

РОССИЙСКАЯ АКАДЕМИЯ НАУК  
ИНСТИТУТ РУССКОГО ЯЗЫКА им. В. В. ВИНОГРАДОВА

А. Я. ШАЙКЕВИЧ,  
В. М. АНДРЮЩЕНКО, Н. А. РЕБЕЦКАЯ

СТАТИСТИЧЕСКИЙ СЛОВАРЬ  
ЯЗЫКА РУССКОЙ ГАЗЕТЫ  
(1990-е годы)

Том 1



ЯЗЫКИ СЛАВЯНСКИХ КУЛЬТУР  
Москва 2008

ББК 83  
Ш 12

Издание осуществлено при финансовой поддержке  
Российского фонда фундаментальных исследований  
(РФФИ)  
проект № 08-06-07023



**Рецензенты:**

доктор филол. наук *Д. О. Добровольский*,  
доктор филол. наук *С. Е. Никитина*

**А. Я. Шайкевич, В. М. Андリュщенко, Н. А. Ребецкая**

Ш 12      Статистический словарь языка русской газеты (1990-е годы) / Рос. акад. наук. Ин-т русского языка им. В. В. Виноградова. Т. 1. — М.: Языки славянских культур, 2008. — 592 с., разд. паг. (VIII, 578 с.). — (Studia philologica).

ISSN 1726-135X  
ISBN 978-5-9551-0279-5

Настоящий словарь представляет собой композицию трех частотных словарей, за каждым из которых стоит свой корпус текстов: 1) тексты девяти московских и петербургских газет за 1997 г., 2) комплект «Независимой газеты» за 1996—2000 гг., 3) корпус газетных заголовков 1991—2000 гг. Общий объем трех корпусов составляет 50 млн слов текста. В печатной версии словаря представлено 52 тыс. разных слов, в электронной версии словник превышает 140 тыс. разных слов, там же читатель найдет и соответствующий обратный словарь.

Во втором томе словаря будут даны таблицы распределения лексики по газетам, жанрам и темам; электронная версия включит таблицы бинарных словосочетаний.

**ББК 83**

ISBN 978-5-9551-0279-5

© Авторы, 2008  
© Языки славянских культур, 2008

## Введение

Жанр частотных словарей когда-то рассматривался лингвистами как сугубо прикладное направление. Один из грандиознейших словарей подобного рода, созданных в докомпьютерную эру, вышел в свет под характерным заголовком — «Словарь учителя...» (Thorndike E. L., Lorge I. The Teacher's Word Book of 30,000 words. N. Y., 1944). Первый частотный словарь русского языка (Josselson H. The Russian Word Count. Detroit, 1953) был адресован преподавателям, он опирался на корпус в 1 млн словоупотреблений (1830–1950) и содержал более 5 тысяч разных лексем. На основе корпуса в 400 тыс. словоупотреблений (литература для детей и юношества) построен «Частотный словарь современного русского языка» Э. А. Штейнфельдт (Таллин, 1963), в нем было представлено 2500 наиболее употребительных слов.

Первый опыт использования компьютера для этих целей материализовался в виде «Частотного словаря русского языка» под ред. Л. Н. Засориной (М., 1977). Словарь построен на базе текстов 1900–1960-х годов общим объемом 1 млн словоупотреблений. Опубликованный словарь включал весь лексический материал (кроме имен собственных). В 1993 г. в Швеции опубликован «Частотный словарь современного русского языка» Л. Леннгрена (Uppsala, 1993). Исходный корпус (1 млн словоупотреблений) содержал в равной доле художественную литературу 1960–1980-х годов и журнально-газетные тексты 1985–1988 годов. К сожалению, в публикацию не включены лексические единицы с частотой 1–9, т. о. в печатном словаре находим около 9 тыс. лексем.

Появление персонального компьютера, сканирующих устройств, развитие Интернета было технологической революцией и, казалось бы, обещало появление все новых и новых частотных словарей. В действительности, прогресс в этой области был не столь быстрым. При росте текстовых корпусов на два (или даже три порядка) автоматизация разметки текста и лемматизации повысили эффективность труда на порядок.

Первым примером такого типа может служить Частотный словарь чешского языка — *Frekvenční slovník češtiny* (pod vedením F. Čermáka a M. Křena). Praha, 2004. Национальный корпус чешского языка, на базе которого создан этот словарь, включает около 100 млн. словоупотреблений. В словаре представлено более 50 тыс. лемм, начиная с частоты 13.

