

УДК 821.161.1, 81'322.5

ПОЭЗИЯ А. А. БЛОКА С ТОЧКИ ЗРЕНИЯ КВАНТИТАТИВНЫХ МЕТОДОВ

Анна Вадимовна Фаттахова

студент образовательной программы «Филология»

Национальный исследовательский университет

Высшая школа экономики – Санкт-Петербург

190121, Санкт-Петербург, Союза Печатников, 16. avfattakhova_1@edu.hse.ru

Исследование посвящено лексико-статистическому анализу поэзии А.А. Блока на материале стихотворений и поэм, написанных поэтом в 1898–1921 гг.: составлен частотный словарь; получены данные о том, какие слова наиболее характерны для его поэзии, и проанализировано, как частотность слов меняется в разные периоды его жизни и творчества. В результате выявлен ряд закономерностей развития образов и мотивов в творчестве поэта; на основании применения статистических методов получены подтверждения мнений известных исследователей поэзии Блока об автобиографичности его поэзии, а также обнаружены некоторые очевидные расхождения с принятыми положениями относительно поэзии Блока.

Ключевые слова: символизм; Блок; анализ поэзии; лексико-статистический анализ; частотный словарь.

Введение

Исследовательский интерес к поэтическим и прозаическим произведениям А.А. Блока был высок уже при жизни поэта и не пропадает до сих пор. Особенности его произведений посвящено большое количество работ [Белый 1994, 2004; Жирмунский 1922; Крохина 2012; Минц 2004; Спивакова 2009; Чуковский 2010 и др.], однако до сих пор не было уделено должного внимания статистике используемой Блоком лексики. В наши дни лексико-статистический анализ стал более доступным, предоставляя возможность производить статистический анализ больших объемов текстовых данных. Предлагаемая работа посвящена лексико-статистическому анализу стихотворений и поэм, созданных Блоком на протяжении всего творческого пути (с 1898 по 1921 гг.).

Цель исследования состоит в составлении частотных словарей лирики Блока и выявлении с их помощью наиболее характерной для его поэзии лексики, а также динамики использования лексики в зависимости от разных периодов его жизни и творчества.

Для интерпретации результатов, полученных с помощью компьютерных методов, было проанализировано несколько работ блоковедов, среди которых Д.Е. Максимов и З.Г. Минц.

«Вырастившая поэзию Блока русская действительность <...> была полна великого брожения напряжения, острейших противоречий, идеологических конфликтов, судорожной борь-

бы за старое и страстного влечения к новому» [Максимов 1981: 17]. Среди особенно важных событий в жизни поэта Д.Е. Максимов выделяет наступление нового века, борьбу двух идеологий (социалистической и капиталистической), революции 1905 и 1917 гг., две войны, серьезно повлиявшие на состояние страны, повсеместную урбанизацию и, как следствие всего перечисленного, – смену идеалов и убеждений многих людей. В стихотворениях Блока прослеживается «ориентация на личность автора» и «тяготение к автобиографизму [там же: 18]: это причина, по которой нельзя оставить без внимания некоторые жизненные события и их взаимосвязь с содержанием сборников поэта.

В начале творческого пути Блок подвержен романтическому влиянию, поэтому в произведениях из первых сборников, по мнению З.Г. Минц, одним из важных мотивов является противостояние поэта (лирического героя) и толпы. В цикле «Стихи о Прекрасной Даме» можно заметить антитезу (свойственную и позднему Блоку), романтическое двоемирие: «противопоставление “земли” и “неба”, материального и духовного» [Минц 1983: 521]. В этом сборнике ощущается влияние на Блока мистицизма и философских идей В.С. Соловьева, среди которых – образ «Души мира» – «женственной по природе духовной субстанции (наиболее близкой к Weltseele Шеллинга и die EwigWeiblichkeit Гёте)» [там же: 535]. Итогом сборника является сомнение в реальности идеала, но еще живущая

надежда на значительные перемены в окружающей поэт действительности.

В 1903–1906 гг. в связи с быстро меняющейся обстановкой в стране, которая волнует поэта, он все чаще затрагивает социальные темы. Лирический герой уходит из «того» мира в «этот», реальный. Блок затрагивает проблему страдания народа, все чаще в его лирике можно заметить мотив бунта, стихии. В статье «Крушение гуманизма» Блок утверждает, что музыка рождается из стихии, которая же является хранителем музыки [Блок 1919: 4–9].

«Родина» (1907–1916 гг.) – цикл стихотворений, посвященных России и созданных вскоре после русской революции. Блок, как и многие писатели, размышляет о будущем родины, переживает, какой путь она изберет. В начале цикла еще нет «духа музыки», упорядоченного и направленного, как воля, только стихия и хаос, передающие настроение людей того времени. Стихия часто связывается Блоком, как и Шопенгауэром [Шопенгауэр 1919: электр. ресурс] с народной силой, особенно в цикле «Фаина» (1906–1908 гг.). З. Г. Минц обращает внимание на важную деталь: по сравнению с ранними сборниками поэта, его «цель перемещается с “небес” на “горестную землю”, неразрывно сплавившись с надеждами на “вочеловечение” идеала» [Минц 1983: 530]. В период между двумя русскими революциями основными для Блока становятся гражданские темы, в частности, проблема долга. Стихотворения начала 1910-х гг. меняют свою тематику, в связи с трагическими событиями в жизни Блока и его путешествием в Италию. В последующий период «дух музыки» все больше прослеживается в циклах стихотворений, вытесняя неконтролируемую стихию. Мир Блока приобретает гармонию и воодушевленность, направленность в будущее.

Если проследить, как меняется стиль А. Блока от тома к тому (в рамках трилогии), можно выделить общие черты. Первый том (центром которого являются «Стихи о Прекрасной Даме») – начало духовного становления, как Блока, так и его лирического героя; воплощение мечты, юности, любви, идеализация окружающей действительности. Второй том – разрушение первого мира, падение героев в «страшный мир» (название первого цикла в третьем томе; центр второго – «Нечаянная Радость» и «Снежная маска»). Третий том – связь с прошлым, древностью (цикл «Родина»), мужественность, жертвенность, частичное возвращение к миру любви из первого тома.

