

А. Я. Шайкевич*
(Москва)

ДИАХРОНИЯ В ЧАСТОТНОМ СЛОВАРЕ

Abstract. Diachrony in the frequency dictionary. This paper is aimed at the analysis of chronological tables in five frequency dictionaries (Russian and French). Information in those tables to a great extent depends on the length of the periods reflected therein. Short-term periods are important for content analysis and sociology, and only seldom — for linguistics. Longer periods give information relevant both to linguistics and humanities at large. The restructuring of the corpus which served for compilation of the frequency dictionary influences the chronological changes as well.

Ключевые слова: лингвостатистика; частотные словари; русский язык; русская проза

Составители частотных словарей обычно подчеркивают прикладную ценность своих трудов. Чаще всего ссылаются на педагогические задачи, особенно на важность учета частоты в преподавании иностранных языков. Впрочем, первый частотный словарь русского языка (Josselson 1953), адресованный преподавателям, опирался на корпус в 1 млн словоупотреблений (1830—1950 гг.) и представлял диахронические данные по трем периодам.

Развитие электроники облегчает создание корпусов текстов, одновременно меняются конечные цели, ради которых создаются частотные словари. 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» (Р., 1971). Соответствующий корпус в 70 млн словоупотреблений включает тексты художественной литературы от 1789 до 1964 г. В итоговых таблицах отражены четыре хронологических среза.

Цель настоящей статьи — понять, какого рода информацию могут дать более или менее продолжительные периоды, представленные в таблицах частотного словаря. Начнем с самого краткого периода в один год, показанного в таблице «Хронологические изменения частот слов в “Независимой газете”» в «Статистическом словаре языка русской газеты (1990-е годы)» (Шайкевич и др. 2008)¹.

¹ В таблицах даются относительные частоты слов: в Табл. 1, 9—13 на 1 млн словоупотреблений, в Табл. 2—8 — на 100 000 словоупотреблений.

Таблица 1

	1996	1997	1998	1999	2000
1. суд	402	411	472	547	627
Кремль	166	187	192	234	347
СМИ	111	165	183	215	292
экстремист, -изм	40	32	69	107	124
Березовский	25	65	104	113	122
холдинг	5	11	33	39	73
2. Ельцин	1212	811	757	605	296
СССР	425	309	260	244	268
Генерал	319	235	239	216	185
Революция	150	132	121	109	110
Коммунизм	76	40	39	35	26
Гениальный	30	27	20	11	14
дурак	31	19	11	11	7
интеллигент	24	24	14	10	10
диссидент	18	17	15	10	8
ДВР	17	16	13	6	2
ГКЧП	13	5	4	3	2
3. Лебедь	480	130	158	59	30
Зюганов	452	100	147	122	110
Явлинский	205	65	100	102	95
Дудаев	119	28	15	12	11
Грачев	75	16	11	14	3
4. секвестр	...	46	8	1	2
Диана	5	24	8	6	5
5. доллар	527	559	756	635	679
Рубль	475	562	717	482	475
6. Примаков	112	159	358	520	88
Косово	1	1	166	401	152
Степашин	11	12	63	327	42
МВФ	92	78	157	247	110
Пушкин	103	103	110	213	76
импичмент	10	13	66	108	10
7. Путин	2	2	27	334	1448
Касьянов	2	1	3	29	180
Гусинский	7	9	14	19	164
подводный	26	25	36	24	116
гимн	16	18	17	16	62
вертикаль	18	15	16	14	61
дзюдо	2	...	1	...	7
8. президент	2638	2376	2474	2086	2605
выборы	1222	551	874	1211	1038
соперник	69	32	44	46	60

	1996	1997	1998	1999	2000
Чечня	725	616	335	560	731
боевик	206	61	48	200	295
9. балет	64	97	93	109	182
Книга	471	265	269	217	244
вы	1100	963	937	730	704
ну	193	161	131	104	98
милый	39	30	25	19	13
10. олигарх	...	4	71	34	120
дефолт	—	-	13	33	22
11. перфоманс	9	5	4	1	...
перформанс	1	1	3	3	2
PR	1	1	4	2	5
пиар			...	1	11

Информация табл. 1 заинтересует прежде всего историков, социологов и политологов, которые отметят рост (1.1) или падение (1.2) частоты, пики частот в отдельные годы (1.3—1.7), двухвершинное распределение, связанное с ритмом президентских выборов (1.8). Слова, объединенные в группе 1.9, свидетельствуют о колебаниях частот в связи со структурными и стилистическими изменениями внутри «Независимой газеты». Лингвист обратит внимание на резкий подъем частоты в 1998 г. у слов *олигарх* и *дефолт*. Наконец, собственно лингвистический интерес представляет освоение недавних заимствований в группе 1.11.

