

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

Макагонов П.П., Маруев С.А.

**Применение критерия распределения связей между
элементами сложной системы для оценки
совершенства целостного текста**

Москва 2020

Аннотация. Известно, что тексты можно рассматривать как сложные социальные системы, и для оценки их совершенства используется закон Ципфа. К критериям совершенства, зрелости и целостности текстов как сложных систем предлагается добавить ранговое распределение связей (PPC) между объектами системы. Предлагается эмпирический закон распределения в текстах служебных слов и слов –связок. Для поиска математической формулировки этого закона рассматриваются примеры сложных систем, модели которых связаны с теорией графов, с обменом сообщениями на форумах в интернете, с распределением городского населения по районам, с транспортными сетями. Для этих систем «на инженерном уровне строгости» показано, что эмпирический закон распределения связей между объектами имеет общее математическое представление (PPC).

Ключевые слова: закон распределения Ципфа-Мандельброта, распределение связей между объектами системы; значимые слова (significant words); служебные слова (syntactic words), слова – связки, коэффициент детерминации (R-squared measure of goodness or Coefficient of determination), золотая пропорция.

Макагонов П.П., профессор кафедры системного анализа и информатики Факультета информационных технологий и анализа данных Института экономики, математики и информационных технологий Российской академии народного хозяйства и государственной службы при Президенте РФ;

Маруев С.А., заместитель директора Института экономики, математики и информационных технологий Российской академии народного хозяйства и государственной службы при Президенте РФ.

Данная работа подготовлена на основе материалов научно-исследовательской работы, выполненной в соответствии с Государственным заданием РАНХиГС при Президенте Российской Федерации на 2019 год.

Оглавление

Введение.....	4
1. Критерий совершенства сложной системы, использующий связи между её объектами.....	1
2. Распределение связей на графах.....	7
3.. Формирование и анализ частотного списка служебных слов и слов – связок.....	8
4. Соотношение служебных и знаменательных слов в целостном тексте. Золотая пропорция.....	12
5. Лингвистические характеристики текстов Библии.....	14
6. Лингвистические характеристики переводов на русский язык смыслов Сур Корана.....	29
7. Проблема составления списка служебных слов в монопольных текстах.....	33
8. Применение распределения слов узкой предметной области на форуме в Интернете.....	37
ЗАКЛЮЧЕНИЕ.....	37
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	38

Введение.

Общепризнанным необходимым признаком совершенства системы является выполнения для ранжированного распределения объектов этой системы закона Ципфа ([1-6], [13-17]) $F(N) = F_1/N^K$ где F_1 -значение функции $F(1)$ – константа, близкая, или совпадающая со значением в первой точке выборки, K – значение, близкое к единице, а N – ранг объекта, то-есть, его номер в выборке при расстановке величин $F(N)$ в порядке убывания. Этот закон часто бывает удобнее представлять, взяв логарифм по любому основанию от гиперболической функции: $\log F(N) = \log F(1) - K \log N$. или

$$\log F(N) = A - K \times \log N. (1)$$

Закон Ципфа справедлив для распределения знаменательных слов. При рассмотрении распределения всех слов (как знаменательных, так и служебных) некоторого целостного текста или коллекции текстов по одной тематике, часто рассматривается в качестве более близкого к совершенству закон Ципфа – Мандельброта:

$$F(N) = \frac{F_1}{(B+N)^K} (2),$$

где B – некоторая положительная константа.

Однако система должна характеризоваться не только своими объектами, но также и связями между ними. Поэтому в совершенной системе должна выполняться не только высокая степень согласованности распределения главной количественной характеристики её элементов, но должен также существовать критерий совершенства связей между этими элементами. Задача поиска такого критерия осложняется, если нет очевидного кандидата на роль параметра, который должен характеризовать связи между элементами, которые определяются как объекты системы.

Для систем, имеющих геометрические или географические характеристики, в качестве объектов обычно рассматриваются локальные образования или точки с характеристиками типа веса, массы или заряда. Тогда связи между объектами определяются характеристиками типа расстояний между этими точками или силы притяжения между точечными массами. В этом случае распределение связей по их количественным характеристикам можно оценивать как соотношение «ранг – размер», где ранг соответствует номеру связи в порядке убывания её значимости. Сами связи могут рассматриваться либо как попарные между объектами системы, либо как отношения каждого объекта ко многим. В исследовании рассматриваются сложные системы, связанные с социальными и гуманитарными науками, но в качестве примеров берутся только те из них, для которых можно сравнительно легко найти статистические количественные характеристики. Примерами таких систем могут быть:

- статические не изменяющиеся во времени системы, например, тексты законченных произведений, не подлежащие переписыванию, или математические модели, созданные для теоретического изучения реальных систем в их фиксированном состоянии;
- динамически изменяющиеся системы, но стабилизирующиеся во времени по главным характеристикам, например, тексты форумов в Интернете, или производственные проекты, имеющие фиксированные границы жизненного цикла;

- динамически развивающиеся системы, о стадии жизненного цикла которых трудно сделать обоснованное заключение, например, системы расселения и/или системы связей между ними.

1. Критерий совершенства сложной системы, использующий связи между её объектами.

Для перечисленных выше примеров систем были проведены исследования [8-11] в которых было показано, распределение (1) на билогарифмическом бланке (по каждой из осей отложены логарифмы сущностей) имеет выпуклую форму, близкую к четверти эллипса. Практика показывает, что для удобства количественного и качественного анализа вместо выпуклой линии в социальных (в том числе лингвистических) исследованиях удобнее работать с аппроксимирующей прямой. Было также показано, что для оценки степени совершенства социальной системы можно использовать степень соответствия распределения связей между элементами системы закону, который на билогарифмическом бланке при соответствующих значениях параметров имеет вид:

$$\text{Log}^G F(N) = A - K * \log^H N \quad (3),$$

где N - ранг элемента системы, а $F(N)$ – параметр, характеризующий связь N -го элемента с другими элементами системы, G , H , A и K -параметры, аппроксимирующие зависимость $F(N)$ от N . Параметры G , H подбираются по критерию максимального соответствия аппроксимирующей прямой исследуемому распределению.

Введем обозначения:

$$U = \log^G F(N) ; V = \log^H N \quad (4),$$

тогда для них зависимость (3) принимает вид

$$U = A - K * V \quad (5).$$

Сам критерий определяется как коэффициентом детерминации R^2 (или R^2) для выборки, сравниваемой с аппроксимацией по формуле (5). Этот критерий работает при оценке аппроксимирующей функции любого вида и потому является наиболее предпочтительным. На основе практики применения примем за нижнюю допустимую границу коэффициента детерминации величину 0,95. При значениях R^2 , лежащих в пределах от 0,95 до 1,0 будем считать, что распределение (3-5) указывает на совершенство системы с точки зрения рангового распределения связей (PPC или по-английски rank links distribution -RLD).

Формулу (5) с обозначениями (3) и (4) можно рассматривать как модифицированный закон Ципфа для связей между элементами системы. Часто показатели степени G и H равны между собой. При этом их значения больше единицы и обычно не превышают 4. При $G=H=1$ формула (5) соответствует закону Ципфа. Чтобы показать, что результаты анализа носят общесистемный характер (для сложных систем), ниже в таблицах №1 и 2 дана сводка результатов исследований, опубликованных в период с 2015 по 2018 год, и дополненных результатами, представленными в данной работе впервые.

В таблице №1 приведен перечень элементов систем и связей между этими элементами. Приведенные понятия требуют детального разъяснения, которое даётся далее.

Отдельные примеры количественных соотношений для связей «ранг – размер», аппроксимированных формулой (3), приведены в таблице №2. Эти соотношения приводятся не обязательно в точном соответствии с представленными ранее в статьях, но могут соответствовать результатам дальнейших исследований. Особенно это замечание относится к возможному изменению результатов при переходе к формуле (3), в которой показатель “Н” рассматривается как произвольное число, тогда как в прежних публикациях в некоторых случаях вводилось ограничение $H=1$.

Таблица №1

	Элементы системы	Связи между элементами системы
1	Совокупность узлов сложного графа	Веса дуг многосвязного графа
2	Множество знаменательных слов текста	Частота употребления слов -связок и служебных слов
3	Участники форума в Интернете	Характеристики активности или объёма сообщений каждого участника
4	Районы города	Численность населения по районам проживания
5	Железные дороги стран	Длина дорог по странам мира
6	Автомобильные дороги стран	Длина дорог по странам мира

Элементы системы в дальнейшем рассмотрении участвуют только своим наименованием и рангом в упорядоченном списке связей. Связи ранжированы по убыванию интенсивности, номер N - это ранг связи, а $F(N)$ – её интенсивность.

	Связи между элементами	Математическое представление интенсивности связей $\text{Log}^G F(N) = A - K * \log^H N$	Коэффициент детерминации
1a	Сила связи линейно расположенных узлов графа [8,10]	Граф с весами узлов $V = 100/N$ $G=1,5; H=1;$	0,9981
1b	Сила связи узлов «двухзвёздного» графа с различными длинами дуг [8,10]	$G=3,25; H=1; A=393,56; K=131,84;$	0,977
2a	Слова -связки и служебные слова сказки Alice's Adventures in Wonderland [8,10]	$G=2,75; H=1; A=15,389; K=7,1786;$	0,9961
2b	Слова -связки и служебные слова сказки The Valiant Little Tailor на английском языке [8,10]	$G=1,75; H=1; A=4,4118; K=2,0685;$	0,9960
2c	Слова -связки и служебные слова текста Homestead Act [8,10]	$G=H=1,25; A=2,5023; K=1,0801$	0,9838
2d	Слова -связки и служебные	The memoirs of the stone: $G=2; H=1.$	0,994

	слова текстов академика А. Ферсмана (число словоупотреблений) [11]	Mineralogy for fun: G=2; H=1.	0,991
3а	Число сообщений участника форума города Уахуапан (Мексика, штат Оахака) [18]	G=1,5; H=1; A=7,1953; K=1,1523	0,9621
3в	Число знаменательных слов в сообщении участника форума как функция ранга участника [18]	G=3,5; H=1,5; A=841,38; K=41,636	0,9847
3с	Слова связки всех сообщений форума Уахуапана (syntactic words in all the posts at the Huajuapán forum) [18]	G=H=1,5; A=21,925; K=0,9648.	0,9917
4а	Численность населения по районам проживания в Чикаго	G=H=3; A=7,0342; K=0,9504	0,9815
4в	Численность населения по районам проживания в городе Калгари (Канада)	G=H=4; A=19,451; K=0,7742	0,9839
5	Длина железных дорог по странам	G=4; H=1; A=731,12; K=335,22	0,9876
6	Длина автомобильных дорог по странам	G=4; H=1; A=2163; K=905,76	0,9878

Данные по пунктам 4, 5 и 6 получены в представленной работе и являются результатами анализа информации, взятой из Википедии по соответствующим поисковым запросам.

Главная трудность формализации указанных зависимостей состоит в выборе способов оценки интенсивности связей. Исключения составляют связи, определяемые длиной мировых транспортных связей. Совершенство этих распределений демонстрируется в строках 5 и 6 таблицы № 2 и далее не обсуждается. Заметим только, что в данном случае трудности могут возникнуть при определении объектов системы, а не связей между ними.

2. Распределение связей на графах.

Для связей на графе опробованы два способа измерения:

В первом варианте взят граф с линейно расположенными узлами и с весами узлов, определяемых по формуле

$$V_i = 100/i. \quad (6)$$

Оценка связей между узлами определяется по силе, вычисляемой по закону притяжения $F(i, j)$., то-есть, по произведению весов узлов (объектов системы), делённому на квадрат расстояния между узлами с номерами i и j :

$$F(i, j) = \frac{V_i V_j}{(i-j)^2} \quad (7)$$

Наилучшее представление (3-5) рангового распределения $F(i, j)$ по убыванию силы связи соответствует параметрам $G=1,5$; $H=1$, при коэффициента детерминации $R^2=0,9981$.

Отметим, что если элементы (объекты) системы упорядочить не по индексу “i”, а случайным образом, то распределение (3) силы связей, представленных выражением (6), и упорядоченных по убыванию ($F(N)$), меняет показатель степени G на 2,

$$\log^2 F(N) = 49,963 - 13,904 \times \log N$$

что при высоком значении коэффициента детерминации может рассматриваться как некий признак ухудшения качества распределения.

Во втором варианте граф представлен системой узлов, образующих дендрит подобный двум звёздным частям, соединённым мостом (см. рисунок №1), а веса узлов распределены по идеальному закону Ципфа, то-есть, по формуле (6)

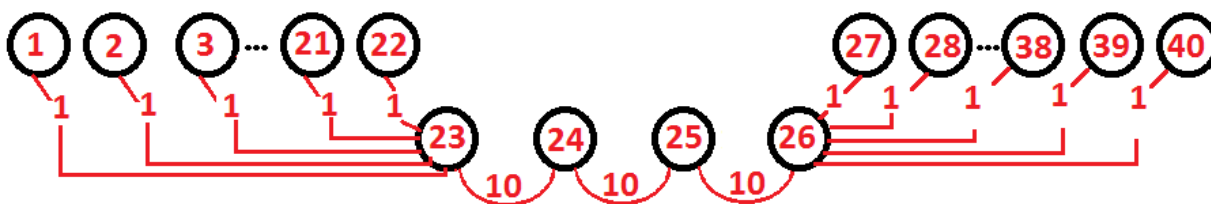


Рисунок №1. «Двухзвёздный» граф с мостами между узлами №№ 23-26.

Такая система возможна, например, при объединении двух систем поселения в одну административную систему. Для силы связей между элементами этой системы получаем распределение:

$$\log^{3,25} F(N) = 393,56 - 131,84 \times \log N$$

Критерий R^2 достигает максимума 0,977 при $G=3,25$.

В работе [8] представлены результаты для других графовых моделей. В частности, там рассмотрены модели двух дендритов: первый с симметричным централизованным распределением дуг и второй с почти линейным распределением. Веса узлов представлены ступенчатым распределением и аппроксимирующим его линейным (на билогарифмическом бланке) распределением Ципфа. Лучшие результаты аппроксимации (3) соответствуют случаю $G=2$, $H=1$ при максимуме критерия R^2 не ниже 0,98. В той же работе показано, что графовые модели могут представлять, например, архитектурный объект, и, значит, иметь социальный аспект. Поэтому вместе с остальными рассмотренными здесь системами они также относятся к социальным или гуманитарным.

3.. Формирование и анализ частотного списка служебных слов и слов – связок

В работе [10] представлен компьютерный анализ двенадцати текстов. К ним можно вернуться для более дательного детального анализа ставших почти очевидными утверждений, что для коротких текстов точность аппроксимации рассматриваемых здесь зависимостей падает. Поэтому для дальнейшей иллюстрации в явном виде используется только три текста из указанного списка (Alice's Adventures in Wonderland Льюиса Керола, сказка The Valiant Little Tailor на английском языке, и текст Homestead Act) и девять новых текстов академиков А Ферсмана и В.Вернадского. Вместе с тем в неявном виде здесь

используется ещё около двухсот текстов безымянных авторов – участников форума города Нуајуарап (Мексика).