Творчество Блока неразрывно связано с таким литературным направлением, как символизм. Главным его понятием является «символ» – об-

раз, который допускает практически неограниченное количество толкований. Символ «должен быть многозначным, должен ощущаться как “бездонно” таинственный, “многосмысленный” (В. Иванов), но, тем не менее, он не может включать в себя <...> бесконечное число интерпретаций» [Минц 2004, 3: 45] (в рамках одной культуры и одной символистской традиции).

«Красота» и «панэстетизм» – «представление об эстетическом как о глубинной сущности мира», «восприятие и осмысление мира в эстетических категориях» [Минц 2004, 2: 459] – важные понятия для понимания взглядов символистов, в особенности Блока. В начале XX в. некоторые символисты отождествляют красоту с хаосом и стихией, а панэстетизм проявляется в противопоставлении понятий, в разделении их на «эстетическое» и «неэстетическое», например: «гармония», она же «космос» – «дистгармония»; «хаос», «музыка» – «стихия», «аполлоническое» – «дионисийское» [Минц 2004, 2: 459–461] Таким образом, для символизма особенное значение имеет двоемирие и контраст.

Блок утверждал, что символизм – это не просто течение в литературе, а философская концепция: «Символистом можно только родиться; <...> Оттого писатели даже с большими талантами не могут ничего поделать с искусством, если они не крещены “огнем и духом” символизма» [Блок 1910].

Материал и методика

Для лексико-статистического анализа были использованы оцифрованные тексты пяти первых томов двадцатитомного собрания сочинений А.А. Блока [Блок 1997]. Выбор издания не случаен: в нем сохраняется именно то разделение стихотворений на три части, которое было проведено самим автором. В первых трех томах располагаются циклы стихотворений Блока (первый том – 1898–1904 гг., второй том – 1904–1908 гг., третий том – 1907–1916 гг.); в четвертом (1897–1915 гг.) – поэтические произведения, не вошедшие ни в один из циклов (не включенные Блоком в «роман-трилогию»), в пятом – поэмы «Двенадцать», «Возмездие», «Скифы» и другие стихотворения 1917–1921 гг., вследствие хронологии не попавшие в более ранние сборники.

Материалом исследования выступает собранный на основе пяти томов стихотворений Блока корпус, состоящий из 103 525 русских и 52 иностранных слов. Иноязычная лексика не учитывалась при подсчете дальнейшей статистики: она была исключена из корпуса, так как ее небольшое количество представляется нерепрезентативным для анализа. В корпус включены все по-

этические произведения Блока, среди которых не только лирические стихотворения, но и поэмы.

Оцифровка текстов была произведена с помощью программы Abby Fine Reader (см. подробнее: <https://pdf.abbyy.com>) произошедшие при этом процессе ошибки были исправлены вручную. Оцифрованные сборники нормализованы: удалены вручную приложения, комментарии, сноски, эпитафии, даты написания стихотворений, их названия и т.п.

С помощью программы морфологического парсинга MyStem 3.1 (<https://yandex.ru/dev/mystem>) была произведена лемматизация всех стихотворений. Получившийся корпус из более чем 100 тысяч слов был проверен на наличие неточностей и нераспознанных форм. Отдельно были исправлено 4 480 глагольных форм, инфинитив которых не соответствовал по виду личной форме. Кроме того, были уточнены некоторые случаи омонимии (например, слово *дали*, которое в одних случаях является существительным *даль*, а в других – глаголом *дать*).

В приложении Microsoft Excel были получены основные статистические характеристики текстов, а также списки частотных лемм. К корпусу

был применен список стоп-слов¹ для русских текстов [Hingston: электр. ресурс], который был адаптирован для текстов стихотворений и поэм Блока: так, из списка стоп-слов были исключены некоторые слова (например, *белый* – важный для Блока цвет, или существительное *жизнь*, важное для понимания контраста в блоковской поэзии). С помощью инструмента Voyant tools (<https://voyant-tools.org>) были визуализированы текстовые представления сборников (облака слов) и построены тренды распределений частот нескольких слов по первым трем стихотворным томам А. Блока.

В Таблицах 1–3 представлена верхняя зона частотных словарей для существительных, прилагательных и глаголов (20 слов) в сравнении с их частотностью в словаре О.Н. Ляшевской, С.А. Шарова [Ляшевская, Шаров 2009: электр. ресурс], полученном на материале Национального корпуса русского языка [Национальный корпус русского языка: электр. ресурс]. На наш взгляд, наиболее характерными для поэзии Блока словами будут являться те, *ipm* которых существенно превышает их *ipm* словаре О.Н. Ляшевской, С.А. Шарова.

Таблица 1

Список 20-ти самых частотных существительных в поэзии Блока

Ранг в словаре суц.	Лемма	Абс. частота	<i>ipm</i> (в корпусе)	Ранг в словаре О.Н. Ляшевской, С.А. Шарова	<i>ipm</i> в словаре О.Н. Ляшевской, С.А. Шарова
1	ночь	463	4472	236	403
2	душа	421	4067	275	357
3	день	417	4028	71	1258
4	сердце	332	3207	432	245
5	сон	326	3149	769	153
6	жизнь	275	2656	66	1390
7	рука	263	2540	74	1201
8	огонь	215	2077	855	140
9	земля	214	2067	184	494
10	небо	200	1932	600	188
11	свет	198	1913	272	360
12	глаз	196	1893	110	858
13	окно	192	1855	370	281
14	мечта	186	1797	2086	57
15	ветер	182	1758	847	140
16	тень	177	1710	1298	94
17	любовь	174	1681	307	324
18	весна	167	1613	1345	91
19	туман	167	1613	2651	43
20	голос	164	1584	221	425

Среди существительных наиболее показательными словами оказались: *ночь* (*ipm* в корпусе Блока – 4 472; *ipm* в словаре – 403) *сон* (3 149 – 153 соответственно), *огонь* (2077 – 140), *небо* (1932 – 188), *мечта* (1797 – 57), *весна* (1613 – 91), *туман*

(1613 – 43). Обратим внимание, что некоторые из этих слов связаны по смыслу (ассоциативно) *ночь & сон & мечта, мечта & туман*. Символизм *ночи* и *мечты* (особенно в ранних сборниках Блока) имеет особое значение, о чем упоминалось ранее.