Обратимся теперь к таблицам частот, где столбцы соответствуют десятилетиям. Примером явится частотный словарь прозы XIX в., разрабатываемый в Институте русского языка РАН (Шайкевич 2012б).

Начнем со слов, у которых хронологические колебания минимальны.

Таблица 2

	1820	1830	1840	1850	1860	1870	1880	1890
за	449	486	443	424	475	476	459	497
по	479	480	464	433	446	483	462	489
дорога	58	54	44	41	43	34	41	48
лошадь	41	43	38	48	41	33	42	37
квартира, комната, дверь, стена, пол, окно, стол	229	256	269	231	251	194	202	250

Такая стабильность вполне ожидаема у предлогов, но удивительна у следующих за ними знаменательных слов — свидетельство замечательного постоянства прозы.

Таблица 3

	1820	1830	1840	1850	1860	1870	1880	1890
1. нежели	29	20	11	5	5	3	3	1
поутру	18	8	7	6	3	3	1	1
другой день	20	14	11	14	12	8	7	9
2. вроде	1	3	5	9	9	9	12	14
вообще	8	9	19	24	24	21	26	23
из-за	11	15	14	19	22	25	30	33
точно	16	27	40	44	45	40	74	99
маленький	21	28	38	45	36	41	43	71
все-таки	6	10	13	28	41	42	35	38
все равно	5	7	10	11	16	16	17	20
тон	3	5	7	13	17	17	17	20
пальто	-	...	2	5	6	4	8	12

Движение частот, отраженное в этой таблице, более или менее совпадает с нашими ожиданиями, но едва ли будет учтено обычными словарями.

Хронологические изменения частоты становятся особенно выразительными при изучении пар (или целых групп) синонимов:

Таблица 4

	1820	1830	1840	1850	1860	1870	1880	1890
1. чтобы, бы, между, ли, же, али, перед, через	1055	1204	1257	1250	1295	1421	1257	1286
чтоб, б, меж, ль, ж, аль, пред, чрез	298	345	250	240	236	225	102	107
2. собой, мной, тобой, рукой, головой	26	61	104	133	138	116	108	146
собою, мною, тобою, рукою, головою	179	151	100	68	65	61	62	55
3. я, ты, мы, вы	4059	3527	3712	3999	3792	3290	2835	3204
мой, твой, наш, ваш	1274	981	722	637	580	576	510	499
4. прежде	55	36	41	42	45	47	32	20
раньше	2	1	2	3	5	7	15	21
5. беспрестанно	17	13	20	18	9	3	2	2
постоянно	2	2	9	12	13	10	10	11
6. довольно	66	38	52	49	42	40	25	24
достаточно	3	2	4	4	9	6	11	8
7. добрый	87	60	55	43	39	49	39	40
хороший	18	17	32	37	35	37	42	50

	1820	1830	1840	1850	1860	1870	1880	1890
8. ибо	70	21	15	8	12	17	10	8
потому что	27	52	75	94	82	74	69	60
так как	11	12	20	22	31	34	32	32
9. весьма	57	16	28	24	26	26	10	9
совершенно	34	30	56	58	57	42	28	29
чрезвычайно	5	10	15	18	10	11	4	5
очень	75	83	146	214	183	162	127	134
10. возразил, -а	16	17	15	34	13	15	7	9
отвечал, -а	132	114	85	140	69	79	29	45
ответил, -а	1	1	2	8	23	47	39	55
11. ныне	6	28	5	1	5	8	3	2
нынче	2	8	8	16	15	14	12	10
теперь	137	139	165	148	203	196	176	186
сейчас	22	15	35	47	66	57	61	72
12. надобно	55	27	32	27	13	9	4	1
надо	12	39	48	71	92	99	86	99
нужно	12	24	44	29	46	32	52	51
придется,								
пришлось	5	7	10	10	12	16	22	21
приходится, -лось	1	1	7	7	8	9	17	16
13. и	4090	3812	4115	3933	4284	4438	4441	5013
да (союз)	169	304	371	304	340	372	272	249
а	619	519	777	791	911	1037	1011	1025
14. вот	103	168	252	225	251	239	223	243
ну	56	63	98	168	197	165	141	184
просто	8	17	34	35	35	31	29	34
значит	19	14	15	32	39	38	33	44
15. ли	232	266	231	213	195	209	153	153