Связи в тексте интерпретируем следующим образом: объектами в тексте признаются только знаменательные слова (significant words), имеющие в узкой предметной области один и тот же смысл даже вне контекста; связями в тексте между объектами (знаменательными словами) признаются те служебные слова, которые подменяют знаменательные слова в пределах абзаца. Это относится к местоимениям, предлогам, артиклям. Эти слова меняют своё значение с нарастанием удаленности от заменяемого слова. То же вспомогательное значение, оцениваемое как связь между знаменательными словами, имеют и вспомогательные глаголы.

Трудность в формировании списка служебных слов состоит в том, что и они, и знаменательные слова формально являются аналогичными элементами языка, отличающимися только деталями. Словарь служебных слов невелик, но относительная частота их употребления выше, чем у знаменательных слов. Поэтому, если заранее составленного списка у исследователя нет, то его можно составить путём анализа нескольких текстов (коллекции текстов) из той же предметной области, которой принадлежит изучаемый текст. Для этого достаточно составить матрицу «слова – тексты» и в ней провести фильтрацию той части частотного списка слов, которая имеет низкие значения отношения стандартного отклонения к среднему значению частоты встречаемости слова в коллекции текстов, или путем надлежащего использования статистической меры TF-IDF (TF — term frequency, IDF — inverse document frequency)

В качестве слов-связок и служебных слов выделяем только те, которые являются производными, или же производными, но содержат части знаменательных слов, которые практически потеряли своё значение в контексте того текстового документа, который мы исследуем.

Примерами производных слов в английском являются, например prepositions, postpositions, conjunctions, particles, articles, pronouns и вспомогательные слова, которые академик В.А. Виноградов рассматривает в «Академической грамматике» (2001) как укоренившиеся взгляды.

Если знаменательные слова или их части, которые содержатся в производных словах связках, не потеряли своё исходное значение, то они не были внесены в список SLW.

Принцип выделения этих слов синтетический, объединяющий логико-семантический и грамматический подходы. В таблице 3 даны примеры элементов, включаемых в частотный список SLW и, включаемых без слов, которые могут быть использованы в текстах как знаменательные.

Таблица 3 – Элементы списка служебных слов

Items included in list of SLW	Items partly included in list of SLW by syntactic words and linkers.
Lastly, Also, Further, Since	Firstly, secondly, finally
Because, Because of	The first point is
As, So, Henc, Therefore, Thus	The following
The, a, an	owing to the fact that
For, of, off, above, in, to	due to the fact that
While	Consequently

Whereas	This means that
He, she, We, they ...	As a result , as a rule, In addition
What, whatever, somewhere, whose,	In spite of the fact
These, us, them, our	For example , For instance
Themselves	it is said, as I see it
And, or	As well as that

Например, в список слов-связок мы включили такие слова, как COULD, MIGHT, BEING, PERHAPS, THROUGH, FEW, MOST, HADN[‘t], NEARER, EXCEPT, LESS, NEITHER, OTHERWISE, AROUND, BELOW, , NEEDN[‘t] AFTERWARDS.

А из слов связок обозначающих обобщение (Summarizing): In short, In brief, In summary In conclusion In a nutshell оставили только “ in”.To summarize To conclude сохраняем только “to”. keep only "to".

В работе [8] были проанализированы частотные списки SWL для двух целостных английских текстов. The Valiant Little Tailor, A fairytale by the Brothers Grimm Alice’s Adventures in Wonderland, by Lewis Carroll. Результаты представлены в таблице №2 (строки 2а и 2в). Там также было показано, что распределение знаменательных слов в тексте «Алисы» более соответствует распределению Ципфа –Мандельброта, а распределение знаменательных слов в тексте о портняжке более соответствует закону Ципфа. Но величина $R^2 = 0,9282$ для распределения знаменательных слов в тексте Керола, не позволяет считать текст совершенным. Возможно, что это связано с малым объёмом при большом числе слов, встреченных один раз (Это так называемые гапаксы). Это ухудшает качество ступенчатых распределений, то-есть распределений, имеющих много редко встречающихся знаменательных слов. В этом отношении авторские сказки могут проигрывать по отношению к народным традиционным сказкам с рационально минимизированным словарём.

Лучшая аппроксимация для силы связей в тексте о портняжке достигается при $G=2,75$. В то же время интересно отметить, что параметр R^2 плавно меняется в пределах 0,9941 -0,9961 в широком диапазоне степени $G= 2,5 - 3$, то-есть имеет пологое поведение при приближении к максимуму.

Аналогичная ситуация возникает при анализе Хомстед акта. (Закона об усадьбах). Закон 1862 года содержит 300 различных знаменательных слов, используемых 604 раза; 100 различных служебных слов использовано 871 раз, и в общей сложности 1475 слов. Таким образом, это довольно короткий текст. Результаты приближения распространения знаменательных слов, и распределения «силы связей» между словами в тексте, представленных в таблице 4. Текст Хомстед акта содержит менее 1800 слов, менее того количества слов, для которого мы определи возможность применения развиваемой методики. При малом числе знаменательных слов аппроксимирующая функция имеет слабый коэффициент детерминации. Тем не менее для служебных слов, которых в тексте больше, чем знаменательных коэффициент детерминации имеет высокое значение.

Таким образом, при невысоком качестве распределения знаменательных слов и при высоком качестве распределения SLW в совокупности получаем свидетельство высокого качества текста за счёт оценки аппроксимации силы связей.

Таблица 4 – Распределения знаменательных слов и распределения «силы связей»

<i>Homestead Act of 1862 (Engl)</i> $Y^{\wedge}G=-K \times X+A$; $Y=LogF$; $X=logN$.				
Analyzed objects	G	K	A	R ²
Significative words distribution	1	0,6365	1,483	0,9293
SLW distribution	1,5	1,5922	3,0834	0,9899

В работе [10] показано, что в научно-популярных трудах всемирно известного академика Ферсмана (Alexander E. Fersman (“The memoirs of the stone” and “Mineralogy for fun”)) параметры аппроксимирующих функций лежат в пределах, указанных в таблице 2 (строка (2с)). Столь же высокий уровень имеет коэффициент детерминации параметры аппроксимирующей функции для служебных слов в работах академика В.И.Вернадского, представленных в таблице 5. (к этой таблице вернёмся при рассмотрении критерия текстов, выражающегося через золотую пропорцию).

Таблица – 5 Критерии совершенства текстов академика В.И.Вернадского.

Работы В.И.Вернадского	Золотая пропорция в отношении служебных и знаменательных слов	G – показатель степени функции, аппроксимирующей распределение служебных слов.	Коэффициент детерминации для аппроксимирующей функции.
Очерк первый. Биосфера в космосе	0,6110	1,75	0,9915
Очерк второй. Область жизни	0,6088	1,75	0,9928
О научном мировоззрении	0,6216	1,75	0,9890
Научная мысль как планетарное явление	0,6185	2,0	0,9897
Несколько слов о ноосфере	0,6431	1,25	0,9830
Проблема времени в современной науке	0,6076	1,5	0,9918
По поводу критических замечаний акад. Д.М.Деборина	0,6776	1,25	0,9885

Аналогом этого случая для можно считать несовершенную систему из двух разнородных текстов или коллекцию текстов по узкой тематике со скрытым присутствием второй тематики.

В таблице, показано, что слияние текстов The Valiant Little Tailor (далее Тейлор) и Alice’s Adventures in Wonderland (далее Алиса) приводит к снижению (хоть и к незначительному) коэффициента детерминации при одновременном повышении степени аппроксимирующей функции.

Таблица 6 – Оценка объединенного текста

Text document	R ² max	G for SLW
---------------	--------------------	-----------

“Alice & “Tailor”	0,9943	2,5
“Alice”	0,9954	2,25
“Tailor”	0,996	1,75

В этой связи более высокие значения коэффициентов G и H можно рассматривать как признак снижения качества теста.

4. Соотношение служебных и знаменательных слов в целостном тексте. Золотая пропорция.

Появление в интернете множества текстов учебной и научной направленности требует оперативной оценки качества этих материалов, которые фактически не проходят никакой экспертизы, но легкодоступны для студенческой среды. Это последнее, в свою очередь, может иметь отрицательные последствия по различным причинам. Для лингвистического анализа такого целостного текста важным моментом является использование всего его словарного объема, а не какого-либо фрагмента.

Если рассматривать целостный текст как систему, то с позиций системного анализа знаменательные слова такого текста можно рассматривать как объекты, а служебные слова – как связи между объектами.

Для текстов, содержащих менее 1800 слов, показатель соизмеримости с «золотым сечением» не работает в силу малого объема (недостаточного насыщения) списка служебных слов.

Как показывает практика и научные изыскания, золотая пропорция, установленная между основными элементами системы, обеспечивает устойчивость и гармонию в технических, экономических, общественных системах, а также архитектуре, искусстве и многих других областях.

Число 1,618 называется золотым числом, а 0,618 – идеальное выражение пропорциональности, в процентном округленном соотношении золотое сечение – это деление какого-либо значения в отношении 62% и 38%.

В природе и во многих сферах человеческой деятельности обнаруживается пропорции золотого сечения, являющиеся условием устойчивости системы. Если пропорция между элементами какой-либо системы соответствует золотому сечению, то она обеспечивает: гармоничность системы, устойчивость, минимальное количество средств на поддержку данной устойчивости.

Для текстов, содержащих не менее 1800 словоупотреблений, получены следующие выводы:

Соотношение ϕ_1 между числом словоупотреблений служебных (b) и знаменательных (a) слов, или ϕ_2 - между «a» и общим числом словоупотреблений (a+b) в целостном тексте близко к «золотой пропорции» ($\phi = 0,618$) для более совершенных (гармоничных) тестов.

Для «совершенных текстов» установлена степень согласованности с предлагаемым распределением со значениями ϕ_1 и ϕ_2 , отклоняющимся от «золотой пропорции» в пределах между 0,73 и 0,535. Для текстов, которые, по всеобщему мнению, не признаны самыми совершенными в данной области знаний или в данном жанре хотя бы на национальном уровне, представленные показатели должны быть хуже.

Более того, выше было показано, что композиция двух «совершенных» текстов разных авторов теряет качество приведенных выше показателей и не соответствует уровню «совершенного» текста. При анализе научных статей это может быть критерием плагиаризма.

В качестве анализируемых в [10] текстов выбраны тексты, общепризнанные как тексты высокого качества и являются таковыми по субъективному мнению их автора, а также тексты, которые находятся в свободном доступе в интернете и распространяются там в качестве образцов учебных материалов для вузов и образцов курсовых работ. Последние должны иметь предположительно менее высокое, но приемлемое качество. К первому типу текстов относятся научно-популярные работы Ферсмана («Воспоминания о камне» и «Занимательная минералогия»), тексты второго типа обозначим условно как «Курсовая работа» и «Курс лекций».

Для проведения анализа каждый текст в отдельности обработан программным комплексом DBD, алгоритмы и возможные результаты использования которого, описаны в работе [10]. В результате применения этого комплекса для каждого текста получены следующие параметры:

- число знаменательных слов и число их словоупотреблений,
- число служебных слов и слов – связок, а также число их словоупотреблений.

По этим параметрам был получен предложенный выше критерий качества тестов: параметр «золотого сечения» ϕ .

«Воспоминания о камне» имеют показатели (8):

$$\phi_1=0,626; \phi_2= 0,589. \quad (8)$$

Для «Занимательной минералогии» результаты несколько хуже:

$$\phi_1=0,647; \phi_2= 0,547; R^2=0,991; G=2. \quad (9)$$

Для «Курса лекций» по минералогии (30):

$$\phi_1=0,697; \phi_2= 0,436. \quad (10)$$

Для «Курсовой работы» по вулканизму (31):

$$\phi_1=0,753; \phi_2= 0,328. \quad (11)$$

Полученные результаты, выписанные в порядке убывания качества по критерию ϕ , при нейтральном поведении параметра G , соответствуют ожидаемым с позиции здравого смысла. Рассмотренные критерии и методика их использования могут быть полезны для экспресс – анализа качества научных текстов.

Отметим, что проверка изложенной методики проводилась на текстах, высокое качество которых является общепризнанным на национальном и даже на мировом уровне. Тексты, относящиеся к наукам о Земле, имеют более узкую аудиторию. Однако, «Воспоминания о камне» и «Занимательная минералогия» (именно в таком порядке), были прочитаны гораздо большим числом молодых людей, выбирающих себе профессию, чем

это можно сказать о трудах В.И.Вернадского. «А.Е. Ферсман, внесший вклад в развитие наук о Земле не меньший, чем В.И. Вернадский, создал вместе со своим учителем современную геохимию, изучающую историю атомов и элементов в земной коре. Этих ученых, посвятивших свою жизнь новой науке, подарившей им мировую известность, А.Н. Толстой назвал «поэтами камня». Вышеуказанные труды А.Е.Ферсмана были выбраны по причине своей научно-популярной направленности, которая и явилась причиной большей популярности по сравнению, например, с его работой «Исследования в области магнезиальных силикатов» (1913).

5. Лингвистические характеристики текстов Библии

Проведём анализ частотного списка слов Библии по тексту Синодального перевода.

Согласно Википедии, Синодальный перевод — устоявшийся термин, обозначающий перевод книг Священного Писания на русский язык, осуществлённый в течение XIX века и утверждённый Святейшим Правительствующим Синодом для домашнего (не богослужебного) чтения. Для богослужений в Русской православной церкви используется перевод Библии на церковнославянский язык, так называемая Елизаветинская Библия.

Выбор для анализа этого текста определяется несколькими причинами:

Поскольку мы утверждаем, что закон Ципфа и закон связей между элементами системы указывает на степень зрелости, или совершенства системы, то для Библии критерии зрелости должны выполняться, можно сказать, безусловным образом.

Есть все основания полагать, что Библия состоит из текстов, написанных разными авторами, но содержание каждого текста в какой-то мере является обращением к определенному кругу лиц, развивающих и дополняющих мысли предшествующих авторов. В этом плане собранию текстов Библии можно уподобить коллекции текстов, возникающих в настоящее время на форумах, обсуждающих заинтересованными лицами фиксированный круг связанных между собой проблем. Поэтому Библию можно рассматривать в некотором роде в качестве идеальной модели для собрания материалов зрелых форумов. Распределение служебных слов, представленное на нижеследующем графике, демонстрирует некоторое отклонение первых четырех слов (И, В, НЕ, НА) от идеальной прямой. Асимптота распределения слов на следующем графике демонстрирует идеальное соответствие распределения слов закону Ципфа – Мандельброта. Показанное далее на графике распределение первых четырёх слов по текстам Библии может служить объяснением сильного отклонения служебного слова «И» в нарративных текстах (рисунок 4).

Последующая таблица распределения служебных слов в первой тысяче рангового распределения позволяет построить само распределение служебных слов (PPC) и показать, что служебные слова удовлетворяют этому распределению с коэффициентом детерминации 0,9784.

