения «хи-квадрата» к табличному его значению при уровне значимости 5% и соответствующем числе степеней свободы. Величина коэффициента ТМВ вычисляется для того, чтобы в дальнейшем не быть ограниченным числом предлогов и других лингвистических единиц. Если мы в будущем возьмем большее количество предлогов и вычислим коэффициент ТМВ, то мы сможем прямо сравнивать их значения. Это тем более важно, что в данном случае коэффициент ТМВ показывает расстояние между авторами. Так, например, из Таб. 4 видно, что расстояние между «Почтой духов» и другой прозой Крылова наименьшее (0,64), а рас-

стояние между «Почтой духов» Крылова и политической прозой М.С.Горбачева – наибольшее (2,70). Расстояние между авторами обратно величине схожести между ними. Следовательно, чем меньше расстояние, тем больше схожесть.

Приведем простой пример определения схожести или различия авторов по частоте встречаемости предлогов. Частота встречаемости предлога «на» в рассказах Крылова составила 117 на 10.000 слов. В письмах сильфов из «Почты духов» Крылова предлог «на» встретился 83 раза на 10.000 слов. Вначале вычислим величину критерия «хи-квадрат». $83 + 117 = 200$. $117 - 83 = 34$. $34 \times 34 = 1156$. $1156 : 200 = 5,78$. $TMB = 5,78 : 3,841 = 1,51$. Так как величина TMB больше единицы, то употребление предлога «на» статистически различно. Можно предположить, что предлог «на» в письмах сильфов описывает пространственную ориентацию совершенно не так, как этот же предлог в рассказах Крылова. Еще больше различается по предлогу «на» (90) «Почта духов» Крылова и «Путешествие из Петербурга в Москву» Радищева (152): $TMB = 4,13$. Это также говорит о различном пространственном ориентировании мышления Крылова и Радищева.

Возьмем частоту употребления предлога «на» в книге Фадеева «Разгром» (129) и в книге Шолохова «Они сражались за Родину» (165) на 10.000 слов. Нужно проверить, является ли расхождение между частотой 129 и 165 статистически значимым. В данном случае значение «хи-квадрат» составит 4,41. Табличное значение при одной степени свободы и 5% уровне значимости составляет 3,841. Таким образом, делим 4,41 на 3,841 и получаем 1,15. Следовательно, коэффициент $TMB = 1,15$. Он превышает единицу. Из этого следует, что Фадеев и Шолохов в этих своих произведениях употребляют предлог «на» по-разному. Если сравнить «Тихий Дон» (книгу вторую) Шолохова по предлогу «на» и его же «Поднятую целину» по выборкам 1-12, то мы получим следующую частоту встречаемости: 164 и 169. «Хи-квадрат» равен 0,08, TMB равен 0,02. Это значительно меньше единицы. Из этого следует, что Шолохов употребляет предлог «на» в этих книгах совершенно одинаково.

Описанная процедура подсчета производится по каждому из взятых нами 25 предлогов. Далее значения «хи-квадрата» складываются. Степень свободы равна $25 - 1 = 24$. Табличное значение «хи-квадрата» при 5% уровне значимости и 24 степенях свободы равно 36,415. Эмпирическое значение суммы «хи-квадрата» получилось 62,63, $TMB = 1,72$. Из этого следует, что текст «Почты духов» Крылова и текст «Путешествия из Петербурга в Москву» Радищева по употреблению 25 предлогов сильно различаются. Таким образом, пространственная ориентация этих двух текстов различна. Разные писатели обычно показывают различную пространственную ориентацию в своих текстах. В то же время пространственная ориентация одного и того же писателя одинакова, что видно из всех приведенных далее таблиц.

Литературная криминалистика

Говоря о различии употребления предлогов в разных текстах, мы фактически вступаем в область литературной криминалистики. Мы пытаемся установить на основе некоторой совокупности объективно постигаемых

свойств какого-либо предмета, в нашем случае текста литературного произведения, его схожесть с другими объектами. В математической лингвистике идентификация литературного объекта происходит по частоте выбранных признаков, присущих данному объекту [Пиотровский и др. 1977: 18-21]. Мы придерживаемся взглядов тех лингвистов, которые рассматривают конкретное художественное произведение как нечто уникальное и неповторимое. Оно имеет особую компоненту внутреннего строения, которая проявляется через различные признаки. Нельзя не согласиться с Г.Я.Мартыненко, который рекомендует исследователям текста брать те структурные признаки, которые отражают закономерности структуры текста и имеют значительную силу, различающую стили [Мартыненко 1983: 58]. При этом, на наш взгляд, следует обратить внимание на то, что эта же сила объединяет некоторые произведения по стилю, поскольку величина признаков будет схожей. Г.Я.Мартыненко построил классификацию произведений русских писателей на основе синтаксических признаков и вывел их типологическую схожесть [Мартыненко 1983: 58-72].

Частота встречаемости определенных предлогов влияет на структуру стиля некоторого текста, которая частично определяется его содержанием, частично же – его формальными свойствами. Количественные характеристики тех или иных свойств произведения (например, частоту встречаемости предлогов) мы будем вслед за В.Фуксом называть *стилевыми* характеристиками [Фукс 1975: 313].

В процессе нашего анализа мы будем обращать внимание на *формальную структуру* текста, но не на его смысловое содержание. Частота встречаемости предлогов или других единиц текста отражает, прежде всего, его формальную структуру. Конечно, нельзя не согласиться с замечанием В.Фука, что здесь всегда имеют место определенные формальные различия: для эпического описания автор выбирает иной язык, чем для драматически заостренного диалога. Научные статьи пишутся в ином стиле, чем поэтическая проза [Фукс 1975: 386-387]. Учитывая это, уже заранее ясно, что научный текст и художественная проза проявят различные количественные характеристики. Более целесообразно сравнивать художественную прозу с художественной прозой.

Резкие переломы стиля наблюдаются часто в произведениях разных писателей. В то же время по коэффициенту ТМВ (как мы увидим далее) можно определить эти переходы даже в пределах одного и того же произведения. Определяя стилевые характеристики, нельзя оставлять такие места без внимания. Каждый раз приходится проверять, являются ли рассматриваемые нами тексты однородными. В нашем случае их можно, скорее всего, назвать приблизительно однородными [Тамбовцев 2010]. Мы покажем, как можно производить такую проверку по коэффициенту ТМВ.

Сильные различия в числовых значениях стилевых характеристик, как верно замечает В.Фукс, можно обнаружить в тех случаях, когда исследуются произведения, относительно которых заранее известно, что их стиль очень неоднороден. В этих случаях нам помогают ориентироваться литературоведческие науки, методы и результаты которых дополняют и уточняют (но, конечно, никоим образом не подменяют) строгий анализ. Мно-

гие еще и сегодня ломают голову над знаменитой проблемой, кто же все-таки написал драмы Шекспира или некоторые анонимные произведения [Тамбовцев 2008].

