

Гейрхетс  
Минадлежность  
Достоевскому:  
К вопросу об атрибуции  
ФМ Достоевскому  
анонимных статей в  
журналах *Время* и *Эпоха*

---

Geir Kjetsaa  
*Attributed to Dostoevsky:  
The problem of attributing to  
Dostoevsky anonymous articles  
in Time and Epoch*

---

SOLUM FORLAG







# Принадлежность Достоевскому:

К вопросу об атрибуции Ф.М.  
Достоевскому анонимных  
статьей в журналах *Время* и *Эпоха*

SLAVICA NORVEGICA IV

Гейр Хетсо

# Приналежность Достоевскому:

К вопросу об атрибуции Ф.М.  
Достоевскому анонимных  
статьей в журналах *Время* и *Эпоха*

Geir Kjetsaa

Attributed to Dostoevsky:  
The problem of attributing to Dostoevsky  
anonymous articles in *Time* and *Epoch*

First published in 1986 in Norway by SOLUM FORLAG A.S.

© Copyright 1986 by Geir Kjetsaa  
Cover design: Jan Pahle

Financially supported by Norges allmennvitenskapelige  
forskningsråd (The Norwegian Research Council for Science and  
the Humanities).

ISBN 82-560-0439-8

All rights reserved. No part of this publication may be reproduced,  
stored in a retrieval system or transmitted in any other form or any  
means, electronic, photocopying, recording or otherwise, without  
the permission in writing from the publishers.

Printed in Norway by  
S. Hammerstads Boktrykkeri, Oslo

# Содержание

1.0.	Задача и методика .....	7
2.0.	Материал .....	16
2.1.	Статьи, принадлежащие Достоевскому .....	16
2.2.	Статьи, приписываемые Достоевскому .....	21
3.0.	Принципы выбора материала .....	29
4.0.	Лингвостатистические параметры .....	31
4.1.	Общее распределение частей речи .....	32
4.2.	Средняя длина слова в буквах .....	42
4.3.	Общее распределение длины слова .....	47
4.4.	Средняя длина предложения в словах .....	51
4.5.	Общее распределение длины предложения .....	55
4.6.	Лексический спектр текста на уровне словаря .....	58
4.7.	Лексический спектр текста на уровне текста .....	60
4.8.	Индекс разнообразия лексики .....	62
5.0.	Подводя итоги .....	67
6.0.	Вместо заключения .....	72
7.0.	Сноски .....	76



## 1.0. Задача и методика

Известно, что художественное творчество шло у Достоевского рука об руку с журналистикой. Первая публицистическая статья написана им в 1845 г., последняя вышла в свет после его смерти. Притом Достоевский редактировал и возглавлял три журнала, имевших широкое влияние на русскую общественную мысль: *Время* (1861–1863), *Эпоха* (1864–1865), *Гражданин* (1873–1874), и издавал свой личный журнал *Дневник писателя* (1876–1877, 1880–1881). В *Полном собрании сочинений* публицистические работы Достоевского составляют несколько томов. На такую длительную и многообразную журналистскую деятельность вряд ли может сослаться какой-либо из великих русских писателей прошлого столетия, за исключением разве что Льва Толстого.

Поэтому нет ничего удивительного в том, что исследователи жизни и творчества Достоевского много внимания уделяли и уделяют его публицистике. Существует ряд специальных работ о политическом направлении издаваемых им журналов, о влиянии Достоевского-публициста на Достоевского-романиста. Двум его журналам посвящены монографии (1). Тем не менее, до сих пор остается открытым один из самых существенных вопросов: какие же статьи действительно были написаны Достоевским?

В отличие от статей, помещенных в *Дневнике писателя*, определение принадлежности Достоевскому статей в журналах *Время*, *Эпоха* и *Гражданин* связано с большими сложностями. Подавляющее большинство этих статей опубликовано анонимно, т.е. либо без подписи, либо под псевдонимами. Впрочем, это относится и к статьям, которые исследователи давно приписывали Достоевскому, более или менее основываясь на документальных данных.

Интересное обсуждение проблем атрибуции Достоевскому анонимных статей *Гражданина* было проведено академиком

В.В. Виноградовым (2). Наша же задача – решение на основе современной техники ЭВМ вопроса о принадлежности Достоевскому ряда анонимных статей в журналах *Время* и *Эпоха*. Эта задача сложилась во время нашего пребывания в Ленинграде в декабре 1977 г. В письме от дирекции Института русской литературы АН СССР (ИРЛИ) и Главной редакции Полного (академического) собрания сочинений Достоевского говорится:

«За последние 50–60 лет ряд видных ученых в СССР и за рубежом приписали Достоевскому значительное число статей и заметок, опубликованных анонимно в журналах *Время*, *Эпоха* и *Гражданин*. Ученые при этом исходили из близости идейно-тематических мотивов или отдельных стилистических приемов этих статей произведениям Достоевского того же периода.

В настоящее время Институт русской литературы ведет работу по авторской атрибуции этих статей в связи с завершением издания Достоевского. Дирекция Института русской литературы и Главная редакция Полного собрания сочинений Достоевского были бы Вам благодарны, если бы Вы согласились, как знаток этого вопроса, провести лингвостилистическое исследование приписанных Достоевскому статей и заметок на основе разработанной Вами методики с применением электронной техники» (3).

Сознавая трудность работы (в ряде случаев речь шла о довольно коротких текстах), мы согласились принять это лестное для себя предложение. Вернувшись в Норвегию, мы разработали проект, исходя из списка статей, присланного нам Пушкинским Домом. В материал бесспорно принадлежащих Достоевскому текстов вошли 26 статей, или около 120 000 текстовых слов, а в материал приписываемых писателю текстов 12 статей, или около 58 000 текстовых слов (4).

Для перевода такого большого материала на язык ЭВМ и для получения необходимых частотных словарей и статистических данных потребовалось полгода. В марте 1979 г. от Норвежского совета по исследованиям в области гуманитарных наук были получены нужные нам средства для проекта: 40 000 норвежских крон, или около 4 000 рублей. Шесть месяцев спустя, не в последнюю очередь благодаря

активному участию со стороны нашего консультанта по ЭВМ И. Фоннесса и нашего ассистента Т.У. Хельгакера, переведенные на язык машины тексты были обработаны на ЭВМ типа DEC-10 при вычислительном центре Университета г. Осло (5).

В фундаментальном труде П. Вашака о методах определения авторства указываются три основных метода атрибуции:

А. Метод документальный и фактический, основанный на информации двух категорий: 1) исходящей от автора – например, рабочие тексты и автографы, переписка, дневники, автобиография и т. д.; 2) данных «неавторских», т. е. исходящих от всех лиц и учреждений, участвующих прямо или вторично в процессе генезиса и фиксации произведений. Утверждается, что «абсолютного документа» как однозначного доказательства авторства не существует.

Б. Метод идеино-тематический, основанный на конфронтации идей, идеиного направления и тематики атрибутированного текста (произведения) и произведений предполагаемых авторов (литературных школ и генераций, периодов времени, и т. д.).

В. Метод языковой и стилистический, исходящий из понятия индивидуального стиля и состоящий в конфронтации лингвостилистических качеств атрибутированного текста и произведений предполагаемых авторов (литературных школ, направлений, периодов времени, и т. д.), с сосредоточением на подсознательных качествах писания (6).

В настоящем исследовании мы будем заниматься вопросом о возможном авторстве Достоевского исключительно исходя из метода В, т. е. из метода языкового и стилистического. При этом важно подчеркнуть, что определение авторства в анонимных статьях в нашу задачу не входит. Вопросов типа: «А если автор статьи не Достоевский, кто тогда мог бы написать ее?» – мы не принимаем. Мы будем давать ответы только на поставленный ИРЛИ вопрос: есть ли, исходя из языковой информации, полученной с помощью ЭВМ, основание утверждать, что спорные, приписываемые Достоевскому статьи действительно написаны им?

Конечно, выделить из большого числа анонимных статей принадлежащие Достоевскому, не располагая документаль-

ными данными, представляется очень затруднительным делом, которое не исключает заблуждений даже со стороны самых крупных специалистов по стилистике (7).

Главным источником установления авторства Достоевского являются данные Н.Н. Страхова в списке, переданном им вдове писателя А.Г. Достоевской (8). В список этот автор включил свыше 20 статей и заметок. В начале 1860-х гг. Страхов был близким другом Достоевского. Надо думать, что, будучи постоянным сотрудником *Времени* и *Эпохи*, он хорошо знал, какие из анонимных статей журналов действительно принадлежали перу писателя. Поэтому авторитетность атрибуции Страхова не подвергалась сомнению исследователей.

Однако имеется и свидетельство самого Достоевского, что за два первых года сотрудничества в журнале *Время* он написал «до ста печатных листов» (9). Хотя свидетельство это и следует считать, по мнению Б.В. Томашевского, преувеличенным, нельзя исключать возможности того, что для *Времени* и *Эпохи* действительно было написано больше статей, чем полагал Страхов.

Поэтому неудивительно, что исследователи жизни и творчества Достоевского давно стремились расширить круг статей, определенных Страховым. Серьезные попытки в этом направлении когда-то были сделаны независимо друг от друга Л.П. Гроссманом (10) и О. фон Шульцем (11). Попытка последнего отличается особенной степенью гипотетичности. Основываясь, главным образом, на идеологических и тематических параллелях с бесспорно принадлежащими Достоевскому статьями, фон Шульц указал на семь статей, по его мнению, несомненно принадлежащих писателю, и на девять статей, вероятность принадлежности которых велика, но окончательной уверенности в этом он не имел.

Вскоре выяснилось, однако, что некоторые из атрибуций фон Шульца были более чем рискованы. Так, в рецензии Н. К. Пиксанова на его работу указывалось, что в состав «несомненно» принадлежащих Достоевскому статей попала одна статья Страхова (12). Само собой разумеется, что подобные ляпсусы сильно подрывают доверие и к остальным утверждениям автора.

В самом деле, работа фон Шульца наглядно показывает слабость и опасность методики определения авторства на основании так называемых «внешних показаний», т.е. на основании одних субъективных впечатлений ученого. Необходимо подчеркнуть, что при решении проблем атрибуции идеологические, тематические, биографические и даже текстуальные сближения не могут иметь безусловной доказательности. При единстве воззрений сотрудников *Времени* и *Эпохи* и при тесной сплоченности их в редакционной работе неудивительно, что они нередко повторяли друг друга. И даже если допустить, что Достоевский принимал активное участие в подготовке анонимных статей к печати, было бы методологически неверно приписывать ему статьи, которые он только редактировал. В этом отношении показательно, что фон Шульц в своем «стилистическом анализе» ограничивается указанием на сходство анонимных статей с заведомо принадлежащими Достоевскому статьями, а о стилистических различиях не упоминает.

Таким образом, методика внешних показаний или одних субъективных стилистических сопоставлений нередко толкает исследователей на опасный путь, где желаемое часто выдается за действительное. Недаром академик В.В. Виноградов, ратуя за перенесение проблемы авторства на почву объективно-исторических и лингвистико-статистических методов, пишет, что «субъективные методы определения автора и реконструкции адекватного авторского текста отжили или, во всяком случае, отживают свой век» (13).

Гораздо более объективен метод «внутренних показаний», включающий в себя лишь сравнение языка и стиля и заранее задающий все параметры, по которым проводится сравнение.

Характерным признаком этого лингвистико-статистического метода определения авторства является использование в качестве идентификаторов объективных характеристик. Установив, что предполагаемый автор имел возможность написать спорный текст (например, что он был в живых, когда текст был написан), мы должны в его бесспорных произведениях найти как можно более инвариантных явлений, по которым можем судить, соответствует ли спорный текст в лингвистическом отношении его бесспорным

произведениям. Если соответствует, то предполагаемый автор является претендентом на авторство спорного текста, а если не соответствует, то мы можем в дальнейшем исключить его из числа претендентов.

Именно понятие *исключения* является во всех определениях спорного авторства коренным понятием. Надо приступить к делу, полагая (подобно Шерлоку Холмсу), что истину можно найти только путем исключения невозможного.

Таким образом, исходя из методики языковых показаний, нельзя доказать, что та или иная статья написана Достоевским. Даже если язык в спорной статье обнаруживает большое сходство с языком бесспорно принадлежащих писателю статей, все же остается по крайней мере теоретическая возможность того, что она написана другим писателем, близким по манере письма Достоевскому (14).

В самом деле, существует только одно несомненное доказательство факта принадлежности произведения писателю, а именно отсутствие сомнений в его авторстве. В случае же возникновения сомнения в авторстве произведения нам остается создать «нулевую» гипотезу о том, что нет существенной разницы между языком, обнаруженным в спорной статье, и языком, обнаруженным в одновременно написанных бесспорных статьях. Если на основании ряда параметров и с помощью математической статистики оказывается, что эту гипотезу следует отвергнуть, то мы имеем право исключить и возможность того, что текст написан интересующим нас писателем.

Можно возразить, что метод «внутренних показаний» или метод адализа «формальных характеристик» текста, позволяющий нам исключить, но не подтвердить авторство произведения, имеет свои ограничения. Но это не значит, что применяемая нами методика может служить лишь вспомогательным средством в руках исследователя, занимающегося проблемой определения авторства. Нет никакого сомнения в том, что мы имеем дело с мощным орудием, имеющим самостоятельную, хотя только отрицательную доказательную силу. На наш взгляд, если спорная статья имеет достаточно значительный объем, скажем, 3 000–4 000 текстовых слов, и при этом существенно отличается по языку и стилю от заведомо принадлежащих Достоевскому

статей, то было бы весьма смело включить ее в собрание сочинений писателя, во всяком случае в основной текст собрания.

Однако прежде чем исключить, надо измерить. Методику внутренних показаний в определении авторства можно рассмотреть как отдельную область стилеметрики, предлагающей, что отдельные элементы языка и стиля писателя поддаются измерению.

Опыт показывает правомерность подобной предпосылки. Так, если сопоставить выборки из *Повестей Белкина* с выборками из *Войны и мира*, то мы сразу обнаружим существенное различие в средней длине предложения. Несомненно, что короткие предложения являются характерным признаком стиля Пушкина, так же, как длинные предложения являются характерным признаком стиля Толстого. Длина слова – еще одна объективная характеристика, по которой возможно идентифицировать стили писателей. Так, большинство слов у Шекспира имеет длину в четыре буквы, между тем, как большинство слов у Бэкона состоит всего лишь из трех. Конечно, подобные факты в свое время плохо подтверждали убеждение некоторых исследователей о том, что пьесы Шекспира написаны Бэкона.

Наряду с этими довольно «грубыми» параметрами современная лингвостатистика имеет в своем распоряжении и ряд других, более инвариантных и поэтому обладающих еще большей «дискриминирующей» силой.

Можно, например, сравнить, какими частями речи писатели открывают свои предложения. Каждый писатель, ежедневно пишущий предложение за предложением, невольно и, чаще всего, бессознательно усваивает особую привычку для образования своих предложений. С особой силой эта привычка оказывается в самом начале предложения. Так, иные писатели начинают свои предложения с предлога, другие с имени существительного и т. д. Из этого следует, что распределение частей речи в первой позиции предложения, позволяющие, как правило, отличить одного писателя от другого, является одним из существенных параметров для решения атрибуции текста.

Далее можно обследовать лексический спектр текста и разнообразие лексики писателя. Именно разнообразие

лексики оказалось важным параметром в нашем исследовании об авторстве *Тихого Дона*. С помощью компьютера было установлено, что если тексты Шолохова и *Тихого Дона* имеют в среднем 374 и 380 разных словоформ на каждые 500 новых текстовых слов, то число разных словоформ у Крюкова составляет только 361. Полученные данные о низкой повторяемости или богатстве словаря Шолохова позволили нам с помощью математической статистики исключить Крюкова как автора всех проанализированных частей *Тихого Дона* (15).