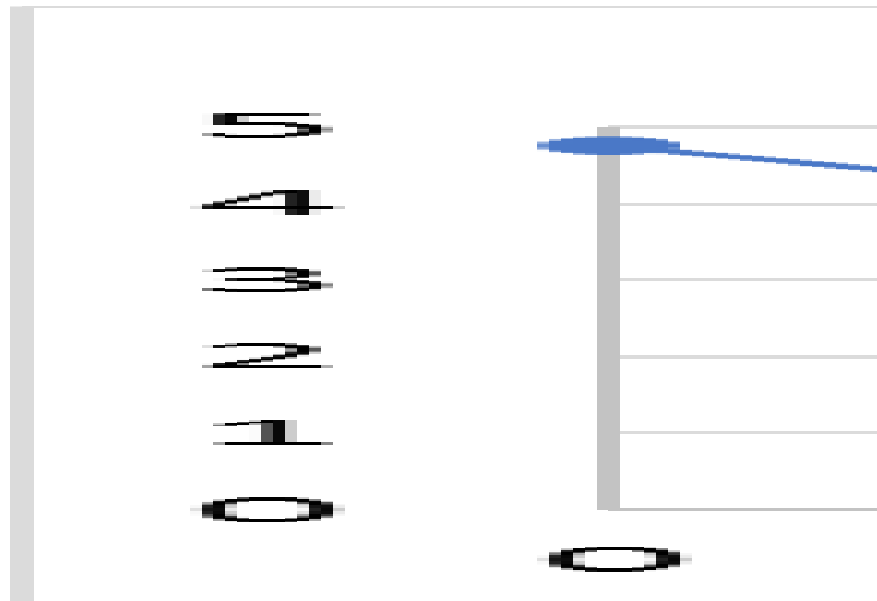


Рисунок 2 – Частотное распределение слов Библии

По оси ординат логарифм частоты словоупотребления, по оси абсцисс – логарифм ранга слова в упорядоченном по убыванию частотном списке.

На рисунке видно, что кроме первых 5 слов, остальные хорошо аппроксимируются кривой Ципфа-Мандельбротта. При этом начиная с 100-ого слова частотное распределение выходит на прямую (рисунок 3) с коэффициентом детерминации 0,9927.

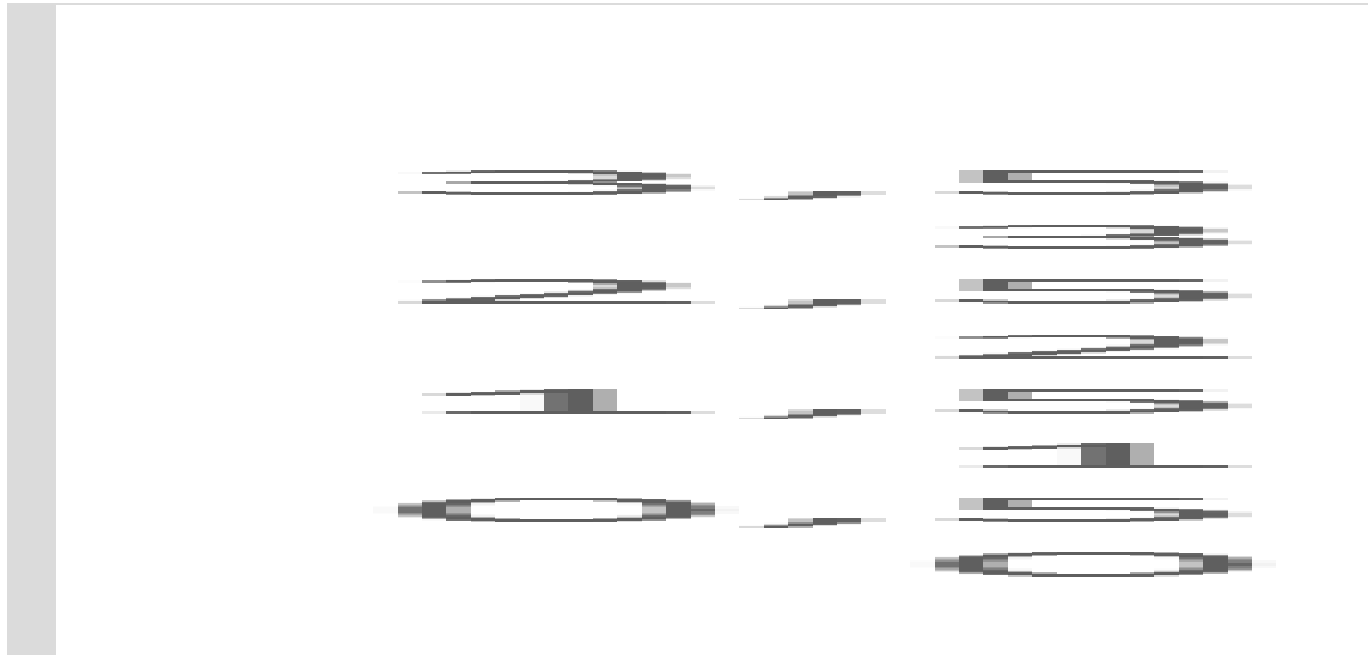


Рисунок 3 – Асимптота частотного распределения слов Библии

По оси ординат логарифм частоты словоупотребления, по оси абсцисс – логарифм ранга слова в упорядоченном по убыванию частотном списке.



Рисунок 4 – Распределение служебных слов по текстам в промиях.

По оси ординат показана частота употребления слов по всему корпусу текстов. По оси абсцисс – номер текста в корпусе в порядке, представленном в таблице 29.

Анализ распределения первых 4 слов с наибольшей частотой показывает, что это служебные слова из которых первое слово «и» сильно отличается по частоте употребления от остальных. По видимому, это можно отнести к особенностям стиля перевода, сделанного в XIX веке, союз «и» выполняет еще и усилительную роль.

Таблица 7 – Количество служебных слов в текстах

Текст	Число слов	и	в	не	на
Книга пророка Иеремии-29190	29190	87	26	22	15
Псалтирь-28684	28684	63	20	18	19
Бытие-27710	27710	105	23	12	10
Книга пророка Иезекииля-27401	27401	94	25	14	19
Книга пророка Исаии-24437	24437	94	22	22	17
Исход-22348	22348	99	23	15	18
Числа-21782	21782	90	29	11	12
Второзаконие-20352	20352	84	23	26	16
Вторая книга Паралипоменон-19051	19051	118	31	10	15
Книга Премудрости Иисуса-18960	18960	81	24	34	12
Первая книга Царств-18239	18239	106	27	16	13
Третья книга Царств-17631	17631	99	27	12	14
От Луки святое благовествование-17526	17526	81	27	21	12
Деяния святых апостолов-16896	16896	76	32	14	12
Первая книга Маккавейская-16672	16672	114	34	8	12
Четвертая книга Царств-16294	16294	114	34	12	14
Левит-16244	16244	79	22	20	27
От Матфея святое благовествование-16212	16212	78	24	23	12
Вторая книга Царств-14760	14760	104	22	12	12
Третья книга Ездры-13965	13965	79	20	21	12
От Иоанна святое благовествование-13840	13840	59	23	31	8
Первая книга Паралипоменон-13685	13685	98	20	5	10

Книга Иисуса Навина-13244	13244	88	26	12	13
Книга Судей израилевых-13124	13124	100	29	14	13
Книга Иова-12991	12991	82	18	27	15
Вторая книга Маккавейская-11039	11039	66	25	11	15
Книга пророка Даниила-10954	10954	104	19	16	14
Притчи Соломона-10414	10414	59	16	26	12
От Марка святое благовествование-10221	10221	88	27	21	11
Откровение святого Иоанна Богослова-8363	8363	120	19	15	23
Вторая книга Ездры-7861	7861	118	28	10	10
Книга Иудифи-7852	7852	82	28	15	18
Книга Неемии-7329	7329	98	25	12	11
Книга Премудрости Соломона-6943	6943	75	23	21	9
Первое послание к Коринфянам святого апостола Павла-6337	6337	47	20	42	6
Послание к Римлянам святого апостола Павла-6295	6295	48	25	33	5
Книга Товита-4990	4990	92	19	21	7
Первая книга Ездры-4971	4971	101	31	8	8
Третья книга Маккавейская-4901	4901	64	27	13	13
Послание к Евреям святого апостола Павла-4626	4626	60	18	30	6
Книга пророка Захарии-4328	4328	98	25	14	16
Второе послание к Коринфянам святого апостола Павла-4312	4312	46	36	36	9
Книга Екклесиаста-3943	3943	84	22	27	9
Книга пророка Осии-3526	3526	83	22	22	12
Плач Иеремии-3449	3449	57	23	33	16
Книга пророка Амоса-3323	3323	80	33	26	20
Первое соборное послание святого апостола Иоанна-2770	2770	58	31	30	3
Книга пророка Варуха-2342	2342	87	22	18	18
Послание к Галатам святого апостола Павла-2114	2114	38	23	36	4
Послание к Ефессянам святого апостола Павла-2090	2090	69	35	16	8
Книга пророка Михея-2078	2078	77	21	21	13
Книга Руфи-1806	1806	92	17	18	9
Песнь песней Соломона-1802	1802	37	17	12	17
Первое соборное послание святого апостола Петра-1651	1651	48	21	19	8
Соборное послание святого апостола Иакова-1609	1609	70	19	36	11
Первое послание к Тимофею святого апостола Павла-1599	1599	67	35	35	6
Послание к Филиппийцам святого апостола Павла-1574	1574	67	32	17	4
Послание к Колоссянам святого апостола Павла-1410	1410	74	41	12	5
Первое послание к Фессалоникийцам-1400	1400	74	24	26	4
Книга пророка Иоилия-1343	1343	105	20	15	19
Книга пророка Малахии-1242	1242	83	16	23	10
Второе послание к Тимофею святого апостола Павла-1155	1155	62	32	18	3

Второе соборное послание святого апостола Петра-1063	1063	66	36	19	11
Книга пророка Софонии-1053	1053	75	24	27	16
Книга пророка Аввакума-1009	1009	75	14	21	25
послание к Титу святого апостола Павла-997	997	55	28	26	5
Книга пророка Ионы-960	960	107	25	16	9
Книга пророка Наума-825	825	85	23	15	21
Второе послание к Фессалоникийцам-764	764	64	24	33	1
Книга пророка Аггея-763	763	79	24	14	22

В таблице 8 представлено начало распределения служебных слов в корпусе текстов, попавших в первую тысячу слов. Это представление показано на рисунке 5. Там же дано распределение квадратов логарифмов этих слов, соответствующее представлению по формуле (3). Это представление показано на рисунке 6

Таблица 8 – Характеристики распределения служебных слов

Слово	N=ранг	F(N) Число употреблений	X=Log(N)	Y=Log(F(N))	X ²	Y ²
И	1	56326	0	4,750709	0	22,56924
В	2	15948	0,30103	4,202706	0,090619	17,66274
НЕ	3	11673	0,477121	4,067182	0,227645	16,54197
НА	4	8805	0,60206	3,944729	0,362476	15,56089
ЕГО	5	8375	0,69897	3,922985	0,488559	15,38981
ЧТО	6	6117	0,778151	3,786538	0,605519	14,33787
ОН	7	5944	0,845098	3,774079	0,714191	14,24367
Я	8	5706	0,90309	3,756332	0,815572	14,11003
С	9	5347	0,954243	3,72811	0,910579	13,89881
ОТ	10	4973	1	3,696618	1	13,66499
ИХ	11	4929	1,041393	3,692759	1,084499	13,63647
ИЗ	12	4445	1,079181	3,647872	1,164632	13,30697
К	13	4252	1,113943	3,628593	1,24087	13,16669
КАК	14	4217	1,146128	3,625004	1,313609	13,14065
А	15	3571	1,176091	3,55279	1,383191	12,62232
ОНИ	16	3515	1,20412	3,545925	1,449905	12,57359
ТЫ	17	3474	1,230449	3,54083	1,514005	12,53748
ИБО	18	3382	1,255273	3,529174	1,575709	12,45507
ВСЕ	19	3099	1,278754	3,491222	1,635211	12,18863
ПО	20	3035	1,30103	3,482159	1,692679	12,12543
НО	21	2871	1,322219	3,458033	1,748264	11,95799
ЧТОБЫ	22	2820	1,342423	3,450249	1,802099	11,90422
ЖЕ	23	2771	1,361728	3,442637	1,854303	11,85175
У	24	2645	1,380211	3,422426	1,904983	11,713
ТО	25	2605	1,39794	3,415808	1,954236	11,66774
КОГДА	26	2463	1,414973	3,391464	2,00215	11,50203
БУДЕТ	27	2452	1,431364	3,38952	2,048802	11,48885
ЗА	28	2400	1,447158	3,380211	2,094266	11,42583
ДЛЯ	29	2311	1,462398	3,3638	2,138608	11,31515
ЕМУ	30	2142	1,477121	3,330819	2,181887	11,09436
МЕНЯ	31	2101	1,491362	3,322426	2,22416	11,03851

ТАК	32	2065	1,50515	3,31492	2,265476	10,98869
О	33	2050	1,518514	3,311754	2,305885	10,96771
ВО	34	2027	1,531479	3,306854	2,345428	10,93528
ТЕБЯ	35	1902	1,544068	3,279211	2,384146	10,75322
МНЕ	36	1879	1,556303	3,273927	2,422077	10,7186
ЕСЛИ	37	1813	1,568202	3,258398	2,459257	10,61716
ВЫ	38	1766	1,579784	3,246991	2,495716	10,54295
БОГ	39	1758	1,591065	3,245019	2,531487	10,53015
ИМ	40	1687	1,60206	3,227115	2,566596	10,41427
ВОТ	41	1683	1,612784	3,226084	2,601072	10,40762
ЭТО	42	1668	1,623249	3,222196	2,634938	10,38255
ПРЕД	43	1598	1,633468	3,203577	2,668219	10,2629
ТЕБЕ	44	1509	1,643453	3,178689	2,700937	10,10407
КТО	45	1494	1,653213	3,174351	2,733112	10,0765
ЕЕ	46	1487	1,662758	3,172311	2,764764	10,06356
КОТОРЫЕ	47	1445	1,672098	3,159868	2,795911	9,984765
ГОВОРИТ	48	1332	1,681241	3,124504	2,826572	9,762527
ДО	49	1329	1,690196	3,123525	2,856763	9,756408
ВАС	50	1324	1,69897	3,121888	2,886499	9,746185
НЕГО	51	1314	1,70757	3,118595	2,915796	9,725637
ВСЕХ	52	1307	1,716003	3,116276	2,944667	9,711174
НИХ	53	1294	1,724276	3,111934	2,973127	9,684135
БЫЛО	54	1290	1,732394	3,11059	3,001188	9,675768
БЫЛ	55	1274	1,740363	3,105169	3,028862	9,642077
МЫ	56	1274	1,748188	3,105169	3,056161	9,642077
КОТОРЫЙ	57	1269	1,755875	3,103462	3,083097	9,631474
ЛИ	58	1253	1,763428	3,097951	3,109678	9,597301
ПОТОМУ	59	1251	1,770852	3,097257	3,135917	9,593003
БУДУТ	60	1173	1,778151	3,069298	3,161822	9,42059
ДА	61	1143	1,78533	3,058046	3,187403	9,351647
НИ	62	1138	1,792392	3,056142	3,212668	9,340006
ВАМ	63	1070	1,799341	3,029384	3,237626	9,177166
ТОГДА	64	1065	1,80618	3,02735	3,262286	9,164846
НИМ	65	1053	1,812913	3,022428	3,286655	9,135073
БОГА	66	1046	1,819544	3,019532	3,31074	9,117572
БЫ	67	1031	1,826075	3,013259	3,334549	9,079728
НАД	68	988	1,832509	2,994757	3,358089	8,968569
ЕСТЬ	69	980	1,838849	2,991226	3,381366	8,947433
СВОЕГО	70	913	1,845098	2,960471	3,404387	8,764387
ИЛИ	71	906	1,851258	2,957128	3,427157	8,744607
ПРОТИВ	72	904	1,857332	2,956168	3,449684	8,738932
ТОТ	73	891	1,863323	2,949878	3,471972	8,701778
БЫЛИ	74	867	1,869232	2,938019	3,494027	8,631956
ОНА	75	864	1,875061	2,936514	3,515855	8,623113
НАС	76	827	1,880814	2,917506	3,53746	8,511838
МОЙ	77	818	1,886491	2,912753	3,558847	8,484132
ПРИ	78	805	1,892095	2,905796	3,580022	8,44365