Таблицы 3 и 4, рассмотренные выше, так или иначе соотносены с диахронической или жанрово-стилистической окраской слов, которая могла быть замечена лингвистами. Такова, например, тенденция сокращения доли притяжательных местоимений в общей частоте местоимений 1 и 2 лица (от 24% в 1820-е гг. до 13% в 1890-е гг.)².

Некоторые тенденции обусловлены не движением языка, а движением литературы. Нижеследующие таблицы ограничены сферой прозы как таковой (без драмы), лингвистический материал выступает в ней лишь как средство, подчиненное задачам разработки темы, построения сюжета, ориентации на тот или иной стилистический регистр.

² Не связано ли это с постепенным избавлением от галлицизмов?

Таблица 5

	1820	1830	1840	1850	1860	1870	1880	1890
1. итак	27	13	11	6	5	4	4	3
между тем	43	51	36	30	31	27	21	16
наконец	83	89	129	82	82	71	50	45
признаюсь	11	9	10	8	4	5	2	1
имя	51	60	26	20	24	26	24	27
дочь, сын	130	92	67	52	57	60	58	62
мальчик, девочка	8	13	29	22	21	27	29	44
2. баба	3	5	18	12	18	13	19	20
мужик	6	8	23	26	33	21	29	24
3. о!	55	48	29	17	27	30	18	24
бедный	50	56	48	23	24	28	25	22
4. сердце	131	163	100	72	56	65	55	60
любовь	77	109	61	58	46	36	37	38
страсть	16	24	15	11	15	9	9	9
слеза	53	52	54	40	32	31	32	36
5. где-то, как-то, кто-то, что-то, почему-то	42	72	107	119	132	132	140	182

Монотонные тренды — не единственная возможность в лексической динамике литературы. Можно представить себе пик частоты слова или группы слов, приходящийся на одно или два десятилетия, или двухвершинное распределение в пределах изучаемого периода, ср.:

Таблица 6

	1820	1830	1840	1850	1860	1870	1880	1890
великий	54	81	20	14	20	37	38	26
русский	62	110	45	24	44	46	37	31
над, под, вокруг	176	242	176	149	171	162	189	226
вода, ветер, земля, берег, море, небо, солнце, огонь	187	248	184	130	111	129	196	275

В русской прозе XIX в. довольно часто наблюдается распределение с одной вершиной, приходящейся на середину столетия, ср.³:

³ Частоты *-то* даны без неопределенных местоимений типа *что-то, где-то*; частота *-с* не включает *да-с, нет-с, ну-с*.

Таблица 7

	1820	1830	1840	1850	1860	1870	1880	1890
1. гостиная	6	9	10	22	14	10	14	12
2. перебил,-а	1	4	9	13	8	8	4	5
прибавил,-а	8	20	22	41	31	31	15	18
3. впрочем	18	29	51	54	43	41	29	21
4. же	331	361	483	483	593	560	456	490
даже	72	84	145	194	189	239	155	140
5. насчет	5	9	8	14	17	20	13	12
про	15	24	35	47	59	80	46	38
6. уж	37	88	147	164	167	132	107	110
-то	46	33	154	205	240	249	176	162
-с	1	19	59	78	97	65	40	33

Создается впечатление, что русская литература, с 1840-х гг. качнувшаяся в сторону разговорности и простонародности, к 1880-м гг. пресытилась этими приемами. (Частица *-с* уходит из языка вообще.)