Таблица 2

Список 20-ти самых частотных глаголов в поэзии Блока

Ранг в словаре глаг.	Лемма	Абс. частота	ipm (в корпусе)	Ранг в словаре О.Н. Ляшевской, С.А. Шарова	ipm в словаре О.Н. Ляшевской, С.А. Шарова
1	быть	818	7901	6	12161
2	знать	226	2183	61	1714
3	идти	220	2125	95	957
4	петь	189	1826	828	143
5	смотреть	189	1826	136	667
6	мочь	186	1797	271	360
7	ждать	177	1710	271	360
8	видеть	145	1401	113	818
9	любить	135	1304	181	503
10	прийти	116	1121	176	523
11	пройти	115	1111	242	398
12	жить	110	1063	126	726
13	забыть	104	1005	478	226
14	стать	102	985	62	1622
15	гореть	98	947	1603	77
16	спать	97	937	486	222
17	плакать	91	879	1165	104
18	глядеть	89	860	554	199
19	звать	89	860	912	131
20	сказать	88	850	42	2397

Из глаголов самую большую разницу в ipm в текстах Блока и в частотном словаре имеют *петь* (1826 – 143), *любить* (1304 – 503), *ждать* (1710 – 360), *гореть* (947 – 77), *звать* (860 – 131). *Петь* – «музыкальный» глагол, является еще одним из множества доказательств музы-

кальности блоковской поэзии. Интересно отметить, что слово *гореть*, как и *огонь*, очень часто повторяется у Блока, создавая контраст с некоторыми очень частотными прилагательными, представленными в Таблице 3 (*снежный, бледный*).

Таблица 3

Список 20-ти самых частотных прилагательных в поэзии Блока

Ранг в словаре сущ.	Лемма	Абс. частота	ipm (в корпусе)	Ранг в словаре О.Н. Ляшевской, С.А. Шарова	ipm в словаре О.Н. Ляшевской, С.А. Шарова
1	белый	206	1990	290	340
2	темный	191	1845	803	146
3	черный	173	1671	293	338
4	ночной	148	1430	1463	85
5	светлый	140	1352	1355	91
6	новый	129	1246	73	1218
7	милый	124	1198	1353	91
8	красный	120	1159	442	241
9	последний	118	1140	130	713
10	синий	110	1063	1498	83
11	тихий	107	1034	1131	107
12	холодный	107	1034	1008	119
13	старый	98	947	264	369
14	далекий	96	927	311	318
15	бледный	94	908	3024	37
16	вечерний	94	908	2891	39
17	полный	94	908	250	391
18	живой	92	889	364	284
19	золотой	88	850	967	124
20	снежный	86	831	4142	25

Среди прилагательных по разнице в *ipm* текстах Блока и частотном словаре особенно выделяются *снежный* (831 – 25), *вечерний* (908 – 39), *бледный* (908 – 37), *ночной* (1430 – 85). В этом списке довольно неожиданным стало слово *бледный*: многое было сказано о контрастности поэзии символистов, которая не принимает «средних» оттенков, поэтому наличие такого цвета удивляет.

Результаты и их интерпретация

На материале исследовательской выборки были получены следующие статистические показатели (данные приводятся для лемм, так как анализ лемм считается наиболее репрезентативным):

- количество типов – 10 061;
- количество токенов – 103 525;
- гапакс (количество уникальных лемм) – 4 193;
- TTR (коэффициент лексического разнообразия) – 0,097184255.

В Таблице 4 приводится список 20-ти самых частотных лемм (слева – до удаления стоп-слов, справа – после их удаления).

Данные Таблицы 5 показывают, что количественно имена существительные преобладают над всеми другими частями речи: существительных в 1,65 раз больше, чем глаголов и в 2,32 раза больше, чем прилагательных.

Таблица 4

Список 20-ти самых частотных лемм в поэзии Блока

До удаления стоп-слов					После удаления стоп-слов				
Ранг	Лемма	Абс. частота	<i>ipm</i>	%	Ранг	Лемма	Абс. частота	<i>ipm</i>	%
1	и	5312	51311	5,13	1	ночь	463	4472	0,45
2	в	3715	35885	3,59	2	душа	421	4067	0,41
3	я	2614	25250	2,52	3	день	417	4028	0,40
4	ты	1641	15851	1,59	4	сердце	332	3207	0,32
5	не	1620	15648	1,56	5	сон	326	3149	0,31
6	на	1268	12248	1,22	6	жизнь	275	2656	0,27
7	мой	868	8384	0,84	7	рука	263	2540	0,25
8	с	830	8017	0,80	8	огонь	215	2077	0,21
9	быть	818	7901	0,79	9	земля	214	2067	0,21
10	как	816	7882	0,79	10	белый	206	1990	0,20
11	что	697	6733	0,67	11	небо	200	1932	0,19
12	но	687	6636	0,66	12	свет	198	1913	0,19
13	мы	611	5902	0,59	13	глаз	196	1893	0,19
14	он	603	5825	0,58	14	окно	192	1855	0,19
15	твой	601	5805	0,58	15	темный	191	1845	0,18
16	она	545	5264	0,53	16	петь	189	1826	0,18
17	все	531	5129	0,51	17	мечта	186	1797	0,18
18	о	531	5129	0,51	18	ветер	182	1758	0,18
19	за	503	4859	0,49	19	ждать	177	1710	0,17
20	ночь	463	4472	0,45	20	тьень	177	1710	0,17

Таблица 5

Статистика по частям речи

Ранг	Часть речи	Абс. частота	<i>ipm</i>	%
1	S	30611	295687	29,57
2	V	18535	179039	17,90
3	A	13198	127486	12,75
4	PR	10182	98353	9,84
5	CONJ	8630	83362	8,34
6	SPRO	7940	76696	7,67
7	ADV	4489	43362	4,34
8	APRO	4265	41198	4,12
9	PART	3538	34175	3,42
10	ADVPRO	1667	16102	1,61
11	INTJ	195	1884	0,19
12	NUM	154	1488	0,15
13	ANUM	121	1169	0,12

Приведем данные из «Частотного словаря русского языка» под редакцией Л.Н. Засориной [Засорина 1977: 927]. Для релевантного сравнения были немного изменены частеречные категории, чтобы они совпадали для обеих выборок: в корпусе Блока разные местоимения объединены в одну группу, в словаре Л.Н. Засориной причастия отнесены к глаголам.