ТВОЕГО	79	799	1,897627	2,902547	3,600989	8,424778
ТВОИ	80	764	1,90309	2,883093	3,621751	8,312227
ТОГО	81	763	1,908485	2,882525	3,642315	8,308948
ТВОЙ	82	761	1,913814	2,881385	3,662683	8,302378
СО	83	737	1,919078	2,867467	3,682861	8,22237
СВОИХ	84	726	1,924279	2,860937	3,702851	8,184958
МЕЖДУ	85	717	1,929419	2,855519	3,722657	8,15399
ВСЕМ	86	711	1,934498	2,85187	3,742284	8,13316
ТАМ	87	685	1,939519	2,835691	3,761735	8,041141
СЕБЯ	88	681	1,944483	2,833147	3,781013	8,026723
СЕБЕ	89	673	1,94939	2,828015	3,800121	7,997669
ПУСТЬ	90	671	1,954243	2,826723	3,819064	7,99036
МОИ	91	652	1,959041	2,814248	3,837843	7,91999
НЕТ	92	652	1,963788	2,814248	3,856463	7,91999
ЕЩЕ	93	621	1,968483	2,793092	3,874925	7,801361
ТВОИХ	94	617	1,973128	2,790285	3,893234	7,785691
НАМ	95	616	1,977724	2,789581	3,911391	7,781761
МОЕГО	96	599	1,982271	2,777427	3,929399	7,7141
НЕМ	97	568	1,986772	2,754348	3,947262	7,586435
ДАЛ	98	560	1,991226	2,748188	3,964981	7,552537
СВОИ	99	549	1,995635	2,739572	3,98256	7,505257
СВОЮ	100	535	2	2,728354	4	7,443914
КО	101	529	2,004321	2,723456	4,017304	7,417211
ТВОЕ	102	525	2,0086	2,720159	4,034475	7,399267
ВЕСЬ	103	514	2,012837	2,710963	4,051514	7,349321
НИМИ	104	512	2,017033	2,70927	4,068423	7,340144

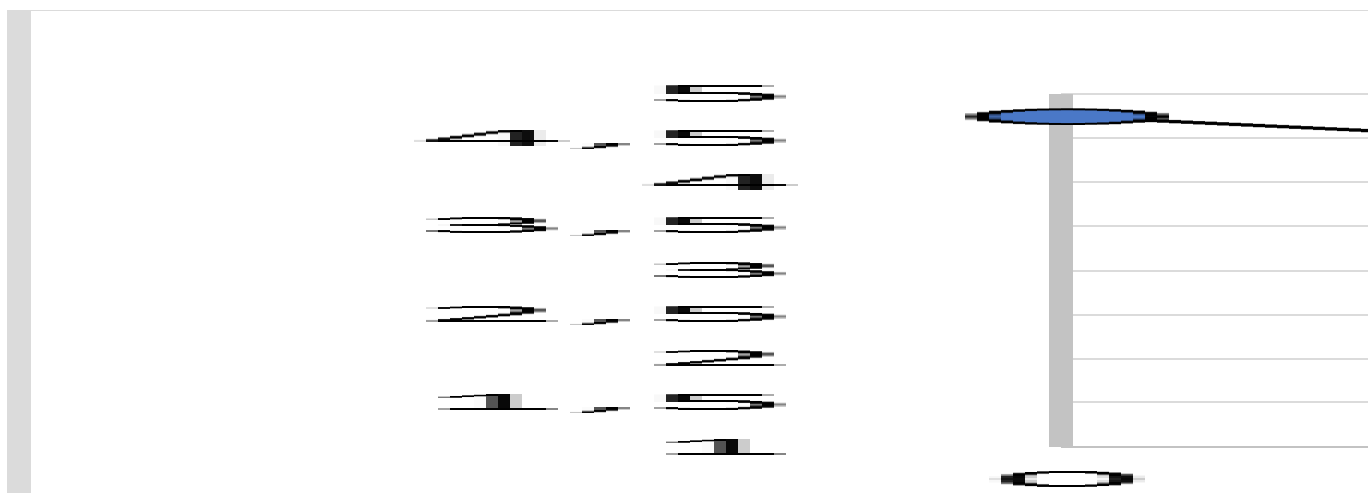


Рисунок 5 – Служебные слова из первой тысячи слов частотного списка слов Библии

Из рисунка следует, что помимо первых пяти служебных слов остальные 302 слова практически идеально ложатся на аппроксимирующую прямую. Коэффициент детерминации для этой совокупности равен 0,9784 (рисунок 6). Поскольку тексты всего корпуса написаны разными авторами, это показывает высокую согласованность словарей либо авторов, либо авторов перевода.

По оси ординат логарифм частоты словоупотребления, по оси абсцисс – логарифм ранга служебного слова в упорядоченном по убыванию частотном списке.

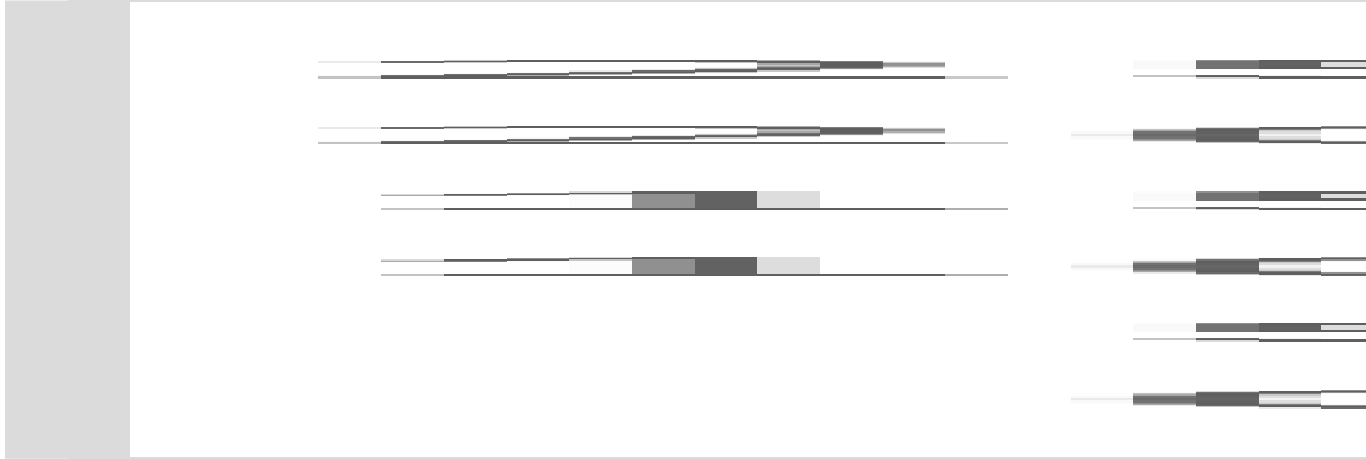


Рисунок 6 – Ранговое распределение связей при $G=H=2$ (по формуле (2))

На рисунке 7 приведен тоновый портрет матрицы распределения первой тысячи слов из корпуса текстов Библии по 70 текстам.

Количество строк и столбцов матрицы меньше вышеуказанных размеров, это объясняется несколькими словами в одной при корреляции этих строк выше 0,8. Коррелируемые слова объединены в одну строку и в последующем их упоминании, они объединены фигурными скобками. Можно утверждать, что совокупности слов в фигурных скобках – это n-граммы.

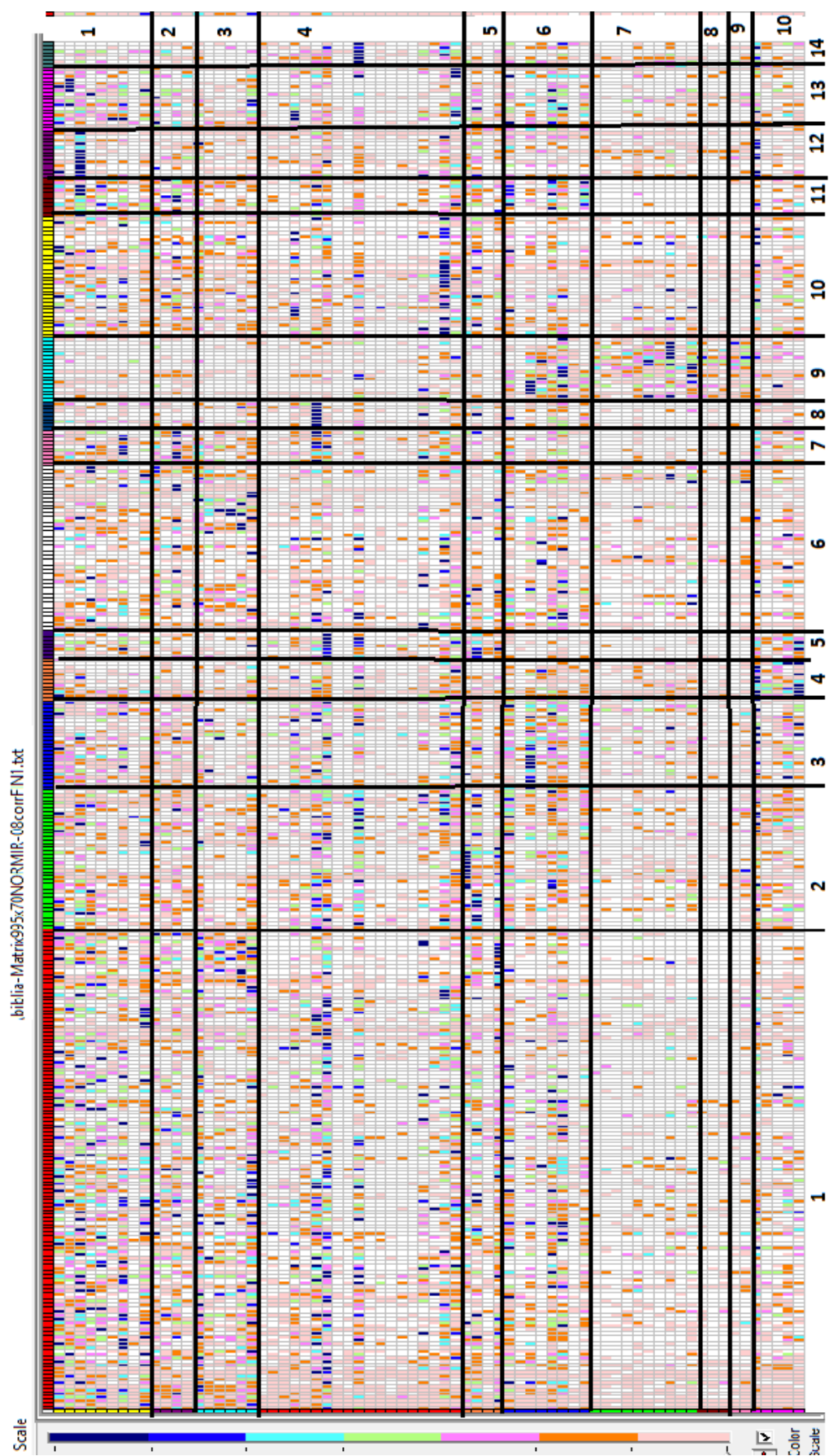


Рисунок 7 – тоновый портрет матрицы распределения первой тысячи слов из корпуса текстов Библии по 70 текстам

Далее построены кластеры текстов и слов и дается краткое описание особенностей сочетания и соответствия кластеров. Кластеры слов за счёт группирования внутри этих кластеров хорошо коррелируемых слов фактически имеют иерархическую структуру. Часть слов при анализ можно опустить по причине их повторяемости без приведения их к нормальной форме по падежу, роду, числу и т.п. Тогда для каждого кластера слов можно выделить сочетания, допускающие формирование концептов и способствующих проведению смысловой интерпретации.

Часть кластеров текстов слабо отличается по словарному составу. На рисунке 7 это кластеры текстов (с первой совокупности столбцов по четвертую). В них входят тексты большого объема, их заголовки начинаются обычно со слова «книга».

Кластер 1

- Бытие-27710,
- Вторая книга Царств-14760,
- Второзаконие-20352,
- Книга Иисуса Навина-13244,
- Книга Судей израилевых-13124,
- Книга Товита-4990,
- Первая книга Царств-18239,
- Песнь песней Соломона-1802,
- Четвертая книга Царств-16294.

Кластер 2

- Вторая книга Маккавейская-11039,
- Книга Иудифи-7852,
- Первая книга Маккавейская-16672,
- Третья книга Маккавейская-4901.

Кластер 3

- Вторая книга Паралипоменон-19051,
- Книга Неемии-7329,
- Вторая книга Ездры-7861,
- Первая книга Ездры-4971,
- Первая книга Паралипоменон-13685,
- Третья книга Царств-17631.

Кластер 4

- Книга пророка Аввакума-1009,
- Книга пророка Амоса-3323,
- Книга пророка Варуха-2342,
- Книга пророка Даниила-10954,
- Книга пророка Захарии-4328,
- Книга пророка Иезекииля-27401,
- Книга пророка Иеремии-29190,
- Книга пророка Иоили-1343,
- Книга пророка Ионы-960,
- Книга пророка Исаии-24437,
- Книга пророка Малахии-1242,
- Книга пророка Михея-2078,

- Книга пророка Наума-825,
- Книга пророка Осии-3526,
- Книга пророка Софонии-1053,
- Откровение святого Иоанна Богослова-8363,
- Плач Иеремии-3449,
- Псалтирь-28684,
- Третья книга Ездры-13965,

Кластер 5 текстов немного более содержит слов из кластера 2 (слов)

- Левит-16244
- Исход-22348
- Книга Руфи-1806
- Числа-21782

Кластер 2 слов

- {ДЕРЕВО ПТИЦЫ},
- РЕКИ, КОНЦА, {ВЕТВИ ДВОРА СТОРОНЫ ЖЕЗЛ},
- ИОРДАН, ДЕТИ, ВЫЙДЕТ, ЕГИПТЕ, {СТОРОНУ ШЛИ УМРЕТ ОДИН ОДНА ПУСТЫНЕ ОДНОГО ДВЕ ПЯТЬ ОДНОМУ ОДНО},
- СОЛНЦА, {ВОЗЬМЕТ СВЯЩЕННИК СВЯЩЕННИКУ ТУК ПРИНЕСЕТ ВОДОЮ ЖЕРТВА НЕЧИСТ ДОЛЖЕН ДОЛЖНА КРОВИ КРОВЬЮ ЖЕРТВЕННИКЕ ТЕЛЬЦА ЯЗВА УВИДИТ ВНЕ ЕШЬТЕ МЯСО ВЕЛИКАЯ БОЛЬШОЙ ЖЕРТВЫ КРОВЬ ДЕЛАЙТЕ СОГРЕШИЛ ПРАВОЙ},
- ОДНУ, СТОРОН, {ГОСПОДНИМ ГОСПОДНИХ},
- {ВЫСОТЫ ЕЗЕКИЯ ЕЗЕКИИ},
- {ГОВОРИТЕ ГОВОРИШЬ},
- ПИТЬ, СКОРО, ПРАХ, ПЛОДЫ, СЕМЯ, РАДОСТЬЮ.