Общие лексические тенденции прозы XIX в. отражаются и в творчестве отдельных писателей. Характерен случай Достоевского (Шайкевич и др. 2003):

Таблица 8

	1844—49	1856—65	1866—72	1873—80
1. а	657	886	842	1050
бедный	76	49	31	24
бледный	34	22	20	13
вдруг	173	154	318	390
все равно	9	9	27	28
давеча	1	29	53	43
даже	183	371	415	376
значит	7	31	22	27
наконец	240	134	129	78
о!	30	30	71	91
сердце	205	95	66	103
слеза	101	46	27	31
2. ведь	157	248	222	201
именно	47	85	78	122
особенно	23	60	55	50
конечно	49	59	92	91
разумеется	25	82	47	37
3. едва	44	15	20	7
чуть	35	52	61	30
4. отвечать	100	115	51	30
ответить	8	11	38	49

	1844—49	1856—65	1866—72	1873—80
5. пока	4	7	24	44
покамест	22	17	11	3
6. отчего	47	22	8	8
почему	28	47	61	63
7. теперь	334	310	309	309
сейчас	35	52	97	95
8. надо	13	112	115	121
надобно	7	11	5	6
нужно	97	20	12	12
9. как будто	200	143	66	36
как бы	21	25	114	140

По большей части движение частот у Достоевского повторяет общие тенденции прозы. Исключением выступает междометие *о!*, максимально представленное в поздних романах писателя. Видно, как постепенно кристаллизуются лексические маркеры Достоевского (*вдруг, давеча, как бы*).

В новейшем частотном словаре (Ляшевская, Шаров 2009), построенном на базе национального корпуса русского языка и охватывающем период 1950—2007 гг., столбцы хронологических таблиц соответствуют двадцатилетиям, ср.:

Таблица 9

	1950—1969	1970—1989	1990—2007
1. ты, твой	6391	6466	7024
вы, ваш	6024	5230	4495
2. он, его	22413	21018	19574
она, ее	9664	9651	10749
3. человек	3237	2800	2476
мужчина	203	268	391
женщина	582	657	804
мужик	100	223	286
баба	139	177	177
4. сказать	5314	4686	3600
говорить	2595	2241	2048
думать, подумать	1988	1827	1667
спросить, спрашивать	1846	1492	1435
5. работа(ть)	1193	1055	961
идти, пойти	2833	2568	1950
стоять, встать	1315	1081	899
сидеть, сесть	1442	1330	1176
6. жена, муж	575	737	1018
7. товарищ	721	528	221
господин	151	254	246

Табл. 9 указывает как на социолингвистические процессы (9.1, 9.3, 9.7), так и на языковые приметы собственно литературных сдвигов, будь то в сфере строения прозы (9.4, 9.5) или в выборе тем и сюжетов (9.2, 9.6). В табл. 10 доминирует именно фактор тем и сюжетов.

Таблица 10

	1950—1969	1970—1989	1990—2007
1. война, бой, фронт	960	690	410
армия, батарея, взвод, военный, госпиталь, дивизия, землянка, окоп, полк, раненый, тревога, тыл, штаб	1440	820	490
боец, командир, командующий, комбат, красноармеец, лейтенант, майор, разведчик, солдат, старшина	1530	1060	520
немец(кий), фюрер, фашист	910	420	170
автомат, бомба, граната, гусеница, орудие, пулемет, пуля, пушка, снаряд, танк	880	440	340
2. лагерь, лагерный, тюрьма	350	280	140
барак, заключенный, зэк, конвой, нары, следователь	380	120	150
3. начальник, начальство	390	290	230
приказ, приказ(ыв)ать	340	250	170
бригадир, выступить, доклад, заведующий, председатель, разрешить, секретарь, указание	580	370	240
4. инженер, рабочий, чертеж	280	140	80
наука(-чный), профессор, ученый	640	500	260
геолог, изобретатель, лаборатория, музей, прибор, сыворотка, экспедиция	460	200	140
5. берег, ветер, волна, камень, куст, лес(ной), озеро, река, снег, солнце, трава	3100	2370	1710
деревня(-енский), изба, колхоз, печь, село, сено, хата	880	620	300
картошка, каша, хлеб	360	300	220

В табл. 11 представлен целый комплекс социальных, литературных и собственно лингвистических процессов.