Христианские богословы и языковеды в огромном количестве и с достойной удивления проницательностью проводили работы над библейскими текстами и, прежде всего, над «Новым заветом». При этом использовались все известные научные методы, разработанные и применяющиеся для исследования исторических текстов. Дискутировался и вопрос о том, написаны ли «Деяния апостолов» и «Евангелие от Луки» одним и тем же автором [Kenny 1986; Ledger 1989].

Установление авторства имеет много общего с криминалистикой. Опишем применение этого метода. Известно деяние (текст), ищется лицо, совершившее это деяние (автор). Эта задача имеет две стороны. Во-первых, нужно исключить всех ложно подозреваемых лиц. Во-вторых, нужно обнаружить истинного виновника. В большинстве случаев исключить несправедливо подозреваемого легче, чем изобличить действительного виновника. В частности, по частоте встречаемости выбранных предлогов мы можем обнаружить, достаточно ли схожи два произведения, чтобы входить в одну генеральную совокупность [Тамбовцев 2003]. Для этого достаточно установления определенного точного порога, который покажет принадлежность двух текстов одной генеральной совокупности. В нашем случае этим порогом является граничное значение величины «хи-квадрат», которое приводится в таблицах. Мы обычно, как это было указано выше, берем 95% надежность, т.е. 5% уровень значимости [Тамбовцев 2003].

Определение списка подозреваемых в плагиате авторов является делом литературоведов. При отсутствии прямых показателей, слухов или намеков на плагиат это является гораздо более трудным делом для специалистов по математической лингвистике. Фактически все, что могут сделать методы математической статистики, – это показать, что два текста настолько близки по выбранным признакам, что входят в одну генеральную совокупность, которую можно понимать как один и тот же текст.

Существенная трудность при исследовании подлинности литературных произведений заключается в том, что в текстах не так легко найти нечто подобное отпечатку пальца в криминалистике. Стилиевые характеристики, претендующие на эту роль, должны удовлетворять следующим трем условиям: во-первых, они на протяжении всей жизни автора должны сохраняться столь же неизменными, как дактилоскопический узор пальца человека; во-вторых, этот литературный «отпечаток пальца» должен однозначно выделять каждого автора среди всех остальных; в-третьих, он должен быть легко опознаваем по форме. После многих исследований, проведенных над целыми комплексами описывающих тексты величин с привлечением текстов множества авторов, сложилось убеждение, что литературного «отпечатка пальца», который удовлетворял бы всем этим трем условиям и был бы универсальным для всех авторов, не существует. Высказывалось и противоположное утверждение, в частности авторами, занимающимися статистикой речи. Однако убедительных доказательств этой точки зрения их работы не содержали. Так, например, выражалось убеждение,

что устойчивые показатели своеобразия авторского стиля можно получить путем статистического анализа распределения предлогов, союзов, частиц, междометий, общих выражений, т.е. тех составляющих речи, которым автор не уделяет особого внимания [Morton 1978]. Но и здесь наши три необходимых условия не выполняются. Выводы, сделанные Грейстоном и Херданом относительно пастырских посланий на основании анализа отношения логарифмов словарного состава и общего количества слов, не являются убедительными. Установление «виновника» оказывается нелегкой задачей и при применении математических методов к анализу стиля. Как в том, так и в другом случае гораздо легче решается задача исключения «ложно подозреваемого» [Фукс 1975: 388-389].

Как известно, к служебным словам относятся предлоги, союзы, частицы и связки [Розенталь и др. 1985: 388-389]. На первом этапе исследования в качестве признака была взята частота встречаемости предлогов. Из всех предлогов исследовались только простые и первообразные предлоги, считающиеся древнейшими по происхождению. Нами были взяты следующие предлоги: *в, к, с, о, у, за, на, от, из, до, ко, по, со, перед, над, под, для, про, об, обо, без, через, ради* и др. Следует подчеркнуть, что, исследуя структуру частоты встречаемости предлогов в тексте, мы тем самым изучаем структуру отношений явлений, признаков и процессов, которые отражены в тексте.

Различные предлоги показывают различное местоположение или направление в пространстве и времени. Так, например, предлог *в* показывает нахождение внутри чего-то или движение вовнутрь. Предлог *про* употребляется при указании на лицо, явление или предмет, на который направлена мысль, речь или действие. По значению он близок предлогу *о* и его варианту *об*, поэтому здесь интересно рассмотреть их разную статистическую дистрибуцию. Предлог *через* употребляется при обозначении какого-либо пространства, среды, предмета или места, с одной стороны которого на другую осуществляется движение. Нельзя не согласиться с И.Фоменко в том, что этот предлог выражает преодоление препятствий и трудностей [Фоменко 1994: 76]. Этот предлог так же хорошо, как и другие предлоги, маркирует стиль писателя, что позволяет отделить один текст от другого. Значения предлогов взяты нами из Словаря русского языка [СРЯ 1988].

Таким образом, частоту встречаемости предлогов следует признать универсальным и информативным признаком. В дальнейшем необходимо установить: 1) степень схожести частоты встречаемости предлогов в одном и том же тексте (произведении) у одного и того же автора; 2) степень схожести встречаемости предлогов в разных текстах (произведениях) у одного и того же автора; 3) степень схожести встречаемости предлогов в разных текстах у разных авторов.

На наш взгляд, степень схожести будет тем больше, чем ближе величины частоты встречаемости каждого из предлогов. Вся совокупность предлогов представляет собой некоторую сеть (решетку), в узлах которых отображаются величины, получаемые в результате порождения текста как следствия мыслительной деятельности писателя. Можно предположить, что каждый человек имеет свою индивидуальную систему порождения

текста. Выбранные нами признаки будут отражать эту индивидуальную систему в некоторых её характеристиках.

Для проверки схожести текста в произведениях одного и того же автора, выражаемой предлогами, мы взяли тексты М.М.Бахтина и других авторов.

Обсуждение результатов исследования

Важно отметить, что мы будем проверять тексты на вхождение в одну и ту же генеральную совокупность. Из всех текстов М.М.Бахтина прежде всего рассмотрим текст его работы «Творчество Франсуа Рабле и народная культура средневековья и Ренессанса». Если этот текст и другие тексты входят в одну генеральную совокупность по критерию «хи-квадрат», то это означает, что они все очень схожи между собой. Как видно из Таб. 1, все величины коэффициента ТМВ различны. Как уже было сказано выше, величина коэффициента ТМВ – это отношение полученного (эмпирического) значения «хи-квадрат» к его теоретическому значению. Если теоретическое значение «хи-квадрат» больше полученного (эмпирического), то данные выборки принадлежат одной генеральной совокупности.