Научной предпосылкой для использования этих параметров является убеждение в том, что каждый писатель отдает предпочтение той или иной речевой конструкции, и что «разница в стиле двух писателей может выявиться как разница в частоте употребления определенных языковых форм» (16). В последние годы ряд исследований подтвердил оправданность данной предпосылки.

Перейдем теперь к интересующей нас проблеме о принадлежности Достоевскому анонимных статей в журналах *Время* и *Эпоха*.

На наш взгляд, целесообразно приступить к этой проблеме следующим образом.

Исходя из большого числа заведомо принадлежащих перу Достоевского статей, мы должны установить, как говорится в криминалистике, «отпечатки пальцев» писателя в его публицистике первой половины 1860-х годов, а потом провести сравнение с «отпечатками пальцев», оставленными в спорных текстах.

При этом необходимо учесть, что «отпечатки пальцев» писателя, в отличие от отпечатков пальцев преступника, безусловно поддаются изменению в зависимости от жанра и времени написания текста.

Последний момент, что писатель в молодости пишет иначе, чем в старости, у нас не играет существенной роли. Как видно из 2.1. и 2.2., все статьи, входящие в наш материал написаны почти одновременно, т.е. в первой половине 1860-х годов, когда выходили *Время* и *Эпоха*. Несомненно, что это обстоятельство значительно облегчает нашу проблему, придавая нашим результатам большую степень надежности.

С другой стороны следует признать, что остается

проблема изменения языка и стиля в зависимости от жанра. Хотя все интересующие нас тексты относятся к публицистике, но язык и стиль могут меняться с переходом писателя от одной темы к другой. Однако с помощью математической статистики мы можем определить отклонения писателя от нормы и учесть их при сравнении его бесспорных текстов со спорными. Только в том случае, если спорный текст перейдет границы отклонения, поставленные самим Достоевским, мы имеем право исключить возможность того, что текст принадлежит перу Достоевского.

## 2.0. Материал

Материалом нашего исследования в основном служил список статей, присланный нам редакцией академического *Полного собрания сочинений* Достоевского. Всего в корпус бесспорно принадлежащих Достоевскому статей вошли 119 107 слов (7 099 предложений), а в корпус приписываемых ему статей 58 039 слов (2 657 предложений).

### 2.1. Статьи, принадлежащие Достоевскому

В основе корпуса бесспорно принадлежащих Достоевскому статей лежат произведения, имеющиеся в списке Н.Н. Страхова, но с прибавлением пяти статей (№ 9, 11, 13, 14 и 17), принадлежность которых Достоевскому в настоящее время считается бесспорной. Отметим, однако, что две из этих статей (№ 14 и, в частности, № 13) все же довольно сильно отличаются от языкового почерка оставленного Достоевским в других бесспорно принадлежащих ему статьях.

Ниже приводится список 26 статей, несомненно принадлежащих Достоевскому. Статьи сопровождаются краткими комментариями, с указанием длины анализируемого компьютером текста.

#### 1. Ряд статей о русской литературе. I. Введение.

*Время*, 1861, № 1, отд. II, стр. 1–34; *Том*, 13, стр. 32–61; *АИ*, 18, стр. 41–70.

Статья помещена без подписи, но принадлежность ее Достоевскому устанавливается списком Страхова.

Слов 12 455. Предложений 664.

#### 2. Ряд статей о русской литературе. II. Г-н -бов и вопрос об искусстве.

*Время*, 1861, № 2, отд. II, стр. 165–205; *Том*, 13, стр. 62–95; *АИ*, 18, стр. 70–103.

Статья опубликована без подписи, но она указана в

списке Страхова. Кроме того, имеется и текстуальное указание на авторство Достоевского (см. *АИ*, 18, стр. 269).

Слов 10 894. Предложений 506.

3. Ряд статей о русской литературе. III. Книжность и грамотность. Статья первая.

*Время*, 1861, № 7, отд. II, стр. 35–52; *Том*, 13, стр. 96–111; *АИ*, 19, стр. 5–20.

Статья помещена без подписи, но принадлежность ее Достоевскому засвидетельствована Страховым.

Слов 5 805. Предложений 345.

4. Ряд статей о русской литературе. IV. Книжность и грамотность. Статья вторая.

*Время*, 1861, № 8, отд. II, стр. 91–130; *Том*, 13, стр. 111–114; *АИ*, 19, стр 21–57.

Статья напечатана без подписи, но она приписана Достоевскому в списке Страхова.

Слов 11 091. Предложений 603.

5. Ряд статей о русской литературе. V. Последние литературные явления. Газета «День».

*Время*, 1861, № 11, отд. II, стр. 64–75; *Том*, 13, стр. 145–154; *АИ*, 19, стр. 57–66.

Статья помещена без подписи, но принадлежность ее Достоевскому удостоверяется списком Страхова.

Слов 3 300. Предложений 194.

6. Петербургские сновидения в стихах и прозе.

*Время*, 1861, № 1, отд. VI, стр. 1–22; *Том*, 13, стр. 154–172; *АИ*, 19, стр. 67–85.

Статья опубликована без подписи, но она приписана Достоевскому в списке Страхова. По указанию его же, стихотворные вставки написаны Д.Д. Минаевым.

Слов 6 617. Предложений 426.

7. Образцы чистосердечия.

*Время*, 1861, № 3, отд. V, стр. 54–70; *Том*, 13, стр. 172–186; *АИ*, 19, стр. 91–104.

Статья помещена без подписи, но принадлежность ее Достоевскому засвидетельствована Страховым.

Слов 3 651. Предложений 243.

8. «Свисток» и «Русский вестник».  
*Время*, 1861, № 3, отд. V, стр. 71–84; *Том*, 13, стр. 186–198; *АИ*, 19, стр. 105–116.  
Статья опубликована без подписи, но она указана в списке Страхова.  
Слов 4 108. Предложений 242.
9. Примечание к «Письму с Васильевского острова в редакцию 'Времени'» Л. К.  
*Время*, 1861, № 4, отд. IV, стр. 151–152; *Том*, 13, стр. 198–199; *АИ*, 19, стр. 117–118.  
Статья напечатана без подписи. Хотя она в список Страхова не включалась, но авторство Достоевского подтверждается отсылкой в статье «Ответ 'Русскому вестнику'», которая имеется в списке.  
Слов 385. Предложений 17.
10. Ответ «Русскому вестнику».  
*Время*, 1861, № 5, отд. V, стр. 15–39; *Том*, 13, стр. 199–220; *АИ*, 19, стр. 119–139.  
Статья помещена без подписи, но она указана в списке Страхова.  
Слов 6 956. Предложений 394.
11. Литературная истерика.  
*Время*, 1861, № 7, отд. IV, стр. 50–57; *Том*, 13, стр. 221–227; *АИ*, 19, стр. 140–146.  
Статья опубликована без подписи. В списке Страхова ее нет. Она приписана Достоевскому Л.П. Гроссманом (см. *Том*, 13, стр. 596).  
Слов 327. Предложений 19.
12. По поводу элегической заметки «Русского вестника».  
*Время*, 1861, № 10, отд. IV, стр. 189–199; *Том*, 13, стр. 227–235; *АИ*, 19, стр. 169–177.  
Статья помещена без подписи, но она имеется в списке Страхова.  
Слов 2 333. Предложений 175.
13. Два лагеря теоретиков. (По поводу «Дня» и кой-чего другого).  
*Время*, 1862, № 2, отд. II, стр. 143–163; *Том*, 13, стр. 235–253; *АИ*, 20, стр. 5–22.

Статья напечатана без подписи. В список Страхова она не включалась. По содержанию и стилистическим признакам она впервые была приписана Достоевскому Л.П. Гроссманом и О. фон Шульцем.

Слов 7 221. Предложений 382.

14. Объявление о подписке на журнал «Время» на 1863 г.  
*Время*, 1862, № 9, стр. 1–12; Том, 13, стр. 506–512; АИ, 20, стр. 206–212.

Хотя в списке Страхова статья не указана, но принадлежность ее Достоевскому устанавливается *Воспоминаниями* Страхова (см. АИ, 20, стр. 393).

Слов 2 560. Предложений 160.

15. Славянофилы, черногорцы и западники. Самая последняя перепалка.

*Время*, 1862, № 9, отд. II, стр. 132–137; Том, 13, стр. 253–258; АИ, 20, стр. 23–27.

Статья опубликована без подписи, но она внесена в список Страхова.

Слов 1 974. Предложений 137.

16. Щекотливый вопрос. Статья со свистом, с превращениями и переодеваньями.

*Время*, 1862, № 10, отд. II, стр. 141–163; Том, 13, стр. 258–277; АИ, 20, стр. 30–49.

Статья помещена без подписи. Авторство Достоевского устанавливается списком Страхова.

Слов 7 200. Предложений 576.

17. Необходимое литературное объяснение по поводу разных хлебных и нехлебных вопросов.

*Время*, 1863, № 1, отд. II, стр. 29–38; Том, 13, стр. 277–286; АИ, 20, стр. 50–58.

Статья напечатана без подписи. В списке Страхова ее нет. Она впервые по содержанию и стилистическим признакам была приписана Достоевскому Л.П. Гроссманом и О. фон Шульцем.

Слов 3 730. Предложений 226.

18. Журнальная заметка. О новых литературных органах и о новых теориях.

*Время*, 1863, № 1, отд. II, стр. 175–188; *Том*, 13, стр. 286–297; *АИ*, 20, стр. 59–70.

Статья опубликована без подписи, но авторство Достоевского устанавливается списком Страхова.

Слов 4 547. Предложений 269.

19. Журнальные заметки. I. Ответ «Свистуну». II. Молодое перо.

*Время*, 1863, № 2, отд. II, стр. 213–226; *Том*, 13, стр. 297–307; *АИ*, 20, стр. 71–82.

Заметки помещены без подписи. Они вошли в список Страхова, но первая сопровождена вопросительным знаком. Однако как Томашевский, так и редакция *АИ*, не сомневаются в принадлежности обеих заметок Достоевскому (см. *Том*, 13, стр. 597; *АИ*, 20, стр. 299–300).

Слов 4 265. Предложений 304.

20. Опять «Молодое перо». Ответ на статью «Современника» «Тревоги 'Времени'» («Современник», март, № 3).

*Время*, 1863, № 3, отд. II, стр. 148–163; *Том*, 13, стр. 309–322; *АИ*, 20, стр. 83–96.

Статья опубликована без подписи, но принадлежность ее Достоевскому засвидетельствована Страховым.

Слов 4 605. Предложений 326.

21. Господин Щедрин или Раскол в нигилистах.

*Эпоха*, 1864, № 5, стр. 274–294; *Том*, 13, стр. 322–340; *АИ*, 20, стр. 102–120.

Статья напечатана без подписи, но она включена в список Страхова. Кроме того, сам Достоевский признавал авторство этой статьи (см. *Гросс*, 23, стр. 280).

Слов 6 155. Предложений 466.

22. Несколько слов о Михаиле Михайловиче Достоевском.

*Эпоха*, 1864, № 6, стр. I–VI; *Том*, 13, стр. 340–343; *АИ*, 20, стр. 121–124.

Статья подписана «Федор Достоевский» и указана в списке Страхова.

Слов 1 353. Предложений 63.

23. Необходимое заявление.

Эпоха, 1864, № 7, отд. IX, стр. 1–4; Том, 13, стр. 343–346; АИ, 20, 125–128.

Статья подписана «Федор Достоевский» и включена в список Страхова.

Слов 881. Предложений 45.

24. Чтобы кончить. Последнее объяснение с «Современником».

Эпоха, 1864, № 9, отд. X, стр. 1–6; Том, 13, стр. 346–350; АИ, 20, 129–132.

Статья помещена без подписи, но она имеется в списке Страхова.

Слов 1 208. Предложений 63.

25. Примечание.

Эпоха, 1864, № 9, стр. 51–55; Том, 13, стр. 350–353; АИ, 20, стр. 133–136.

Статья напечатана за подписью «Ф. Достоевский» и указана в списке Страхова.

Слов 1 484. Предложений 70.

26. Каламбуры в жизни и в литературе.

Эпоха, 1864, № 10, отд. V, стр. 20–32; Том, 13, стр. 354–364; АИ, 20, стр. 137–147.

Статья опубликована без подписи. Авторство Достоевского удостоверяется списком Страхова.

Слов 4 002. Предложений 184.

## 2.2. Статьи, приписываемые Достоевскому

Ниже приводится список 12 статей, которые в разное время различными исследователями приписывались Достоевскому. У Б.В. Томашевского обсуждается и ряд других статей и заметок (см. Том, 13, стр. 610–614). Но во всех этих случаях аргументы против авторства Достоевского представляются вполне убедительными (17).

Что касается выбранных для проверки статей, то следует отметить, что выдвинутые для них аргументы в пользу авторства Достоевского обладают неодинаковой силой. Если

в одних случаях (например, № IV и № VII говорилось об авторстве Достоевского вполне уверенно, то в других же случаях (например, № X и № XII) исследователями допускалось, что статьи эти могли быть написаны и другими лицами.

Далее мы вернемся к конкретному обсуждению проблемы авторства Достоевского для указанных 12 статей. Здесь же приведем их названия, сопровождая каждое краткими комментариями и указанием длины текста.

I. Письмо Постороннего критика в редакцию нашего журнала по поводу книг г-на Панаева и «Нового поэта».

*Время*, 1861, № 1, отд. II, стр. 46–64; *АИ*, 27, стр. 129–145.

Статья помещена за подпись «Посторонний критик». Приписал ее Достоевскому О. фон Шульц, оговорившись, правда, что «пока не имеет для этого вполне достаточных доказательств» (*Schoultz*, pp. 7–8). Атрибуция фон Шульца была отвергнута Б.В. Томашевским, который считал свидетельство С.А. Венгерова о принадлежности статьи А.Ф. Писемскому «достаточно авторитетным» (Том, 13, стр. 609). Со ссылкой на атрибуцию Венгерова статья указана И.Ф. Масановым в числе принадлежащих Писемскому в *Словаре псевдонимов русских писателей, ученых и общественных деятелей* (т. II, Москва 1957, стр. 374). Не сомневается в авторстве Писемского и А.С. Долинин (см. *Ф.М. Достоевский. Письма. I*, Москва-Ленинград 1928, стр. 523). Того же мнения В.С. Нечаева (см. *Журнал М.М. и Ф.М. Достоевских «Эпоха». 1864–1865*, Москва 1975, стр. 234), отмечая, однако, что в редакторской книге журнала *Время* имени Писемского нет (там же, стр. 261). В.А. Туниманов, комментируя статью для т. 27 академического *Полного собрания сочинений* Достоевского, в котором она перепечатывается впервые, отвергает свидетельство Венгерова о

принадлежности статьи Писемскому. Нелогично было бы предположить, пишет исследователь, что Писемский, будучи в то время редактором *Библиотеки для чтения*, скрывшись за псевдонимом «Посторонний критик», избрал новый журнал *Время* трибуной для нападок на собственный журнал. Зато, продолжает Туниманов, есть многое в этой статье, которое объединяет ее с polemическими статьями Достоевского. «Совокупность приведенных аргументов», заключает учёный, «позволяет с большой долей уверенности включить «Письмо постороннего критика» в раздел 'Dubia'» (*АИ*, 27, стр. 400).