Таблица 11

	1950—1969	1970—1989	1990—2007
1. деньги, баксы, богатый, доллар, купить, продать, рубль, сумма	670	910	1380
2. банк, бизнес, босс, клиент, офис, партнер, секретарша, товар, фирма	126	107	470
3. Бог, божий, господь, грех, молитва, монастырь, священник, церковь	470	695	865
гроб, могила, покойник, похороны(-ить), прах, ритуал	160	270	340
4. бар, бокал, бутылка, вино, водка, выпить, выпивка, закуска, застолье, кафе, коньяк, кофе, пиво, пьяный, ресторан, рюмка, шампанское	830	1120	1750
5. квартира, балкон, ванная, гостиная, лифт, подъезд, прихожая, спальня, туалет, уборная, этаж коммуналка	440	630	1200
	—	5	14 ⁴
6. вообще	240	337	456
проблема	37	70	153
ситуация	19	57	121
абсолютно, буквально, вариант, демонстрировать, естественно, коллега, конкретный,-о, контроль, минимум, нормальный, плюс, практически, реакция, специально, тип, элита	350	580	940
7. дерьмо, задница, козел, сволочь, сука, тварь, хрен	111	173	313
б..., г..., ж..., х...	12	38	89

Упомянутый выше «Dictionnaire des fréquences» дает следующее разбиение на хронологические столбцы:

⁴ Замечательный пример эволюции немаркированного (и потому почти не упоминаемого в литературе) случая в нечто социально маркируемое.

Таблица 12

	1789—1848	1849—1879	1880—1918	1919—1964
1. avoir	18453	18501	18424	18152
de, du	64445	62031	63010	63254
2. cependant	482	407	287	278
pourtant	203	312	363	402
3. manière	416	236	214	259
façon	112	212	226	265
4. ouvrage	259	97	69	75
œuvre	172	230	332	340
travail	279	319	336	350
5. face	151	275	316	249
visage	151	219	268	441
6. mère	473	726	504	479
maman	25	86	93	123
7. tu,toi	3080	3978	3701	3904
vous	6697	7128	5483	4679
8. adieu	236	188	85	57
bonheur, heureux,				
malheur(eux)	1322	837	724	657
fortune	225	149	110	62
douleur, triste	480	402	375	249
larme, pleurer	358	349	367	214
9. ciel	508	439	397	304
dieux	332	373	288	165
10. baron, comte, cour, duc,				
duchesse, empereur, mar-				
quis, noble, prince, roi	2682	1565	966	719
nom, nommer	842	688	590	478
11. beau	965	982	888	741
bon, doux	703	559	470	353
âme, cœur, sein	1983	1429	1523	1113
12. nation, peuple	825	371	330	281
13. non, pas	6701	7787	8720	9704
mais	4098	4083	4431	5169
déjà	501	482	605	708
chose	1111	1283	1352	1390
14. bureau	46	75	91	105
camarade	46	47	109	107
certain	380	380	426	696
exactement, justement	58	66	155	213
chance	36	57	69	110
geste	64	121	223	242
type	24	64	69	136

	1789—1848	1849—1879	1880—1918	1919—1964
15. comprendre, expliquer, penser, pensée, savoir conscience	2923 126	3110 137	3682 217	4309 334
16. problème réalité, réel	25 152	37 137	65 234	173 398
17. blanc, bleu, noir, rouge arbre, bois, eau, froid, jardin bras, jambe, main cheveu, lèvres, œil, yeux, tête fenêtre, maison, mur manger, table dame, femme, fille, madame mari enfant	676 1120 1173 1795 711 270 1976 179 664	1469 1511 1738 2371 1065 535 2647 253 859	1305 1476 1665 2387 1117 576 2381 210 812	966 1088 1352 1864 898 421 1537 117 613
18. argent guerre, paix anglais, français état, gouvernement liberté, libre objet, sujet	202 434 546 838 483 607	126 295 385 483 311 304	161 367 396 456 311 332	322 598 580 564 417 577

Создается впечатление, что в эволюции французской прозы (в отличие от русских таблиц) перепады частот играли меньшую роль. Большею оказывается лексическая стабильность. Ясно виден контраст двух средних периодов с предшествующим и последующим. Впрочем, выделенные здесь периоды оказываются слишком продолжительными, что сужает круг потенциальных интерпретаций. Кажется, что оптимальная периодизация для подобного рода таблиц должна включать два уровня. На нижнем (более дробном) уровне периоды составляли бы 10—15 лет⁵, на высшем уровне границы между периодами определялись бы эмпирически по максимальным лексическим контрастам между периодами низшего уровня.

Частотные словари русского языка еще далеки от того, чтобы отразить лексическую эволюцию за один или два века. Однако уже сейчас можно приступить к сравнению частотных словарей разных периодов⁶. Первый опыт в

⁵ В порядке гипотезы предложим для русской прозы такие границы между периодами нижнего уровня: 1821, 1841, 1856, 1881, 1905, 1917/18, 1931, 1940, 1956, 1990/91.