Таб. 1

Упорядоченный ряд расстояний по ТМВ между текстом М.М.Бахтина «Творчество Франсуа Рабле и народная культура средневековья и Ренессанса» и другими произведениями этого же и других авторов:

- 1) Бахтин М.М. Формы времени и хронотопа (0,41);
- 2) Бахтин М.М. Творчество Достоевского (0,52);
- 3) Сталин И.В. Марксизм и вопросы языкознания (1,14);
- 4) Медведев П.Н. Формальный метод в литературоведении (1,18);
- 5) Гинзбург Л.Я. О психологической прозе (1,30);
- 6) Волошинов В.Н. Статьи (1,37);
- 7) Волошинов В.Н. Марксизм и философия языка (1,42);
- 8) Карамзин Н.М. История государства Российского (1,56);
- 9) Кассиль Л.А. Дело вкуса. Эстетика (1,70);
- 10) Радищев А.Н. Путешествие из Петербурга в Москву (1,97);
- 11) Горбачев М.С. Как это было (2,31);
- 12) Солженицын А.И. Бодался теленок с дубом (2,50);
- 13) Радищев А.Н. О человеке, о его смертности и бессмертии (2,53);
- 14) Булгаков М.А. Мастер и Маргарита (2,63);
- 15) Кассиль Л.А. Будьте готовы, Ваше высочество (2,76);
- 16) Достоевский Ф.М. Идиот (2,90);
- 17) Тургенев И.С. Новь (3,02);
- 18) Чехов А.П. Рассказы (3,07).

В данном случае таблица демонстрирует, что тексты М.М.Бахтина «Творчество Франсуа Рабле и народная культура средневековья и Ренессанса» и «Формы времени и хронотопа» – это фактически один и тот же текст с точки зрения математической статистики, так как ТМВ меньше единицы (ТМВ = 0,41). Следующим по этому коэффициенту идет текст М.М.Бахтина «Творчество Достоевского» (ТМВ = 0,52). Здесь величина коэффициента ТМВ тоже меньше единицы. Следовательно, текст М.М.Бахтина «Творчество Франсуа Рабле и народная культура средневековья и

Ренессанса» и текст Бахтина «Творчество Достоевского» тоже весьма похожи статистически, т.к. эти тексты принадлежат одной генеральной совокупности.

Этот же критерий может показать, что два произведения статистически различаются. Мы покажем это на примере произведений В.Н.Волошинова и П.Н.Медведева, которые находятся под подозрением в плагиате. Взято два текста Волошинова и один текст Медведева: Медведев «Формальный метод в литературоведении», Волошинов «Статьи» (без его работы «Марксизм и философия языка»), ТМВ = 1,37; Волошинов «Марксизм и философия языка», ТМВ = 1,42. Величины ТМВ показывают, что тексты Медведева и Волошинова не принадлежат одной и той же генеральной совокупности. Это означает, что тексты написаны разными авторами.

Несколько более похож на текст Бахтина «Творчество Франсуа Рабле и народная культура средневековья и Ренессанса» текст Медведева «Формальный метод в литературоведении» – ТМВ = 1,18. В то же время величина ТМВ больше единицы. Это говорит о том, что это статистически разные тексты. Следовательно, по данному параметру плагиата ни со стороны Волошинова, ни со стороны Медведева нет.

Данный метод установления статистической близости текстов работает на текстах и других авторов. Так, возьмем для примера текст Радищева «Путешествие из Петербурга в Москву» и текст «Почты духов», изданный Крыловым. После опубликования второго из названных текстов в русском обществе ходили слухи, что Радищев написал за Крылова текст «Почты духов». Авторство текста «Путешествие из Петербурга в Москву» никогда не подвергалось сомнению. Возьмем его в качестве эталона.

Наши расчеты показали, что текст «Путешествия...» имеет значительное отличие от текста «Почты духов» – его ТМВ составил 1,72. Еще меньше на текст «Почты духов» по распределению предлогов похож текст Радищева «О человеке» (ТМВ = 2,77). С надежностью 95% следует говорить о том, что Радищев не писал текста «Почты духов» [Тамбовцев и др. 2008а; 2008б].

Проанализируем расстояния по коэффициенту ТМВ между текстом Волошинова «Марксизм и философия языка» и текстами других авторов, прежде всего П.Н.Медведева.

Таб. 2

Расстояния по коэффициенту ТМВ между текстом Волошинова «Марксизм и философия языка» и текстами других авторов:

- 1) Волошинов В.Н. Статьи (без текста «Марксизм и философия языка») (0,87);
- 2) Медведев П.Н. Формальный метод в литературоведении (0,89);
- 3) Бахтин М.М. Творчество Достоевского (1,14);
- 4) Сталин И.В. Марксизм и вопросы языкознания (1,23);
- 5) Гинзбург Л.Я. О психологической прозе (1,34);

- 6) Антропология. Сборник текстов многих авторов (1,35)¹;
- 7) Бахтин М.М. Творчество Франсуа Рабле (1,42);
- 8) Педагогика. Сборник текстов многих авторов (1,44);
- 9) Виноградов В.В. Проблемы авторства и теории стилей (1,59);
- 10) Бахтин М.М. Формы времени и хронотопа (1,84);
- 11) Успенский Б.А. Поэтика (2,03);
- 12) Чернышевский Н.Г. Эстетические отношения (2,09);
- 13) Кассиль Л.А. Дело вкуса. Эстетика (3,08);
- 14) Горбачев М.С. Как это было (3,59);
- 15) Жолковский А.К. Михаил Зощенко (3,61);
- 16) Кассиль Л.А. Будьте готовы, Ваше высочество (5,24);
- 17) Довлатов С.Д. Зона (5,31).

Величина ТМВ (0,87) указывает на то, что между текстом Волошинова «Марксизм и философия языка» и текстом других его статей нет статистического различия, из чего следует, что Волошинов сам писал эти два текста.

Настораживает то, что схожесть между текстом Волошинова «Марксизм и философия языка» и текстом Медведева «Формальный метод в литературоведении» значительная (ТМВ = 0,89). Это говорит о том, что оба текста каким-то образом повлияли друг на друга или же имеет место плагиат. В то же время текст Волошинова «Марксизм и философия языка» достаточно сильно отличается от текстов других авторов. Если взять тексты других статей Волошинова, то необходимо отметить их значительное отличие от текста Медведева (ТМВ = 1,18). Это говорит о том, что плагиата не наблюдается.

Для того чтобы развеять сомнения в том, что в текстах одного и того же писателя предлоги встречаются примерно одинаково, т.е. однородно, необходимо взять несколько писателей с несколькими прозаическими произведениями. Мы проанализировали в этом плане нескольких писателей, как классиков, так и современников. В группу писателей-классиков вошли Пушкин, Гончаров, Достоевский, (Л.Н.)Толстой, Тургенев. В группу современных писателей были включены Акунин, Емец, Довлатов, Катаев, Фадеев и др. Результаты отражены в Таб. 10. Расстояние между «Капитанской дочкой» Пушкина минимально (0,53) до другого произведения Пушкина «Повести Белкина». Как можно увидеть из всех других таблиц, это достаточно большая схожесть. В тоже время из упорядоченного ряда расстояний по возрастанию видно, что самым близким Пушкину по упо-

¹ Напр., следующие работы: 1) Вереш П.Т. Общественные и этнографические группы манси в XVII-XIX вв. // Congress Quartus Internationalis Fenno-Ugristarum. Budapestini Habitus, 9-15 Septembris 1975, Pars V. Budapest: АК, 1983; 2) Марк К. Антропология прибалтийско-финских народов. Таллин, 1975; 3) Хить Г.Л. Дерматоглифика народов СССР. М., 1983; 4) Бубрих Д.В. Историческое прошлое карельского народа в свете лингвистических данных // Известия Карело-финской научно-исследовательской базы Академии наук СССР, № 3 (1948). Тексты не издавались отдельной книгой и были собраны автором в виде файлов. То же касается упоминаемого под номером 8 сборника текстов по педагогике.