Одновременно меняются конечные цели, ради которых создаются статистические словари языка. Во главу угла ставится уже не прикладная задача отбора лексики для студентов-иностранцев, но фундаментальная задача описания языка во всем разнообразии жанров, стилей и периодов развития. Первым замечательным образцом такого рода был «*Dictionnaire des fréquences* (P.: Didier, 1971). Соответствующий корпус в 70 млн словоупотреблений включал тексты художественной литературы от 1789 до 1964 г. В итоговых таблицах отражены 4 хронологических среза.

Настоящий словарь должен положить начало серии статистических словарей для разных периодов и разных жанров русского языка. Замысел возник еще в начале 1990-х гг. В те «тощие» для науки годы в Отделе машинного фонда Института русского языка им. В. В. Виноградова РАН удалось вручную ввести в компьютер заголовки нескольких газет 1991–93 гг. («Вечерняя Москва», «Известия», «Комсомольская правда», «Московская правда», «Независимая газета», «Правда», «Сегодня»). Позднее к ним были добавлены заголовки электронных корпусов, упоминаемых ниже. Общий объем корпуса заголовков составляет 1,5 млн словоупотреблений. В первом томе нашего словаря соответствующие частоты даны в правом столбце.

К началу 1997 г. ситуация начала меняться. Появилась возможность подписки на электронные издания газет, что, впрочем, потребовало продолжительной неустанной борьбы с опечатками. Сама эта работа стала возможной благодаря финансовой поддержке РФФИ (гранты 97-06-80100, 00-06-80230, 04-06-80094). Так возник корпус девяти газет 1997 года (в основном, второй половины года): «Известия» (2146 тыс. словоупотреблений), «Литературная газета» (1043), «Московский комсомолец» (1887), «Независимая газета» (3536), «Новая газета» (419), «Правда-5» (737), «Российские вести» (706), «Санкт-петербургские ведомости» (1614), «Сегодня» (2423). В этот корпус были добавлены тексты из националистических изданий 1991–93 гг. («Завтра» и т. п. — 231), из «Московских новостей» 1995 г. (201) и «Литературной газеты» 1994 г. (45). общий объем результирующего корпуса составил 14987 тыс. или круглым счетом 15 млн словоупотреблений. В настоящем словаре этот корпус условно именуется «Корпусом 1997», соответствующие частоты приводятся в левом столбце.

## II

Компакт-диски с электронными комплектами «Независимой газеты» стали появляться с 1998 г. Стало возможным появление третьего газетного корпуса — текстов «Независимой газеты» за 1996–2000 гг. общим объемом 35 млн словоупотреблений. Соответствующие данные образуют средний столбец наших таблиц. На материале этого последнего корпуса можно впервые начать анализ краткосрочной лексической динамики, что методически важно для составителей частотных словарей, кроме того, материал этот может заинтересовать историков и политологов.

Таким образом, публикуемый словарь представляет собой композицию трех частично пересекающихся частотных словарей. В левый столбец включено 10% материала среднего столбца (вторая половина 1997 г.), большая часть материала правого столбца присутствует в двух других столбцах. Суммирование данных трех столбцов, следовательно, неправомерно.

### Словник

В нашем материале представлено 811 тысяч разных графических слов. Чтобы сделать осмысленным это богатство форм, необходимо свести его к меньшему числу вокабул. В громадном большинстве случаев (но не во всех) вокабула совпадает с леммой, т. е. со словарной формой представления слова. О принципах лемматизации речь пойдет ниже, пока же остановимся на проблеме ограничения объема словника. Для печатной версии словаря вводим следующие правила:

1. Слова русского языка (за исключением имен собственных), их аффиксальные производные и сложные слова со слитным написанием включаются в словарь, начиная с частоты 13.

2. а) Энклитические частицы с дефисным написанием (*-де, -ка, -кась, -от, -с, -таки, -тко, -то*) отделяются от предшествующего слова и включаются в словарь без частотных ограничений. Также разделяются слова с конечным *-другой*.

б) Разделение слова с дефисом не проводится, если частица *-то* присоединяется к местоименному слову с начальным *к-* (*з-*) или *ч-*.

в) Сложные числительные с дефисным написанием разделяются на два слова. Также разделяются дефисные написания с названиями месяцев.

г) Не разделяются на части графические слова *вообще-то, все-таки, наконец-то, ну-ка, ну-с, ну-тка, опять-таки, потому-то, столько-то, с какой-то, так-то, такой-то, то-то, туда-то, тут-то*.