Кластер 6 текстов чуть сильнее представлены кластеры слов 3 и 9

- От Матфея святое благовествование-16212,
- Книга пророка Аггея-763,
- От Иоанна святое благовествование-13840,
- Первое соборное послание святого апостола Иоанна-2770,
- От Луки святое благовествование-17526,
- Деяния святых апостолов-16896,
- Второе соборное послание святого апостола Петра-1063,
- От Марка святое благовествование-10221.

Кластер 3 (повторится далее)

{ИИСУС ГОВОРЮ ОТВЕТ УЧЕНИКИ ИСТИННО ИОАНН ЧАС ГОВОРИЛИ УВИДЕВ УСЛЫШАВ ПЕТР УЧЕНИКОВ ОТКУДА СУББОТУ ВОН ПРИХОДИТ},
 СМЕРТЬ, СТОЯЛ, ЗНАЕШЬ, СЕМЕНИ, ОТВЕЧАЛ, {ЖИЗНЬ ВИДИТ},
 ОТВЕЧАЛИ, ЗНАЛИ, ПОЧЕМУ, КАМЕНЬ, ЗНАЮ ЗНАЛ, ИМЕТЬ, НАПИСАЛ,
 ВИНО, ИМЕЛИ, ГРЕХА, РАДОСТЬ, ДВЕРИ, СВЕТ, СКАЖУ, {ДРУГ ДРУГА ДРУГУ},
 ХОЧЕТ, ИДИ,

Кластер 9 слов

ХРИСТОС, СВЯТЫМ, {МИР МИРА МИРЕ ЗНАЕМ},
 СВИДЕТЕЛЬСТВО, ЗНАЯ, ЗНАЕТЕ, ЗАПОВЕДЬ, ДУХОМ, {ИИСУСА ДУХА МЕРТВЫХ ПРИНЯЛИ},

БРАТИЯ, {ХРИСТЕ ИИСУСЕ ХРИСТА},
 ЛЮБВИ, БЛАГОДАТЬ, ПЛОТИ, ВЕРЕ, БОЖИЕЙ, ГОСПОДЕ, {ПАВЕЛ МУЖИ},
 Кластер 1 слов

{СКАЗАЛ СЫНА ВЗЯЛ СКАЗАЛИ ПРИШЕЛ ПОШЕЛ ГОВОРИЛ ПРИШЛИ
 ПОШЛИ ВЫШЕЛ МЕСТО СТАЛ ДОЧЬ КОТОРОМ ГОСПОДИН ВОШЕЛ ВСТАЛ
 ПОЙДИ ХЛЕБА СКАЗАВ ВЫШЛИ ДОРОГЕ ЕЛИ СКАЖИТЕ ПОЙДУ ВОШЛИ СЮДА
 ОДЕЖДУ ОТЕЦ СКАЗАЛА ГОСПОДИНА ГОСПОДИНУ ОТЦУ ОТЦА ЖИЛ ЖЕНУ
 СЕРЕБРО МАТЕРИ ЛЕТ УВИДЕЛ СТАЛО НАШЕЛ ИОСИФ ИАКОВ АВРААМ РОДИЛА
 СКОТ ИОСИФА АВРААМА АВРААМУ ИАКОВУ ФАРАОНУ ФАРАОН БРАТЬЯМ
 ДОЧЕРЕЙ БРАТУ БЛАГОСЛОВИЛ БРАТА БРАТ ЖЕНЫ ЖЕНЕ ЖЕНА ГОЛОД ПОЛЕ
 ПОЛЯ ДОЧЕРИ СТАДА ВОДА ДНЕЙ СИМ ДЕТЕЙ МЕСТЕ СДЕЛАТЬ МЕСТА
 ОСТАВИЛ ВЕЛИКИМ},

{ДОМ ДОМА ДОМЕ ПОСТАВИЛ ЛИЦЕМ ДЕСЯТЬ ГОДА ЦАРЯ ОСТАВИЛИ
 ИЕРУСАЛИМА ОТЦАМИ ИЗРАИЛЬСКИЙ ВОЦАРИЛСЯ ДЕЛАЛ ИЗРАИЛЬСКИХ
 ЦАРСТВОВАЛ ЦАРЕМ ЦАРЬ ГОД ГОСПОДНЕМ ИУДЕЙСКИЙ ОЧАХ ИЗРАИЛЬТЯН
 ГОРОДЕ ПРОРОКОВ ИЗРЕК ПРОРОК КОНЕЙ ЦАРЕЙ ЦАРСКОГО ЦАРЮ},

{СДЕЛАЛ ДЕРЕВА ТРИ ДВА СДЕЛАЛИ ЖЕРТВЕННИК РАБОВ ЗОЛОТА
 ЗОЛОТОМ ВСЁ СОВЕРШИЛ СТОРОНЕ ВОЗЬМИ ЕГИПЕТСКОЙ ШЕСТЬ ОДНОЙ
 СДЕЛАЙ ФАРАОНА ЧИСТОГО ЗОЛОТЫХ ПРОСТЕР СКИНИЮ ЗАВТРА ДВУХ
 ОДЕЖДЫ ВОЗЬМИТЕ ВХОДА СЕДЬМОЙ УМЕРЕТЬ ВЕЩИ ГОВОРЯ ПОВЕЛЕЛ
 ВСЕСОЖЖЕНИЯ ГОЛОВУ СДЕЛАЕТ ГОСПОДУ ГОСПОДОМ ТРИДЦАТЬ
 ШЕСТЬДЕСЯТ ИМЕНА ЧЕТЫРЕСТА СОРОК ПЯТЬДЕСЯТ СЫНОВ ПРИНЕСЛИ},

{ГОРОДА ГОРОД ВЗЯЛИ ИДЕТ ЖИТЕЛИ МЕЧОМ ГОРОДОВ ПОСТУПИЛ
 ДОЛИНЕ ИЗРАИЛЮ ИЗРАИЛЯ ИЗРАИЛЬ},

{ПОСЛАЛ ИДТИ ВОЗВРАТИЛСЯ НАВСТРЕЧУ СДЕЛАЛСЯ УВИДЕЛИ
 УСЛЫШАЛИ ПОСЛАЛИ УЗНАЛ},

{ДНЯ ГОСПОДЕНЬ ПОСТАВИЛИ СОШЕЛ},

{ИЕРУСАЛИМ ИЕРУСАЛИМЕ СВЯЩЕННИКИ СВЯЩЕННИКОВ СОСУДЫ
 ИУДЕЕ ОТЦОВ СОБРАНИЕ ИУДЕЮ},

ГОСПОДНЯ, {ЛЮДИ УМЕР УБИЛ ПАЛ ШЕЛ СЛУГИ СЛУГ УБЕЖАЛ СЫНОМ
 САУЛА ДАВИДУ ФИЛИСТИМЛЯН ДАВИД ПОРАЗИЛ РАБ СЛОВУ РАБА ХОДИЛ
 НАЗАД РАБУ ИЗРАИЛЕМ},

{НАРОДОВ НАРОДАМ ИЗРАИЛЕВА ДОМУ ВОРОТА КРАЯ ИЗРАИЛЕВУ
 МЕРЗОСТИ УЗНАЮТ ГОРАХ МЕЧ СДЕЛАЮ ЯРОСТИ ПРИВЕЛ СКАЖИ ДЕЛАЛИ
 ИЗРАИЛЕВ МЕЧА ОТДАМ СЛОВО ГОСПОДНЕ ИСТРЕБЛЮ ПОШЛЮ},

{ГОЛОС ЖИВУЩИХ},

БОЖИЙ, СЫНУ, МНОЖЕСТВО, СКАЗАТЬ, {ХРАМ ХРАМА},

КОВЧЕГ, ОВЕЦ, РУКА, СЕРЕБРА, ЧАСТЬ, {ЖИТЕЛЕЙ ХОДИЛИ СЛУЖИЛИ},

{ИМЕНЕМ СЕМ},

РАЗ, {ОТДАЛ ДАЙ},

РУКОЮ, ХОРОШО, ПРЕДАЛ, {ИЗБРАЛ УСТРОИЛ},

ОГНЕМ, ПРОРОКА, ЗОЛОТО, ЗАВЕТА, {БОЖИЕ ВЗЯВ ПРИДЯ НАШЛИ ВИДЯ
 НАДЛЕЖАЛО},

ХОДИТЬ, ЖИВ, ВИНА, СЛЫШАЛ, ГОЛОВЫ, ИЗРАИЛЬТЯНЕ, СТЕНЫ,
 САМОГО, БЛАГОСЛОВЕН, ДОБРОЕ, СТАРЕЙШИНЫ, ВСТАНЬ, ЖЕНЩИНЫ,

БЛИЖНЕГО, ХОТЕЛ, ХРАМЕ, БОЙСЯ, ГЛАЗ, ОТКРЫЛ, ТРЕХ, ИЗРАИЛЕ, КАМНИ, ПУТЕМ, ПОЙДЕМ, УГОДНО, СЛЫШАЛИ, ДОМОМ, ДАР, КОЛЕН, БЕДСТВИЯ, КНИГЕ, ОСТАЛОСЬ, ИДИТЕ, МОГУ, ПРИЧИНЕ, СТАРЕЙШИН, ЯВИЛСЯ, ВСЛЕД, СЕМИ, ЯЗЫЧНИКОВ, ЗАКОНЕ, ЖЕНОЮ, БЛАГОВОЛЕНИЕ, ЖИВОТНЫХ, ВОЗВРАТИЛИСЬ, ДОЖДЬ, КАМНЕЙ, КОЛЕСНИЦЫ, ПРИВЕЛИ, ПОСТУПАТЬ, ПРАВУЮ, БОЙТЕСЬ, ГОЛОДА, ЗЛОЕ, СОВЕТ, УТРОМ, СОГРЕШИЛИ, ВОЗВРАТИТСЯ, ПОЛОЖИЛИ, {ПРИНОШЕНИЯ ОВНА ВСЕСОЖЖЕНИЕ ЖЕРТВУ ПОРОКА ГРЕХ ХЛЕБНОЕ ЕЛЕЕМ МУКИ ПРИНОСИТЬ ЖЕРТВЕННИКА СТАНА ПРИНОШЕНИЕ ААРОНУ СОБРАНИЯ ИЗРАИЛЕВЫМ РОДЫ СЫНАМ ПОЛА ААРОНА ИЗРАИЛЕВЫХ СКИНИИ ААРОН МОИСЕЙ МОИСЕЯ МОИСЕЮ РАБОТЫ ПУСТЫНЮ ОБЛАКО МУЖЕСКОГО СЕМЬДЕСЯТ ЧАСТИ СКОТА},

СДЕЛАЛИСЬ, {ИЕРЕМИИ ИЕРЕМИЯ БЕДСТВИЕ ВАВИЛОНСКОГО ВАВИЛОН ВАВИЛОНСКИЙ ИУДЕЙСКОГО ПОСЫЛАЛ ИУДЕЙСКИХ ПЛЕН ПРОРОКИ ИУДЫ ГОРОДАХ СТРАНЫ КНЯЗЕЙ КНЯЗЬЯ},

{ОБО САМАРИИ ВВЕЛ},

ГРЕХОВ, ПОСТУПАЛИ, ЦАРИ, ИСТРЕБИЛ, САВАОФ, ДЕЛО, МОЛИТВУ, ОЧАМИ, {СЫНЫ ИЗРАИЛЕВЫ СЫНАМИ СТАН},

{ТЫСЯЧ ТЫСЯЧИ ВОЙНЫ ОВНОВ АГНЦЕВ ПОКОЛЕНИЕ ОТПРАВИЛИСЬ РАСПОЛОЖИЛИСЬ ОБЩЕСТВО ВЫШЕ СВЕРХ ТЕЛЦОВ СТАНОМ ВОЙНУ ТРИСТА ШЕСТЬСОТ ДВЕСТИ СТО КАЖДОГО МЕСЯЦА ДВАДЦАТИ ГОСПОДНЮ ДВАДЦАТЬ ПЕРВЫЙ ВТОРОЙ СЫН ТРЕТИЙ},

СЕМЬ, ПЛЕМЕНИ, ТАЛАНТОВ, НАЧАЛЬНИК, ЖИВЕТ, ЖЕН, {МЕСЯЦ СВЯЩЕННИКА},

НАЧАЛЬНИКИ, ВОЛОВ, ДВЕНАДЦАТЬ, ТЫСЯЧУ, ПРИНЕС, {БЫВШИХ БЫВШИЕ ПРИКАЗАЛ},

{ИУДЕЕВ ИУДЕИ},

Кластеры 7, 8 и 9 текстов по -существованию образуют один кластер по словарному составу. Они соответствуют кластеру 9 слов. В кластеры попали относительно короткие тексты Посланий Апостолов, включающие не более 6300 слов.

Кластер 9 слов

ХРИСТОС, СВЯТЫМ, {МИР МИРА МИРЕ ЗНАЕМ},

СВИДЕТЕЛЬСТВО, ЗНАЯ, ЗНАЕТЕ, ЗАПОВЕДЬ, ДУХОМ, {ИИСУСА ДУХА МЕРТВЫХ ПРИНЯЛИ},

БРАТИЯ, {ХРИСТЕ ИИСУСЕ ХРИСТА},

ЛЮБВИ, БЛАГОДАТЬ, ПЛОТИ, ВЕРЕ, БОЖИЕЙ, ГОСПОДЕ, {ПАВЕЛ МУЖИ},

Кластер 7 текстов

- Первое послание к Фессалоникийцам-1400,
- Первое соборное послание святого апостола Петра-1651,
- Второе послание к Фессалоникийцам-764,
- Послание к Галатам святого апостола Павла-2114,
- Послание к Евреям святого апостола Павла-4626,
- Послание к Ефессянам святого апостола Павла-2090,
- Послание к Колоссянам святого апостола Павла-1410,
- Послание к Римлянам святого апостола Павла-6295,
- Соборное послание святого апостола Иакова-1609,

- Первое послание к Коринфянам святого апостола Павла-6337.

Кластер 8 текстов

- послание к Титу святого апостола Павла-997,
- Второе послание к Тимофею святого апостола Павла-1155,
- Первое послание к Тимофею святого апостола Павла-1599.