треблению предлогов является А.И.Солженицын (1,02), а самым далеким – политический текст М.С.Горбачева (5,18). Расстояния между Пушкиным и другими писателями показывают схожесть стилей по употреблению предлогов (см. Таб. 3).

Расстояния (т.е. величины коэффициента ТМВ) между текстом романа Л.Н.Толстого «Война и мир» (т. 1) и другими произведениями Л.Н.Толстого подтверждают то, что все они входят в одну генеральную совокупность (см. Таб. 11). Самым близким к этому тексту является текст второго тома романа «Война и мир» (0,26). Дальше всех, как это было и с другими писателями, стоит политический текст М.С.Горбачева (5,40).

Данные Таб. 13 показывают, что расстояния у текстов одного и того же писателя могут быть еще меньше. Так, расстояние между текстом Тургенева в романах «Накануне» и «Новь» составило всего 0,09. Еще меньшее расстояние (Таб. 14) демонстрируют тексты Достоевского «Идиот» и «Бесы» – 0,08. Это говорит о большом сходстве текстов по употреблению предлогов.

Очень близко друг к другу стоят произведения Д.Емца. Так, его «Таня Гроттер и молот Перуна» отстоит от «Таня Гроттер и посох волхвов» всего на 0,09. Как видно из Таб. 21, «Третий всадник мрака» и «Билет на Лысую гору» имеют расстояние 0,16. В отношении текстов Д.Емца можно сказать, что они явно написаны одним автором и в одном стиле. Сильное единообразие стиля проявляется в почти одинаковых частотах предлогов.

Интересно отметить, что одно из произведений Катаева-сына «Две сказки» сильно напоминает произведение Катаева-отца «Белеет парус одинокий» (ТМВ = 0,59). Это может говорить о том, что Катаев-сын пытался копировать прозу своего отца. Это может также говорить о том, что Катаев-отец произвел тотальную редакторскую правку этого произведения своего сына. Другое произведение Катаева-сына «Доктор велел мадеру пить» совершенно непохоже на стиль Катаева-отца (ТМВ = 2,83). При этом нужно иметь в виду, что схожесть текстов Катаева «Сын полка» и «Растратчики» составила 0,84.

Близко располагаются и произведения А.И.Солженицына (Таб. 24).

Нельзя не заметить, что близость разных произведений одного автора характерна как для писателей-классиков, так и для современных писателей. Это видно из всех таблиц. Из этого следует, что наш метод верно показал разницу стилей в письмах сильфов и в других прозаических произведениях И.А.Крылова. Наше предыдущее исследование по установлению авторства показало, что предположение противников Крылова о том, что несколько глав в «Почте духов» написаны не И.А.Крыловым, а Радищевым, неверно. Мы можем с надежностью 95% утверждать, что текст писем в «Почте духов» не был написан Радищевым. Следовательно, можно достоверно утверждать, что Радищев не подходит в качестве «подозреваемого» даже для писем сильфов. Также можно уверенно утверждать, что Карамзин и другие писатели не причастны к написанию текста «Почты духов». Письма гномов из «Почты духов» и проза Крылова написаны одним автором, т.е. самим Крыловым. Следовательно, наше статистическое исследование позволит литературоведам привести надежную аргументацию

в пользу авторства И.А.Крылова большей части текста «Почты духов», так как из 86.092 слов «Почты духов» 40.533 слова писем гномов и 14.864 слова писем других существ, т.е. всего 55.397 слов, написаны самим Крыловым. Это составляет 64% всего текста «Почты духов».

Остаются сомнения по поводу того, писал ли Крылов письма сильфов сам. Мы установили, что между текстами писем сильфов, с одной стороны, и текстами писем гномов (1,29) и прозы Крылова (1,69), с другой стороны, существует значительная разница. В тоже время между письмами сильфов и текстами Радищева или других писателей разница еще больше. Как видно из значений ТМВ, маловероятно, что Радищев писал письма сильфов. Принимая во внимание исследование М.В.Разумовской и наше исследование, становится понятно, почему в отношении писем сильфов целым рядом писателей и исследователей литературы высказывались большие сомнения. По статистическому распределению предлогов, как установило наше исследование, письма сильфов сильно отличаются от известных текстов Крылова. М.В.Разумовская сличила тексты маркиза Д'Аржана с письмами сильфов и установила, что письма сильфов – это фактически русские переводы указанных французских текстов. Писатели достаточно различаются с точки зрения пространственной ориентации, которая выражена в частоте различных предлогов. Письма гномов и рассказы Крылова отличаются от писем сильфов тем, что пространственная ориентация у них разная. При переводе на русский язык пространственная ориентация Д'Аржана в какой-то мере сохранилась. Это и вызвало разницу в употреблении предлогов. Если же И.А.Крылов переводил не сам, а пользовался услугами переводчика, то это также объясняет разницу текстов. В этом случае на стиль маркиза Д'Аржана наложился еще и стиль переводчика с французского на русский язык [Тамбовцев и др. 2008а].

Выводы

1) Данные всех таблиц показывают, что произведения одного и того же автора обычно слабо различаются по коэффициенту ТМВ, что может говорить о близости пространственно-временных структур, отображаемых предлогами. На этом можно построить поиск автора или же определить плагиат.

Тексты Бахтина, с одной стороны, и Волошинова и Медведева, с другой, статистически различны. Из этого следует, что тексты, подписанные Бахтиным, писал сам Бахтин. Тексты, подписанные Волошиновым или Медведевым, отличаются от текстов Бахтина достаточно сильно, поэтому мы не можем говорить о том, что Волошинов или Медведев украли текст Бахтина.

2) В то же время может вызывать подозрение в плагиате то, что текст Медведева «Формальный метод в литературоведении» очень похож (ТМВ = 0,89) на текст Волошинова «Марксизм и философия языка».

ПРИЛОЖЕНИЯ

Таб. 3

Частота встречаемости предлогов в текстах разных авторов на 10.000 слов.