Исследование показало, что, несмотря на заметные отличия в долях между распределениями частот (см. Рис. 1), результаты теста χ^2 обнаружили отсутствие статистически значимых различий между распределением частот частей речи. Однако стоит отметить, что процент существительных в поэтическом корпусе на 5,75% выше, чем в Частотном словаре, а глаголов, наоборот, на 2,1% меньше. Таким образом, можно сделать

вывод, что в стихотворениях Блока частотность существительных выше частотности глаголов (ср.: в Частотном словаре под редакцией Л.Н. Засориной разница между существительными и глаголами всего 3,82%, в то время как в корпусе Блока этот разрыв составляет 11,67%).

Обратим также внимание на конкретные слова и проследим развитие образов по сборникам. Для начала приведем три облака слов, созданные по каждому из томов «трилогии вочеловечивания» (см. Рис. 2).

Из приведенных выше визуализаций мы видим, какие слова проходят в качестве лейтмотива через все сборники поэта – это слова *день, ночь, сон, ду-*

ша, сердце, темный, рука и др. Однако нельзя сказать, что облака слов очень похожи, т. к. можно заметить значительные различия. К примеру, в первом томе одним из наиболее частых слов является *день* (абсолютная частота – 114; ipm – 5698), которое немного реже встречается в третьем томе (абсолютная частота – 81; ipm – 3634) и значительно реже – во втором (абсолютная частота – 48; ipm – 1955). При этом лемма *ночь*, которая примерно одинакова по частоте во всех томах, становится наиболее частотной именно в третьем. Кроме того, во втором томе лемма *ночь* представлена почти в таком же количестве, как и лемма *тёмный* (102; 4154 – 79; 3217 соответственно).

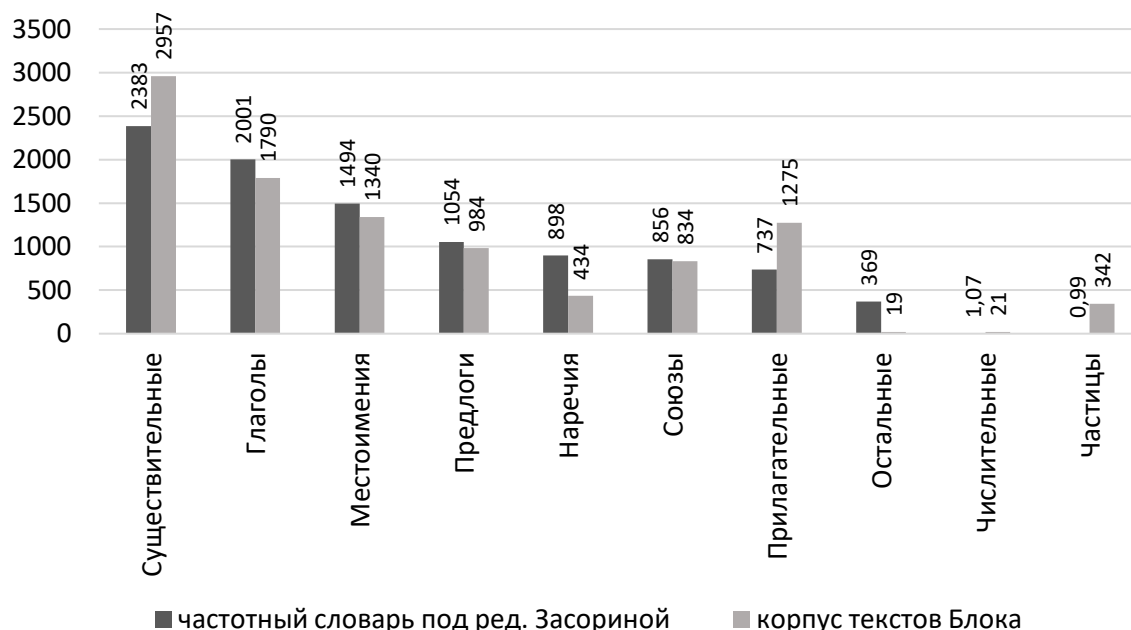


Рисунок 1. Распределение частей речи в текстах Блока на фоне распределения частей речи в Частотном словаре под ред. Л.Н. Засориной



Рисунок 2. Облако слов разных томов (периодов): а) первый том; б) второй том; в) третий том

В первом томе, как и в третьем, наиболее часто из «телесных» слов встречается *сердце*, в то время как во втором на первый план выходит *рука*. Однако во всех частотных списках также присутствуют *душа* и *голос* (а также глагол *петь*), как представители «нематериального», которое находится в постоянной борьбе с материальным. Победа «духовного» незаметна на

примере такой визуализации, однако довольно наглядно видна благодаря графику, построенному с помощью Voyant tools (см. Рис. 3). На этом и последующих графиках представлены частотности употребления слов, соединенные кривой для наглядности (lemmas 1, lemmas 2, lemmas 3 – лемматизированные корпуса стихотворений первого, второго и третьего томов соответственно).