Слов 5 804. Предложений 338.

II. Гаваньские чиновники в домашнем быту, или Галерная гавань во всякое время дня и года. (Пейзаж и жанр) Ивана Генслера. «Библиотека для чтения». Ноябрь и декабрь 1860.

*Время*, 1861, № 2, отд. II, стр. 139–150. *АИ*, 27, 145–155.

Рецензия напечатана без подписи. Предположительно внесена Б.Ф. Егоровым в «Библиографию критики и художественной прозы Ап. Григорьева» (см. Ученые записки Тартуского государственного университета, вып. 98, 1960, стр. 236). К мнению Егорова присоединяется академик В.В. Виноградов, отмечая, однако, в рецензии правку Достоевского (см. *Русскую литературу*, 1964, № 2, стр. 77). А.А. Григорьев числится автором статьи и в росписи В.С. Нечаевой, хотя с вопросительным знаком (*Журнал М.М. и Ф.М. Достоевских «Эпоха». 1864–1865*, Москва 1975, стр. 235). В *АИ* внесена В.А. Тунимановым в отдел «Dubia».

Слов 3 328. Предложений 203.

III. Противоречия и увлечения «Времени».

*Время*, 1861, № 8, отд. IV, стр. 135–142; *АИ*, 27, стр. 155–161.

Статья опубликована без подписи. Несмотря на

то, что статья, по мнению Л.П. Гроссмана, близко напоминает «идейный строй и литературную манеру Достоевского», она при отсутствии более точных данных не была включена им в собрание сочинений писателя (см. *Гросс*, 22, стр. XXV-XXVI). Б.В. Томашевский тоже пишет, что «достаточных указаний на принадлежность ее Достоевскому не имеется» (*Том*, 13, стр. 611). По мнению В.С. Нечаевой, перед нами статья, несомненно принадлежащая редакции журнала (см. *Журнал М.М. и Ф.М. Достоевских «Эпоха»*. 1864–1865, Москва 1975, стр. 263). У Нечаевой автором статьи числится М.М. Достоевский, но с вопросительным знаком (там же, стр. 240). В *АИ* включена В.А. Тунимановым в раздел «*Dubia*».

Слов 1 988. Предложений 123.

#### IV.

Выставка в Академии художеств за 1860–1861 год.

*Время*, 1861, № 10, отд. II, стр. 147–168; *Том*, 13, стр. 529–547; *АИ*, 19, 151–168.

Статья помещена без подписи. Принадлежность статьи Достоевскому доказывалась уже Л.П. Гроссманом, включившим ее в собрание сочинений писателя (*Гросс*, 22, стр. 108–141). Полную уверенность в написании ее Достоевским выражал и О. фон Шульц. Более осторожен Б.В. Томашевский, включивший статью в отдел статей, приписываемых Достоевскому. Предположение Томашевского о возможности ее написания Платоном Кусковым (*Том*, 13, стр. 608) убедительно опровергнуто В.С. Нечаевой, выдвинувшей, в свою очередь, предположение о том, что первые 5 страниц написаны Достоевским, а следующие 22 страницы Я.П. Полонским (см. *Журнал М.М. и Ф.М. Достоевских «Время»*. 1861–1863, Москва 1972, стр. 263–264). Подробно обсуждается вопрос об авторстве этой статьи Г.М. Фридлендером, включившим ее – хотя и с оговорками о возможности участия в ее написании М.М. Достоевского – в основной

корпус академического издания. Поскольку данная статья безусловно самая значительная из спорных статей в нашем корпусе, мы в дальнейшем будем уделять ей особое внимание.

Слов 7 229. Предложений 380.

V. Ряд статей о русской литературе. Вопрос об университетах.

*Время*, 1861, № 11, отд. II, стр. 76–104; *АИ* 19, 187–210.

Статья напечатана без подписи. Она является второй из двух статей, помещенных под общим заголовком «Ряд статей о русской литературе. Статья пятая». Принадлежность Достоевскому первой из этих статей удостоверяется списком Н.Н. Страхова (см. настоящее исследование, 2.1., № 5). Вторая же статья обычно приписывалась М.М. Достоевскому. Признавая вероятность принадлежности первой главки второй статьи М.М. Достоевскому, В.С. Нечаева приписывала вторую главку статьи Ф.М. Достоевскому (см. *Журнал М.М. и Ф.М. Достоевских «Время». 1861–1863*, Москва 1972, стр. 144–148). Г.М. Фридлендер, включивший статью в отдел «Приложения» к т. 19 академического *Полного собрания сочинений* писателя, пришел к следующему выводу: «Скорее всего Ф.М. Достоевский принимал значительное (если не основное) участие в написании начала и конца статьи, материал для средней ее части (стр. 188–202) подготовлен кем-то другим (М.М. Достоевским? Н.Н. Страховым?); хотя и она, по-видимому, несет на себе определенные следы редактуры писателя» (*АИ*, 19, стр. 339).

Слов 5 571. Предложений 320.

VI. Николай Александрович Добролюбов. Некролог.

*Время*, 1861, № 11, отд. IV, стр. 31–32.

Статья помещена без подписи. На возможную принадлежность ее Достоевскому впервые ука-

зала Е.Н. Дрыжакова (см. ее доклад «Достоевский и Добролюбов» в кн.: *Первая всесоюзная межвузовская конференция по проблемам изучения и преподавания литературной критики в высшей школе*, Ленинград 1974, стр. 53–60).

Слов 261. Предложений 15.

VII. Рассказы Н.В. Успенского.

*Время*, 1861, № 12, отд. II, стр. 174–183, *Том*, 13, 547–555; *АИ*, 19, 178–186.

Статья опубликована без подписи. Она впервые была приписана Достоевскому Л.П. Гроссманом, включившим ее в собрание сочинений писателя (*Гросс*, 22, стр. 145–160). Принадлежность статьи Достоевскому доказывает и О. фон Шульц, (*Schoultz*, р. 41 sq.). Однако, по мнению Б.В. Томашевского, включившего ее только в отдел «*Dubia*», вопрос не следует считать решенным. С другой стороны, почти полную уверенность в принадлежности статьи Достоевскому высказывает Г.М. Фридлендер, включивший ее в основной корпус *АИ*.

Слов 3 443. Предложений 167.

VIII. Полемический случай с «Основой» и «Сионом».

*Время*, 1861, № 12, отд. IV, стр. 114–116.

Статья помещена без подписи. Упоминая статью как близко напоминающую «идейный строй и литературную манеру Достоевского», Л.П. Гроссман все же при отсутствии более точных данных не считал возможным включить ее в собрание сочинений писателя (см. *Гросс*, 22, стр. XXV–XXVI). Б.В. Томашевский тоже считает, что нет достаточных указаний на принадлежность статьи Достоевскому (см. *Том*, 13, стр. 611). Следуя за Томашевским, В.С. Нечаева пишет, что принадлежность ее Достоевскому «не доказана» (см. *Журнал М.М. и Ф.М. Достоевских «Эпохи»*. 1864–1865, Москва 1975, стр. 265).

Слов 816. Предложений 26.

- IX. Дворянин, желающий быть крестьянином.  
*Время*, 1861, № 12, отд. IV, стр. 117–123.  
Статья напечатана без подписи. Отмечая в статье сходство с журнальными приемами Достоевского, Б.В. Томашевский все же пишет, что нет достаточных оснований включать ее в собрание сочинений писателя. Со ссылкой на Томашевского В.С. Нечаева также заключает, что принадлежность статьи Достоевскому «не доказана» (см. *Журнал М.М. и Ф.М. Достоевских «Эпоха». 1864–1865*, Москва 1975, стр. 265).  
Слов 2 091. Предложений 74.
- X. Политическое обозрение.  
*Эпоха*, 1864, № 9, стр. 1–26.  
Статья помещена без подписи. Судя по редакторской книге журнала, обозрение это принадлежит А.А. Головачеву. На возможную принадлежность начала его Достоевскому впервые указала В.С. Нечаева (см. *Журнал М.М. и Ф.М. Достоевских «Эпоха». 1864–1865*, Москва 1975, стр. 80–82, 270).  
Слов 8 946. Предложений 380.
- XI. Наши домашние дела.  
*Эпоха*, 1864, № 12, стр. 1–28.  
Статья напечатана без подписи. На возможное участие в ней Достоевского впервые указала В.С. Нечаева. Считая, что первые шесть страниц обозрения принадлежат Достоевскому, она приписывает остальную его часть А.У. Порецкому, написавшему в Эпохе ряд статей с тем же заголовком (см. *Журнал М.М. и Ф.М. Достоевских «Эпоха». 1864–1865*, Москва 1975, стр. 44–47, 259, 271).  
Слов 8 391. Предложений 277.
- XII. Политическое обозрение. Общий обзор главнейших политических событий прошедшего года.  
*Эпоха*, 1864, № 12, стр. 1–32.  
Статья опубликована без подписи. Исходя из

сведений редакторской книги журнала, В.С. Нечаева автором статьи считает К. Немшевича, хотя и с вопросительным знаком. Вместе с тем она предполагает участие в ней Достоевского (см. *Журнал М.М. и Ф.М. Достоевских «Эпоха». 1864–1865*, Москва 1975, стр. 83–84, 259, 271).

Слов 10 171. Предложений 354.

### 3.0. Принципы выбора материала

Надежность лингвостатистических методов в значительной степени увеличивается объемом и однородностью материала, выбранного для сопоставления. В идеале следует выбирать только авторскую речь, исключая разговоры и мысли персонажей. Очень строго соблюдался этот принцип в выборе материала для скандинавского исследования авторства *Тихого Дона* (18). Единицей выбора материала для сопоставления в этом исследовании служил абзац. Но если в каком-нибудь абзаце были малейшие признаки, кроме авторской, речи персонажа, – абзац этот заменялся другим, содержащим только авторскую речь.

Необходимо признать, что на этот раз перед нами задача гораздо более трудная. Дело в том, что некоторые из присланных для проверки спорных текстов состоят всего лишь из двух-трех страниц, что сильно усложняет решение задачи. Надеясь кое-что сказать и о возможности написания этих текстов Достоевским, мы решились менее критически относиться к выбору материала, включая в него и разговоры и вопросы. Представляется, что это решение для нашей цели имеет как отрицательные, так и положительные последствия. Например, допущение в корпус Достоевского разговоров и вопросов явно угрожает однородности материала, так как текст, в котором содержатся реплики, довольно резко отличается от текста, состоящего из сплошной авторской речи (так, средняя длина предложения будет в текстах с репликами неизбежно короче). Иными словами, создается опасность установления слишком «широкого» стилистического почерка для Достоевского, позволяя нам приписать ему тексты, ему не принадлежащие. С другой стороны, отличительной чертой публицистики Достоевского начала 1860-х годов являются как раз разговоры с читателями,

содержащие много вопросительных и восклицательных обращений. Для исследователя, желающего установить языковой почерк Достоевского-публициста, трудно обойтись без этого материала. Поэтому мы решились выбрать как можно больше предложений из текстов Достоевского, придерживаясь, разумеется, тех же правил при выборе материала из приписываемых ему текстов.

Итак, для перевода текстов на язык компьютера мы решились включить все предложения, написанные прозой и явно принадлежащие перу Достоевского, даже в тех случаях, когда он подражает другим писателям (см., например, Том, 13, стр. 304) или создает так называемые «романы-карикатуры» (см., например, Том, 13, стр. 334). Предложения, в которых содержатся, цитаты, принимались с условием, чтобы чужой материал не превышал 10 слитно написанных слов. Таким образом, включались предложения типа: «'Придирка, да еще смешная!' скажут нам просветители» (Том, 13, стр. 114). Включались в материал и предложения, в которых содержатся заглавия книг, картин, статей, за исключением случаев перечислений, таких заглавий (например, Том, 13, стр. 540). С другой стороны, исключались из материала все предложения, взятые Достоевским у других писателей, а также сноски, стихотворные вставки и заголовки статей.

## 4.0. Лингвостатистические параметры

На основании предварительных исследований языковой почерк Достоевского-публициста мы сочли целесообразным установить, исходя из следующих 15 параметров:

1. Общее распределение частей речи в первых двух и в последних трех позициях предложения.
2. Распределение частей речи в первой позиции предложения.
3. Распределение частей речи во второй позиции предложения.
4. Сочетание частей речи в первых двух позициях предложения.
5. Распределение частей речи в третьей с конца позиции предложения.
6. Распределение частей речи в предпоследней позиции предложения.
7. Распределение частей речи в последней позиции предложения.
8. Сочетание частей речи в последних трех позициях предложения.
9. Средняя длина слова в буквах, вычисляемая на основании выборок размером в 500 текстовых слов.
10. Общее распределение длины слова.
11. Средняя длина предложения в словах, вычисляемая на основании выборок размером в 30 предложений.
12. Общее распределение длины предложения.
13. Лексический спектр текста на уровне словаря.
14. Лексический спектр текста на уровне текста.
15. Индекс разнообразия лексики.

Далее мы будем давать отчет о применении этих параметров и о полученных результатах.

## 4.1. Общее распределение частей речи

Одним из самых важных параметров в стилеметрике является общее распределение частей речи в языке писателя. В самом деле, этот параметр считается настолько важным, что у одних исследователей он используется даже как единственный параметр при «дискриминации» различных писателей (19). Вообще давно установлено, что один писатель часто отличается от другого именно своим предпочтением отдельных частей речи. Так, широкое употребление глаголов невольно придает стилю писателя стремительный, динамический характер, тогда как частое применение имен существительных является типичной чертой стиля писателя, занимающегося философскими и абстрактными темами. В качестве примеров можно привести богатый глаголами язык народных сказок в отличие от научной прозы, или язык народного поэта А.В. Кольцова в отличие от языка поэта-мыслителя Е.А. Баратынского (20).

Особенно выделяется своеобразие писателя в распределении и сочетании частей речи в начале и в конце предложения. В результате указанные выше параметры 1–8 играют большую роль и в определении авторства произведения. Так, применением этих параметров отчетливо отличались друг от друга такие писатели, как К.М. Симонов и К.Г. Паустовский (21). Весьма показательными были эти параметры также в нашем исследовании об авторстве *Тихого Дона* (22). Не менее полезными оказались они в настоящем исследовании.

Принимая эти параметры в нашем исследовании, мы пользовались ручным кодированием предложений, а потом обрабатывали полученный материал с помощью ЭВМ. При этом мы старались применять определения частей речи, предлагаемые в авторитетной *Грамматике русского языка*, изданной АН СССР. Правда, эти определения были приняты с некоторыми отклонениями, вызванными тем обстоятельством, что в нашем распоряжении было всего только 10 цифр (0–9), из которых две были предназначены для обозначения запятой (цифра 9) и терминатора, т.е. точки, многоточия, вопросительного и восклицательного знаков (цифра 8). В результате некоторые категории пришлось расширить. Так, в категорию «Наречие» вошли, кроме наречий, и

частицы, а также междометия. Но беда здесь невелика. Главное, чтобы все тексты, бесспорные и спорные, были одинаково обработаны.

Ниже приводятся цифры, обозначающие категории или части речи:

0 = Имя прилагательное.

Например: бел(ый), второй, измученный (вид).

## 1 = Предлог.

Например: в(о), благодаря.

## 2 = Наречие.

Например: весело (играть), даже, ну.

3 = Союз.

Например: и, если, чтобы(ы).

#### 4 = Местоимение.

Например: он, этот, который.

5 = Имя существительное (во всех функциях кроме функции подлежащего).

Например: человек, Иванов, больной, семья, несколько.

6 = Имя существительное (в функции подлежащего, Сущ./Подл.).

