

**ЛУЧШАЯ  
КНИГА  
О КНИГАХ**

WALL STREET JOURNAL

Ben Blatt

Nabokov's



Favorite



Word Is



*Mauve*



WHAT THE NUMBERS REVEAL  
ABOUT THE CLASSICS, BESTSELLERS,  
AND OUR OWN WRITING

Бен Блатт

Любимое



СЛОВО



Набокова —



*лиловый*



ЧТО МОЖЕТ РАССКАЗАТЬ  
СТАТИСТИКА О НАШИХ  
ЛЮБИМЫХ АВТОРАХ

**БОМБОРА™**

Москва 2019

УДК 82.09  
ББК 83.3  
Б68

Ben Blatt

NABOKOV'S FAVORITE WORD IS MAUVE:  
What the Numbers Reveal About the Classics, Bestsellers,  
and Our Own Writing

Copyright © 2017 by Benjamin Blatt  
Simon & Schuster, Inc., the original publisher

**Блатт, Бен.**

Б68 Любимое слово Набокова — лиловый. Что может рас-  
сказать статистика о наших любимых авторах / Бен Блатт ;  
[пер. с англ. Нестеркиной Н. С.]. — Москва : Эксмо, 2019. —  
288 с.

ISBN 978-5-04-096734-6

Великие писатели и авторы бестселлеров часто дают советы о том, как правильно писать, как важно найти свой уникальный стиль и быть оригинальным. Но всегда ли они сами пользуются своими советами? Как часто они используют клише и вставляют слова-паразиты? Следуют ли трехактной структуре? Как прописывают персонажей? Сколько обладателей Пулитцеровской премии вообще окончили университет? Эта книга даст вам ответы на все эти вопросы, но не совсем в обычной форме — внутри вы найдете схемы, графики и диаграммы, открывающие все секреты изнанки мира литературы. Они будут полезны не только любителям книг и желающим узнать больше о своих любимых авторах, но и начинающим писателям, которые имеют шанс научиться на ошибках лучших!

УДК 82.09  
ББК 83.3

ISBN 978-5-04-096734-6

© Нестеркина Н.С., перевод  
на русский язык, 2018  
© Оформление. ООО «Издательство  
«Эксмо», 2019

*Моей матери, Фейт Минард.  
И моим друзьям на 44 Bow Street.*

# Оглавление

Введение	9
<b>1.</b> Используй умеренно	17
<b>2.</b> Он написал, она написала	47
<b>3.</b> В поисках отпечатков	77
<b>4.</b> Пиши по примеру	101
<b>5.</b> Греховные удовольствия	123

<b>6.</b> Британия и Америка	139
<b>7.</b> Клише, повторы и любимые слова	163
<b>8.</b> Как судить книгу по обложке	197
<b>9.</b> О началах и концах	215
Эпилог	235
Благодарности	239
Примечания	241



# Введение

**А**лександр Гамильтон, Джеймс Мэдисон или Джон Джей? На протяжении более 150 лет историки пытались выяснить, кто же автор 12 записок из «Федералиста» — основополагающего документа в истории американского марша на пути к демократии. Несмотря на всемирную известность сочинения в качестве символа американской истории, авторы отдельных статей долгое время оставались неизвестными. Вопрос о том, кто же из отцов-основателей сочинил эти эссе, спровоцировал такую длинную дискуссию, что она стала своеобразной салонной игрой, популярной среди историков. Так кто же на самом деле написал те принципы, на основе которых построено американское государство?

Ответ прятался в тексте «Федералиста» — но, чтобы найти его, филологи должны были не внимательно читать, а внимательно считать. Им нужно было внимать цифрам.

Эта загадочная история началась в конце 1787 года, когда в нью-йоркских газетах была опубликована серия эссе, продвигающих идею ратификации конституции, подписанная именем Публий. Попытка скрыть настоящие имена авторов за патриотичным псевдонимом казалась смехотворной. На самом деле примерно из 4 миллионов людей, проживающих

в США в 1787 году, можно было без особого труда вычеркнуть всех претендентов, кроме троих.

Ни для кого не было секретом, что Гамильтон, Мэдисон и Джей были настоящими авторами, но никто из них не хотел признаваться в авторстве конкретных эссе. Каждый преследовал свои политические амбиции — позднее они доросли до министра финансов, президента и председателя Верховного суда соответственно. Иными словами, у них были серьёзные причины вести себя осмотрительно. В итоге их чрезмерная осторожность привела к тому, что тайна авторства так и осталась неразгаданной, ещё долгие годы волнуя профессоров истории и «диванных» энтузиастов.

Вы можете подумать, что в то время исследователи и проницательные политики могли своими силами определить авторство статей. В конце концов, ведь было всего три потенциальных кандидата, каждый со своим политическим курсом и стилем общения. Это все равно что сейчас в *New York Times* появилась бы анонимная публикация, сочинённая Баракком Обамой, Хиллари Клинтон и Берни Сандерсом. Или же неподписанный манифест за авторством Джорджа Буша-младшего, Джона Маккейна и Дональда Трампа. Кандидаты могут выступать на одной стороне, но они точно не идентичны.