3. Имена собственные включаются в словарь, если их частота превышает 29. Этот же порог действителен для аббревиатур организаций, компаний и фирм, всевозможных мероприятий с цифровым компонентом (*Ту-130, Формула-1* и т. п.).

4. Леммы, записанные латиницей, включаются в словарь, начиная от частоты 5.

5. В словарь включены и числа, если их частота превышает 99.

### Лемматизация

Строгое разделение частей речи потребовало бы изощренной программной системы или ручного обследования миллионов контекстов. Особенно трудоемко разделение прилагательных и причастий, в несколько меньшей мере прилагательных и существительных. Далекое не все эти задачи имеют однозначное решение.

При составлении данного словаря мы следовали гибкой политике, иногда жертвуя грамматической чистотой в пользу лексической семантики. Синтетические компаративы и суперлативы выделялись в особые вокабулы, равно соотношенные с прилагательными и наречиями. Деадъективные формы на *-о* считались особыми вокабулами (в них может быть скрыт какой-то процент кратких форм среднего рода прилагательных).

Стремление к сильной семантической дифференциации заставляло иногда расщеплять традиционные словарные статьи обычных словарей. Так, разделены формы числа существительных *Бог* и *боги, выбор* и *выборы, круг* и *круги, многое* и *многие, новости* и *новость, рамка* и *рамки, сведение* и *сведения, цвет, цветок* и *цветы*. Трактуются как особая вокабула форма *образом*. Отдельными вокабулами стали *болит* (*болело*), *давай(те)*, *кажется, придется* (*пришлось*), *приходится* (*приходилось*), *разумеется, следует* (*следовало*), *сообщается, спрашивается, стоит*

(*стоило* — как модальное слово), *хватает* (*хватало*) и *хватит* (*хватило*). При таком расщеплении учитывались как семантические, так и частотные соображения. Большинство выделенных глагольных вокабул встречаются намного чаще суммы всех остальных форм соответствующего глагола.

Особыми вокабулами являются многие краткие формы прилагательных: *велик, намерен, нужен, обязан, принято, равен, согласен, таков, убежден*. Довольно часто в отдельные вокабулы выделялись потенциальные причастия (частично адъективированные или субстантивированные): *действующий, заведующий, интригующий, командующий, кричащий, курящий, любящий, манящий, мертвящий, моющий, мыслящий, мятущийся, обжигающий, облитерирующий, ободряющий, объемлющий, окружающий, освежающий, отдыхающий, отравляющий, отягчающий, ошеломляющий, палящий, подавляющий, подобающий, подрастающий, подходящий, порочащий, последующий, потрясающий, правительствующий, правящий, практикующий, предержащий, предстоящий, предшествующий, предвидущий, проходящий, приходящий, прилегающий, прилежащий, приличествующий, проезжающий, проникающий, присутствующие, профилирующий, процветающий, пьянящий, решающий, руководящий, соответствующий, страждущий, странствующий, текущий, трудящийся, тяжущийся, угрожающий, удручающий, ужасающий, укоряющий, упреждающий, успевающий, устрашающий, учащийся, фашиствующий, хрустящий, хулиганствующий, цветущий, чарующий; минувший, прошедший; обвиняемый, предполагаемый; направленный, обеспеченный, обреченный, обусловленный, основанный, построенный, предназначенный, представленный, предусмотренный, расположенный, распространенный, связанный, сокращенный, указанный.*

В какой-то мере разделение на вокабулы проводилось в рамках одной части речи исключительно по семантическим соображениям. Таковы *лев, лев* (болг.) и *Лев*; *среда* и *среда* (день недели), *статья* и *статья* (журн.).

### Словосочетания

Отличительной чертой настоящего словаря можно считать включение в него в качестве вокабул не только отдельных слов, но и словосочетаний. Нам известен только один частотный словарь, дающий сочетания слов (*collocations*). Это Allen S. e. a. *Frequency Dictionary of Present-day Swedish based on newspaper material*. Uppsala, 1970–75. Статистике словосочетаний посвящен 3-й том словаря, где приводятся все бинарные словосочетания с частотой 2 и более (включая самые тривиальные грамматические сочетания). Словарь Аллена построен на корпусе 1 млн словоупотреблений.