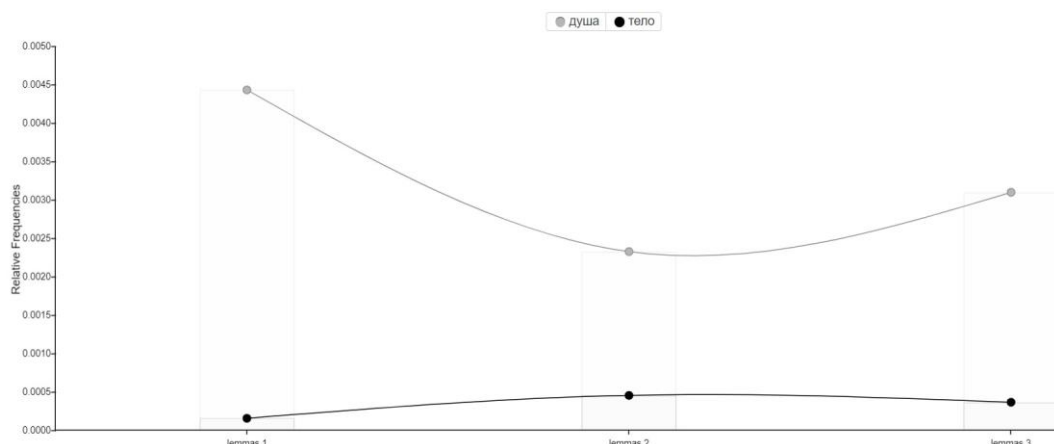


Рисунок 3. График распределения частот «контрастных» слов *душа* и *тело* в лемматизированных корпусах 1–3 томов

Неоднократно упомянутая приверженность Блока к контрастам проявляется, в том числе, благодаря использованию цветов. Рассмотрим распределение белого и черного на протяжении трех томов (поиск осуществлялся по лемматизированным корпусам). На Рисунке 4 можно видеть, что два цвета, как два начала, «борются»

друг с другом: сначала белого цвета было значительно больше, чем черного (в 1-ом томе *ipm* *белого* – 2549, *черного* – 1200), однако, достигнув своего пика во втором томе (*ipm* – 2932), частота *белого* начинает снижаться, пересекает возрастающую линию *черного* и опускается еще ниже (в 3-ем томе *ipm* *белого* – 1301, *черного* – 2153).

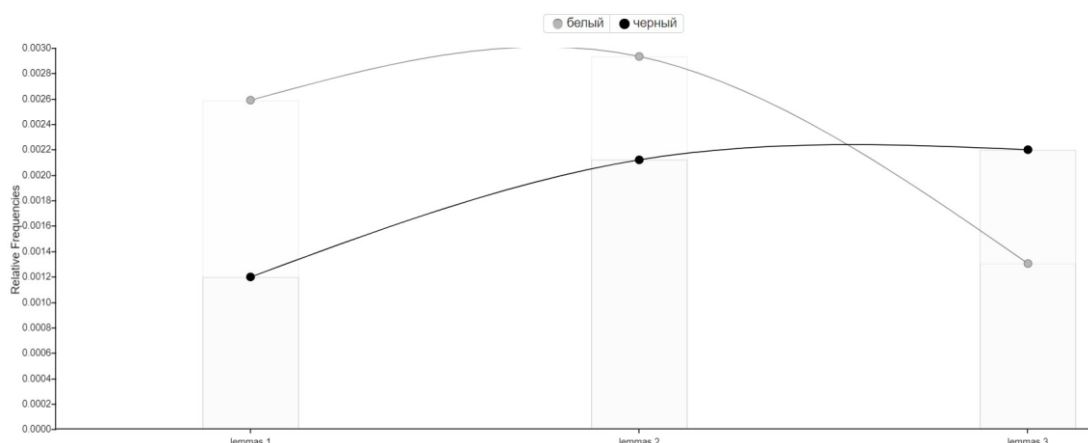


Рисунок 4. График распределения частот «контрастных» слов *белый* и *черный* в лемматизированных корпусах 1–3 томов

Еще несколько антитез, характерных Блоку, наглядно представлены на Рисунках 5 и 6 (антитезы *жизнь* – *смерть*, *небо* – *земля*). Результаты получились довольно неожиданными: первый сборник раннего Блока, еще не до конца отошедшего от соловьевского мистицизма, получился *дневным* и очень *земным*. Опираясь на исследования З.Г. Минц, мы ранее считали, что лирический герой Блока (а также идеалы поэта, такие как Прекрасная Дама) ко второму сборнику должны «спуститься с неба на землю», но статистические показатели демонстрируют обратную картину: частотность *неба* во втором сборнике возрастает, в то время как частотность *земли* стремительно падает. Это нарушает и еще одну гипотезу: циклы, в которых Блок пе-

реживает за судьбу родины, обращается к истории, социальным темам и долгу, получились самыми *небесными*, хотя должны быть *земными*.

Интересной (но более ожидаемой, в этом случае) деталью показалось нам возрастание «жизнерадостности» Блока к концу своего «романа в стихах». Частотность леммы *жизнь* значительно превышает частотность леммы *смерть* в третьем томе. Это может быть объяснено тем, что в третьем томе Блок примеряется с тем, от чего ушел ранее – *юности*, *любви*, *неба* (хотя последнее под сомнением, благодаря полученной нами частотности) и ввиду второй революции находится в одухотворенном состоянии, надеется на счастливое будущее для России.

Позднему периоду творчества блоковской поэзии характерна музыкальность: невозможно не отметить увеличение частоты «музыкальных» слов (см. Рис. 7) на примере слов *музыка* (1 том – $ipm = 50$, 3 том – $ipm = 314$) и *петь* (1 том – $ipm = 1649$, 3 том – $ipm = 2557$).

Продолжая темы контрастов, вновь обратимся к тому, что в первом сборнике достаточно частотной леммой оказалось существительное

день. По трендам мы можем проследить, что лемма *день* в первом томе находится на пике частотности ($ipm = 5698$), затем он совершает резкое падение во втором ($ipm = 1955$) и в третьем немного поднимается ($ipm = 3634$), но все равно занимает значительно более низкий уровень по частоте, чем в первом томе. В свою очередь, частота употребления слова *ночь* плавно растет от первого к третьему тому.