Кластер 9 текстов

- Послание к Филиппийцам святого апостола Павла-1574,
- Второе послание к Коринфянам святого апостола Павла-4312

Кластер 10 текстов обогащен словами из кластеров слов 3 и 4 (приведены ниже)

- Книга Премудрости Иисуса-18960,
- Книга Екклесиаста-3943,
- Книга Премудрости Соломона-6943,
- Книга Иова-12991,
- Притчи Соломона-10414

Кластеры 3 и 4 слов

Кластер 3 слов

ИИСУС ГОВОРЮ ОТВЕТ УЧЕНИКИ ИСТИННО ИОАНН ЧАС ГОВОРИЛИ
УВИДЕВ УСЛЫШАВ ПЕТР УЧЕНИКОВ ОТКУДА СУББОТУ ВОН ПРИХОДИТ,
СМЕРТЬ, СТОЯЛ, ЗНАЕШЬ, СЕМЕНИ, ТАКИЕ, ОТВЕЧАЛ, ЖИЗНЬ ВИДИТ,
ОТВЕЧАЛИ, ЗНАЛИ, ПОЧЕМУ, КАМЕНЬ, ЗНАЮ ЗНАЛ, ИМЕТЬ, НАПИСАЛ, ВИНО,
ИМЕЛИ, ГРЕХА, РАДОСТЬ, ДВЕРИ, СВЕТ, СКАЖУ, ДРУГ ДРУГА ДРУГУ, ХОЧЕТ,
ИДИ,

Кластер 4

ГРЕХИ, ЖИЗНИ, ЧЕЛОВЕК ЧЕЛОВЕКА ЧЕЛОВЕКУ, ДЕЛАЕТ, БОГАТСТВО,
ГОВОРИ БЫВАЕТ СТРАХ МУДРОСТИ МУДРОСТЬ, ОГНЯ, ЛЮБИТ, НАЧАЛА,
СЛОВОМ, НАДЕЖДА, ЗНАЕТ

Кластеры слов 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 ничем особенно не примечательны.

Это не должно нас разочаровывать, так как Библия не должна иметь много разных тем.

Кластер 5 слов

ДОМЫ, УТРА, ГЛАЗАХ, ЛЕВ, ОСТАНЕТСЯ, ПОЭТОМУ, ВОПЛЬ,
ОТКРОВЕНИЯ.

Кластер 6 слов

АНГЕЛ, {НЕБА ВЕЛИКИЙ ГОЛОСОМ ГРОМКИМ НЕБО ВЛАСТЬ ДАНО},
ВОСТОКУ, НОГ, ДУХ, СЕЛ, СОБРАЛИСЬ, БУДЬТЕ, ДАЛИ, ТЕЛО ТЕЛА,
НАРОДЕ, ПРАЗДНИК, ЧИСЛО, {МУЖ МУЖА},
{КОЛЕНА ИОРДАНА УДЕЛ ИИСУСУ ПРЕДЕЛ ПЛЕМЕНАМ ПРЕДМЕСТЬЯ
ЖРЕБИЙ},
АНГЕЛА, {КРЕПОСТЬ РАДОВАТЬСЯ},
ЛЮБОВЬ, {ИСТИНЫ ИСТИНЕ},
ВЕРОЮ, ПРЕСТОЛЕ, {СВЯТАГО СВЯТЫЙ},
ПРЕБЫВАЕТ, ВОЙДЕТ, ЧРЕВА, ВОСЕМЬ, {БРАТЯ БРАТЬЕВ},
{БРАТЬЯМИ РОДИЛ СЫНОВЬЯ},

ИМЕЛ, КАМНЯМИ, СЫНОВЕЙ, ТЫСЯЧА, СКАЗАНО, ЖИЛИ, ЛЕВИТОВ
ЛЕВИТЫ, {СОЛОМОНА ПОСТРОИЛ СОЛОМОН},

ИЛИЯ, {ВОЙСКО МУЖЕЙ ВОЙСКА КРЕПОСТИ ОТПРАВИЛСЯ СТРАНУ
СТРАНЕ СРАЖЕНИЕ СОБРАЛ ИУДА СИМОН},

ИОНАФАН, БОЖИЯ, {НАПИСАНО БОЖИЮ},

ВЕЛИКОГО

Кластер 7 слов, ЛЮДЬМИ, ТАКОЙ, ПАЛИ, ВЕЛИКОЕ, ЛЮДЯМ, МЕСТАХ,
СВЯТИЛИЩЕ СВЯТИЛИЩА СОЮЗ ГОДУ ОБРАТИЛСЯ, ИСКАТЬ, ЭТИХ, ИМИ.

Кластер 8 слов, ВОРОТ ЛОКТЕЙ КРУГОМ, ОДНИ, ГОРОДУ, ЧЕЛОВЕЧЕСКИЙ,
ЧЕТЫРЕХ, СУДИТЬ, ЧЕТЫРЕ, ПОДНЯЛ.

Кластер 9 слов, ХРИСТОС, СВЯТЫМ, МИР МИРА МИРЕ ЗНАЕМ,
СВИДЕТЕЛЬСТВО, ЗНАЯ, ЗНАЕТЕ, ЗАПОВЕДЬ, ДУХОМ, ИИСУСА ДУХА МЕРТВЫХ
ПРИНЯЛИ, БРАТИЯ, ХРИСТЕ ИИСУСЕ ХРИСТА, ЛЮБВИ, БЛАГОДАТЬ, ПЛОТИ,
ВЕРЕ, БОЖИЕЙ, ГОСПОДЕ, ПАВЕЛ МУЖИ.

Кластер 10 слов, СОН, НОГИ, СВЯТЫХ, ПРИЗВАЛ ОТПУСТИЛ, ГОЛОСА,
ЗАКОН, ДАНИИЛ, ЦАРСТВО, ДАРЫ, ВЕЛИК, ВЕНЕЦ, НОЧЬ, ПОВЕЛЕНИЕ, ЗАКОНА,
НАСЛЕДИЕ, ВОЗЗВАЛ, СИЛЬНО, ГНЕВА, ГОСПОДЬ ЗЕМЛЮ БОГ НАРОД НАРОДА
НАРОДУ ДЕНЬ ЗЕМЛЕ ЗЕМЛИ РУКАХ РУКИ БУДЕШЬ ЗЕМЛЯ РУКУ ГОРЕ ЛИЦЕ
ЖИТЬ ГОРЫ НАРОДОМ ЗАВЕТ ВЫВЕЛ СЛУЖИТЬ ДАМ ЕГИПТА ГЛАЗАМИ ВОДУ
ЕГИПЕТ ПОЙДЕТ СИЮ СКАЖУТ РУК ЛИЦА НАРОДАМИ БОГА БОГУ БОГОМ
ГОСПОДА ВРЕМЯ ДНИ ЛЮДЕЙ СМЕРТИ СЛОВ ЮНОСТИ СЛОВА ГЛАЗА ЗЛО ОТЦЫ
ЗНАТЬ СЕРДЦА ГНЕВ РУКЕ ОГОНЬ ВОД НАРОДЫ СЛАВА МОИМИ ПРЕДО СУД
НЕБЕСА ПРАВДЫ ПРАВДА РАДИ СИОНА ГНЕВЕ ЯРОСТЬ БЕЗЗАКОНИЕ
БЕЗЗАКОНИЯ ИМЯ ДАЛ ВОДЫ УСЛЫШАЛ ИАКОВА ПОЛОЖИЛ НАШЕМУ ОЧИ
ХЛЕБ ПЛОТЬ НОЧЬЮ ДНЕМ СОТВОРИЛ ВИДЕТЬ СКОРБИ ПОСРЕДИ ВРАГИ ВОВЕК
ПЕСНЬ БОЖЕ НАЧАЛЬНИКУ РОД МИЛОСТЬ ГОСПОДИ ВРАГОВ ХОРА ПСАЛОМ
СПАСЕНИЕ ПРАВДЕ ПОМОЩЬ УСЛЫШЬ СПАСЕНИЯ ДУША ДУШУ МИЛОСТИ
ИЗБАВИЛ ЧУДЕСА ПРАВДУ ВЕЧНО ТВОЕЮ СЕРДЦЕМ НЕБЕС ЖИЛИЩЕ БЛАЖЕН
СИЛОЮ СИЛА СИЛУ ИМЕНИ ПРЕСТОЛ ДАВИДА КОСТИ, РАНО, УСТ УСТА ПУТИ
УХО УСТАМИ ПУТЬ ПИЩУ СЕРДЦЕ НЕЧЕСТИВЫХ ДУШЕ ЯЗЫК ЗЛА ДУШИ
ДОБРО УСТАХ СЛОВАМ СЛАВУ ДЕЛА ВСЕГДА ДЕЛАХ ВЕКА ВСЕВЫШНЕГО
СЛАВЕ, СИЛЫ, ЦАРСТВА, РАБАМИ, ПОЙДИТЕ, СЕСТРА, ЗАКОНУ,
НАВУХОДНОСОР, ДЕЛАМ, ВЕКИ ПОЙТЕ, ПАМЯТЬ, ФИЛИСТИМЛЯНЕ САМУИЛ
САУЛ САУЛУ, РАБЫ, ЖИВЫХ.

Кластер слов 11

РУКАМИ, ДОБЫЧУ, МОЛИЛСЯ, МАТЬ, ПОГУБИТЬ, ПРЕДЕЛЫ, ВОЙТИ,
ВЗЯТЬ, ДАВАТЬ, НАЧАЛ СПРОСИЛ НАЧАЛИ, РОДА.

Кластер слов 12

СМОТРИ, БЛАГОСЛОВЕНИЕ, ЗАКОНЫ, ПЛОД, БОГОВ ОТЦАМ БОГАМ
ГЛАСА, ВСЯКОМ, ЕЮ, АМИНЬ, ВСЯКУЮ, ДЕЛЕ, ДЕЛАЙ, ГЛАС, {СКАЖЕТ ДАСТ
ВИДЕЛИ ДАТЬ ЗАПОВЕДИ ДАЕТ ЗАПОВЕДАЛ КЛЯЛСЯ ИСПОЛНЯТЬ СЕГОДНЯ
СРЕДЫ ЗЕМЛЕЮ ВЛАДЕНИЕ СЛУШАТЬ}

Кластер слов 13

ЖЕНЩИНА, МНОГИЕ МНОГИХ, ДРУГИХ, ОБА, МОГ, ВРЕМЕНИ, ЗНАМЕНЕНИЯ, ОСТАЛСЯ, ОДНОМ, ИМЕЯ, НИКОГО, МИРОМ, ЕЛ, ЗАЧЕМ, ИОАВ, НИКОГДА, ПОКАЗАЛ, ВИДЕЛ КОНЕЦ ВРЕМЕНА

Кластер слов 14, ДЕЛ, СЛАВЫ, ЖИВУЩИЕ, ГОР, РАДОСТИ, СУДА.

Кластеры текстов 1, 2, 3, 5 и 6 богаты словами но практически во всех кластерах слов. Кластер 6 содержит тексты Евангелий.

Кластер 6 текстов

- От Матфея святое благовествование-16212,
- Книга пророка Аггея-763,
- От Иоанна святое благовествование-13840,
- Первое соборное послание святого апостола Иоанна-2770,
- От Луки святое благовествование-17526,
- Деяния святых апостолов-16896,
- Второе соборное послание святого апостола Петра-1063,
- От Марка святое благовествование-10221.

Более детальная интерпретация результатов кластеризации предполагает обращение к специалистам по Библии.

Интересно, что для Библии не соблюдается правило Золотой пропорции.

В таблице 9 показано, что $\phi^2 = 0,507$ вместо 0,62

Таблица 9.

Служебные слова	316850
Общее число словоупотреблений	642634
Знаменательные слова	325784
Доля служебных слов	0,493048921
Доля знаменательных слов	0,506951079

6. Лингвистические характеристики переводов на русский язык смыслов Сур Корана

Критерии совершенства текстов позволяют выявить наиболее совершенный (целостный) с лингвистической точки зрения перевод одного и того же документа. Так было проведено сравнение переводов смыслов Корана на русский язык, имеющих разное авторское толкование, и сравнить качество переводов с точки зрения методики, основанной на использовании предлагаемого сочетания критериев. Хотя Коран состоит из 114 сур (глав) разной длины (от трех до 286 стихов (аятов)) и организованы (за небольшим исключением) по убыванию длины сур, восстановить хронологию их написания затруднительно. Однако, рассмотрение каждого перевода в отдельности позволяет судить о степени совершенства каждого перевода на русский язык.

Из 15 наиболее известных переводов на русский язык в качестве примеров текстов для анализа выбраны переводы Кулиева Э. Р., Османова М.-Н. О., Пороховой В. М. и Крачковского И. Ю.

Рассмотрим сначала перевод Э.Р.Кулиева. На последующих трех рисунках (с 8 по 10) показано ранговое распределение слов всех 114 Сур Корана (перевод смыслов Э.Р. Кулиева) в билогарифмическом масштабе (на билогарифмической шкале).

На рисунке 8 дано распределение всех слов. Коэффициент детерминации в этом случае равен 0,9755. Исключив из рассмотрения первые по частоте 31 слово, мы не улучшили качество линейной аппроксимации распределения, так как на интервале $\log(N)$ от 3,5 до 4,1 расположены слова, встречающиеся 3, 2 и 1 раз (на одной и той же из трех ординат) с общим числом разных слов 9221 из 12377 слов (см. рисунок 9). Такое количество слов, расположенных трех горизонтальных ступенях, дает ощутимую невязку. Исключив эти слова из аппроксимируемой области, получаем распределение, представленное на рисунке 10, аппроксимируемое прямой с коэффициентом детерминации 0,9975.

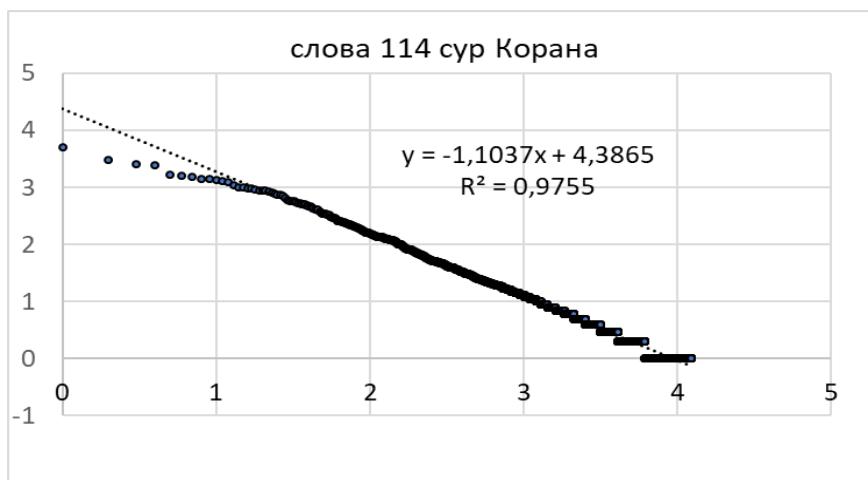


Рисунок 8



Рисунок 9.



Рисунок 10.

Распределение служебных слов в данном переводе представлено при аппроксимации логарифмов частот и рангов на рисунке 11. Представление в виде РРС по формулам (3-5) дано на рисунке 12.