В таблицу 1 настоящего словаря включено несколько десятков словосочетаний, таких, как: *в качестве, в конечном счете, в конце концов, в связи с, в противном случае, в случае, в соответствии с, в состоянии, в сущности, в течение, в том числе, в угоду, в ходе, в частности, вверх тормашками, во избежание, вряд ли, до свидания, до сих пор, друг друга, за неимением, за счет, и так далее, и тому подобное, исполняющий обязанности, к сожалению, круглый стол, между тем, может быть, на плаву, на скаку, на протяжении, населенный пункт, настоящее время, по крайней мере, по поводу, по сравнению с, по существу, повестка дня, потому что, права человека, с поличным, сломя голову, стало быть, судя по, Счетная палата, табель о рангах, так как, так называемый, так что, таким образом, тем не менее, тем самым, то есть, только что, тот самый, точка зрения, Филькина грамота, что касается.*

Во втором томе настоящего словаря будут представлены все бинарные словосочетания с частотой 3 и более. Приведем несколько (лемматизированных) примеров с левым прилагательным. Числа указывают на частоту словосочетания, числа в скобках - частота прилагательного.

**кричащий** [82] заголовок 4, проблема 3, противоречие 3  
**моющий** [21] средства 11  
**мятущийся** [46] душа 9  
**мыслящий** [292] люди 27, политик 9, человек 9, лидер 5, часть 5 интеллигенция 3  
**отравляющий** [88] вещество 50, газ 5  
**ошеломляющий** [87] успех 15 впечатление 8 победа 3

Обратимся к более частым словам:

## IV

### действующий [5766]:

ныне 30, сейчас 15, сегодня 11, уже 5;  
законодательство 457, конституция 380, закон 178, соглашение 18, устав 11, документ 10,  
договор 9, УК 9, программа 8, контракт 5, редакция 5;  
система 54, модель 22, механизм 21, норма 20, порядок 18, режим 11, ставка 11, правило 10,  
схема 8, бюджет 8, санкции 4, тариф 4;  
президент 400, губернатор 231, председатель 174, глава 107, мэр 35, премьер 18,  
руководитель 13, премьер-министр 9, лидер 9, министр 9, градоначальник 5;  
власть 96, политик 49, правительство 37, депутат 30, парламент 22, кабинет 14,  
руководство 7, состав 6;  
орган 35, структура 13, круглый стол 5;  
сотрудник 15, член 12, генерал 7, офицер 6, чиновник 5;  
чемпион 10;  
храм 11, церковь 4;  
предприятие 18, газопровод 8, шахта 7, АЭС 6, банк 5, месторождение 5, выставка 5,  
объект 4, станция 4, блок 4, скважина 4;  
субъект 7;  
фактор 7;  
армия 34;  
лицо 415.

Приведенные примеры сгруппированы по смыслу и в своей совокупности семантически характеризуют данное прилагательное. Обратим внимание на сравнительно новое словоупотребление — *действующий чемпион*. Мы видим здесь также кандидата во фразеологизмы - *действующая армия* и безусловный фразеологизм - *действующее лицо*.

Прилагательное *действующий* лишено каких-либо эмоциональных коннотаций. Напротив, прилагательное (или причастие) *правящий* ярко окрашено стилистически, и правые компоненты словосочетаний это демонстрируют.

### правящий [2504]:

партия 491, коалиция 336, элита 255, режим 123, круги 119, класс 113, верхушка 60,  
альянс 41, слой 39, большинство 38, группа 20, группировка 18, блок 18, династия 15,  
бургомистр 12, олигархия 11, архиерей 10, кабинет 10, лагерь 10, ныне 10, номенклатура 9,  
семья 8, команда 7, бюрократия 3, епископ 6, курс 6, структура 6, власть 5, президент 5, клан 4,  
монарх 4, истеблишмент 3, пока 3, триумвират 3.

Исследователь-лингвист многое почерпнет из статистики, казалось бы, самых простых сочетаний. Вот два примера предложного управления:

на Украине	2504	в Украине	142
на Украину	618	в Украину	25
на Охотном	213	в Охотном	48

Мы видим, что «Независимая газета» в 1990-х годах держится традиции предложного управления словом *Украина*, но явно отказывается от традиционного московского *в Охотном ряду*.