7 = Глагол.

Например: видеть, увидя, (он был) убит.

8 = Терминатор (точка, восклицательный и вопросительный знаки, многоточие).

9 = Запятая.

Само кодирование – двухцифровое: первая цифра обозначает позицию, занимаемую словом в предложении, а вторая цифра – категорию и часть речи в соответствии с принятыми определениями. Таким образом, в кодированной форме каждое предложение принимает вид телефонного номера, состоящего из пяти двухцифрованных чисел.

Приведем для ясности несколько примеров принимаемого нами кодирования:

1      2      3      4      5

/1/ Для Европы Россия – одна из загадок Сфинкса.

1 5 1 5 5

Предл. Сущ. Предл. Сущ. Сущ.

1            2  
3            4            5  
/2/            Любовь города берет.  
                  6            5            7  
                  Сущ./Подл. Сущ. Глагол  
Кодирование: 16 25 36 45 57

1    2            3    4    5  
/3/            Это, положим, так.  
                  4    9            7    9    2  
                  Местоим. Точ.    Гл. Точ. Нареч.  
Кодирование: 14 29 37 49 52

1            2  
3            4    5  
/4/            Мы серьезно.  
                  4            2    8  
                  Местоим. Нареч. Терминатор  
Кодирование: 14 22 34 42 58

Как явствует из примеров, в длинных предложениях кодируются только первых два (1, 2) и последних три (3, 4, 5) слова. В предложениях, содержащих пять слов, кодируются все слова, а в редких случаях, когда предложения (без запятых) состоят лишь из трех или четырех слов, имеет место двойное кодирование отдельных слов. То же самое относится к предложениям, состоящим из двух слов, причем в таких случаях терминатор (.!? . .) рассматривается как часть речи в последней (5) позиции предложения. Наконец следует отметить, что предложения в одно слово при применении данного параметра не учитывались. Не принимались также немногочисленные предложения, имевшие в интересующих нас позициях другие знаки препинания, кроме запятых и терминаторов. Исключается, например, предложение «Книги на лицо; можно сравниться», имеющее

точку с запятой в третьей с конца позиции предложения. В результате число предложений, которые обрабатывались для параметров 1–8 было несколько ниже (6 669), чем число предложений, рассматриваемых для других параметров (7 099).

Разумеется, полученные «телефонные номера» легко обрабатывались компьютером. В качестве вывода информации нам были предоставлены списки с сообщением о том, какие части речи употребляются в текстах и в какие сочетания они входят.

Возьмем для примера «Петербургские сновидения в стихах и прозе», – текст, бесспорно принадлежащий Достоевскому. Как явствует из Таблицы 1, в этом тексте в первой (1) позиции предложения имеется 44 случая употребления глагола (7), см. коду 17, а в последней (5) позиции предложения 60 случаев употребления глагола (7), см. коду 57.

*Таблица 1*

*Общее распределение частей речи в первых двух и в последних трех позициях предложения. «Петербургские сновидения в стихах и прозе»*

<i>Кода</i>	<i>Число</i>
10:	9
11:	27
12:	93
13:	89
14:	98
15:	23
16:	14
17:	44
20:	23
21:	28
22:	73
23:	13
24:	76
25:	45
26:	20
27:	68

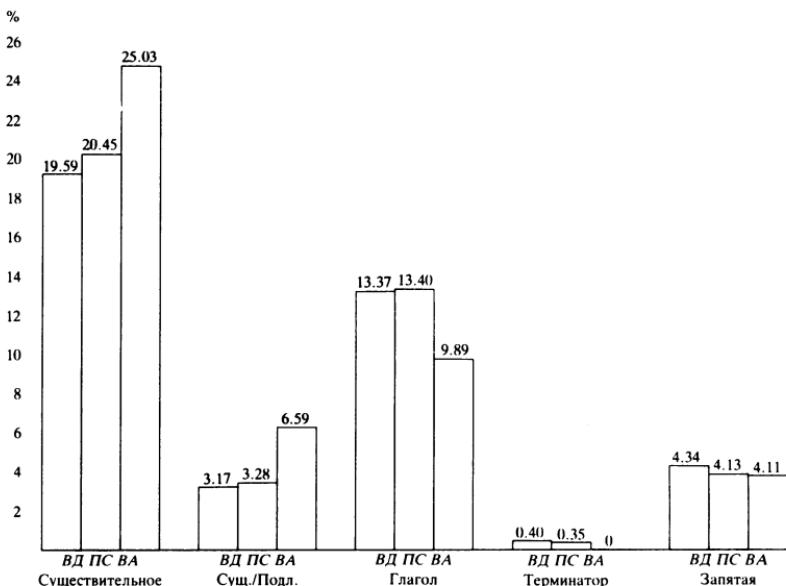
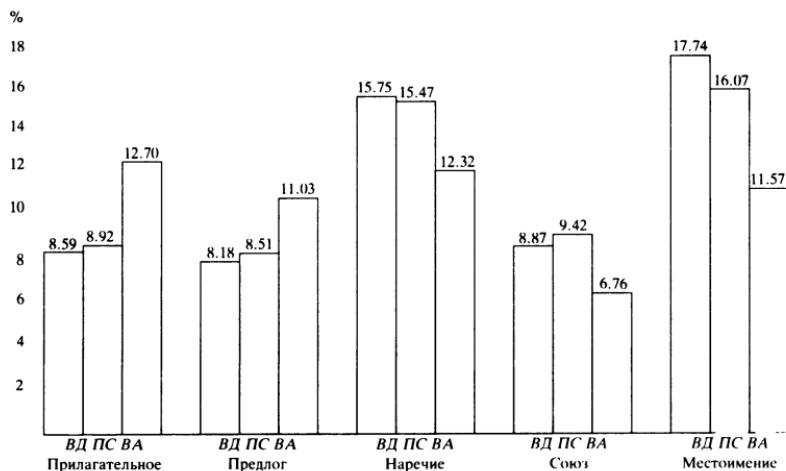
29:	51
30:	29
31:	66
32:	62
33:	49
34:	56
35:	53
36:	6
37:	53
39:	23
40:	89
41:	48
42:	54
43:	36
44:	69
45:	50
46:	2
47:	41
49:	8
50:	27
52:	25
54:	20
55:	235
56:	23
57:	60
58:	7
Итого:	1 985

Получив таким образом информацию об общем распределении частей речи во всех текстах, принадлежащих Достоевскому, мы можем провести сопоставление с соответствующей информацией, извлеченной из приписываемых ему текстов.

Дискриминирующая сила рассматриваемого параметра представляется вполне удовлетворительной. Так, на Рисунке 1 указано общее распределение частей речи (в %) в трех текстах. В первой колонке представлена совокупность 26 бесспорно принадлежащих Достоевскому текстов («Весь Достоевский» = *ВД*). Потом, во второй колонке, представлен бесспорно принадлежащий Достоевскому текст «Петер-

бургские сновидения в стихах и прозе» (ПС). И, наконец, в третьей колонке, представлен спорный текст «Выставка в Академии художеств за 1860–1861 год» (ВА).

Рисунок 1  
Общее распределение частей речи



ВД = Весь Достоевский  
ПС = «Петербургские сновидения в стихах и прозе»  
ВА = «Выставка в Академии художеств за 1860–1861 год»

Бросается в глаза, что *ПС* имеет профиль гораздо ближе к *ВД*, чем *ВА*. Так, в *ВА* употребляется значительно больше имен прилагательных, предлогов и имен существительных, тогда как активность наречий, союзов, местоимений и глаголов заметно ниже, чем в *ВД* и *ПС*. Например, соотношение «глагол : имя существительное» измеряется в *ВД* и *ПС* коэффициентом 0,59 и 0,56, а в *ВА* коэффициентом 0,31. И соотношение «местоимение : имя существительное» в *ВД* измеряется величиной 0,78, а в *ВА* 0,37. Ниymi словами, публицистика Достоевского дает в среднем 78 местоимений на 100 имен существительных, а *ВА* всего 37. Нельзя не предположить, что за такими цифрами стоят не только изменения функционирования элементов языковой структуры и ее разновидностей, но и различия психологии творчества, различия видения писателем действительности и способов ее речевого отображения и изображения (23). В каком-то смысле речь Достоевского по сравнению с речью автора «Выставки в Академии художеств за 1860–1861 год» заметно более «действенна» и «динамична», и заметно менее «предметна».

Теперь возникает вопрос: можно ли с помощью теории вероятности исключить возможность того, что общее распределение частей речи в первых двух и в последних трех позициях предложения в *ВА* производится писателем, имеющим общее распределение частей речи в первых двух и в последних трех позициях предложения Достоевского? Другими словами, есть ли у нас основание думать, что профиль *ВА* и профиль *ВД* имеют общее происхождение?

Математическая статистика имеет в своем распоряжении орудие (критерий «хи-квадрат»), позволяющее нам установить вероятность происхождения текстов *ПС* и *ВА* из совокупности текстов *ВД*. Словом, с помощью сравнения текстов *ПС* и *ВА* с текстами *ВД* мы можем проверить, написаны ли эти тексты Достоевским или нет.

Сравнивая различные категории или части речи (в %) в сопоставляемых текстах неодинаковой длины, мы пользовались формулой:

$$x^2 = \frac{n_1 \cdot n_2}{10\ 000} \sum \frac{(p_1 - p_2)^2}{f_1 + f_2},$$

где  $n$  обозначает длину текста, и  $f$  и  $p$  частоту каждой категории в числах и процентах (24). В Таблице 2 проведено сравнение между *ВА* и *ВД*, причем  $k$  обозначает число категорий.

Таблица 2  
Сравнение между *ВА* ( $f_1$ ), и *ВД* ( $f_2$ )

Кат.	$f_1$	$f_2$	$p_1 \%$	$p_2 \%$	$p_1 - p_2$	$(p_1 - p_2)^2$	$f_1 + f_2$	$\frac{(p_1 - p_2)^2}{f_1 + f_2}$
0	235	2 866	12,70	8,59	4,11	16,8921	3 101	0,0054473
1	204	2 726	11,03	8,18	2,85	8,1225	2 930	0,0027722
2	228	5 253	12,32	15,75	-3,43	11,7649	5 481	0,0021465
3	125	2 959	6,76	8,87	-2,11	4,4521	3 084	0,0014436
4	214	5 915	11,57	17,74	-6,17	38,0689	6 129	0,0062113
5	463	6 535	25,03	19,59	5,44	29,5936	6 998	0,0042289
6	122	1 056	6,59	3,17	3,42	11,6964	1 178	0,0099290
7	183	4 457	9,89	13,37	-3,48	12,1104	4 640	0,0026100
8	0	132	0	0,40	-0,40	0,1600	132	0,0012121
9	76	1 446	4,11	4,34	-0,23	0,0529	1 522	0,0000348
$k=10$	1 850	33 345	100,00	100,00	—	—	35 195	0,0360357
	$n_1$	$n_2$						

Получается:

$$\frac{n_1 \cdot n_2}{10 \ 000} = \frac{1 \ 850 \cdot 33 \ 345}{10 \ 000} = 6 \ 168 \ 825$$

$$\sum \frac{(p_1 - p_2)^2}{f_1 + f_2} = 0,0360357$$

$$x^2 = 6 \ 168 \ 825 \cdot 0,0360357$$

$$x^2 = 222,30 \text{ (с помощью машины 222,31).}$$

Критические значения «хи-квадрат» при 9 степенях свободы ( $df = k_0 = k-1 = 10-1 = 9$ ) указаны в статистической таблице:

$$x^2_{0,05;9} = 16,92; x^2_{0,01;9} = 21,67; x^2_{0,001;9} = 27,88$$

Оказывается, что если *ПС* по сравнению с *ВД* имеет числовое значение «хи-квадрат» в 5,21, то *ВА* имеет числовое значение «хи-квадрат» в 222,31. Последний результат намного превышает критическое значение коэффициента в 21,67 (при 99% надежности и 9 степенях свободы). Из этого следует, что если в *ПС* авторство Достоевского подтверждается, то оно в *ВА* явно отпадает. Таким образом, мы можем отвергнуть «нулевую гипотезу» о том, что между языком *ВА* и *ВД* нет существенной разницы. В самом деле разница настолько велика, что – судя по параметру «Общее распределение частей речи в первых двух и в последних трех позициях предложения» – текст *ВА* не может быть написанным Достоевским.

Необходимо учесть, однако, что и среди текстов, бесспорно принадлежащих Достоевскому, имеются такие, которые по сравнению с *ВД* превышают критическое значение «хи-квадрат» на уровне 0,01 ( $\chi^2 = 21,67$ ). В Таблице 3 указаны числовые значения «хи-квадрат» для всех интересующих нас текстов.

Таблица 3

*Общее распределение частей речи в первых двух и в последних трех позициях предложения. Числовые значения «хи-квадрат» по сравнению с ВД.*

Текст	$\chi^2$
1	40,08
2	25,59
3	22,67
4	36,70
5	12,47
6	5,21
7	33,23
8	18,42
9	9,31
10	10,60
11	10,78
12	19,44
13	75,91
14	46,66

15	11,29
16	22,09
17	13,08
18	28,05
19	13,65
20	26,16
21	26,04
22	10,55
23	11,15
24	5,54
25	14,51
26	31,83
I	16,39
II	57,49
III	30,03
IV	222,31 >max
V	38,87
VI	20,88
VII	16,13
VIII	11,22
IX	44,41
X	268,11 >max
XI	253,71 >max
XII	396,08 >max

Как видно из таблицы, результаты для отдельных текстов Достоевского показывают известные отклонения от сборного корпуса *ВД*. Самый высокий коэффициент (75,91) имеет текст № 13, т.е. не включавшаяся в список Н.Н. Страхова статья «Два лагеря теоретиков». Из этого следует, что на сей раз нельзя исключить возможности написания Достоевским статей с коэффициентами ниже 75,91. Однако, как показывает таблица, тексты IV, X, XI и XII имеют коэффициенты, намного превышающие данный коэффициент. Все эти статьи имеют распределение частей речи в начале и в конце предложения настолько отличное от «подлинного» Достоевского, что предположение об авторстве писателя надо отвергнуть.

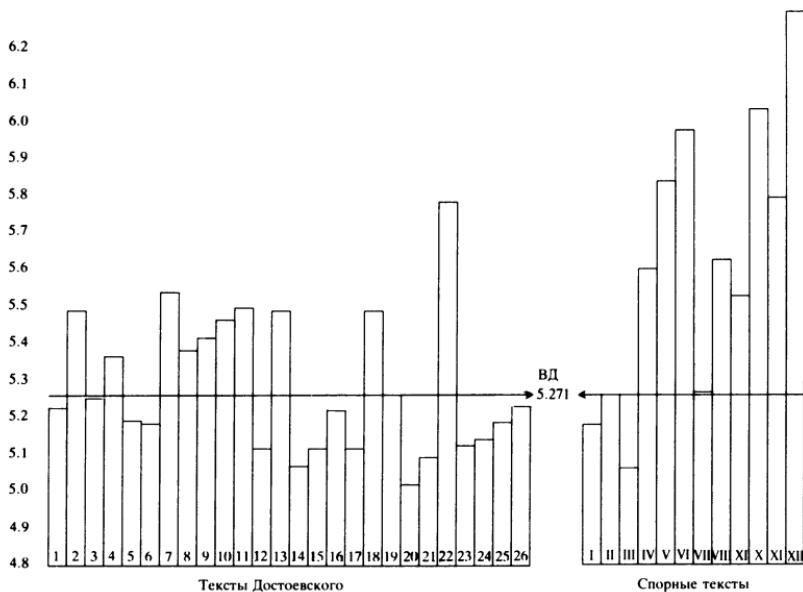
Подобные сопоставления были предприняты по параметрам 2–8. Результаты вычислений показаны в Таблице 14.