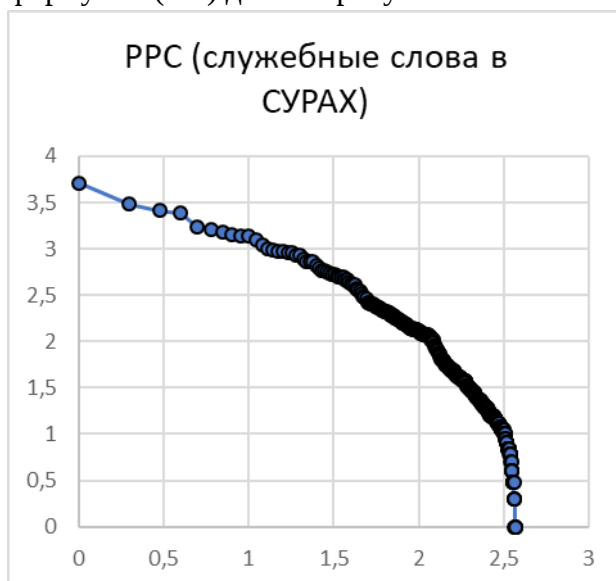


Рисунок 11.

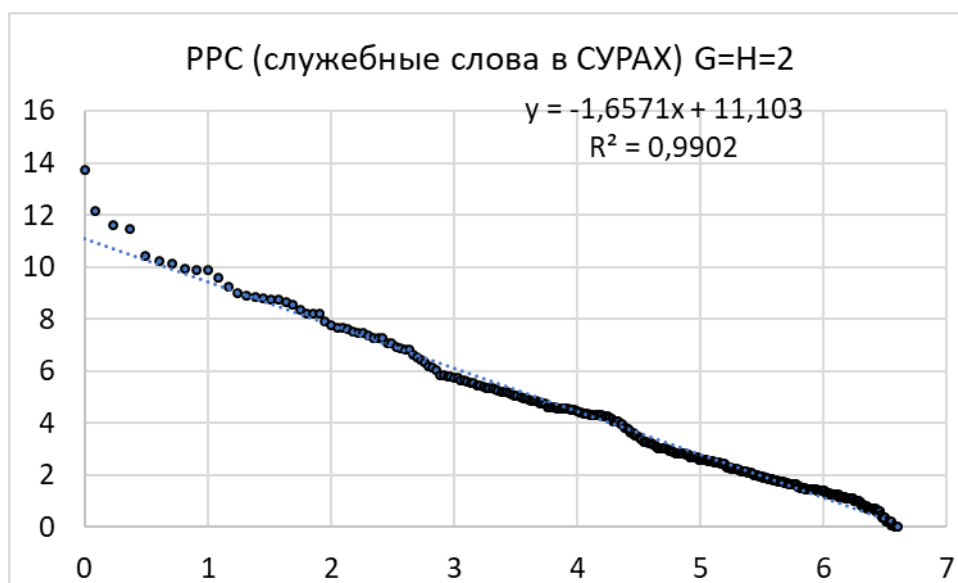


Рисунок 12

Критерий золотого сечения резко не соблюдается. Знаменательных слов только 46 процентов, а служебных 54%, и это при нормальном соотношении 62% к 38%.

Проведя аналогичные расчеты для частотных списков слов и служебных слов в переводах [Османова М.-Н. О.](#), Пороховой В. М. и Крачковского И. Ю. получаем характеристики аналогичных распределений, представленные в таблице 10.

Таблица 10.

Авторы переводов -толкований	Аппроксимация полного распределения слов текста и R^2	Аппроксимация без первых 31 слова и последних, встречающихся 3, 2 или 1 раз, и R^2	Аппроксимация распределения служебных слов для $G=H=2$, и R^2
Кулиев Э. Р.	$y = -1,1037x + 4,3865$ $R^2 = 0,9755$	$y = -1,0904x + 4,3698$ $R^2 = 0,9975$	$y = -1,6571x + 11,103$ $R^2 = 0,9902$
Османов М.-Н. О.	$y = -1,0483x + 4,1971$ $R^2 = 0,9691$	$y = -1,0798x + 4,3165$ $R^2 = 0,9967$	$y = -1,6803x + 11,171$ $R^2 = 0,9905$
Порохова В. М.	$y = -1,0511x + 4,2453$ $R^2 = 0,9709$	$y = -1,0546x + 4,2807$ $R^2 = 0,9965$	$y = -1,67x + 11,327$ $R^2 = 0,9888$
Крачковский И. Ю.	$y = -1,0845x + 4,2488$ $R^2 = 0,9709$	$y = -1,1125x + 4,3562$ $R^2 = 0,996$	$y = -1,7604x + 11,358$ $R^2 = 0,9895$

Оценки золотой пропорции по ϕ^2 сведены в таблицу 11.

Таблица 11. Соотношение служебных и знаменательных слов.

	Кулиев Э. Р.	Османов М.-Н. О.	Порохова В. М.	Крачковский И. Ю.
Всего словоупотребле ний	119358	118697	127844	111620
служебных	64724	65186	71834	65858
знаменательных	54634	53511	56010	45762
%служебных	54,28	54,9	56,2	59,0
% знаменательных (ϕ_2)	45,8	45,1	43,8	41,0

Полученные результаты, выписанные в порядке убывания качества по критерию ϕ , при нейтральном поведении параметра G , соответствуют ожидаемым с позиции здравого смысла. Рассмотренные критерии и методика их использования могут быть полезны для экспресс – анализа качества научных текстов.

Отметим, что проверка изложенной методики проводилась на текстах, высокое качество которых является общепризнанным на национальном и даже на мировом уровне.
 $\phi_2 =$

7. Проблема составления списка служебных слов в моноязычных текстах.

Служебные слова в текста Библии и Корана выделялись вручную. Общей методики их выделения пока нет. Их количество сопоставимо. Число служебных слов, выделенных из первой тысячи ранжированного списка слов Библии равно 309. Общее число служебных слов, выделенных из первых 2500 слов ранжированного списка четырёх текстов Корана равно 561.

Таблица 12. Совпавшие служебные слова в Коране и в Библии

№№	В Коране	В Библии	Служебное слово	№№	В Коране	В Библии	Служебное слово
1	1	1	И	122	0,002275	0,006569	МОЯ
2	0,392033	0,283137	В	123	0,007057	0,006409	БЫЛА
3	0,459168	0,20724	НЕ	124	0,004084	0,006409	КАЖДЫЙ
4	0,18755	0,156322	НА	125	0,003671	0,006391	МНОГО
5	0,1458	0,148688	ЕГО	126	0,003516	0,006356	МОИХ
6	0,324018	0,1086	ЧТО	127	0,023525	0,00632	СРЕДИ
7	0,239407	0,105529	ОН	128	0,00274	0,006019	ТВОИМ
8	0,081612	0,101303	Я	129	0,014425	0,005859	ТАКЖЕ
9	0,117132	0,09493	С	130	0,024714	0,005841	МОЖЕТ

10	0,11571	0,08829	ОТ	131	0,002223	0,005841	ТЕПЕРЬ
11	0,210299	0,087508	ИХ	132	0,008117	0,005663	НИКТО
12	0,17297	0,078916	ИЗ	133	0,014735	0,005646	СВОЕМУ
13	0,170695	0,075489	К	134	0,001577	0,005575	ТВОЮ
14	0,0805	0,074868	КАК	135	0,039009	0,005504	ТЕМ
15	0,127498	0,063399	А	136	0,002973	0,005468	ПОСЕМУ
16	0,358349	0,062405	ОНИ	137	0,009332	0,005362	БЕЗ
17	0,10449	0,061677	ТЫ	138	0,011892	0,005308	ВСЕГО
18	0,014839	0,060043	ИБО	139	0,005429	0,005291	ДАБЫ
19	0,043378	0,055019	ВСЕ	140	0,002895	0,00522	СВОЕМ
20	0,060052	0,053883	ПО	141	0,003387	0,005202	САМ
21	0,078949	0,050971	НО	142	0,014166	0,005184	ЧЕГО
22	0,061138	0,050066	ЧТОБЫ	143	0,023369	0,005166	ПЕРЕД
23	0,150738	0,049196	ЖЕ	144	0,0129	0,005131	ВАШИХ
24	0,050668	0,046959	У	145	0,001215	0,005131	ТВОЯ
25	0,195331	0,046249	ТО	146	0,004084	0,005113	НАШЕГО
26	0,108187	0,043728	КОГДА	147	0,002327	0,00506	МОЕЙ
27	0,055089	0,043532	БУДЕТ	148	0,067989	0,00506	ТЕ
28	0,094383	0,042609	ЗА	149	0,002689	0,005024	ВСЯКИЙ
29	0,137528	0,041029	ДЛЯ	150	0,00486	0,004918	ЕЙ
30	0,044722	0,038029	ЕМУ	151	0,00181	0,004882	МОЮ
31	0,030298	0,037301	МЕНЯ	152	0,003206	0,004882	СИЕ
32	0,045084	0,036662	ТАК	153	0,023654	0,004865	ТОМ
33	0,090634	0,036395	О	154	0,010521	0,004829	НЕЙ
34	0,027299	0,035987	ВО	155	0,018561	0,00474	ВАМИ
35	0,044283	0,033768	ТЕБЯ	156	0,002714	0,004723	ТВОЕМУ
36	0,030944	0,033359	МНЕ	157	0,002637	0,004687	МОИМ
37	0,130393	0,032188	ЕСЛИ	158	0,017088	0,004634	НАШИ
38	0,256237	0,031353	ВЫ	159	0,00318	0,004598	ДРУГОЙ
39	0,116356	0,029951	ИМ	160	0,018897	0,004438	БУДЕТЕ
40	0,024921	0,02988	ВОТ	161	0,079958	0,00435	ТЕХ
41	0,082413	0,029613	ЭТО	162	0,016054	0,004314	ВМЕСТЕ
42	0,00897	0,028371	ПРЕД	163	0,013417	0,004261	ОБ
43	0,039966	0,02679	ТЕБЕ	164	0,009642	0,004243	НИЧЕГО
44	0,163612	0,026524	КТО	165	0,029755	0,004243	ЧЕМ
45	0,022904	0,0264	ЕЕ	166	0,010315	0,004243	ЧТОБ
46	0,127679	0,025654	КОТОРЫЕ	167	0,000931	0,004119	ВСЕЙ
47	0,039811	0,023595	ДО	168	0,007057	0,004101	ЭТОТ
48	0,148773	0,023506	ВАС	169	0,002611	0,004083	ЧЕРЕЗ
49	0,045834	0,023328	НЕГО	170	0,005222	0,003995	СТАЛ
50	0,015123	0,023204	ВСЕХ	171	0,00698	0,003941	ОНО
51	0,090324	0,022973	НИХ	172	0,001008	0,003924	ОПЯТЬ
52	0,037769	0,022902	БЫЛО	173	0,026161	0,003924	ПОТОМ
53	0,035106	0,022618	БЫЛ	174	0,008686	0,003853	НЕЕ
54	0,240648	0,022618	МЫ	175	0,003412	0,003817	МНОГИЕ
55	0,025515	0,02253	КОТОРЫЙ	176	0,004576	0,003764	РАДИ
56	0,024843	0,022245	ЛИ	177	0,026006	0,003728	КОГО

57	0,009177	0,02221	ПОТОМУ	178	0,005972	0,003693	КОТОРАЯ
58	0,043327	0,020825	БУДУТ	179	0,029496	0,003533	ТОМУ
59	0,010004	0,020293	ДА	180	0,007781	0,003515	ХОТЯ
60	0,059354	0,020204	НИ	181	0,006437	0,00348	ВМЕСТО
61	0,106377	0,018997	ВАМ	182	0,001267	0,003444	ВСЕМИ
62	0,020086	0,018908	ТОГДА	183	0,023137	0,003444	ЭТОГО
63	0,042318	0,018695	НИМ	184	0,0076	0,003409	НАШИХ
64	0,097924	0,018304	БЫ	185	0,006178	0,003373	НАМИ
65	0,033477	0,017541	НАД	186	0,00243	0,003284	ДОЛЖНЫ
66	0,033762	0,017399	ЕСТЬ	187	0,008272	0,003249	КОТОРЫМ
67	0,020655	0,016209	СВОЕГО	188	0,002197	0,003249	МОЕМУ
68	0,072745	0,016085	ИЛИ	189	0,01215	0,003213	ДАЖЕ
69	0,062999	0,015819	ТОТ	190	0,006101	0,003196	НАШ
70	0,044464	0,015393	БЫЛИ	191	0,032521	0,00316	КРОМЕ
71	0,021818	0,015339	ОНА	192	0,010987	0,003	ВАШ
72	0,032288	0,014682	НАС	193	0,005377	0,002929	БОЛЬШЕ
73	0,018768	0,014523	МОЙ	194	0,000853	0,002894	ВСЮ
74	0,008402	0,014292	ПРИ	195	0,005739	0,002894	ЭТИ
75	0,01458	0,014185	ТВОЕГО	196	0,005455	0,002841	КОТОРОЙ
76	0,001086	0,013564	ТВОИ	197	0,022232	0,00277	ЭТОМ
77	0,073495	0,013546	ТОГО	198	0,003567	0,002752	ВСЯКОГО
78	0,018044	0,013511	ТВОЙ	199	0,023783	0,002734	РАЗВЕ
79	0,013572	0,013085	СО	200	0,014218	0,002681	ВАШЕГО
80	0,019052	0,012889	СВОИХ	201	0,003826	0,002379	КАЖДОГО
81	0,01882	0,012729	МЕЖДУ	202	0,001241	0,002361	ВСЕГДА
82	0,008479	0,012623	ВСЕМ	203	0,001603	0,002361	КУДА
83	0,019259	0,012161	ТАМ	204	0,00349	0,002308	КАКОЕ
84	0,018277	0,01209	СЕБЯ	205	0,000905	0,002219	МОЕМ
85	0,019957	0,011948	СЕБЕ	206	0,008841	0,002166	САМИ
86	0,019673	0,011913	ПУСТЬ	207	0,006876	0,002113	БУДЬ
87	0,004912	0,011575	МОИ	208	0,024145	0,002077	НЕУЖЕЛИ
88	0,032159	0,011575	НЕТ	209	0,002275	0,002059	ТОЙ
89	0,010935	0,011025	ЕЩЕ	210	0,002714	0,001988	СКОРО
90	0,002171	0,010954	ТВОИХ	211	0,004472	0,001971	ВАШИМ
91	0,029005	0,010936	НАМ	212	0,004059	0,001917	ДРУГИЕ
92	0,006178	0,010635	МОЕГО	213	0,007549	0,0019	МОГЛИ
93	0,007264	0,010084	НЕМ	214	0,002792	0,001829	КАКОЙ
94	0,013779	0,009747	СВОИ	215	0,001008	0,001811	ВСЯ
95	0,008557	0,009498	СВОЮ	216	0,001577	0,001811	ТУ
96	0,005791	0,009392	КО	217	0,001241	0,001811	ЭТУ
97	0,001965	0,009321	ТВОЕ	218	0,010651	0,00174	ДРУГИХ
98	0,033994	0,00909	НИМИ	219	0,003981	0,00174	МОГ
99	0,050849	0,008753	ТОЛЬКО	220	0,004033	0,001722	ВАШЕ
100	0,00305	0,008628	БУДУ	221	0,002275	0,001722	САМОГО
101	0,003309	0,008504	МОЕ	222	0,005299	0,001669	ИМИ
102	0,013287	0,008486	СВОЕЙ	223	0,003335	0,001669	НАШИМ
103	0,018018	0,008433	СВОИМ	224	0,004343	0,001598	ТАКОЕ