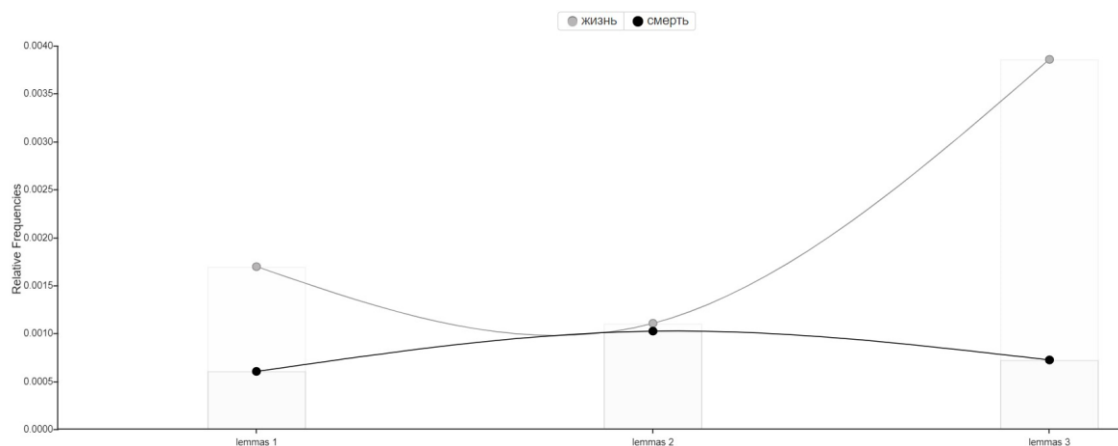


Рисунок 5. График распределения частот «контрастных» слов *жизнь* и *смерть* в лемматизированных корпусах 1–3 томов

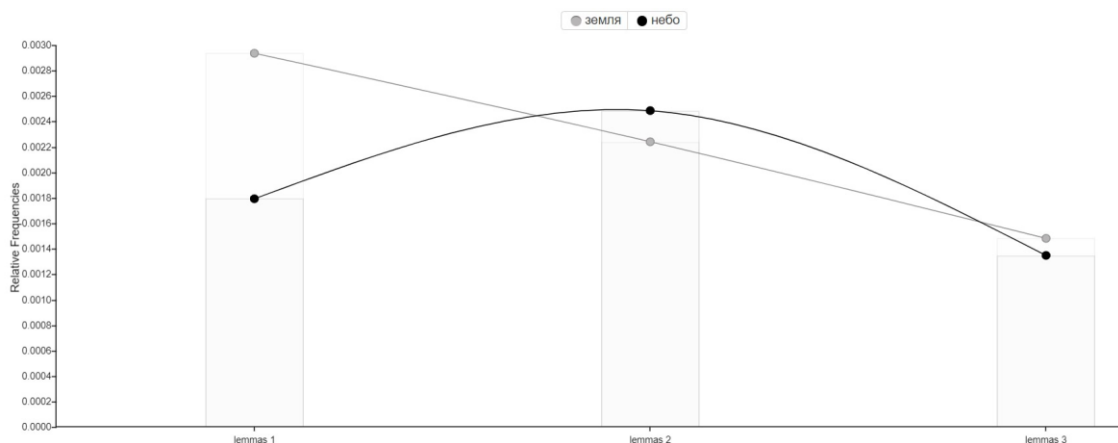


Рисунок 6. График распределения частот «контрастных» слов *земля* и *небо* в лемматизированных корпусах 1–3 томов

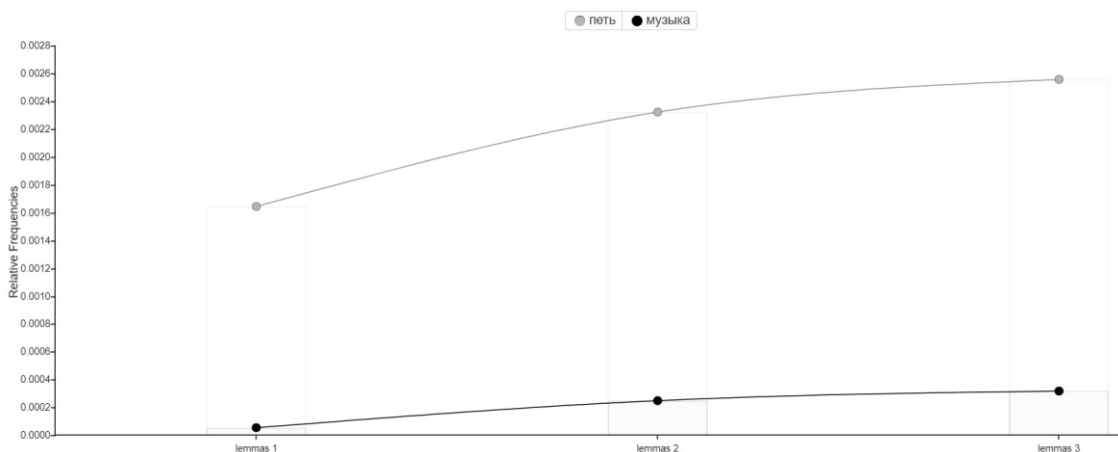


Рисунок 7. График распределения частот «музыкальных» слов в лемматизированных корпусах 1–3 томов

Продолжая темы контрастов, вновь обратимся к тому, что в первом сборнике достаточно частотной леммой оказалось существительное *день*. По трендам мы можем проследить, что лемма *день* в первом томе находится на пике частотности (*ipm* – 5698), затем он совершает резкое падение во втором (*ipm* – 1955) и в третьем немного поднимается (*ipm* – 3634), но все равно занимает значительно более низкий уровень по

частоте, чем в первом томе. В свою очередь, частота употребления слова *ночь* плавно растет от первого к третьему тому.

Отдельно стоит остановиться на следующем моменте: по словам З.Г. Минц, «образы ветра, метели прошли через всю поэзию Блока, став в ней своеобразными опорными символами динамизма жизни» [Минц 1983, 527]. Однако статистические данные не совсем это подтверждают (см. Рис. 9).