## 4.2. Средняя длина слова в буквах

Принимая в нашем исследовании параметр «длина слова», мы определяем «слова» как последовательности букв, отделенные друг от друга пробелом или знаком препинания. Таким образом, в предложении «Г-н -бов не столько критик, сколько публицист» (АИ, 18, стр. 81) компьютером регистрируются 7 слов.

На Рисунке 2 представлены данные о средней длине слова.

Рисунок 2  
Средняя длина слова в  
буквах



Как видно из рисунка, тексты Достоевского имеют в среднем 5,271 буквы в каждом слове. Конечно, наблюдаются известные отклонения от этого среднего значения, особенно в коротком тексте 22, но в общем тексты Достоевского оставляют впечатление однородности. Зато заметно отличаются от текстов Достоевского некоторые из спорных текстов, в частности длинные тексты X, XI и XII.

Для того, чтобы провести тест исключительности, нужно

располагать данными о средней длине слова в выборках одинаковой длины. Целесообразной представляется выборка в 500 текстовых слов, которая, по нашему опыту, несет достаточно большую информацию о колебаниях языковых частот. В результате не поддаются измерению тексты 9, 11 и VI, в которых длина текста меньше, чем 500 словоформ.

В Таблице 4 приводятся статистические данные, необходимые для решения стоящей перед нами проблемы о возможной принадлежности спорных текстов Достоевскому. Применяются в таблице следующие статистические символы:

- $n$  = число выборок;  
 $\bar{m}$  = средняя длина слова в выборке (или в выборках, если текст состоит из нескольких выборок);  
 $s^2$  = дисперсия;  
 $s$  = среднее квадратичное отклонение.

Таблица 4  
*Средняя длина слова. Статистические данные*

Текст	$n$	$\bar{m}$	$s^2$	$s$
1	24	5,214	0,094	0,307
2	21	5,456	0,037	0,193
3	11	5,261	0,059	0,243
4	22	5,325	0,092	0,303
5	6	5,172	0,159	0,398
6	13	5,174	0,089	0,298
7	7	5,501	0,082	0,285
8	8	5,349	0,088	0,296
9	—	—	—	—
10	13	5,427	0,045	0,213
11	—	—	—	—
12	4	5,087	0,027	0,164
13	14	5,429	0,033	0,182
14	5	5,033	0,013	0,115
15	3	5,107	0,013	0,115
16	14	5,227	0,042	0,205

17	7	5,120	0,037	0,193
18	9	5,444	0,035	0,186
19	8	5,295	0,048	0,219
20	9	5,020	0,030	0,172
21	12	5,078	0,055	0,235
22	2	5,793	0,335	0,578
23	1	5,166	—	—
24	2	5,198	0,043	0,206
25	2	5,144	0,277	0,526
26	8	5,217	0,035	0,187
<i>ВД</i>	225	5,277	0,078	0,278
 I	11	5,181	0,032	0,178
II	6	5,280	0,018	0,136
III	3	5,049	0,012	0,111
IV	14	5,539	0,057	0,239
V	11	5,703	0,137	0,370
VI	—	—	—	—
VII	6	5,229	0,046	0,214
VIII	1	5,542	—	—
IX	4	5,493	0,017	0,131
X	17	5,936	0,052	0,228
XI	16	5,690	0,074	0,271
XII	20	6,169	0,025	0,159

Так как выборки Достоевского достаточно однообразны (можно даже говорить о нормальном распределении), мы имеем право провести тест исключительности с помощью известного критерия Стьюдента. Формула этого критерия такова:

$$t = \frac{\bar{m}_1 - \bar{m}_2}{sd} \quad \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}$$

В этой формуле  $\bar{m}_1$  и  $\bar{m}_2$  – сравниваемые средние частоты,  $n_1$  и  $n_2$  – число выборок в двух различных сериях, и  $sd$  – несмешенная оценка среднего квадратичного отклонения в двух сериях выборок, вычисляемая для сравнения двух средних частот по формуле:

$$sd = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$$

Для примера будем сравнивать *ВА* с *ВД*:

$$\begin{aligned} ВА: \bar{m}_1 &= 5,539 (n_1 = 14); s_1^2 = 0,057 \\ ВД: \bar{m}_2 &= 5,277 (n_2 = 225); s_2^2 = 0,078 \end{aligned}$$

$$sd = \sqrt{\frac{13 \cdot 0,057 + 224 \cdot 0,078}{237}} = 0,277$$

$$t = \frac{5,539 - 5,277}{0,277} \sqrt{\frac{14 \cdot 225}{14 + 225}} = \frac{0,262 \cdot 3,630}{0,277} = 3,434$$

Итак, выборочная величина равна 3,434 (более надежное вычисление получается с помощью машины: 3,443). Этую величину надо сравнить с статистической таблицей, учитывая степени свободы, т.е. в нашем случае  $14+225-2=237$ . В статистической таблице указано, что критическое значение критерия Стьюдента при 237 ( $\approx \infty$ ) степенях свободы равно 1,960 (0,05), 2,576 (0,01) и 3,291 (0,001). Получается, что выборочная величина при сравнении *ВА* с *ВД* намного превышает найденные критические значения на всех уровнях. Поэтому можно отклонить гипотезу о равенстве средних, т.е. исключить возможность того, что *ВА* принадлежит перу Достоевского. В Таблице 5 приводятся числовые значения критерия Стьюдента для всех интересующих нас текстов.

*Таблица 5*  
*Средняя длина слова. Критерий Стьюдента*

Текст	<i>t</i>
1	-1,052
2	2,876 max

3	-0,185
4	0,759
5	-0,905
6	-1,296
7	2,089
8	0,711
9	-
10	1,906
11	-
12	-1,363
13	2,013
14	-1,952
15	-1,053
16	-0,660
17	-1,483
18	1,785
19	0,179
20	-2,741
21	-2,434
22	2,589
23	-
24	-0,401
25	-0,669
26	-0,608
I	-1,127
II	0,028
III	-1,417
IV	3,443 > max
V	4,877 > max
VI	-
VII	-0,423
VIII	-
IX	1,541
X	9,505 > max
XI	5,737 > max
XII	14,107 > max

Как явствует из таблицы, отсутствуют результаты для текстов 9, 11, 23, VI и VIII, из-за недостаточной длины этих

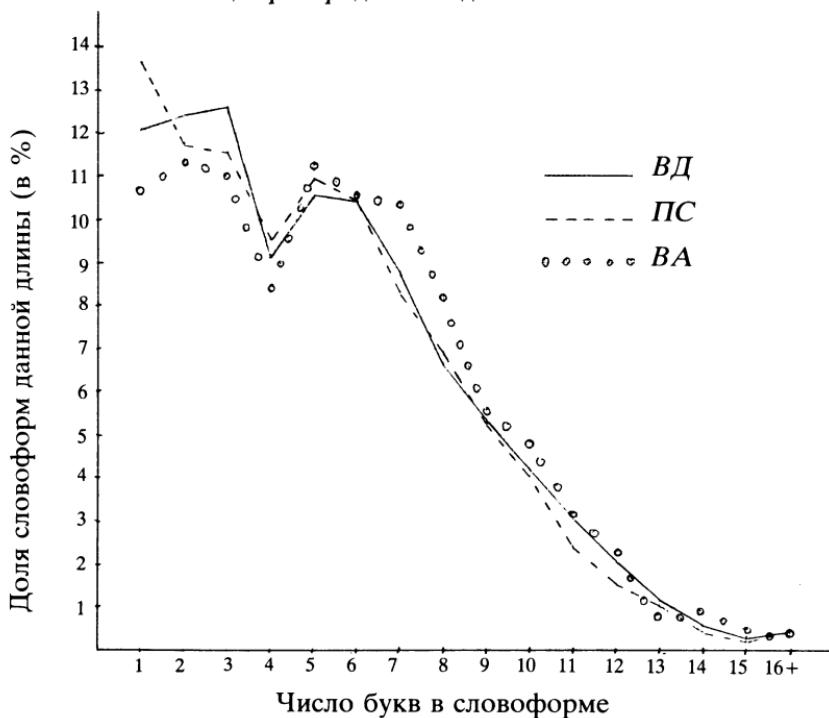
текстов. Что касается текстов Достоевского, то все они однородны с *ВД* на уровне 0,01, за исключением текстов 2 и 20 (последний текст очень короток). Самый высокий коэффициент имеет текст 2: 2,876. На этом основании можно, исходя из данного параметра, исключить возможность принадлежности Достоевскому следующих спорных текстов: IV, V, X, XI и XII.

### 4.3. Общее распределение длины слова

С помощью компьютера мы получили данные об общем распределении длины слова во всех текстах, т.е. сколько в каждом тексте слов, имеющих по 1, 2, 3 . . . 16 и более буквами.

Весьма явственно обнаруживается разница между *ВД* и некоторыми спорными текстами в зрительной форме. На Рисунке 3 мы проводили сравнение между *ВД*, *ПС* и *ВА*.

Рисунок 3  
Общее распределение длины слова



Как видно из рисунка, кривые для *ВД* и *ПС* значительно ближе друг к другу, чем кривые для *ВД* и *ВА*. В частности, бросается в глаза, что в *ВД* и *ПС* содержится гораздо больше коротких слов – в 1–4 буквы, чем в *ВА*. Зато в *ВА* больше длинных слов – в 5–12 букв, чем в *ВД* и *ПС*.

Опять-таки возникает вопрос: можно ли, исходя из цифр, лежащих в основе полученных кривых, исключить возможность написания *ВА* Достоевским?

Статистик подошел бы к этой проблеме, ставя перед собой следующий вопрос: какова вероятность того, что распределения длин слов в буквах в *ВА* и в *ВД* взяты из одной и той же «генеральной совокупности» и могут рассматриваться как управляемые одними и теми же закономерностями?

Для выяснения этого вопроса удобно применить непараметрический критерий Колмогорова-Смирнова, измеряющий разницу между накопленными относительными частотами в сравниваемых текстах по формуле:

$$\lambda = d_{\max} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}},$$

где  $d_{\max}$  обозначает максимальную разницу между накопленными относительными частотами, и  $n_1$  и  $n_2$  – количество слов в сопоставляемых текстах (25). В Таблице 6 мы при помощи этого критерия сравнили распределение длин слова в *ВА* с распределением длин слова в *ВД*.

Таблица 6  
Сравнение между *ВА* ( $f_1$ ) и *ВД* ( $f_2$ )  
Критерий Колмогорова-Смирнова

$m$	$f_1$	$f_2$	$p_1$	$p_2$	$p_1^*$	$p_2^*$	$d = p_1^* - p_2^*$
1	773	14 460	0,1069	0,1214	0,1069	0,1214	-0,0145
2	821	14 842	0,1136	0,1246	0,2205	0,2460	-0,0255
3	797	15 028	0,1103	0,1262	0,3308	0,3722	-0,0414
4	608	10 896	0,0841	0,0915	0,4149	0,4637	-0,0488
5	811	12 554	0,1122	0,1054	0,5271	0,5691	-0,0420
6	763	12 395	0,1056	0,1041	0,6327	0,6732	-0,0405
7	747	10 359	0,1033	0,0870	0,7360	0,7602	-0,0242
8	589	7 931	0,0815	0,0666	0,8175	0,8268	-0,0093

9	402	6 353	0,0556	0,0533	0,8731	0,8801	-0,0070
10	348	5 117	0,0481	0,0429	0,9212	0,9230	-0,0018
11	221	3 596	0,0306	0,0302	0,9518	0,9532	-0,0014
12	162	2 451	0,0224	0,0206	0,9742	0,9738	0,0004
13	56	1 423	0,0077	0,0119	0,9819	0,9857	-0,0038
14	66	738	0,0091	0,0062	0,9910	0,9919	-0,0009
15	31	390	0,0043	0,0033	0,9953	0,9952	0,0001
16+	34	574	0,0047	0,0048	1,0000	1,0000	0,0000
	7 229	119 107	1,0000	1,0000	—	—	—
	<i>n<sub>1</sub></i>	<i>n<sub>2</sub></i>					

Символы в Таблице 6 обозначают:

*m* – частота букв в слове;

*f<sub>1</sub>* и *f<sub>2</sub>* – число слов с частотой *m*;

*p<sub>1</sub>* и *p<sub>2</sub>* – соответствующие относительные частоты;

*p<sub>1</sub><sup>\*</sup>* и *p<sub>2</sub><sup>\*</sup>* – накопленные частоты;

*d* – разница между накопленными частотами;

*n<sub>1</sub>* и *n<sub>2</sub>* – объемы текстов.

В последней колонке Таблицы 6 мы находим максимальное значение, *d<sub>max</sub>*, которое в нашем примере = -0,0488. Получается результат:

$$\lambda = -0,0488 \sqrt{\frac{7\ 229 \cdot 119\ 107}{7\ 229 + 119\ 107}} = 4,030$$

Оказывается, что полученное эмпирическое значение  $\lambda$  в 4,030 превышает критическое значение даже на уровне значимости 0,01 (1,628). Следовательно, вероятность того, что разница между кривою *ВА* и кривою *ВД* случайна, – меньше, чем в 1%, т.е. разницу между распределениями частот слов разной длины в двух текстах надо признать существенной. Но если на основании этого параметра можно исключить *ВА*, как текст, принадлежащий Достоевскому, то нельзя исключить *ПС*, ибо соответствующее сравнение между этим текстом и *ВД* дает эмпирическое значение  $\lambda$  только в 1,278, что ниже критического значения даже на

уровне значимости 0,05 (1,358). В Таблице 7 приводятся числовые значения критерия Колмогорова-Смирнова для всех интересующих нас текстов.

Таблица 7

*Общее распределение длины слова. Критерий Колмогорова-Смирнова*

Текст	$\lambda$
1	0,814
2	2,133
3	1,141
4	0,762
5	1,834
6	1,278
7	2,986
8	1,261
9	0,873
10	2,103
11	1,220
12	1,892
13	3,761
	max
14	1,334
15	1,780
16	1,235
17	2,024
18	1,619
19	0,781
20	3,341
21	2,441
22	2,453
23	0,794
24	0,628
25	0,976
26	1,122
I	1,239
II	0,908
III	1,346
IV	4,030 > max

V	4,014> max
VI	1,578
VII	0,677
VIII	1,478
IX	1,464
X	7,932> max
XI	5,504> max
XII	10,955> max

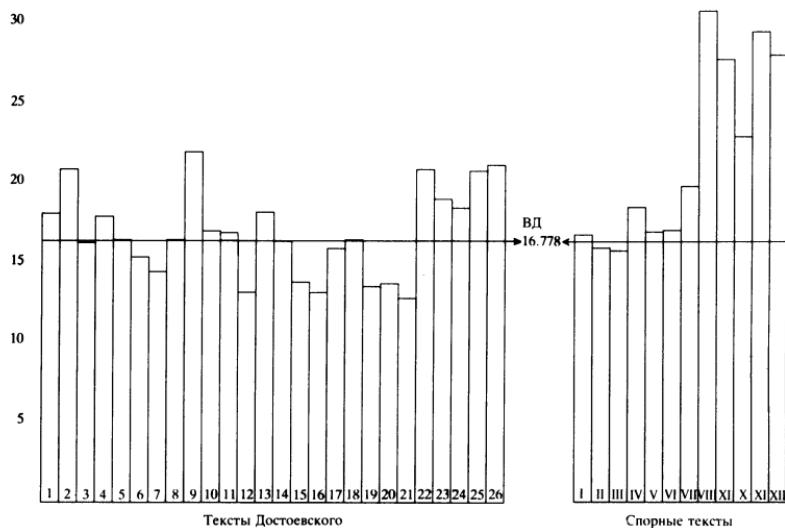
Как видно из таблицы, некоторые тексты Достоевского показывают довольно высокие числовые значения  $\lambda$ . Самый высокий коэффициент опять-таки имеет непохожий на почерк Достоевского текст 13 – «Два лагеря теоретиков» (3,761). Поэтому, исходя из параметра «Общее распределение длины слова», позволительно исключить возможность принадлежности Достоевскому следующих спорных текстов: IV, V, X, XI и XII.