БРА = Бахтин М.М. «Творчество Франсуа Рабле»
 БФВ = Бахтин М.М. «Формы времени и хронотопа»
 БДО = Бахтин М.М. «Проблемы творчества Достоевского»
 ВОЛ = Волошинов В.Н. «Марксизм и философия языка»
 МДВ = Медведев П.Н. «Формальный метод в литературоведении»
 ПушК = Пушкин А.С. «Капитанская дочка»
 ЧехОс = Чехов А.П. «Остров Сахалин»
 ЛерГ = Лермонтов М.Ю. «Герой нашего времени»
 Горб = Горбачев М.С. «Как это было»

	БРА	БФВ	БДО	ВОЛ	МДВ	ПушК	ЧехОс	ЛерГ	Горб
на	93	80	68	68	65	124	197	144	153
до	17	13	14	14	9	11	24	10	12
в	366	402	301	310	313	201	288	175	427
к	51	48	65	63	62	64	46	48	64
с	97	98	91	65	86	117	94	87	127
за	15	16	15	7	8	52	31	53	28
из	39	25	38	33	39	35	52	34	29
по	27	31	27	27	31	35	65	42	64
во	22	20	13	15	13	13	16	7	14
о	32	25	32	28	53	30	28	17	65
от	34	36	32	39	56	32	10	23	24
перед	7	3	9	3	3	7	3	5	5
над	4	3	4	8	1	3	1	3	3
под	7	6	4	4	4	12	13	2	9
для	36	39	36	51	43	7	26	13	45
между	13	14	12	31	20	10	10	10	37
у	25	38	24	8	15	35	40	50	19
про	1	0	2	0	0	5	2	3	0
без	5	6	9	8	10	7	12	6	8
через	4	8	2	2	2	10	7	10	5
со	10	11	8	8	6	17	10	4	14
ко	2	2	1	1	1	10	2	5	1
об	6	6	4	16	6	5	4	8	18
обо	0	0	0	0	0	2	0	2	0
ради	0	0	0	0	1	1	0	0	0

Таб. 4

Расстояние между авторами. Сходство в употреблении предлогов в «Почте духов» И.А.Крылова по сравнению с другими текстами Крылова и других авторов по величине коэффициента ТМВ. Весь текст «Почты духов» – 86.092 слова.

№	Крылов: «Почта духов». Текст всех писем	Величина ТМВ
1.	Крылов: «Проза. Рассказы»	0,64
2.	Чернышевский: «Что делать?»	1,68
3.	Радищев: «Путешествие из Петербурга в Москву»	1,72
4.	Карамзин: «Письма русского путешественника». Т. 1	1,76
5.	Карамзин: «История государства Российского». Т. 2	1,98
6.	Доценко: «Команда Бешеного»	2,05
7.	Пушкин: «Капитанская дочка»	2,15
8.	Тургенев: «Новь»	2,23
9.	Толстой: «Война и мир». Т. 1	2,23
10.	Фадеев: «Разгром»	2,23
11.	Вайнеры: «Я – следователь»	2,37
12.	Горький: «Мать»	2,56
13.	Емец: «Билет на Лысую гору»	2,65
14.	Крюков: «Рассказы». № 1	2,68
15.	Радищев: «О человеке»	2,77
16.	Чехов: «Остров Сахалин»	2,94
17.	Лермонтов: «Герой нашего времени»	2,98
18.	Катаев: «Белеет парус одинокий»	3,05
19.	Булгаков: «Мастер и Маргарита»	3,19
20.	Горбачев: «Как это было»	3,70
21.	Шолохов: «Тихий Дон». Кн. 1. № 1-14	4,73
22.	Проханов: «Идущие в ночи»	4,74

Таб. 5

Расстояние между авторами. Сходство в употреблении предлогов в письмах гномов из «Почты духов» И.А.Крылова по сравнению с другими текстами Крылова и других авторов по величине коэффициента ТМВ. Объем писем гномов – 40.533 слова.

№	Крылов: Письма гномов из «Почты духов»	Величина ТМВ
1.	Крылов: «Проза. Рассказы»	0,60
2.	Крылов: Письма других существ из «Почты духов»	0,98
3.	Крылов: Письма сильфов из «Почты духов»	1,29
4.	Карамзин: «Письма русского путешественника». Т. 1	1,48
5.	Кассиль-Поляновский: «Улица младшего сына»	2,02
6.	Радищев: «Путешествие из Петербурга в Москву»	2,08
7.	Лермонтов: «Герой нашего времени»	2,09

Таб. 6

Расстояние между авторами. Сходство в употреблении предлогов в письмах сильфов из «Почты духов» И.А.Крылова по сравнению с другими текстами Крылова и других авторов по величине коэффициента ТМВ. Письма сильфов – 27.672 слова или 32,14% от всего текста «Почты духов».

№	Крылов: Текст писем сиффов из «Почты духов»	Величина ТМВ
1.	Крылов: Письма гномов из «Почты духов»	1,29
2.	Крылов: «Проза. Рассказы»	1,69
3.	Радищев: «Путешествие из Петербурга в Москву»	1,94
4.	Радищев: «О человеке, о его смертности»	1,97
5.	Карамзин: «Письма русского путешественника». Т. 2	2,12
6.	Карамзин: «История государства Российского». Т. 2	2,44
7.	Доценко: «Команда Бешеного»	3,24
8.	Тургенев: «Новь»	3,38
9.	Фадеев: «Разгром»	3,56
10.	Пушкин: «Капитанская дочка»	3,65
11.	Горький: «Мать»	3,77
12.	Акунин: «Пелагия и черный монах»	3,95
13.	Вайнеры: «Я – следователь»	3,98
14.	Довлатов: «Зона»	4,04
15.	Лимонов: «Это я – Эдичка»	4,10
16.	Горбачев: «Как это было»	4,38
17.	Емец: «Билет на Лысую гору»	4,50
18.	Лермонтов: «Герой нашего времени»	4,70
19.	Булгаков: «Мастер и Маргарита»	4,73
20.	Катаев: «Белеет парус одинокий»	4,74
21.	Чехов: «Остров Сахалин»	5,96

Таб. 7

Расстояние между авторами. Сходство в употреблении предлогов в письмах других существ из «Почты духов» И.А.Крылова по сравнению с другими текстами Крылова и других авторов по величине коэффициента ТМВ. Объем писем других существ – 14.864 слова.

№	Крылов: Письма других существ из «Почты духов»	Величина ТМВ
1.	Крылов: «Проза. Рассказы»	0,83
2.	Крылов: Письма гномов из «Почты духов»	0,98
3.	Радищев: «Путешествие из Петербурга в Москву»	1,53
4.	Радищев: «Дневник, письма»	1,87
5.	Радищев: «О человеке, о его смертности»	2,52

Таб. 8

Расстояние между авторами. Сходство в употреблении предлогов в прозе (рассказах) И.А.Крылова по сравнению с другими текстами Крылова и других авторов по величине коэффициента ТМВ. Объем выборки – 34.758 слов.

№	Крылов: «Проза. Рассказы»	Величина ТМВ
1.	Крылов: Письма гномов из «Почты духов»	0,60
2.	Крылов: Весь текст «Почты духов»	0,64
3.	Крылов: «Письма других существ»	0,83

4.	Кассиль: «Будьте готовы, Ваше высочество»	1,47
5.	Кассиль-Поляновский: «Улица младшего сына»	1,56
6.	Крылов: «Письма сильфов»	1,69

Таб. 9

Расстояние между авторами. Сходство в употреблении предлогов в тексте Н.М.Карамзина «История государства Российского» (т. 2) по сравнению с другими текстами Карамзина и других авторов по величине коэффициента ТМВ. Объем выборки – 62.539 слов.