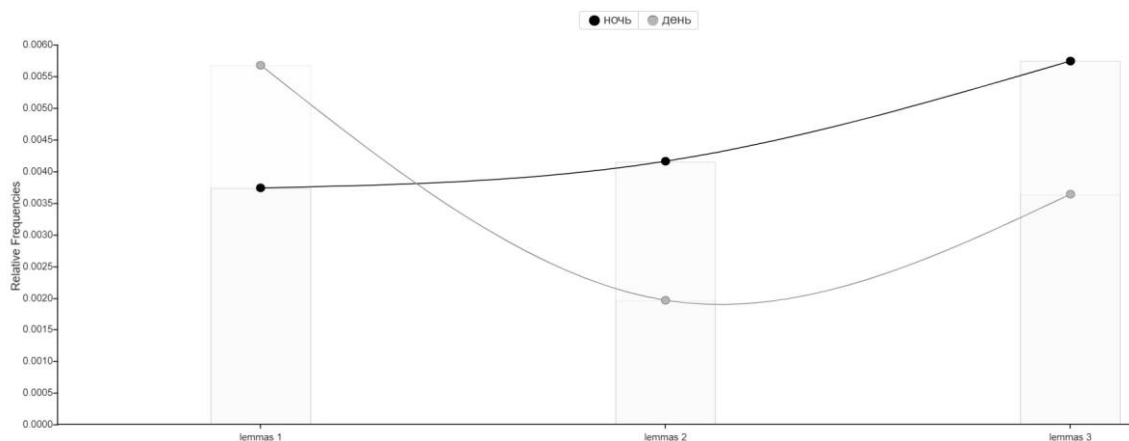


Рисунок 8. График распределения частот «контрастных» слов *день* и *ночь* в лемматизированных корпусах 1–3 томов

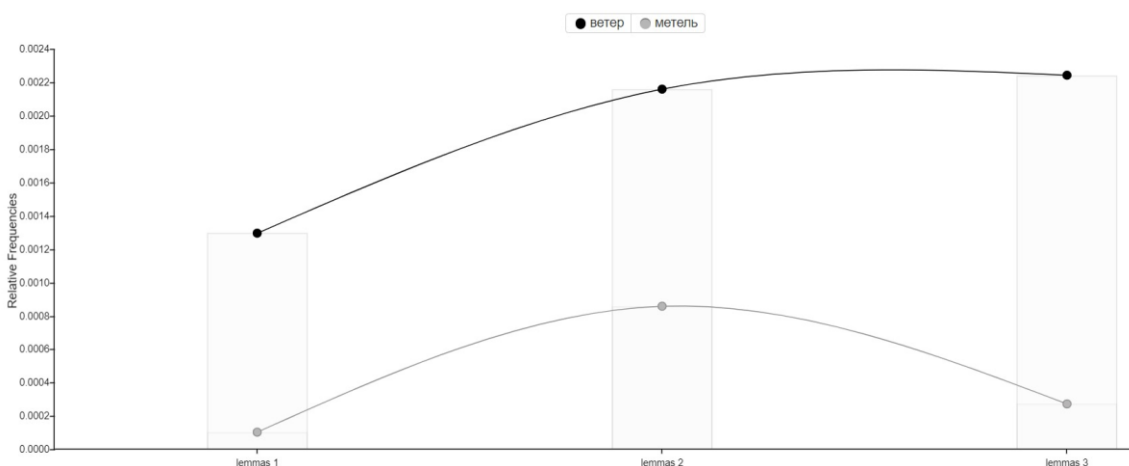


Рисунок 9. График распределения частот слов *ветер* и *метель* в лемматизированных корпусах 1–3 томов

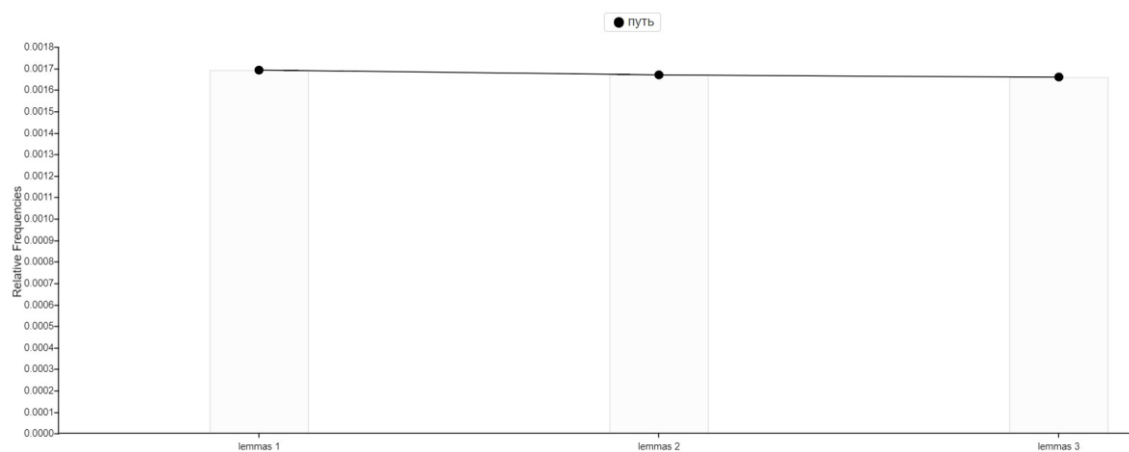


Рисунок 10. График распределения частоты слова *путь* в лемматизированных корпусах 1–3 томов

Из графика, представленного на Рис. 9 видно, что лемма *ветер* действительно достаточно частотна и проходит через все три сборника, в отличие от *метели*, частота употребления которой является более или менее значительной лишь во втором томе (в котором одним из ключевых является цикл «Снежные маски»). А вот тема пути, судя по всему, является для Блока действительно одной из самых важных и стабильных на протяжении всего его творчества (см. Рис. 10).

Заключение

По сравнению с «Частотным словарем современного русского языка» [Ляшевская, Шаров 2009: электр. ресурс] выявлены следующие особенности корпуса Блока: наиболее показательными существительными для его поэзии оказались слова *ночь, сон, огонь, небо, мечта, весна, туман*. Ряд наиболее частотных для русского языка слов – *человек, время, работа* не представлен в верхней зоне частотного словаря Блока. Из глаголов, самыми репрезентативными стали *петь, любить, ждать, гореть и звать*. При сравнении с частотным словарем О.Н. Ляшевской, С.А. Шарова было обнаружено, что в последнем наибольшую частотность употребления имеют глаголы с базовым значением (*мочь, сказать, говорить, знать, стать, идти, иметь, видеть*). Многие из этих слов добавлены в список стоп-слов, либо не вошли в верхнюю границу словаря, показывая отличие поэтической речи от обычной. Анализ самых частотных прилагательных показал, что для поэзии Блока характерны прилагательные *снежный, вечерний, бледный и ночной*, в то время как в частотном словаре О.Н. Ляшевской, С.А. Шарова одни из первых строк среди прилагательных занимают *русский, общий, высокий, лучший*, не вошедшие в верхнюю зону частотного словаря Блока.