104	0,031228	0,008309	БЫТЬ	225	0,005144	0,00158	БУДЬТЕ
105	0,009539	0,008184	КОТОРОЕ	226	0,000724	0,001562	БЫВАЕТ
106	0,032004	0,008184	ПОСЛЕ	227	0,006334	0,001562	ВСЯКОЙ
107	0,003697	0,007776	СЕЙ	228	0,002999	0,001491	НИКОГО
108	0,017398	0,007723	КОТОРЫХ	229	0,003231	0,001474	МОГУТ
109	0,003464	0,007563	БУДЕШЬ	230	0,020241	0,001403	КОМУ
110	0,007523	0,007563	КОТОРОГО	231	0,003904	0,001385	ЭТОЙ
111	0,001525	0,007563	МНОЮ	232	0,005041	0,001349	ВСЯКУЮ
112	0,009591	0,007492	ГДЕ	233	0,003981	0,001314	ВАШЕЙ
113	0,002223	0,007368	ТВОЕЙ	234	0,00667	0,001296	ПОЧЕМУ
114	0,009022	0,007297	СВОЙ	235	0,00137	0,001278	МОИМИ
115	0,006928	0,007048	СВОЕ	236	0,002223	0,001225	НАШЕМУ
116	0,017243	0,006924	ВАШИ	237	0,003309	0,001225	ОТКУДА
117	0,00411	0,006835	ПОД	238	0,003231	0,001207	ЕЮ
118	0,005713	0,006764	БОЛЕЕ	239	0,002378	0,001207	ОБО
119	0,010547	0,006693	ПРЕЖДЕ	240	0,005661	0,00119	ТАКИМ
120	0,01502	0,006658	УЖЕ	241	0,001137	0,001172	ВСЯКАЯ
121	0,013494	0,006604	НЕМУ	242	0,001861	0,001172	ЭТА

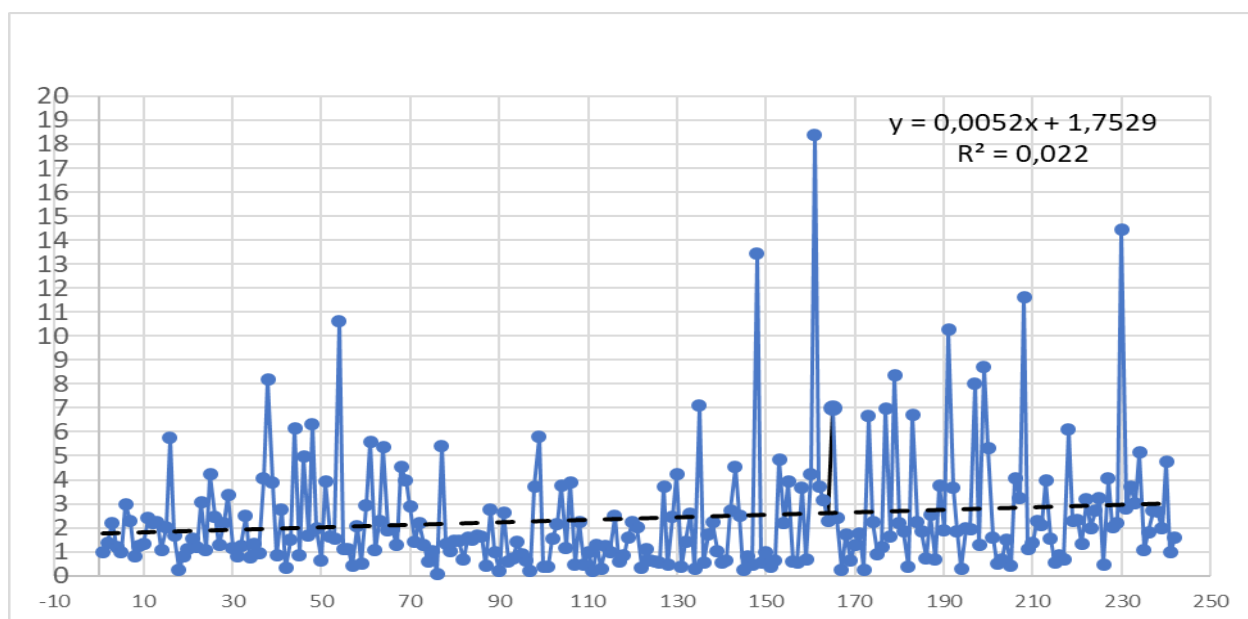


Рисунок 13. Отношение частот совпавших служебных слов в Коране и в Библии.
($F(K)/F(B)$)

В среднем совпадающие служебные слова встречены в Коране в 2,38 раза чаще, чем в Библии.

Сравнение связей в религиозно-духовных и научных текстах также показывает, что разнообразие слов, обозначающих связи в религиозно-духовных текстах выше, чем в научной литературе. В научной литературе больше знаменательных слов, то-есть в научных текстах имеется меньше ссылок на предыдущий контент, по сравнению с религиозными текстами.

8. Применение распределения слов узкой предметной области на форуме в Интернете.

Распределение слов узкой предметной области на форуме исследовано нами не только позиции верификации PPC модели, но и с точки зрения оценки силы взаимодействия между участниками Форума.

В качестве связей между участниками форума, обменивающимися сообщениями в Интернете, могут рассматриваться разные сущности.

Нами были выбраны тексты, которыми обменивались участники. В качестве интенсивности связей рассматривалась:

- Число сообщений участника форума города Уахуапан (Мексика, штат Оахака) через 10 недель со дня открытия (Distribution graphs «Rank vs Messages (posts) number» for the Huajuapán community forum during a period of the first 10 weeks from a launching point of the forum life cycle);
- Число знаменательных слов в сообщении участника форума как функция ранга участника (message on each topic a number of significant words as a function of message author's index (rank));
- Слова-связки всех сообщений форума Уахуапана; (syntactic words in all the posts at the Huajuapán forum).

Высокое качество текстов, следующее из приведенных оценок коэффициента детерминации для всех трёх представленных характеристик степени связей на форуме, можно объяснить тем, что тексты форума являются плодом коллективного творчества многих участников, заинтересованных в успешном результате обсуждения. Подробности описаны в работе [9].

Следует отметить, что подобную модель впервые можно было бы наблюдать на примере коллективного творчества авторов Библии, и получить при этом больше информации в результате анализа такого рода взаимодействия.

ЗАКЛЮЧЕНИЕ

Изложенный материал посвящен в первую очередь формальному определению совершенства или зрелости текста. Под этим мы понимаем его тематическое единство. В основу анализа положено соответствие распределения элементов системы и связей между ними закону Ципфа. Новизна исследования заключается в анализе не только распределения элементов, но и связей между ними. В качестве характеристики связи между знаменательными словами используются распределения служебных слов. Словарь служебных слов как правило беднее словаря знаменательных слов, но частота употребления служебных слов много выше, чем у знаменательных. Поэтому для систем, представленных в виде текстов, в частотном списке которых не делается различий между служебными и знаменательными словами, удобно применять для описания совместного распределения всех слов закона Ципф-Мандельброта.

На примере художественных, научных и юридического текстов показано, что распределение служебных слов аппроксимируется кривой (3) ранговым распределением служебных слов, которая имеет право на самостоятельное название PPC или распределение типа G-H (G-H distribution).

Для текстов заведомо разных авторов распределение служебных слов близко к ранговому распределению связей элементов сложной системы, если тексты посвящены

обсуждению общей темы. Этот результат подтвержден на материалах форума, где обсуждаются проблемы маленького населенного пункта. Сходные результаты получены для корпуса текстов Синодального перевода Библии. Кластеризация, полученная по распределениям знаменательных слов и служебных слов позволила выделить группы текстов по сходными характеристиками. Этот результат нуждается в обсуждении со специалистами в области филологии и истории.

Полученные результаты закладывают основу для формальной автоматизированной оценки текстов различного назначения.

Нелингвистические модели пригодны также и для самостоятельного использования в соответствующих приложениях. Это демонстрируется последующими публикациями, включающих анализ проблем самоорганизации городских агломераций ([19,20]

показано, что проблемы монофункциональных городов в периоды стагнации экономики помимо федеральных и региональных программ решает население путем самоорганизации, в том числе на основе трудовой маятниковой суточной миграции в ближайшие города с более приемлемыми экономическими условиями. Это является начальной причиной агломерационных процессов в областях с большим числом моногородов. Для оценки зрелости городских агломераций рассматриваются экспериментальные модели ранговых распределений городов системы (региона) и критерии зрелости таких систем. На основе выработанных критериев исследуются ранговые распределения моногородов и в целом систем городов Свердловской и Челябинской областей на современном этапе их развития. Оценивается качество параметров, с помощью которых можно измерять интенсивность притяжения городов в миграционных процессах.

В статье Ноздриной Н.Н., Макагонова П.П. и Минченко М.М. [20] анализируется динамика пространственной структуры городской системы расселения Московской области. Показано, что распределение населения по городам Московского региона соответствует закону рангового распределения связей, что является признаком трансформации простой городской системы расселения в агломерационную.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1.. (1913). *Auerbach, F.* Das Gesetz der Bevölkerungskonzentration, Petermanns Geographische Mitteilungen. LIX: 73-76.
- 2.. *Adamic, L.A and Huberman, B.A.* 'Zipf"s law and the Internet', Glottometrics 3, 2002, 143150.
- 3.. (2007) *Danilevskiy I.V.* Zipf-Pareto"s law, new quantum technologies and philosophy of unconsciousness. Quantum Magic, Volume 4, Issue 1, pp. 1201-1209, 2007.
- 4.. (1970) *Hill, B. M.* Zipf's law and prior distributions for the composition of a population. Journal of the American Statistical Association, 65:1220-1232, 1970.
- 5.. (2003) *Kali, R.* "The city as a giant component: a random graph approach to Zipf's law," Applied Economics Letters 10: 717-720 (4), 2003.

- 6.. (1992) *Li Wentian*. "Random Texts Exhibit Zipf's-Law-Like Word Frequency Distribution". *IEEE Transactions on Information Theory* 38 (6): 1842–1845. doi:10.1109/18.165464. 1992.
- 7.. (2001) *Makagonov, P.* Management of the urban territory development. Moscow, RANEPa, 2001, 352 pp. ISBN 5 -8081-0080-1.
- 8.. (2015) *Makagonov, P.* Model of the Empirical Distribution Law for Syntactic and Link Words in "Perfect" Texts. Chapter 2 in «Modern Computational Models of Semantic Discovery in Natural Language», 2015, IGI Global ISBN13: 978-1-46668-6908, ISBN10: 1466686901, edited by Jan Žižka and František Dařena pp. 27-47.
- 9.. (2018) Pavel P. Makagonov, Celia Bertha Reyes Espinoza, Oleg M Dzikun Local Forum as a Knowledge Source. Chapter 5 in *Techno-Social Systems for Modern Economical and Governmental Infrastructures*. IGI Global (pages 100-124) ISBN13: 9781522555865| ISBN10: 1522555862|EISBN13: 9781522555872|DOI: 10.4018/978-1-5225-5586-5
- 10.. (2015-a) Makagonov, P., "Quality evaluation of texts on geosciences based on linguistic criteria", XII International Conference New Ideas In Earth Sciences S-XVI, Numerical Methods and Mathematical Modeling, pp.280-281, 2015 Макагонов П.П. / Оценка качества текстов в науках о земле с помощью лингвистических критериев //«Новые идеи в науках о Земле», Материалы XII Международной научно-практической конференции. – Москва. – 2015. <https://cloud.mail.ru/public/Vh7tiAF22PFz/eng.pdf>
- 11.. (2015-b) Makagonov Pavel P. The evaluation of quality of a complex system based on the distribution of links between its objects. In the *Modeling and data analysis* №1, 2015. ISSN 2219-3758, Moscow, pp.27-40.
- 12.. (2018) Nozdrina, N. N., Makagonov, P. P., Minchenko, M. M. Morphology of urban settlement system of the Moscow region // *System modeling of socio-economic processes: Materials of the 41st International Scientific School-Seminar n.a. Academician S. S. Shatalin*, September 30 – October 4, 2018, Nizhniy Novgorod – Voronezh: Department of Social Studies of the Russian Academy of Science, Central Economic Mathematical Institute of the Russian Academy of Science, Voronezh State University, 2018.
- 13.. (2011) *Neumann, Peter G.* "Statistical metalinguistics and Zipf/Pareto/Mandelbrot", SRI International Computer Science Laboratory, accessed and archived 29 May 2011.
- 14.. (1980) *Orlov, Y. K.* (1980). Invisible Harmony. In: "Mysl' i chislo" („Thought and Number"), # 3, "Znanie", Moscow. pp. 70-106.
- 15.. (1935). *Zipf George K.* The Psychobiology of Language. Houghton-Mifflin. (<http://citeseer.ist.psu.edu/context/64879/0>).
- 16.. (1949). *Zipf, G. K.* Human Behavior and the Principle of Least Effort. Cambridge, Massachusetts: Addison-Wesley. p.1.
- 17.. Makagonov, Pavel. "Implementación de la Ley de Zipf para el análisis visual de patrones del ambiente urbano." REPOSITORIO NACIONAL CONACYT (2004).
- 18.. Alexander Trousov, Sergey Maruev. *Techno-Social Systems for Modern Economical and Governmental Infrastructures / Advances in Finance, Accounting, and Economics*: IGI Global.- 2018.-C. 1-286.// Local Forum as a Knowledge Source Pavel P. Makagonov, Celia Bertha Reyes Espinoza, Oleg M Dzikun -c.- 2018.-C. 100-125
- 19.. (2019) Pavel P. Makagonov¹, Lyudmila V. Tokun², Liliana Chanona Hernández³, Edith Adriana Jiménez Contreras⁴ Systemic Criteria for the Evaluation of the Role of Monofunctional Towns in the Formation of Local Urban Agglomerations: 1 Russian Presidential

Academy of National Economy and Public Administration, Russia, 2 State University of Management, Finance and Credit Department, Russia 3 Instituto Politécnico Nacional, Escuela Superior de Ingeniería Mecánica y Eléctrica, Mexico 4

Instituto Politécnico Nacional, Escuela Superior de Cómputo, Mexico. В журнале *Computación y Sistemas*, Vol. 23, No. 4, 2019, pp. 1457–1467. doi: 10.13053/CyS-23-4-3324ISSN 2007-9737 (Scopus, Web of Science).

20.. (2019) Ноздрина Надежда Николаевна, Макагонов Павел Петрович, Минченко Михаил Михайлович, Особенности развития городской системы расселения Московской области Features of the Development of the Urban Settlement System in the Moscow Oblast. В журнале «Проблемы прогнозирования», №6 2019 и её перевод помещен в журнале *Studies on Russian Economic Development*, 30(6), 682-693 DOI 10.1134/S1075700719060133 (Scopus) Nozdrina, N.N., Makagonov, P.P. & Minchenko, M.M. Features of the Development of the Urban Settlement System in the Moscow Oblast. <http://link.springer.com/article/10.1134/S1075700719060133>

21.. Переводы Корана на русский язык https://ru.wikipedia.org/wiki/%D0%9F%D0%B5%D1%80%D0%B5%D0%B2%D0%BE%D0%B4%D1%8B_%D0%9A%D0%BE%D1%80%D0%B0%D0%BD%D0%B0_%D0%BD%D0%B0_%D1%80%D1%83%D1%81%D1%81%D0%BA%D0%B8%D0%B9_%D1%8F%D0%B7%D1%8B%D0%BA