#### 4.4. Средняя длина предложения в словах

Пользуясь в нашем исследовании параметром «длина предложения», мы определяем «предложения» как последовательности слов, отличающиеся друг от друга терминаторами (.!? . . .).

На Рисунке 4 представлены данные о средней длине предложения.

Рисунок 4  
Средняя длина предложения в словах



Как явствует из рисунка, тексты Достоевского имеют в среднем 16,778 слов в каждом предложении. Несмотря на отдельные отклонения, тексты эти оставляют впечатление однородности. Зато заметно отличаются от текстов Достоевского некоторые из спорных текстов, в которых средняя длина предложения достигает 30 и более слов.

Для того, чтобы провести тест исключительности с помощью критерия Стьюдента, как мы это сделали для средней длины слова (см. 4.2.), мы попросили у компьютера данные о средней длине предложения в выборках размером в 30 текущих предложений. На основании полученной информации в Таблице 8 приводятся необходимые статистические данные для проведения теста, а в Таблице 9 даются числовые значения критерия Стьюдента для всех интересующих нас текстов.

Таблица 8  
Средняя длина предложения. Статистические данные

Текст	$n$	$\bar{m}$	$s^2$	$s$
1	22	18,797	25,968	5,096
2	16	21,777	12,850	3,585

3	11	16,785	6,609	2,571
4	20	18,390	13,890	3,727
5	6	17,394	16,665	4,082
6	14	15,645	8,165	2,857
7	8	15,150	6,516	2,553
8	8	17,046	15,182	3,896
9	—	—	—	—
10	13	17,708	6,929	2,632
11	—	—	—	—
12	5	13,613	16,395	4,049
13	12	18,808	2,956	1,719
14	5	16,513	15,808	3,976
15	4	13,875	16,255	4,032
16	19	12,437	11,124	3,335
17	7	16,905	5,544	2,355
18	8	17,142	2,718	1,649
19	10	14,023	12,052	3,472
20	10	14,890	12,891	3,590
21	15	13,184	13,716	3,704
22	2	21,283	37,845	6,152
23	1	21,733	—	—
24	2	19,567	3,920	1,980
25	2	19,867	6,241	2,498
26	6	21,222	9,237	3,039
ВД	226	16,843	18,116	4,256

I	11	17,136	10,620	3,259
II	6	16,183	2,876	1,696
III	4	16,217	11,375	3,373
IV	12	19,119	6,967	2,639
V	10	17,277	14,711	3,836
VI	—	—	—	—
VII	5	21,213	18,458	4,296
VIII	—	—	—	—
IX	2	27,433	28,880	5,374
X	12	23,569	12,242	3,499
XI	9	30,771	47,407	6,885
XII	11	28,530	5,887	2,426

*Таблица 9*  
*Средняя длина предложения. Критерий Стьюдента*

<i>Текст</i>	<i>t</i>
1	2,018
2	4,522
3	-0,045
4	1,572
5	0,313
6	-1,038
7	-1,117
8	0,133
9	—
10	0,724
11	—
12	-1,680
13	1,590
14	-0,172
15	-1,384
16	-4,397
17	0,038
18	0,197
19	-2,063
20	-1,428
21	-3,247
22	1,465
23	—
24	0,903
25	1,002
26	2,501
 I	0,225
II	-0,378
III	-0,293
IV	1,832
V	0,316
VI	—
VII	2,271
VIII	—
IX	3,499

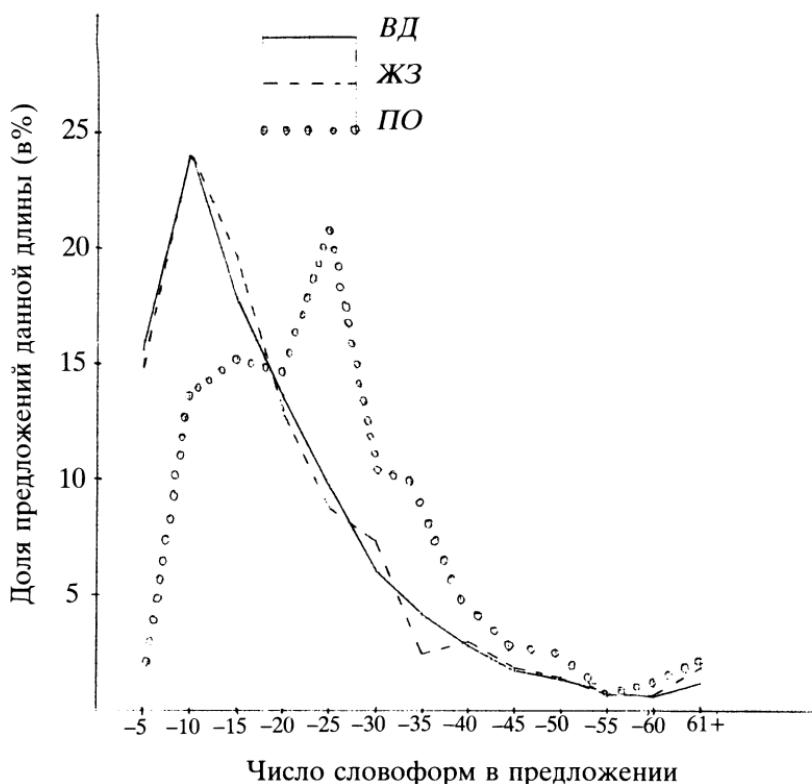
X	5,375> max
XI	9,370> max
XII	9,024> max

Оказывается, что параметр «Средняя длина предложения» обладает меньшей дискриминирующей силой, чем параметр «Средняя длина слова». Если с помощью параметра «Средняя длина слова» мы сумели исключить возможность принадлежности Достоевскому пяти спорных текстов (см. 4.2.), то мы на этот раз можем исключить только три текста, а именно X, XI и XII.

#### 4.5. Общее распределение длины предложения

С помощью компьютера была получена информация об общем распределении длины предложения по интервалам в 1–5, 6–10, 11–15 . . . 61+ предложения. В качестве примера мы на Рисунке 5 опять-таки провели сравнение, на этот раз между текстами Достоевского (*ВД*), бесспорно принадлежащей ему статьей «Журнальная заметка. О новых литературных органах и о новых теориях» (*ЖЗ*) и приписываемой ему статьей «Политическое обозрение» (*ПО*). При этом мы отложили на горизонтальной оси возможные интервалы предложений, а на вертикальной указали, какую долю (в %) от общего числа предложений составляют предложения, длина которых укладывается в каждый из этих интервалов.

Рисунок 5  
Общее распределение длины предложения



Как видно из рисунка, если кривые для *ВД* и *ЖЗ* почти покрывают друг друга, то кривая для *ПО* явно держится особняком. Бросается в глаза, в частности, что в *ВД* и в *ЖЗ* содержится гораздо больше предложений в интервалах 1–5, 6–10 и 11–15 слов, чем в *ПО*. С другой стороны, в *ПО* значительно больше длинных предложений. Характерно, что если самый частый интервал в *ВД* и в *ЖЗ* – 6–10 предложений, то самый распространенный в *ПО* интервал – 20–25 предложений. В Таблице 10 приводятся числовые значения критерия Колмогорова-Смирнова для всех текстов.

**Таблица 10**  
**Общее распределение длины предложения. Критерий**  
**Колмогорова-Смирнова**

Текст	$\lambda$
1	1,268
2	3,468
3	0,869
4	1,283
5	0,531
6	1,053
7	1,396
8	0,369
9	0,511
10	1,425
11	1,048
12	1,511
13	3,752
14	0,556
15	1,058
16	3,763
	max
17	0,541
18	0,174
19	1,618
20	2,060
21	3,017
22	1,546
23	1,032
24	0,668
25	1,946
26	1,663
I	1,421
II	0,758
III	0,570
IV	3,023
V	0,832
VI	0,683
VII	1,727
VIII	1,886

IX	2,621
X	5,202> max
XI	5,406> max
XII	6,801> max

Как явствует из таблицы, среди текстов Достоевского самый высокий коэффициент получает текст 16 – «Щекотливый вопрос». В результате вычисления позволительно исключить возможность принадлежности Достоевскому спорных текстов X, XI и XII.

#### 4.6. Лексический спектр текста на уровне словаря

Можно исследовать стилистический почерк писателя и с точки зрения лексического спектра текста, т.е. с точки зрения распределения частот слов в тексте. Лексический спектр устанавливается на основе частотного словаря текста. При этом для сравнения необходимо извлекать данные из выборок одинаковой длины. Поэтому с помощью компьютера нам были предоставлены частотные словари на каждые 500 словоформ текста. Для текстов 9, 11 и VI, не имеющих 500 слов, данных не получалось. В других текстах все словоформы распределились по отдельным группам на основе частоты встречаемости ( $m$ ). Потом компьютер определил их число в каждой группе ( $f$ ) и покрываемость текста ( $mf$ ). В первом случае мы получили распределение частот на уровне словаря (см. данную главу), а во втором случае – на уровне текста (см. 4.7.). В обоих случаях, если текст состоит из нескольких выборок, мы суммировали частоты встречаемости в тексте. Возьмем для примера текст 1:

$m$	$f$	$mf$
1	5 740	5 740
2	929	1 858
3	308	924
4	170	680
5	71	355
6	61	366
7	34	238

8	20	160
9	19	171
10+	88	1 508
Итого	7 440	12 000

Если в выборках, составляющих текст 1, имеется всего только 88 словоформ с частотой 10 и более, то эти высокочастотные, «строевые» слова покрывают довольно большую часть текста (1 508 из 12 000, т.е. 12,57%).

В Таблице 11 даются числовые значения критерия Колмогорова-Смирнова.

*Таблица 11  
Лексический спектр текста на уровне словаря. Критерий  
Колмогорова-Смирнова*

Текст	$\lambda$
1	0,210
2	0,356
3	0,519
4	0,211
5	0,707
6	1,318
7	1,528
8	0,356
9	—
10	0,437
11	—
12	0,576
13	1,103
14	0,473
15	1,055
16	0,275
17	0,580
18	0,425
19	0,616
20	0,340
21	0,618
22	0,673
23	0,309

24	0,269
25	0,329
26	0,918
I	0,775
II	0,877
III	0,387
IV	2,366 > max
V	0,559
VI	—
VII	0,529
VIII	0,344
IX	1,253
X	2,031 > max
XI	1,222
XII	3,552 > max

Как видно из таблицы, все тексты Достоевского однородны на уровне значимости 0,01 (1,628). Самый высокий коэффициент имеет текст 7 (1,528), что позволяет нам исключить возможность принадлежности Достоевскому текстов IV, X и XII.

#### 4.7. Лексический спектр текста на уровне текста

Конечно, рассматриваемый параметр имеет много общего с параметром «Лексический спектр текста на уровне словаря». Как явствует из Таблицы 12, в которой даются числовые значения критерия Колмогорова-Смирнова, применение этого параметра дает право исключить возможность принадлежности Достоевскому текстов IV, X, XI и XII.

*Таблица 12  
Лексический спектр текста на уровне текста. Критерий  
Колмогорова-Смирнова*

Текст	$\lambda$
1	1,071
2	0,758

3	1,116
4	2,106
5	0,634
6	2,289
7	2,261
8	0,677
9	—
10	1,991
11	—
12	1,497
13	2,915
14	3,009 max
15	2,334
16	0,714
17	2,487
18	0,678
19	0,988
20	1,944
21	1,883
22	1,644
23	0,482
24	0,757
25	0,600
26	2,565
 I	1,272
II	2,228
III	0,710
IV	5,304> max
V	2,504
VI	—
VII	1,249
VIII	1,116
IX	1,871
X	6,535> max
XI	3,074> max
XII	9,499> max

## 4.8. Индекс разнообразия лексики

В лингвостатистике принято измерять богатство лексики текстов при помощи так называемого индекса разнообразия, или индекса TTR (англ. type-token ratio), т.е. при помощи отношения числа разных словоформ (types) к числу словоупотреблений (tokens).

Таким образом, при помощи индекса разнообразия обследуется степень повторяемости в словаре писателя. При этом подразумевается, что писатель, употребляющий те же самые слова, имеет, или во всяком случае демонстрирует, менее богатый словарь, чем писатель, употребляющий по мере развития текста все новые и новые слова.

Понятно, что для решения нашей проблемы индекс разнообразия лексики станет параметром исключительной важности. Ведь все, умеющие читать Достоевского в оригинале, обратили внимание на то, что он любит повторять одни и те же слова и выражения (26). Возьмем для примера следующую фразу из *Кроткой*: «Глупо, глупо, глупо и глупо!». В этом предложении употребляется пять слов, но четыре из них одинаковы; или, выражаясь другими словами, только два из них различны. Подобными примерами богат и язык Достоевского-публициста. Так, Достоевский очень любит анафору, причем повторяются не только слова (Знают . . . Знают . . . Знают . . .; АИ, 18, 42; Неужели . . . Неужели . . . Неужели . . . АИ, 18, 55), но и предложения (Мы хотим . . . Мы хотим . . .; АИ, 18, 68; Вы хотите . . . Вы хотите . . .; там же). Более того, Достоевский нередко начинает свои предложения с повторения одних и тех же слов: «Ум, ум, самая тревожная боязнь за свой ум, – вот в чем главное дело!» (АИ, 18, 52); «Право, право, право жаль, что вы черногорцы!» (АИ, 20, 26). В таких случаях писатель как бы выступает оратором, старающимся зачаровать своих слушателей магией слов. Несомненно, что высокая повторяемость слов составляет одну из самых отличительных черт в языке Достоевского (27). Разумеется, компьютер поможет выявить эту особенность стиля писателя.

Известно, что число словоупотреблений (длина текста) увеличивается быстрее, чем число разных словоформ

(объем словаря). Из этого следует, что прямое сравнение значений индекса разнообразия разных текстов возможно только с условием одинакового объема текстов. Поскольку в наш материал входит много коротких текстов, мы заставили компьютер предоставить нам списки с указанием числа разных словоформ на каждые 500 новых текстовых слов, т.е. на приблизительно каждые две новые страницы текста.

Оказывается, что *ВД*, состоящий из 225 выборок, имеет в среднем только 307 разных словоформ на каждые 500 текстовых слов. Для сравнения можно отметить, что части I, II и IV *Тихого Дона* имеют в среднем 380 разных словоформ на каждые 500 текстовых слов. Это, конечно, большая разница, доказывающая, что Достоевский и Шолохов в этом отношении обладают очень различными стилистическими почерками: словарь Достоевского необычайно ограничен, тогда как словарь Шолохова необычайно богат.

Так как в текстах Достоевского порции разных словоформ обладают удивительной стабильностью (можно даже говорить о нормальном распределении), то с помощью критерия Стьюдента на основании рассматриваемого параметра исключается возможность принадлежности Достоевскому текстов IV, X, XI и XII (см. Таблицу 13).

Таблица 13  
Индекс разнообразия лексики. Критерий Стьюдента.