## Хронология

В словарь включена хронологическая таблица текстов «Независимой газеты». Ниже приводятся несколько примеров различных тенденций динамики частоты:

	1996	1997	1998	1999	2000
<b>Монотонное нарастание частоты</b>					
Евросоюз/ЕС	504	676	997	974	1663
компания	2904	3656	4225	4117	4759
холдинг	37	77	229	270	509
СМИ	776	1155	1289	1506	2041
Кремль	1164	1338	1341	1636	2427
чиновник	938	1225	1298	1069	1308
номинация	194	291	314	331	341
суд	2813	2880	3384	3827	4389
(ген)прокуратура	907	986	1190	1370	1957
экстремист/-зм	283	224	480	749	889
фигурант	18	18	36	69	53
Березовский	173	458	725	790	854
<b>Монотонное падение частоты</b>					
СССР	2976	2160	1821	1712	1874
оппозиция	2736	2440	2108	1723	1668
революция	1052	927	841	763	772
джаз	160	144	43	79	68
<b>Пик частоты в 1996 г.</b>					
Ельцин	8485	5678	5302	4238	2069
консилиум	51	14	9	17	6
теннис	93	54	58	57	40
Гайдар	573	467	518	313	148
КС	499	497	339	147	192
Лебедь	3360	913	1105	414	212
Зюганов	3165	698	1032	837	763
<b>Пик частоты в 1997 г.</b>					
секвестр	2	323	53	10	14
Диана	33	164	58	40	38
празднование	152	273	211	200	245
<b>Пик частоты в 1998 г.</b>					
доллар	3692	3916	5290	4443	4752
рубль	3327	3943	5019	3377	3337
правительство	9458	10810	13916	10652	9455
отставка	1172	1356	1846	1750	1222
<b>Резкий подъем частоты в 1998 г.</b>					
олигарх	3	27	496	238	842
<b>Пик частоты в 1999 г.</b>					
МВФ	651	557	1098	1729	781
дефолт			90	230	136
Примаков	788	1115	2497	3637	614
Степашин	79	86	439	2299	297
импичмент	82	94	454	758	70
одномандатный	43	40	108	391	109
Косово	8	8	1164	2812	1068
Пушкин	723	722	773	1488	541

## VI

	Пик частоты в 2000 г.				
Путин	11	14	189	2336	10133
Касьянов	11	7	23	203	1260
Гусинский	49	66	96	118	1151
Кадыров	6	3	36	46	325
политтехнолог	1			12	63
святой	643	581	531	560	874
гимн	114	128	118	113	435
вертикаль	128	102	115	101	424
дзюдо	12	3	4	3	46
погибнуть	1146	964	1135	1242	1937
подводный	179	176	256	171	815
	Пики частоты в четные годы				
олимпиада	240	50	155	86	153
	Снижение частоты в 1997-98 гг.				
президент	18469	16641	17313	14600	18239
выборы	8553	3855	6117	8485	7263
соперник	484	222	311	320	418
патриарх	423	342	273	337	437
Чечня	5077	4314	2349	3920	4418
боевик	1442	425	335	1397	2068
солдат	711	581	551	759	735
бандит	234	277	196	440	425
	Изменения частот в связи с внутригазетными структурными и стилистическими обстоятельствами				
балет	449	680	650	761	1273
наука	2350	1936	1578	1321	1264
кино	1100	1146	661	705	841
книга	3300	1851	1886	1592	1700
архитектура	420	507	158	125	189
г-н	3805	4040	3171	1432	982
вы	7701	6741	6560	5107	4927
мой	4503	4046	3554	2828	2735
бы	15055	13341	13345	12092	12167
ни	9313	7738	8367	7938	7846
странно/ый	1108	970	835	828	761
корова	218	80	124	86	84
ну	1348	1127	919	725	687
оный	87	63	36	54	32
милый	272	209	175	117	88
совесть	351	356	258	193	205
	Освоение заимствований				
перфоманс	63	32	29	9	1
перформанс	5	9	18	20	17
рг	4	5	30	17	33
пиар			1	4	74
пиаровский			5	17	50



### Возможности внешнего и внутреннего сравнения

В настоящее время доступен только один источник для внешнего сравнения нашего корпуса: это электронный частотный словарь, составленный С. А. Шаровым на базе строящегося Национального корпуса русского языка. Корпус быстро растет, растет и частотный словарь. В нашем распоряжении есть версия начала 2006 г. с базой в 16 млн. словоупотреблений, которую можно сравнить с нашим левым столбцом. (Корпус Шарова включает преимущественно тексты художественной литературы, в том числе и старой). Вот относительные частоты некоторых слов в обоих словарях (на миллион словоупотреблений):