В 1804 году наконец было найдено решение. 11 июля 1804 года Гамильтон оставил записку своему другу Эгберту Бенсону, называя автора каждого эссе. Гамильтон готовился к дуэли с Аароном Берром, назначенной на следующий день, и, оценивая свои шансы на выживание, осознавал огромную историческую значимость «Федералиста». Он решил, что знание об истинных авторах не должно погибнуть вместе с ним.

Это событие должно было положить конец всей загадочной истории. У нации любопытных наблюдателей не было оснований сомневаться в информации от первоисточника. Тем не менее спустя 13 лет, вскоре после своего второго президентского срока, Мэдисон опубликовал собственный список авторов — отличный от того, что был написан Гамиль-

тоном. Мэдисон претендовал на авторство тех же 12 эссе, о написании которых ранее заявлял Гамильтон.

Начался новый виток горячих дебатов, и разногласия среди историков бушевали ещё более века. В 1892 году будущий сенатор Генри Кэбот Лодж написал статью, в которой поддерживал Гамильтона, в то время как известный историк Э. Г. Бурн выступал на стороне Мэдисона.

Большинство историков пытались определить авторов, опираясь на политическую идеологию, которая излагалась в каждом эссе. Стал бы Мэдисон поддерживать центральную банковскую систему на этих условиях? Поддержал бы с такой лёгкостью Гамильтон ограничение полномочий конгресса? Или же подобное скорее мог написать Джон Джей?

Только в 1963 году, два столетия спустя, ключ к загадке был наконец найден. Ответ подсказали профессор Гарварда Фредерик Мостеллер и Дэвид Уоллес из Чикагского университета. Однако в отличие от многих профессоров, которые пытались решить эту задачу до них, Мостеллер и Уоллес вовсе не были историками. Они не прославились благодаря научным работам по ранней истории Америки и не публиковали статей об известных исторических фигурах. Мостеллер и Уоллес были статистиками.

Одна из наиболее известных работ Мостеллера касалась Мировой серии<sup>1</sup> и рассматривала вопрос, достаточно ли семи игр, чтобы статистически определить лучшую бейсбольную команду. А за несколько лет до решения проблемы, кто же автор статей «Федералиста», Уоллес опубликовал работу под названием «Границы нормальных приближений распределений Стьюдента и хи-квадрат», что, вполне возможно, для вас звучит как полнейшая чушь. Точно так же в 1963 году для профессоров истории звучала мысль о том, что вероятностные функции могут решить историческую загадку.

Методология Мостеллера и Уоллеса, применённая для решения проблемы авторства, не имела ничего общего с по-

---

<sup>1</sup>Решающая серия игр в ежегодном чемпионате США по бейсболу.

литикой или идеологиями. Вместо этого они стали первыми статистиками, использовавшими такие параметры, как частота и вероятность употребления слова.

Процесс вычислений был довольно сложным, полным уравнений с факториалами, экспонентами, суммированиями, логарифмами и *t*-распределениями. Но суть этих методов была поразительно проста:

- Подсчитать частоту появления общеупотребительных слов в эссе, которые, как мы знаем, написали Гамильтон или Мэдисон.
- Подсчитать частоту употребления этих же слов в эссе, где автор неизвестен.
- Сравнить эти показатели частотности, чтобы определить автора спорных статей.

Ещё до того, как были применены вероятностные уравнения, результаты статистического подхода в ретроспективе стали казаться совершенно очевидными. В «Федералисте» Мэдисон использовал слово *в то время как* (*whilst*) практически в половине статей, где его авторство было подтверждено, но ни разу не употребил слово *пока* (*while*). Гамильтон же, напротив, использовал слово *while* примерно в трети своих эссе, но никогда не употреблял *whilst*.

Однако Мостеллер и Уоллес не стали полагаться в своём анализе только на одно слово, что было бы статистически ненадежным. Вместо этого они систематически выбирали десятки базовых слов и затем выясняли частотность употребления каждого из них в эссе со спорным авторством. Оказалось, что употребление многих слов, имеющих совсем не политическое значение, у двух авторов разительно отличалось. Например, Мэдисон использовал слово *also* (*также*) в два раза чаще Гамильтона, а Гамильтон обращался к слову *according* (*в соответствии*) намного чаще, нежели Мэдисон.

Критерий фальсифицируемости сыграл на руку Мостеллеру и Уоллесу. Они могли продемонстрировать на статьях,

чьё авторство было достоверно подтверждено, что их методы действительно могут определить автора с превосходной точностью. Проанализировав 12 эссе, чьё авторство оспаривалось, Мостеллер и Уоллес пришли к выводу, что автором всех 12 является Джеймс Мэдисон.