Текст	<i>t</i>
1	0,723
2	- 0,281
3	- 1,117
4	0,905
5	0,096
6	2,069
7	- 1,785
8	- 0,428
9	-
10	1,551
11	-
12	- 0,969

13	3,377	max
14	- 1,101	
15	- 2,204	
16	0,536	
17	- 0,560	
18	0,562	
19	- 0,769	
20	- 1,533	
21	- 1,726	
22	1,665	
23	-	
24	0,176	
25	0,216	
26	- 2,566	
 I	1,096	
II	2,082	
III	0,614	
IV	- 5,536> max	
V	2,558	
VI	-	
VII	- 1,181	
VIII	-	
IX	0,982	
X	6,596> max	
XI	3,405> max	
XII	10,247> max	

Очень наглядно можно показать разницу между бесспорно принадлежащими Достоевскому текстами, с одной стороны, и сомнительно принадлежащими ему текстами, с другой стороны, с помощью доверительных интервалов. Для того, чтобы установить доверительные интервалы, нужно сравнивать абсолютную среднюю ошибку, пользуясь формулой:

$$\varepsilon = t \frac{s}{\sqrt{n}},$$

где  $t$  найдется в таблице Стьюдента в соответствии с уровнем значимости и степенями свободы ( $n = n-1$ ). Если выберем уровень значимости в 0.99 (99%), мы для ВА получим:

( $n = 14$ ,  $n_o = 13$ );  $t_{0,99;13} = 3,012$ :

$$3,012 \frac{18,082}{3,742} = 14,555$$

Доверительный интервал получается так:

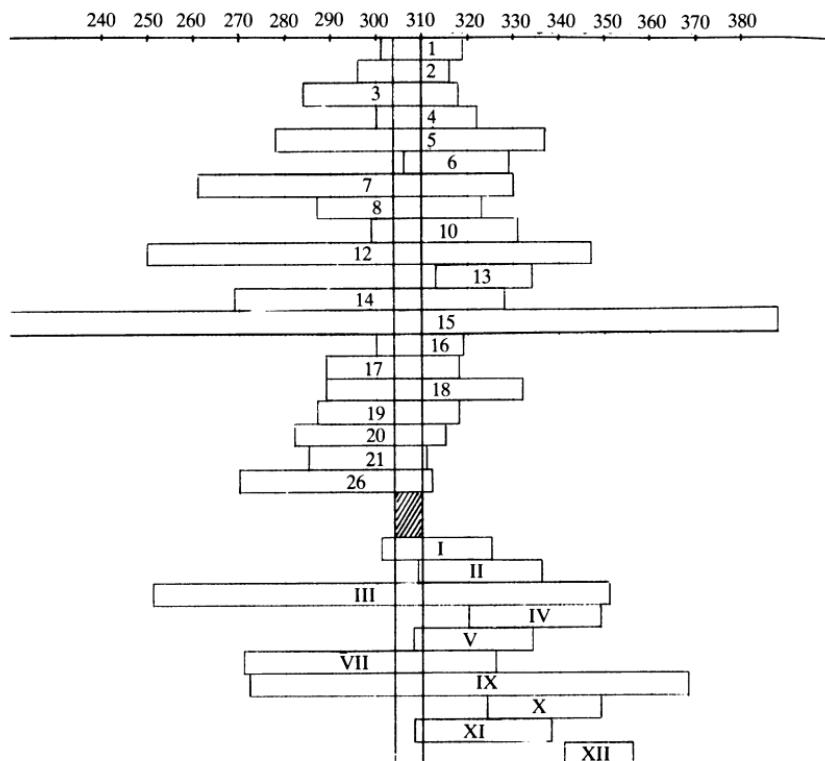
$$\bar{m} \pm \varepsilon = 334 \pm 14,555,$$

т.е. внутри доверительных пределов 320 . . . 349. Это значит, что в 99 из 100 экспериментов средняя выборочная частота будет находиться в интервале 320–349 разных словоформ на каждые 500 словоупотреблений. Видимо, это слишком много для того, чтобы допустить возможность написания этого текста Достоевским, чьи тексты, в среднем, имеют доверительные пределы лишь в 304 . . . 310 разных словоформ.

На Рисунке 6 приводятся результаты для интересующих нас текстов.

Рисунок 6

Доверительные интервалы для среднего значения длины.  
Число разных словоформ на каждые 500 словоупотреблений



В текстах, состоящих из менее, чем 3 выборок (9, 11, 22–25, VI, VIII) доверительные интервалы, если они вообще смогут быть составлены, будут слишком широкими, чтобы позволить исключения. А что касается других текстов, то вероятность авторства Достоевского можно измерить дистанцией от столбца, составленного на основании всех текстов писателя (304 . . . 310). Как явствует из рисунка, все бесспорно принадлежащие Достоевскому тексты более или менее покрыты столбцом, опять-таки за интересным исключением не имеющегося в списке Н.Н. Страхова текста 13. Гораздо менее «достоевский» профиль, однако, показывают некоторые спорные тексты, обладающие значительно более богатым словарем. Приписывать эти тексты перу Достоевского не дает права теория вероятности.

## 5.0. Подводя итоги

Применение указанных 15 параметров при исследовании текста названных статей привело к результатам, отраженным суммарно в Таблице 14. Здесь «плюс» обозначает совпадение (или близость) данного характера соответствующему параметру авторизованных текстов Достоевского, «минус» – противоположный, отрицательный результат сопоставления, а «ноль» – практическую невозможность получения в данном случае вполне определенного ответа.

Таблица 14  
Результаты

Параметр	Тексты											
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
1	+	+	+	–	+	+	+	+	+	–	–	–
2	+	+	+	–	+	+	+	+	+	–	–	–
3	+	+	+	–	–	+	+	+	+	–	–	–
4	+	+	+	–	+	+	+	+	+	–	+	–
5	+	–	+	–	+	+	+	+	+	–	–	–
6	+	+	+	–	+	+	+	+	+	–	–	–
7	+	+	+	–	+	+	+	+	+	–	–	–
8	+	+	+	–	+	+	+	+	+	–	–	–
9	+	+	+	–	–	0	+	0	+	–	–	–
10	+	+	+	–	–	+	+	+	+	–	–	–
11	+	+	+	+	+	0	+	0	+	–	–	–
12	+	+	+	+	+	+	+	+	+	–	–	–
13	+	+	+	–	+	0	+	+	+	–	+	–
14	+	+	+	–	+	0	+	+	+	–	–	–
15	+	+	+	–	+	0	+	0	+	–	–	–

Постараемся коротко прокомментировать полученные результаты.

### *Текст I*

Письмо Постороннего критика в редакцию нашего журнала по поводу книг г-на Панаева и «Нового поэта».

Результат наших вычислений – 15 плюсов. Это означает, что у нас нет возможности установить уловимые по нашим 15 параметрам стилистические различия между этим текстом и текстами, составляющими корпус Достоевского. Как отмечено выше (1.0.), исходя из методики языковых показаний, нельзя доказать, что эта статья действительно написана Достоевским, но нет и убедительных стилистических аргументов против его авторства. Поэтому можно согласиться с В.А. Тунимановым, что статья эта заслуживает включения в раздел «*Dubia*».

### *Текст II*

Гаваньские чиновники в домашнем быту, или Галерная гавань во всякое время дня и года. (Пейзаж и жанр) Ивана Генслера.

Результат наших вычислений – 14 плюсов – также весьма показателен, давая исследователям, на наш взгляд, право включить и ее в раздел «*Dubia*». Конечно, исключить полностью ее принадлежность Ап. Григорьеву мы можем только на основании анализа большого числа одновременно написанных им статей. Но это – задача будущего.

### *Текст III*

Противоречия и увлечения «Времени».

Результат наших вычислений предельный и здесь – 15 плюсов. Возражая против включения этой статьи в *Полное собрание сочинений* Достоевского, Л.П. Гроссман писал: «Имея постоянно в виду необходимость строжайшего отбора при определении автора анонимных материалов, мы не включили в это *Собрание* целого ряда статей, близко напоминающих идейный строй и литературную манеру Достоевского, но не представляющих достаточных указаний на его авторство» (Гросс, 22, стр, XXV). Конечно, осторожность ученого вполне правомерна. Отметим, однако, что язык и стиль данной статьи намного ближе к корпусу статей Достоевского, чем статья «Два лагеря теоретиков», авторство которой тот же Гроссман считал несомненным

(Гросс, 23, стр. 123). На основании большого стилистического сходства текста III с принадлежащими Достоевскому статьями, мы склоняемся к включению его в раздел «*Dubia*».

#### *Текст IV*

Выставка в Академии художеств за 1860–1861 год.

Результат наших вычислений – 2 плюса и 13 минусов – говорит против принадлежности статьи Достоевскому, особенно если принять во внимание большую длину ее: 7 229 слов. Правда, в статье высказываются взгляды об искусстве, о живописи, о дагерротипизме, которые во многом совпадают с положениями бесспорно принадлежащих Достоевскому статей. Однако нельзя сказать, что взгляды эти отличаются особой оригинальностью. Напротив, как свидетельствует комментарий к этой статье (*АИ*, 19, стр. 319–321), они высказывались и другими сотрудниками почвеннических журналов начала 1860-х гг. На наш взгляд, включение этой статьи Г.М. Фридлендером в корпус *АИ* (см. 19, стр. 151–168) представляется весьма рискованным. В лучшем случае можно было бы включить ее в раздел «*Dubia*», как это в свое время сделал Б.В. Томашевский, но не в основной корпус произведений писателя.

#### *Текст V*

Ряд статей о русской литературе. Вопрос об университетах.

Результат наших вычислений – 12 плюсов и 3 минуса – можно признать достаточно красноречивым. С точки зрения нашей методики, можно включить статью в отдел «*Dubia*», как это и сделано (см. *АИ*, 19, стр. 187–210). Особенno важно учесть длину статьи: 5 571 слово. При таком объеме ее наши параметры должны действовать надежно. Отметим, что и данная статья намного ближе к корпусу текстов Достоевского, чем «бесспорно» принадлежащая ему статья «Два лагеря теоретиков».

#### *Текст VI*

Николай Александрович Добролюбов. Некролог.

Результат наших вычислений – 10 плюсов и 0 минусов – с первого взгляда может показаться положительным, но на

самом деле это не так. Оказывается, что длина текста (261 слово) так мала, что в 5 из 15 случаев наши параметры не действуют. Судя по другим экспериментам (28), объективные характеристики текста в принципе не дают возможности установить авторство текстов небольшого объема (100–300 словоупотреблений). В данном случае мы можем только подтвердить эти итоги: исходя из нашей методики, при сравнении короткого текста с большим корпусом текстов будет слишком легко получить «положительные» результаты. Поэтому установить с достаточной надежностью принадлежность этой статьи Достоевскому, ориентируясь лишь на объективные характеристики текста, не представляется возможным.

### *Текст VII*

Рассказы Н.В. Успенского.

Результат наших вычислений – 15 плюсов при 0 минусов – предельно положительно надежен, особенно если принять во внимание большой объем статьи. В этом случае результаты нашей методики полностью совпадают с результатами методики внешних показаний, выдвинутыми Г.М. Фридлендером и А.И. Батюто. Можно спорить лишь о том, следует ли эту статью включить в основной корпус произведений писателя, как это сделано в *АИ*, или в раздел «*Dubia*», как это сделал Б.В. Томашевский.

### *Текст VIII*

Полемический случай с «Основой» и «Сионом».

Результат наших вычислений – 12 плюсов при 0 минусов – может показаться говорящим скорее в пользу авторства Достоевского. Но и здесь нужно принять во внимание краткость текста – всего только 816 слов. Трудно поэтому дать надежный ответ на вопрос о принадлежности статьи Достоевскому.

### *Текст IX*

Дворянин, желающий быть крестьянином.

Результат наших вычислений – 15 плюсов при 0 минусов – свидетельствует о близости стилистическому почерку Достоевского. Во всяком случае нельзя на основании нашей ме-

тодики исключить возможность написания статьи писателем. Правда, длина текста – 2 091 слово – маловата для того, чтобы с достаточной уверенностью включить ее в основной корпус его произведений. Думается, что статья заслуживает включения в раздел «Dubia».

#### *Текст X*

Политическое обозрение.

При большой длине статьи (8 946 слов) и при полученном результате (15 минусов) вряд ли может быть сомнение в том, что эта статья отношения к Достоевскому не имеет. Включить такую глубоко чуждую по языку и стилю Достоевского статью даже в раздел «Dubia» было бы, на наш взгляд, непростительной ошибкой.

#### *Текст XI*

Наши домашние дела.

Результат наших вычислений – 2 плюса при 13 минусах – явно не говорит в пользу авторства Достоевского. Даже если допустить, что писатель в качестве редактора оставил свой отпечаток на начале статьи, этого, разумеется, недостаточно для включения ее даже в раздел «Dubia».

#### *Текст XII*

Политическое обозрение. Общий обзор главнейших политических событий прошедшего года.

На основании наших вычислений – сплошных минусов по всем параметрам – можно, при большой длине статьи (10 171 слово), с полной уверенностью исключить возможность принадлежности ее Достоевскому.

## 6.0. Вместо заключения

В начале предисловия к подготовленному им двадцать второму, дополнительному, тому *Полного собрания сочинений Ф.М. Достоевского* в издании «Просвещение» Л.П. Гроссман писал в 1918 г.:

«В существующие 'Полные собрания сочинений Ф.М. Достоевского' включено далеко не все, написанное им. При отсутствии критического издания его сочинений большинство читателей до сих пор лишено возможности знакомиться с целыми отделами анонимной публицистики Достоевского и с многочисленными вариантами различных печатных редакций, сквозь которые часто проходили его художественные произведения, прежде чем получить свой окончательный текст.

Ввиду непрерывного роста и углубления критического и читательского интереса к творчеству Достоевского, когда каждая написанная им строка приобретает первостепенную ценность, кажется совершенно недопустимым, чтобы многие сотни напечатанных им страниц продолжали оставаться вне читательского кругозора. Нам показалось необходимым извлечь их из позабытых книжек старых журналов и снова обратить на общее пользование» (см. Гросс, 22, стр. XX).

Конечно, нельзя не согласиться с точкой зрения Гроссмана. Но само усердие исследователя ознакомить читателей с «анонимной публицистикой Достоевского» чревато и великой опасностью, а именно опасностью включения в собрание сочинений писателя не принадлежащих ему статей.

К сожалению, в истории любой, в том числе русской, литературы можно указать на ряд случаев, когда исследователи, ограничиваясь лишь указанием на совпадение или близость идей, приписывали анонимные сочинения тому или иному крупному писателю. Именно таким образом,

отмечает и академик В.В. Виноградов, приписано Белинскому много анонимных статей в тех журналах, в которых он был руководящим критиком, и те же принципы применялись и по отношению к публицистическим статьям Салтыкова-Щедрина и Герцена, которые, подобно Достоевскому, были редакторами журналов, где сплошь и рядом появлялись анонимные, идейно близкие к редакции статьи (29). «Методика сопоставления или отождествления содержания анонимного сочинения с тематикой произведений подходящего писателя, – указывает академик Виноградов, – самый слабый пункт современных атрибуций, направленных по пути бездоказательного пополнения сокровищницы творчества передовых представителей русской литературы XIX, а отчасти и XX века всякого рода сомнительными анонимными сочинениями» (30).