По результатам проведенного исследования не было выявлено данных, которые бы кардинальным образом отличались от полученных ранее блоковедами, в частности, было подтверждено, что поэзия Блока автобиографична: в лирике зачастую встречаются темы, важные для самого поэта и прослеживаются изменения в ориентирах, во взгляде на мир и ситуацию в стране (недаром для Блока так важна тема *пути*, проходящая через всю его поэзию, что было доказано статистически).

Однако было выявлено несколько несоответствий с выдвинутыми ранее предположениями. Например, при переходе от первого тома стихотворений Блока ко второму, по мнению З.Г. Минц, лирический герой «спускается с неба на землю»; результаты подсчета частотности лемм *небо* и *земля* показывают, что происходит

скорее обратный процесс, так как *irm* слова *земля* в первом томе значительно выше *irm* слова *небо*, а во втором томе – наоборот. Также, в отличие от другого утверждения З.Г. Минц, не было обнаружено, что образы *ветра* и *метели* проходят через всю поэзию Блока.

Для дальнейшего развития исследования представляется целесообразным проведение более детального анализа конкретных частей речи (например статичных vs динамичных), динамики образов (например образа Петрограда) и иных важных особенностей поэзии Блока.

Благодарности

Публикация подготовлена в результате проведения исследования по проекту «Текст как Big Data: моделирование конвергентных процессов в языке и речи цифровыми методами», выполненного в рамках Программы фундаментальных исследований НИУ ВШЭ в 2023 г.

Примечание

¹ Стоп-слова — слова, не несущие смысловой нагрузки: знаки, символы, служебные части речи (предлоги, союзы, междометия) и проч.

Список источников

Блок А.А. Андрей Белый и Александр Блок: Переписка. 1903–1919 // Собрание сочинений: в 12-и т. / под общ. ред. С.С. Лесневского М.: Прогресс-Плеяда, 2001. Т. 12. С. 15–21.

Блок А.А. Дневники и записные книжки (фрагменты). 1921. [Электронный ресурс]. URL: http://az.lib.ru/b/blok_a_a/text_0090.shtml (дата обращения: 01.11.2023).

Блок А.А. Полное собрание сочинений и писем: в 20-и т. / ред. Ю.К. Герасимов. М.: Наука, 1997.

Ляшевская О.Н., Шаров С.А. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М.: Азбуковник, 2009. [Электронный ресурс]. URL: <http://dict.ruslang.ru/freq.php> (дата обращения: 01.11.2023).

Национальный корпус русского языка. [Электронный ресурс]. URL: <https://ruscorpora.ru> (дата обращения: 01.11.2023).

Hingston W. 10000-russian-words.txt. [Электронный ресурс]. URL: <https://github.com/hingston/russian/blob/master/10000-russian-words.txt> (дата обращения: 01.11.2023).

Список литературы

Белый А. Александр Блок: pro et contra / сост., вступ. ст., примеч. Н.Ю. Грякаловой. СПб.: Рус. христ. гум. ин-та, 2004. 736 с.

Белый А. Символизм как миропонимание / сост., вступит. ст. и прим. Л.А. Сугай. М.: Республика, 1994. 528 с.

Блок А. А. О современном состоянии русского символизма. 1910. [Электронный ресурс]. URL: http://az.lib.ru/b/blok_a_a/text_1910_o_sostoyanii_simvolizma.shtml (дата обращения: 01.11.2023).

Блок А.А. Крушение гуманизма. 1919. [Электронный ресурс]. URL: http://dugward.ru/library/blok/blok_krushenie_gumanizma.html (дата обращения: 01.11.2023).

Жирмунский В.М. Поэзия Александра Блока. Петербург: Картонный домик, 1922. 103 с.

Частотный словарь русского языка / под. ред. Л.Н. Засориной. М.: Русский язык, 1977. С. 926–927.

Крохина Н.П. Софья Вл. Соловьёва и А. Блока // Соловьёвские исследования. 2012. Вып. 4(36). С. 74–82.

Максимов Д.Е. Поэзия и проза Ал. Блока. Л.: Советский писатель: ленинградское отделение, 1981. 525 с.

Миц З.Г. Александр Блок // История русской литературы: в 4-х т. Л.: Наука, 1983. Т. 4. С. 520–548.

Миц З.Г. Александр Блок и русские писатели // Миц З.Г. Блок и русский символизм: Избранные труды: в 3-х кн. СПб.: Искусство-СПб, 2004. Кн. 2. 775 с. Кн. 3. 478 с.

Стивакова Е.М. Язык цвета в идиостиле А. Блока: автореферат дис. ... канд. филол. наук. Владивосток, 2009. 23 с.

Чуковский К.И. Александр Блок как человек и поэт: введение в поэзию Блока М.: Русский путь, 2010. 181 с.

Шопенгауэр А. О сущности музыки: Выдержки из соч. Шопенгауэра под. Историко-эстет. секции и со вступ. статьей К. Эйгес. Петербург: Музыкальное государственное изд-во, 1919. [Электронный ресурс]. URL: <https://opentextnn.ru/music/interpretacija-teksta-muzyki/shopengauer-a-o-sushhnosti-muzyki> (дата обращения: 01.11.2023).

THE POETRY OF A. A. BLOCK IN TERMS OF QUANTITATIVE METHODS

Anna V. Fattakhova

Student, Program in Philology

National Research University Higher School of Economics – St Petersburg

The study is dedicated to the lexical-statistical analysis of A.A. Blok's poetry based on the poems written by the poet during 1898–1921. A frequency dictionary has been compiled. It was revealed the most typical words of his poetry, and how the frequency of words changes in different periods of his life and work was analyzed. As a result, a number of patterns in the development of images and motifs in the poet's work have been identified; Based on the application of statistical methods, confirmation of the opinions of famous researchers of Blok's poetry about the autobiographical nature of his poetry was obtained, and some obvious discrepancies with most common positions regarding Blok's poetry were also discovered.

Keywords: symbolism; Blok; poetry analysis; statistical analysis of lexis; frequency dictionary.