№	Карамзин: «История государства Российского». Т. 2	Величина ТМВ
1.	Карамзин: «История государства Российского». Т. 1. Ч. 1	0,79
2.	Карамзин: «История государства Российского». Т. 1. Ч. 2	0,99
3.	Крылов: Весь текст «Почты духов»	1,98
4.	Пушкин: «Капитанская дочка»	3,94
5.	Вайнеры: «Я – следователь»	3,95
6.	Достоевский: «Идиот»	4,17
7.	Тургенев: «Новь»	4,28
8.	Емец: «Билет на Лысую гору»	4,34
9.	Катаев: «Белеет парус одинокий»	4,45
10.	Горький: «Мать»	5,06
11.	Лермонтов: «Герой нашего времени»	5,53

Таб. 10

Расстояние между авторами. Сходство в употреблении предлогов в «Капитанской дочке» Пушкина с другими текстами Пушкина и других авторов по величине коэффициента ТМВ. Объем выборки – 34.471 слово.

№	Пушкин: «Капитанская дочка»	Величина ТМВ
1.	Пушкин: «Повести Белкина»	0,53
2.	Пушкин: «Дубровский»	0,63
3.	Солженицын: «В круге первом»	1,02
4.	Тургенев: «Новь»	1,09
5.	Достоевский: «Идиот»	1,14
6.	Емец: «Билет на Лысую гору»	1,26
7.	Горький: «Мать»	1,39
8.	Булгаков: «Мастер и Маргарита»	1,39
9.	Чернышевский: «Что делать?»	1,68
10.	Веллер: «Великий последний шанс»	1,79
11.	Шолохов: «Тихий Дон». Кн. 1. Выборки № 1-14	2,20
12.	Бажов: «Сказы»	2,40
13.	Проханов: «Идущие в ночи»	3,46
14.	Горбачев: «Как это было»	5,18

Таб. 11

Расстояние между авторами. Сходство в употреблении предлогов в романе Л.Н.Толстого «Война и мир» (т. 1) с другими текстами Толстого и других авторов по величине коэффициента ТМВ. Объем выборки – 270.139 слов.

№	Толстой: «Война и мир». Т. 1	Величина ТМВ
1.	Толстой: «Война и мир». Т. 2.	0,26
2.	Толстой: «Воскресенье»	0,29
3.	Толстой: «Война и мир». Т. 3.	0,43
4.	Толстой: «Юность»	0,50
5.	Толстой: «Отрочество»	0,54
6.	Толстой: «Казачи»	0,63
7.	Толстой: «Поликушка»	0,69
8.	Толстой: «Крейцерова соната»	0,73
9.	Толстой: «Война и мир». Т. 4	0,74
10.	Толстой: «Смерть Ивана Ильича»	0,88
11.	Толстой: «Анна Каренина»	0,95
12.	Крюков: «Рассказы». Выборка 3	1,11
13.	Крюков: «Рассказы». Выборка 1	1,15
14.	Емец: «Билет на Лысую гору»	1,32
15.	Булгаков: «Мастер и Маргарита»	1,92
16.	Лимонов: «Это я – Эдичка»	2,10
17.	Горбачев: «Как это было»	5,40

Таб. 12

Расстояние между авторами. Сходство в употреблении предлогов в романе И.С.Тургенева «Дворянское гнездо» с другими текстами Тургенева и других авторов по величине коэффициента ТМВ. Объем выборки – 95.720 слов.

№	Тургенев: «Дворянское гнездо»	Величина ТМВ
1.	Тургенев: «Накануне»	0,65
2.	Тургенев: «Новь»	0,69
3.	Тургенев: «Дым»	0,75
4.	Веллер: «Великий последний шанс»	1,14
5.	Достоевский: «Униженные и оскорбленные»	1,25
6.	Толстой: «Анна Каренина»	1,47
7.	Шолохов: «Поднятая целина». № 1-12	1,70
8.	Шолохов: «Донские рассказы»	3,81
9.	Шолохов: «Тихий Дон». Кн. 1. № 1-14	3,88

Таб. 13

Расстояние между авторами. Сходство в употреблении предлогов в романе И.С.Тургенева «Новь» с другими текстами Тургенева и других авторов по величине коэффициента ТМВ. Объем выборки – 80.649 слов.

№	Тургенев: «Новь»	Величина ТМВ
1.	Тургенев: «Накануне»	0,09
2.	Тургенев: «Отцы и дети»	0,18
3.	Тургенев: «Рудин»	0,19
4.	Тургенев: «Дым»	0,22
5.	Тургенев: «Дворянское гнездо»	0,69
6.	Тургенев: «Записки охотника»	0,84
7.	Пушкин: «Капитанская дочка»	1,09
8.	Горький: «Мать»	1,16
9.	Крюков: «Рассказы». Ч. 3	1,34
10.	Крюков: «Рассказы». Ч. 1	1,41
11.	Веллер: «Великий последний шанс»	1,46
12.	Вайнеры: «Я – следователь»	1,59
13.	Шолохов: «Тихий Дон». Кн. 3	1,90
14.	Шолохов: «Тихий Дон». Кн. 2	1,95
15.	Булгаков: «Мастер и Маргарита»	2,21
16.	Чехов: «Остров Сахалин»	2,79
18.	Проханов: «Идущие в ночи»	4,66
19.	Горбачев: «Как это было»	5,60

Таб. 14

Расстояние между авторами. Сходство в употреблении предлогов в романе Ф.М.Достоевского «Идиот» с другими текстами Достоевского и других авторов по величине коэффициента ТМВ. Объем выборки = 213.658 слов.

№	Достоевский: «Идиот»	Величина ТМВ
1.	Достоевский: «Бесы»	0,08
2.	Достоевский: «Преступление и наказание»	0,28
3.	Достоевский: «Игрок»	0,29
4.	Достоевский: «Подросток»	0,38
5.	Достоевский: «Вечный муж»	0,58
6.	Достоевский: «Униженные и оскорбленные»	0,65
7.	Достоевский: «Скверный анекдот»	0,71
8.	Достоевский: «Двойник»	0,73
9.	Достоевский: «Записки из подполья»	0,77
10.	Веллер: «Великий последний шанс»	1,10
11.	Довлатов: «Чемодан»	1,13
12.	Пушкин: «Капитанская дочка»	1,14
13.	Лермонтов: «Герой нашего времени»	1,27
14.	Емец: «Билет на Лысую гору»	1,39
15.	Крюков: «Рассказы». Выборка 1	1,54
16.	Акунин: «Пелагия и черный монах»	1,66
17.	Шолохов: «Тихий Дон». Кн. 2	1,82
18.	Гончаров: «Фрегат Паллада». Т. 2	1,89
19.	Булгаков: «Мастер и Маргарита»	2,28

20.	Проханов: «Идущие в ночи»	4,84
21.	Горбачев: «Как это было»	5,75

Таб. 15

Расстояние между авторами. Сходство в употреблении предлогов в романе И.А.Гончарова «Обыкновенная история» с другими текстами Гончарова и других авторов по величине коэффициента ТМВ. Объем выборки – 97.998 слов.