Таким образом, при установлении автора анонимных произведений необходимо соблюдать строжайшую осторожность. Запомним слова В.Я. Брюсова о чрезмерном усердии исследователей приписывать анонимные произведения Пушкину:

«Наводнять сначала журналы сомнительными сенсациями с именем Пушкина, а потом сочинения Пушкина сомнительными страницами есть подлинный грех перед русским обществом. За эфемерную и легко добываемую славу 'открывателей' должны будут горько расплачиваться русские читатели» (31).

Полностью присоединяясь к мнению академика Виноградова, что ссылка «на совпадение идей, выраженных в анонимном и псевдонимном сочинении, с мировоззрением, со взглядами того или иного автора может быть дополнительной лишь в том случае, если будет убедительно показана общность стилистических принципов воплощения одинаковых идей» (32), мы в настоящем исследовании постарались установить некоторые стилистические принципы публицистики Достоевского и обследовать, насколько похожи на эти принципы язык и стиль приписываемых ему статей. На наш взгляд, этот метод, в котором все параметры были оговорены до начала предпринятой работы, представляет собой едва ли не единственный объективный подход к проблемам установления спорного авторства. Конечно,

можно возразить, что применяемые нами 15 параметров раскрывают далеко не все стилистические принципы писателя. Но несомненно и то, что все эти широко употребляемые параметры, пользующиеся в стилеметрике репутацией большой надежности, раскрывают некоторые весьма существенные стилистические принципы.

В этом отношении важно отметить, что выбранные нами объективные параметры *действуют вместе*, позволяя нам в *своей совокупности* исключить возможность принадлежности Достоевскому ряда анонимных статей. Хотя некоторые параметры коррелятивны, но это вряд ли единственное объяснение. Предпринятое нами исследование указывает на то, что раз установлена определенная тенденция к исключению возможности принадлежности Достоевскому какого-нибудь текста, она будет усиливаться при привлечении новых параметров. Уже с помощью первых восьми параметров, касающихся распределения и сочетания частей речи в начале и в конце предложения, была отмечена явная тенденция к исключению нескольких текстов, а с прибавлением еще семи параметров, и при этом достаточно различных, эта тенденция лишь только усилилась. Есть основание думать, что с прибавлением новых параметров тенденция эта будет еще более очевидной.

Возьмем для примера наличие в бесспорных текстах Достоевского высокочастотных, «ключевых» слов. Под термином «ключевое слово» мы подразумеваем слово, которое употребляется писателем значительно чаще, чем у других писателей. При этом в исследованиях о спорном авторстве лучше всего пользоваться не «полнозначными» существительными, глаголами или прилагательными, а так называемыми «строевыми» словами, такими, которые употребляются часто и независимо от темы и жанра произведения.

В настоящем исследовании мы условно определили «ключевое слово» как слово, встречающееся в корпусе подлинных произведений Достоевского в среднем более чем один раз на каждую тысячу текстовых слов (т.е. с относительной частотой  $0.100+$ ) и минимум в три раза чаще, чем в научно-публицистических и газетно-журнальных материалах, применяемых в *Частотном словаре русского*

языка (33). Интуитивно предположилось, что именно с таким строгим определением легче всего найти самые характерные для Достоевского ключевые слова.

С помощью предоставленного компьютером словаря языка Достоевского, в котором слова расположены в порядке убывания их частот, было найдено 7 таких слов, которые несомненно принадлежат к излюбленным словам Достоевского-публициста. Каждое из них употребляется писателем с частотой от семи до трех раз большей, чем в словаре Засориной:

*Таблица 15  
Ключевые слова*

Тексты Достоевского			Тексты Засориной		
119 107 слов			Около 500 725 слов		
	Частота	Отн. частота		Частота	Отн. частота
1. даже	532	0,446	701	0,140	
2. ведь	372	0,312	293	0,059	
3. потому	333	0,279	357	0,071	
4. уж	254	0,213	223	0,045	
5. именно	224	0,188	252	0,050	
6. хоть	223	0,187	129	0,026	
7. слишком	126	0,106	91	0,018	

Взятые в своей совокупности, эти слова употребляются Достоевским-публицистом более чем 17 раз на каждую тысячу текстовых слов, а в современных нам публицистических и журнальных текстах обычно только 4 раза. Оказывается, что и в иных спорных текстах такие слова встречаются гораздо реже, чем в бесспорных текстах писателя. Ниже приводится список спорных текстов с указанием частоты этих 7 слов на каждую тысячу текстовых слов. Для сравнения в скобках даются результаты, полученные с помощью наших 15 параметров:

*Таблица 16  
Ключевые слова в спорных текстах*

1.	V:	18,67	(12+, 3-)
2.	III:	18,11	(15+, 0-)
3.	VII:	14,81	(15+, 0-)

4.	I:	12,41	(15+, 0-)
5.	II:	11,12	(14+, 1-)
6.	VI:	7,66	(10+, 0-)
7.	IX:	7,17	(15+, 0-)
8.	IV:	7,15	( 2+, 13-)
9.	XII:	6,59	( 0+, 15-)
10.	X:	6,04	( 0+, 15-)
11.	VIII:	3,68	(12+, 0-)
12.	XI:	2,98	( 2+, 13-)

Как видно из таблицы, в текстах, имеющих много минусов, т.е. там, где, по нашим параметрам, вероятность принадлежности текста Достоевскому сводится к нулю, ключевые слова писателя используются редко.

Приведем еще один аргумент в пользу своих выводов. Давно установлено, что в случае, если какое-то слово имеет два варианта, то писатели предпочитают (и притом обычай достаточно устойчив) один из этих вариантов. Надо думать, что закон этот действует и у Достоевского. Возьмем для примера союз *чтобы* с очень распространенным у Достоевского вариантом *чтоб*. Оказывается, что в бесспорно принадлежащих Достоевскому статьях имеются 252 случая употребления формы *чтоб* при 34 случаях употребления формы *чтобы*. Итак, у Достоевского-публициста форма *чтоб* в среднем употребляется в 7-8 раз чаще, чем форма *чтобы*. Показательно, что форма *чтобы* в 10 из 26 текстов Достоевского вообще отсутствует. Так, в известной статье «Г-н -бов и вопрос об искусстве» (10 894 слова) мы находим 28 случаев употребления формы *чтоб* и ни одного случая употребления формы *чтобы*. В этом отношении небезынтересно отметить, что в спорных статьях, и опять-таки в самых спорных, изобилует форма *чтобы* при почти полном отсутствии формы *чтоб*. Ниже приводится список спорных текстов и указания соотношения *чтоб:чтобы*. Для сравнения в скобках опять-таки даются результаты, полученные с помощью наших 15 параметров:

**Таблица 17**  
**Соотношения чтобы: чтобы в спорных текстах**

чтоб:чтобы			
1.	I:	17: 0	(15+, 0-)
2.	III:	8: 0	(15+, 0-)
3.	II:	5: 0	(14+, 1-)
4.	VII:	11: 1	(15+, 0-)
5.	V:	21: 5	(12+, 3-)
6.	X:	20: 7	( 0+, 15-)
7.	IX:	4: 6	(15+, 0-)
8.	IV:	3:15	( 2+, 13-)
9.	XI:	1:22	( 2+, 13-)
10.	VIII:	0: 0	(12+, 0-)
11.	VI:	0: 2	(10+, 0-)
12.	XII:	0: 8	( 0+, 15-)

Создается почти та же картина, что и показанная выше. В текстах с многими плюсами изобилует форма *чтоб*, тогда как в текстах, где преобладают минусы, эта излюбленная Достоевским форма представляет редкость. Имеется в нашем материале только одно исключение, а именно текст X, но и здесь соотношение форм *чтоб* и *чтобы* значительно ниже ( $20:7=2,86$ ), чем в среднем у Достоевского ( $252: 34 = 7,41$ ).

Итак, думается, что нет надобности применять в нашем исследовании большее число параметров. При наличии достаточно больших текстов электронная вычислительная техника, идя рука об руку с математической статистикой, в состоянии предоставить исследователю мощное и адекватное орудие для подхода к проблеме атрибуции текстов.

## 7.0. Сноски

1. В.С. Нечаева, 1) *Журнал М.М. и Ф.М. Достоевских «Время». 1861–1863*, Москва 1972; 2) *Журнал М.М. и Ф.М. Достоевских «Эпоха». 1864–1865*, Москва 1975.
2. См. В.В. Виноградов, 1) *О языке художественной литературы*, Москва 1959; 2) *Проблема авторства и теория стилей*, Москва 1961; 3) «Из анонимного фельетонного наследия Достоевского», в кн.: В.В. Виноградов, *Исследования по поэтике и стилистике*, Ленинград 1972, стр. 185–211.
3. Письмо от 8 декабря 1979 г., подписанное Директором ИРЛИ А.С. Бушминым и Зам. Главного редактора *Полного собрания сочинений Достоевского* Г.М. Фридлендером.
4. В 2.1. и 2.2. приводится полный перечень текстов, вошедших в материал исследования.
5. Предварительные результаты исследования были опубликованы автором в журнале *Scando-Slavica*, tomus 26, 1980, pp. 19–31. В том же году мы прочитали доклады о нашем проекте на двух международных конференциях, в Кембридже и в Бергамо: 1) Association for Literary and Linguistic Computing (см. *ALLC Journal*, Volume 2, 1981, pp. 25–33) и 2) International Dostoevsky Society (см. *Dostoevsky Studies*, Volume 1, 1980, pp. 73–88. Кроме того, автор по приглашению АН СССР выступил с докладом в Пушкинском Доме; русское резюме исследования опубликовано в сб.: *Достоевский. Материалы и исследования*, 6, Ленинград 1985, стр. 207–224.
6. См.: P. Vašák, *Metody určování autorství*, Praha 1980, p. 201.
7. Об исключительной трудности установления авторства анонимных статей в *Гражданине* свидетельствует тот факт, что из приписываемых Достоевскому академиком В.В. Виноградовым статей, как установила редакция

- академического издания *Полного собрания сочинений* Достоевского, рецензия на роман А. Пальма Алексей Слободин, а возможно, и статья о Соборянах Н. Лескова принадлежат В.П. Мещерскому.
8. См. Ф.М. Достоевский, *Полное собрание сочинений*, т. 18, Ленинград 1978, стр. 209. Далее ссылки на это издание обозначаются *АИ* (академическое издание), с указанием тома и страницы.
  9. См. Ф.М. Достоевский, *Полное собрание художественных произведений*. Т. 1–13. Под ред. Б. Томашевского и К. Халабаева. Москва-Ленинград 1926–1930. Т. 13 (1930): *Статьи за 1845–1878 годы*, стр. 561. Далее ссылки на это издание обозначаются *Том*, с указанием тома и страницы.
  10. См. Ф.М. Достоевский, *Полное собрание сочинений*. Т. 1–23, Петербург-Петроград 1911–1918. Т. 22–23 (1918): *Забытые и неизвестные страницы*. Собрал и comment. Л.П. Гроссман. Далее ссылки на это издание обозначаются *Гросс*, с указанием тома и страницы.
  11. Oskar von Schoultz, «Ein Dostojewskij-Fund». *Commentationes humanarum litterarum*, I: 4, Helsingfors 1924 (далее – *Schoultz*).
  12. *Печать и революция*, 1925, № 2, стр. 249–251.
  13. В.В. Виноградов, *Проблема авторства и теория стилей*, стр. 159.
  14. Так, недавно в прессе сообщалось о том, что ЭВМ «нашла» новую пьесу Шекспира. Однако в действительности доказано лишь то, что данная пьеса (*The Brooke of Sir Thomas More*) по языку и стилю настолько близка к шекспировским пьесам, что нельзя исключить ее принадлежность перу Шекспира.
  15. См. Geir Kjetsaa et al., *The Authorship of «The Quiet Don»*, Oslo-Atlantic Highlands 1984.
  16. G. Herdan, *Language as Choice and Chance*, Groningen, 1956, p. 13.
  17. Сюда относится, например, отрывок «Не тронь меня», дважды приписанный Достоевскому Л.П. Гроссманом (см. *Творчество Достоевского. 1821 – 1881 – 1921. Сборник статей и материалов под редакцией Л.П. Гроссманом*, Одесса 1921, стр. 111–120; Л.П. Гроссман,

*Семинарий по Достоевскому. Материалы, библиография и комментарии*, Москва-Петроград 1922, стр. 82–92).

18. См. Geir Kjetsaa et al., op. cit.
19. См., например, Charles Muller, *Initiation à la statistique linguistique*, Paris 1968; Б.Н. Головин, *Язык и статистика*, Москва 1971.
20. О частом в творчестве Баратынского употреблении существительных по сравнению с другими частями речи, см. Гейр Хетсо, *Евгений Баратынский. Жизнь и творчество*, Осло – Берген – Тромсё 1973, стр. 547–548.
21. См. Bengt Beckman, *Försök med en metod för mätning av skillnader i författarstil*. Stockholms Universitet, Slaviska institutionen, 1966.
22. См. Geir Kjetsaa et al., op. cit.
23. См. Б.Н. Головин, «Опыт вероятностно-статистического изучения некоторых явлений истории русского литературного языка XIX–XX вв.», *Вопросы языкоznания*, 1965, № 3, стр. 137–146.
24. См. J. Tuldava, «Statistikilised testid keeleteaduses» (Методы проверки статистических гипотез в языкоznании). В кн.: *Tartu riiklik ülikool, Linguistica II*, Tartu 1970, pp. 125–196, в частности pp. 177–178.
25. См. Ю.А. Тулдава, «О статистической структуре текста», *Советская педагогика и школа*, т. IX, Тарту 1974, стр. 5–33, в частности, стр. 25–26.
26. Эта тенденция заметна еще в переводе писателем *Евгении Гранде Бальзака*, см. В.С. Нечаева, *Ранний Достоевский. 1821–1849*, Москва 1979, стр. 117 и сл.
27. Интересные наблюдения над повторами писателя имеются в кн.: Е.А. Иванчикова, *Синтаксис художественной прозы Достоевского*, Москва 1979. Кстати, готовя свои произведения к переизданию, Достоевский иногда убирал повторения однокоренных или одинаковых слов. Так, в повести *Слабое сердце* фраза «Смотри, брат, смотри, что мне пишет» была заменена фразой «Смотри, брат, что мне пишет» (АИ, 2, 415).
28. См. В.И. Батов, Ю.А. Сорокин, «Атрибуция текста на основе объективных характеристик. (Итоги эксперимента)», *Известия Академии Наук СССР, Серия литературы и языка*, 1975, том 34, № 1, стр. 76–78.

29. В.В. Виноградов, *О языке художественной литературы*, Москва 1959, стр. 306–307.
30. Там же, стр. 305–306.
31. В. Брюсов, *Мой Пушкин*, Москва-Ленинград 1929, стр. 194.
32. В.В. Виноградов, *О языке художественной литературы*, Москва 1959, стр. 311.
33. См. Л.Н. Засорина (ред.), *Частотный словарь русского языка*, Москва 1977.

## SLAVICA NORVEGICA

Published by the Norwegian Association of Slavists.

Editors: Erik Egeberg (Tromsø), Siri Sverdrup Lunden (Oslo) and Per Restan (Bergen).

1. Geir Kjetsaa, Sven Gustavsson, Bengt Beckman & Steinar Gil: The Authorship of *The Quiet Don*. (1984)
2. Alf Grannes, Arnvid Lillehammer & Egil Pettersen: Documents russes sur la pêche et le commerce russes en Norvège au XVIII<sup>e</sup> siècle. (1984)
3. Geir Kjetsaa: Dostoevsky and His New Testament. (1984)
4. Гейр Хетса: Принадлежность Достоевскому: К вопросу об атрибуции Ф.М. Достоевскому анонимных статей в журналах *Время и Эпоха*



