	Шаров	1997		Шаров	1997
а	9808	6204	в	29028	41058
весь	8605	3243	год	2042	4058
и	36359	30908	для	2250	3634
к	5740	5002	который	3289	5264
не	20690	14502	Москва	421	1604
но	5673	4333	новый	840	1933
у	5033	2926	о	3518	4896
я	15631	4001	первый	1293	1966
баба	138	13	по	5735	7796
бабушка	120	35	политический	141	702
барахло	11	1	правительство	123	1225
бежать	222	60	президент	144	1737
беззвучно	16	1	российский	150	2074
белый	493	220	Россия	359	3033
берег	240	109	страна	467	1716
береза	24	4	активный	45	115
беспокоиться	45	13	аналогичный	18	80
бесшумно	24	1	аргумент	20	60
бить	146	51	арестовать	38	89
блестеть	46	3	аспект	13	33
Бог	377	149	банк	138	674
болтать	37	4	безопасность	56	380
боль	192	59	безусловно	29	76
большой	1631	855	Белоруссия	9	187
бормотать	46	3	беспредел	3	14
босой	24	3	бизнес	43	164
ботинок	54	11	биржа	7	53
бояться	338	112	благодаря	55	137
брат	272	119	ближайший	81	277
бровь	72	6	более	618	1299
бродить, брести	71	17	большинство	112	310
бросить, бросать	344	114	бомж	10	26
брюки	57	13	борьба	134	383
бумажка	53	15	будущий, -ее	202	411
бутылка	170	59	бывший	174	492
бык	80	10	бюджет	33	551
валяться	52	8	бюджетник	1	29
варить	64	5	валюта	21	94
ваш	668	343	валютный	11	89
вдоль	100	26	ведомство	17	167
вдруг	713	154	версия	40	140
ведро	56	10	верховный	26	109
ведь	884	533	весьма	120	285

## VIII

ведьма	33	5	ветеран	18	77
вежливо	37	7	вещание	3	27
везти	61	30	взаимоотношение	9	69
веревка	51	9	визит	31	246
веселый, -о	209	51	вице-премьер	-	267
ветер	209	60	вклад	13	115
вечер	391	258	включая	30	87
взять, братъ	1003	515	владелец	42	111
видеть	1062	362	власть	279	1047
вниз	158	23	военный	239	606
вода	618	266	возвращение	44	106
воздух	276	103	возглавить, -лять	25	176
волноваться	70	18	вооруженный	41	193
вон	146	21	вопрос	511	1229
вот	2205	859	воскресенье	37	89

Общее впечатление от сравнения двух списков: В корпусе Шарова намного больше глаголов, местоимений 1 и 2 лица, слов, обозначающих действие, имеющих отношение к психологии, зрению, слуху, природе. Множество других маркеров корпуса подтверждают это общее впечатление. Не только *бежать*, но и *броситься, быстро, быстрый, вернуться, возвращаться*. Не только *бык*, но также *амбар, баран*. Наряду со словом *белый* видим также *алый, багровый, бурый*. На советское прошлое (на войну, в частности) указывают *батальон, бинокль, боец, бригадир, взвод, авоська*. Рядом со словом *Бог* находим *ад, ангел, вера, верить*. Некоторые слова корпуса пахнут XIX веком, ср. *ах, барин, бледный, борода, братец, взгляд, взглянуть, вздохнуть, возражать/-зить, воскликнуть, возле*.

В газетном корпусе 1997 г. преобладают имена, конкретные лексемы легко объясняются современной тематикой российских газет. Такой эффект объясняет слишком большим стилистическим и хронологическим расстоянием между двумя корпусами. Примером внутреннего сравнения газет могут служить стилистически резко различающиеся «Московский комсомолец» и «Независимая газета» (относительная частота на 1 млн словоупотреблений).

	МК	НГ		МК	НГ
аж	116	12	амбиция	19	43
алкоголик	24	3	бандформирование	3	20
апартаменты	20	3	большевизм	2	8
баба	32	8	брифинг	14	25
бандит	123	49	буржуазия	2	7
бардак	8	2	глобализация	1	19
блокбастер	4	1	глобальный	37	71
бойфренд	8	1	декларация	44	74
болельщик	41	11			
бомж	42	10			
босс	41	6			
бутик	14	1			
вдруг	250	100			
вырезвитель	15	1			
граффити	11	1			
девушка	292	51			

Второй том словаря покажет внутреннюю дифференциацию лексического материала по газетам, жанрам и темам. Как уже указывалось выше, в электронной версии будут представлены бинарные словосочетания.