№	Гончаров: «Обыкновенная история»	Величина ТМВ
1.	Гончаров: «Обломов»	0,35
2.	Гончаров: «Обрыв»	0,54
3.	Крюков: «Рассказы». Ч. 1	1,32
4.	Шолохов: «Поднятая целина». Кн. 1. Выборка № 13-24	1,47
5.	Шолохов: «Тихий Дон». Кн. 3	2,30

Таб. 16

Расстояние между авторами. Сходство в употреблении предлогов в романе М.Горького «Мать» с другими текстами Горького и других авторов по величине коэффициента ТМВ. Объем выборки – 92.760 слов.

№	Горький: «Мать»	Величина ТМВ
1.	Горький: «Дело Артамоновых»	0,45
2.	Горький: «Фома Гордеев»	0,66
3.	Горький: «Челкаш»	0,88
4.	Довлатов: «Зона»	1,10
5.	Достоевский: «Идиот»	1,15
6.	Тургенев: «Новь»	1,16
7.	Веллер: «Великий последний шанс»	1,22
8.	Пушкин: «Капитанская дочка»	1,39
9.	Шолохов: «Они сражались за Родину»	1,72
10.	Емец: «Билет на Лысую гору»	1,40
11.	Крюков: «Рассказы». Ч. 1	1,90
12.	Катаев: «Белеет парус одинокий»	2,02
13.	Вайнеры: «Я – следователь»	2,17
14.	Булгаков: «Мастер и Маргарита»	2,32
15.	Гончаров: «Фрегат Паллада». Т. 2	2,41
16.	Серафимович: «Железный поток»	2,41
17.	Чехов: «Остров Сахалин»	3,01
18.	Горбачев: «Как это было»	6,29

Таб. 17

Расстояние между авторами. Сходство в употреблении предлогов в романе Б.Акунина «Пелагия и черный монах» с другими текстами Акунина и других авторов по величине коэффициента ТМВ. Объем выборки – 86.246 слов.

№	Акунин: «Пелагия и черный монах»	Величина ТМВ
1.	Акунин: «Пелагия и белый бульдог»	0,23
2.	Акунин: «Декоратор»	0,28
3.	Акунин: «Турецкий гамбит»	0,29
4.	Акунин: «Кладбищенские истории»	0,49
5.	Акунин: «Любовник смерти»	0,77
6.	Крюков: «Рассказы». Выборка 1	1,16
7.	Вайнеры: «Я – следователь»	1,27
8.	Гончаров: «Фрегат Паллада». Т. 2	1,30
9.	Веллер: «Великий последний шанс»	1,44
10.	Шолохов: «Поднятая целина». Выборка № 1-12	1,48
11.	Шолохов: «Тихий Дон». Кн. 1. Выборка № 1-14	1,64
12.	Лермонтов: «Герой нашего времени»	1,77
13.	Лимонов: «Это я – Эдичка»	2,37
14.	Тургенев: «Новь»	3,41
15.	Горбачев: «Как это было»	5,23

Таб. 18

Расстояние между авторами. Сходство в употреблении предлогов в произведении М.А.Булгакова «Мастер и Маргарита» с другими текстами Булгакова и других авторов по величине коэффициента ТМВ. Объем выборки – 47.127 слов.

№	Булгаков: «Мастер и Маргарита»	Величина ТМВ
1.	Булгаков: «Театральный роман»	0,86
2.	Булгаков: «Записки юного врача»	0,90
3.	Булгаков: «Белая гвардия»	0,92
4.	Фадеев: «Разгром»	1,20
5.	Чехов: «Остров Сахалин»	1,68
6.	Шолохов: «Тихий Дон». Кн. 1. Выборки № 1-14	1,77
7.	Серафимович: «Железный поток»	2,09
8.	Тургенев: «Новь»	2,21
9.	Достоевский: «Идиот»	2,28

Таб. 19

Расстояние между авторами. Сходство в употреблении предлогов в произведении братьев Вайнеров «Я – следователь» с другими текстами Вайнеров и других авторов по величине коэффициента ТМВ. Объем выборки – 45.551 слово.

№	Вайнеры: «Я – следователь»	Величина ТМВ
1.	Вайнеры: «Петля и камень на зеленой траве»	0,50
2.	Вайнеры: «Евангелие от палача»	0,64
3.	Чехов: «Остров Сахалин»	1,15
4.	Акунин: «Пелагия и черный монах»	1,27

5.	Доценко: «Команда Бешеного»	1,31
6.	Фадеев: «Разгром»	1,35
7.	Тургенев: «Новь»	1,59
8.	Достоевский: «Идиот»	1,65
9.	Шолохов: «Они сражались за Родину»	1,85
10.	Горький: «Мать»	2,17

Таб. 20

Расстояние между авторами. Сходство в употреблении предлогов в произведении С.Д.Довлатова «Зона» с другими текстами Довлатова и других авторов по величине коэффициента ТМВ. Объем выборки – 47.127 слов.

№	Довлатов: «Зона»	Величина ТМВ
1.	Довлатов: «Чемодан»	0,74
2.	Довлатов: «Заповедник»	0,76
3.	Довлатов: «Иностранка»	0,87
4.	Фадеев: «Разгром»	1,01
5.	Горький: «Мать»	1,10
6.	Доценко: «Команда Бешеного»	1,11
7.	Катаев: «Белеет парус одинокий»	1,27
8.	Толстой: «Война и мир». Т. 1	1,29
9.	Тургенев: «Новь»	1,33
10.	Пушкин: «Капитанская дочка»	1,44
11.	Достоевский: «Идиот»	1,53
12.	Булгаков: «Мастер и Маргарита»	1,66
13.	Шолохов: «Поднятая целина». Выборка № 1-12	1,83
14.	Гончаров: «Фрегат Паллада». Т. 2	2,04
15.	Лимонов: «Это я – Эдичка»	2,24
16.	Горбачев: «Как это было»	6,56

Таб. 21

Расстояние между авторами. Сходство в употреблении предлогов в произведении Д.Емца «Третий всадник мрака» с другими текстами Емца и других авторов по величине коэффициента ТМВ. Объем выборки – 66.915 слов.

№	Емец: «Третий всадник мрака»	Величина ТМВ
1.	Емец: «Билет на Лысую гору»	0,16
2.	Емец: «Маг полуночи»	0,22
3.	Емец: «Магический контрабас»	0,26
4.	Емец: «Локон Афродиты»	0,27
5.	Крюков: «Рассказы». Ч. 2	1,07
6.	Шолохов: «Тихий Дон». Кн. 1. Выборка № 1-14	1,53

Таб. 22

Расстояние между авторами. Сходство в употреблении предлогов в произведении Д.Емца «Билет на Лысую гору» с другими текстами Емца и других авторов по величине коэффициента ТМВ. Объем выборки – 66.915 слов.