⁶ Общие принципы сравнения частотных словарей еще не разработаны.

этом роде отражен в публикации (Шайкевич 2012а), где частотный словарь русской прозы 1851—80 гг. сравнивался со словарем Ляшевской и Шарова. Обнаружено примерно две тысячи слов, чья частота в середине XIX в. статистически значимо превышала их частоту во второй половине XX в. Некоторые контрасты такого типа показаны в следующей таблице.

Таблица 13

	1851—80	1950—2007		1851—80	1950—2007
вы, ваш	9460	4900	ах	512	158
вздор	116	10	али	132	10
ерунда	2	42	весьма	241	86
возле	81	216	видать	243	76
подле	96	11	впрочем	432	231
довольно	446	159	возразить	175	66
достаточно	52	103	да (союз)	3790	1289
благодарить	169	40	вообще	227	395
бы	3909	2968	нормальный	7	103
ли	2150	1288	вариант	2	53
мать	714	554	проблема	1	116
мама	67	447	барин	261	11
наоборот	13	106	лошадь	438	130
напротив	176	36	извозчик	86	13
вверх	51	128	кучер	105	9
вниз, -у	114	255	водитель	—	73
вернуться	161	442	машина	31	602
воротиться	118	6	верста	143	26
бандит	2	59	километр	...	77
вроде	95	270	ванная	...	50
всерьез	1	36	фильм	—	75
выяснить(ся)	7	95	трусы	—	29

Следует помнить, что хронологические таблицы частотных словарей фиксируют факт изменения (или сохранения) частот. Интерпретация фактов — дело дальнейшего анализа. Конечно, есть случаи, когда объяснение лежит на поверхности. Тривиальными окажутся частоты у слов *верста* и *километр*. К чисто технологическим причинам сведется динамика частот слов *лошадь*, *извозчик*, *кучер*, *водитель*, *машина*. Слом социальной системы объяснит падение частоты слова *барин* и всей группы 10 в табл. 12. Общеευропейская экспансия абстракций станет фактором, объединяющим в одну группу слова с растущей частотой: *вроде*, *вообще*, *нормальный*, *вариант*, *проблема*, *type*, *problème*. Параллелизм социального развития у народов Европы

позволит в одном ракурсе рассматривать сокращение частоты *вы* и *vous*, появление «детских» синонимов (*мама* и *taman*) у слов *мать* и *mère*.

Менее очевидны (а иногда и загадочны) изменения частот у слов, не связанных к определенной тематике. Почему сокращается частота слов *бы* и *ли*? Чем объяснить рост частоты у *вниз* и *вверх*? Что стоит за перетасовкой частот в группах синонимов?

В любом случае для ответов на такие вопросы потребуется обращение к исходному корпусу текстов, что не всегда осуществимо в настоящее время. В более отдаленной перспективе можно ожидать прогресса и в самих частотных словарях, желательно включать в них не только слова, но и какие-то словосочетания, чаще разводить омонимию и полисемию.

Л и т е р а т у р а

- Ляшевская, Шаров 2009 — *Ляшевская О. Н., Шаров С. А.* Частотный словарь современного русского языка, М., Азбуковник, 2009.
- Шайкевич 2012а — *Шайкевич А. Я.* Лексика русской прозы XIX в. на фоне Национального корпуса русского языка // Вестник РГГУ. Сер. «Филологические науки. Языкознание / Московский лингвистический журнал». 2012. 14. 8(88). С. 110—129.
- Шайкевич 2012б — *Шайкевич А. Я.* Лексическая динамика русской прозы XIX в. (Поколения прозаиков). М., 2012.
- Шайкевич и др. 2003 — *Шайкевич А. Я., Андриющенко В. М., Ребецкая Н. А.* Статистический словарь языка Достоевского, М.: Языки славянских культур, 2003.
- Шайкевич и др. 2008 — *Шайкевич А. Я., Андриющенко В. М., Ребецкая Н. А.* Статистический словарь языка русской газеты (1990-е годы), М.: Языки славянских культур, 2008.
- Dictionnaire des fréquences. P.: Didier, 1971.
- Josselson 1953 — *Josselson H. H.* The Russian word count and frequency analysis of grammatical categories of Standard Literary Russian. Detroit: Wayne University Press, 1953.