В письменном изложении результатов своего исследования математики были предельно осторожны, возможно, из-за опасений разозлить историков, которые на протяжении нескольких поколений чесали головы в поисках ответа. Однако цифры, представленные в эксперименте, говорили сами за себя; они свидетельствовали в пользу абсолютной надёжности их метода. Для всех текстов, где авторство было подтверждено, метод срабатывал безупречно; для всех эссе с неизвестным автором результаты были постоянными. Заявление Гамильтона об авторстве было необоснованным.

Сегодня, когда было проведено бесчисленное количество как статистических, так и других исследований, выводы Мостеллера и Уоллеса — что Мэдисон и был автором — стали общепринятыми и среди статистиков, и среди историков. Мостеллер и Уоллес шли впереди своего времени. Их исследование, включающее некоторые сложные формулы, опиралось исключительно на подсчёт слов. Благодаря современным компьютерам подсчёт слов и частоты их употребления стал обыденным явлением. Но в 1963 году это было ошеломляющей новинкой.

Тогда слова подсчитывались вручную: например, чтобы узнать, сколько раз слово *ироп* появилось в эссе, они штудировали страницу за страницей. Чтобы в полной мере понять, через что прошли Мостеллер и Уоллес (или как минимум их ассистенты по исследованию), я распечатал все страницы «Федералиста» и начал считать, сколько раз употреблялось слово *ироп*. Через полчаса я осилил только одну восьмую часть — примерно 40 страниц — и насчитал 37 случаев использования слова *ироп*. Спустя некоторое время мои глаза начали болеть, а голова стала совсем тяжёлой. Задача «най-

ди *ирон*» стала похожа на изошрённую версию детской игры «Найди Вальдо».

Я перестал притворяться, что на дворе стоит 1963 год. Вместо этого я сделал несколько подсчётов, возможных только в 21 столетии: я полез в Google, забил в строку «Федералист полный текст», кликнул по ссылке первого результата в поиске и открыл файл в Microsoft Word. После двухминутных манипуляций и запроса «Найти всё» для слова *ирон* я получил 46 случаев употребления в той части, которую до того просмотрел вручную. Компьютер проделал работу не только на 28 минут быстрее, но и в разы точнее, чем мои уставшие глаза.

Ещё более потрясающее открытие: количество времени, необходимое человеку, чтобы изучить текст «Федералиста» в поисках другого слова, будет исчисляться часами, тогда как время, затраченное компьютером, для любого слова практически не увеличится. Подобное исследование всех работ Шекспира, Библии, «Моби Дика» или даже собрания английской литературы было бы невыполнимой задачей для Мостеллера и Уоллеса. Сегодня же любому подростку по силам подсчитать количество упоминаний слова в большом тексте, используя компьютер.

Спустя 50 лет после того, как Мостеллер и Уоллес опубликовали своё исследование, использование компьютерных технологий для работы с текстом находит всё более широкое применение. Google использует анализ текста в своих поисковых запросах и для подбора рекламы. Исследователи анализируют текст, чтобы понять, почему твиты становятся виральными, а СМИ часто используют похожие заголовки с минимальными изменениями, чтобы повысить количество просмотров. Однако применение в технических компаниях — это лишь один из возможных вариантов.

Моллестер и Уоллес использовали статистику, чтобы найти ответ на один-единственный вопрос об авторстве. Успех их эксперимента оказался куда более значимым. Все писатели обладают своим уникальным стилем, постоянным

и предсказуемым. Оказывается, не только политики 18 века оставляют свои стилистические отпечатки. Авторы всех книг, будь они популярными и признанными или же, напротив, неизвестными и критикуемыми, воспроизводят свои слова и структуру текста на протяжении десятков лет их писательского пути.

Вопрос, на который ответили Моллестер и Уоллес, касался довольно ограниченной сферы. Однако анализ текста может ответить на множество различных вопросов, которыми долгие годы задавались любопытные писатели и читатели. Правда ли, что Эрнест Хемингуэй использовал меньше наречий, чем другие писатели? Как уровень навыка чтения влияет на популярность книги? Правда ли, что мужчины и женщины пишут по-разному? Следуют ли писатели своим собственным советам и так хороши ли эти советы? Что, помимо очевидной орфографии, отличает американских и британских романистов? Какие любимые слова у наших любимых авторов, от Владимира Набокова до Э. Л. Джеймс?

Пока в научных кругах растёт интерес к изучению структуры письма успешных авторов, по-прежнему появляются вопросы, на которые нам ещё предстоит ответить. Эти вопросы одинаково интересны и полезны как для типичного читателя литературного мейнстрима, так и для начинающего писателя. Вас, скорее всего, не интересует распределение Пуассона или программы для парсинга<sup>1</sup>, используемые для определения частей речи, но, вероятно, вы бы хотели узнать, как пишет ваш любимый автор — и что это может сказать о вас как о читателе.

Аналитический подход к писательскому мастерству может быть удивительным и информативным, а зачастую даже весёлым. Кроме того, он позволит нам больше узнать об авторах, которых мы читаем каждый день, и о словах, что мы сами используем. Именно об этом мы поговорим в этой кни-

---

<sup>1</sup> Парсинг (parsing) — синтаксический анализ текста в лингвистике и информатике.