№	Емец: «Билет на Лысую гору»	Величина ТМВ
1.	Емец: «Третий всадник мрака»	0,16
2.	Крюков: «Рассказы». Ч. 2	1,06
3.	Довлатов: «Зона»	1,28
4.	Достоевский: «Идиот»	1,39
5.	Серафимович: «Железный поток»	1,70
6.	Шолохов: «Тихий Дон». Кн. 1	1,76
7.	Карамзин: «История государства Российского». Т. 2	4,34

Таб. 23

Расстояние между авторами. Сходство в употреблении предлогов в произведении В.П.Катаева «Белеет парус одинокий» с другими текстами Катаева и других авторов по величине коэффициента ТМВ. Объем выборки – 70.343 слова.

№	Катаев: «Белеет парус одинокий»	Величина ТМВ
1.	Катаев: «Растратчики»	0,36
2.	Катаев: «Сын полка»	0,55
3.	Катаев-сын: «Две сказки»	0,59
4.	Катаев: «Алмазный мой венец»	0,65
5.	Пушкин: «Капитанская дочка»	1,24
6.	Толстой: «Война и мир». Т. 1	1,32
7.	Достоевский: «Идиот»	1,63
8.	Тургенев: «Новь»	1,74
9.	Катаев-сын: «Доктор велел мадеру пить»	2,83

Таб. 24

Расстояние между авторами. Сходство в употреблении предлогов в произведении А.И.Солженицына «Бодался теленок с дубом» с другими текстами Солженицына и других авторов по величине коэффициента ТМВ. Объем выборки – 164.318 слов.

№	Солженицын: «Бодался теленок с дубом»	Величина ТМВ
1.	Солженицын: «Раковый корпус»	0,39
2.	Солженицын: «В круге первом»	0,49

Таб. 25

Расстояние между авторами. Сходство в употреблении предлогов в произведении А.А.Фадеева «Разгром» с другими текстами Фадеева и других авторов по величине коэффициента ТМВ. Объем выборки – 47.127 слов.

№	Фадеев: «Разгром»	Величина ТМВ
1.	Фадеев: «Последний из Удэге»	0,33
2.	Фадеев: «Молодая гвардия»	0,44
3.	Фадеев: «Разлив»	0,65
4.	Фадеев: «Черная металлургия»	0,76
5.	Веллер: «Великий последний шанс»	1,09
6.	Серафимович: «Железный поток»	1,16
7.	Булгаков: «Мастер и Маргарита»	1,20
8.	Шолохов: «Поднятая целина». Выборка № 1-12	1,27
9.	Вайнеры: «Я – следователь»	1,35
10.	Горбачев: «Как это было»	5,26

ЛИТЕРАТУРА

- БПМ 1993 – Бахтин под маской / Маска вторая: П.Н.Медведев. Формальный метод в литературоведении. М., 1993.
- Васильев 1991 – *Васильев Н.Л.* М.М.Бахтин или В.Н.Волошинов? // Литературное обозрение. 1991. № 9.
- Мартыненко 1983 – *Мартыненко Г.Я.* Многомерный синтактико-статистический анализ художественной прозы. Структурная и прикладная лингвистика. Л., 1983.
- Марусенко и др. 2001 – *Марусенко М.А., Бессонов Б.А., Аникин М.А., Мясоедова Н.Е.* В поисках потерянного автора. Этюды атрибуции. СПб., 2001.
- Пешков 1995 – *Пешков И.В.* Хроника. Факты. Информация. Один вопрос вокруг конференций // Диалог. Карнавал. Хронотоп. 1995. № 3.
- Пиотровский и др. 1977 – *Пиотровский Р.Г., Бектаев К.Б., Пиотровская А.А.* Математическая лингвистика. М., 1977.
- Разумовская 1978 – *Разумовская М.В.* «Почта духов» И.А.Крылова и романы маркиза Д'Аржана // Русская литература. 1978. № 1.
- Розенталь и др. 1985 – *Розенталь Д.Э., Теленкова М.А.* Словарь-справочник лингвистических терминов. М., 1985.
- СРЯ 1988 – Словарь русского языка. Т. 1-4. М., 1988.
- Тамбовцев 2003 – *Тамбовцев Ю.А.* Типология функционирования фонем в звуковой цепочке индоевропейских, палеоазиатских, урало-алтайских и других языков мира: компактность подгрупп, групп, семей и других языковых таксонов. Новосибирск, 2003.
- Тамбовцев 2010 – *Тамбовцев Ю.А.* Фонотипологическая близость лингвистических объектов по критерию «хи-квадрат» // Вестник НГУ. Серия: Информационные технологии. Т. 8. 2010. Вып. 3.

- Тамбовцев и др. 2008а – Тамбовцев Ю.А., Тамбовцева А.Ю., Тамбовцева Л.А. Типология распределения некоторых лингвистических единиц в тексте как показатель авторства текста // Вестник Омского университета. 2008. № 2.
- Тамбовцев и др. 2008б – Тамбовцев Ю.А., Тамбовцева А.Ю., Тамбовцева Л.А. Приложил ли Радищев свою руку к написанию части писем из «Почты духов» Крылова? // *Studia Slavica Hungarica*. № 53. 2008.
- Фоменко 1994 – Фоменко И.В. Служебные слова как возможность реконструкции авторского мироощущения // Новое литературное обозрение. 1994. № 8.
- Фоменко и др. 1996 – Фоменко В.П., Фоменко Т.Г. Авторский инвариант русских литературных текстов (предисловие академика А.Т.Фоменко) // Фоменко А.Т. Новая хронология Греции. Античность и средневековье. Т. 2. М., 1996.
- Фукс 1975 – Фукс В. По всем правилам искусства: точные методы в исследованиях литературы, музыки и изобразительного искусства. М., 1975.
- Brainerd 1974 – Brainerd B. Weighing Evidence in Language and Literature: A Statistical Approach. Toronto, 1974.
- Brownlee 1949 – Brownlee K.A. Industrial Experimentation. Oxford, 1949.
- Burrows 1978 – Burrows J.F. Computation into Criticism: A Study of Jane Austin's Novels and an Experiment in Method. Oxford, 1978.
- Butler 1985 – Butler Ch. Statistics in Linguistics. Oxford, 1985.
- Herdan 1966 – Herdan G. The Advanced Theory of Language as Choice and Chance. Berlin, 1966.
- Holmes 1985 – Holmes D. The Analysis of Literary Style – A Review // The Journal of the Royal Statistical Society. Series A. 148. 1985.
- Kenny 1986 – Kenny A. A Stylometric Study of the New Testament. Oxford, 1986.
- Ledger 1989 – Ledger G.R. Re-Counting Plato: A Computer Analysis of Plato's Style. Oxford, 1989.
- Morton 1978 – Morton A.Q. Literary Detection. Oxford, 1978.
- Tambovtsev et al. 2007 – Tambovtsev Y., Tambovtseva L., Tambovtseva J. The frequency of linguistic units as an indicator of typological similarity between texts // *California Linguistic Notes*. Vol. XXXII. 2007. № 2.