

STUDIA PHILOLOGICA



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

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

ДИСТРИБУТИВНО-
СТАТИСТИЧЕСКИЙ АНАЛИЗ
ЯЗЫКА РУССКОЙ ПРОЗЫ
1850—1870-х гг.

Том 2



Издательский Дом ЯСК
Москва 2016

УДК 811.161.1
ББК 81.2 Рус
Ш 17



Издание подготовлено при финансовой поддержке
Российского гуманитарного научного фонда (РГНФ)
проект № 16-04-16022

Рецензенты:

д. ф. н. А. Ф. Журавлев, д. ф. н. С. А. Крылов

Шайкевич А. Я., Андриющенко В. М., Ребецкая Н. А.

Ш 17 Дистрибутивно-статистический анализ языка русской прозы 1850—1870-х гг. Т. 2. М.:
Издательский Дом ЯСК, 2016. — 848 с. — (Studia philologica.)

ISBN 978-5-9907947-7-1

В соответствии с двояким характером исследования в целом настоящий том включает две части. В части 4 дистрибутивно-статистический анализ (ДСА) проводится на материале 15 миллионов слов, непосредственно примыкающих друг к другу в текстах русской прозы 1850—1870-х гг. Тот же формальный метод прилагается к английскому корпусу (4 миллиона слов романов Диккенса). И в том, и в другом случае ДСА приводит к открытию дистрибутивных классов и некоторых грамматических конструкций. Часть 5 представляет собой частотный словарь русского корпуса, показывающий распределение слов по основным жанрам, по микрожанрам и по текстам 25 писателей.

В существенно расширенном виде том 2 представлен на компакт-диске.

УДК 811.161.1
ББК 81.2 Рус

ISBN 978-5-9907947-7-1

© А. Я. Шайкевич, В. М. Андриющенко, Н. А. Ребецкая, 2016
© Издательский Дом ЯСК, 2016

Часть 4

**ДИСТРИБУТИВНО-СТАТИСТИЧЕСКИЙ АНАЛИЗ
В МИНИМАЛЬНОМ ИНТЕРВАЛЕ**

Напомним основное направление дистрибутивно-статистического анализа (ДСА), сформулированное в т. 1 нашей монографии.

Ставится задача описания языка по большому корпусу текстов, в котором представлены цепочки букв, разделенные пробелами (= графические слова). Эта задача решается в несколько этапов, каждому из которых соответствует свой интервал текста, т. е. задаваемый способ членения корпуса. Первый этап анализа (микроинтервал) предполагает изучение комбинаторики букв в рамках графического слова.

На следующем этапе (в минимальном интервале) изучается комбинаторика графических слов внутри бинарных словосочетаний, т. е. пар слов, непосредственно предшествующих одно другому.

Семантическая информация в ДСА не используется.

4.1. РАБОТА ДСА В МИНИМАЛЬНОМ ИНТЕРВАЛЕ НА МАТЕРИАЛЕ КОРПУСА РУССКОЙ ПРОЗЫ

4.1.0. СИСТЕМАТИЗАЦИЯ МОРФОЛОГИЧЕСКИХ ДАННЫХ

Эффективной работе ДСА в минимальном интервале должна предшествовать окончательная систематизация морфологических данных, полученных на микроинтервале. Вернемся к результатам п.п. 2.2.4 и 2.2.5 (т. 1, с. 124—127), где приведен список 43 агрегированных парадигм с соответствующими наборами суффиксов. Этот список может быть пересмотрен с учетом двух важных процедур.

1. Следует учесть ограничения на комбинаторику букв, отмеченные в табл. 1.8 и 1.9 (т. 1, с. 47—48).

2. Следует допустить интерпретацию набора ассоциированных суффиксов как отражение нескольких парадигм, надстраиваемых одна над другой в результате деривации основ.

В качестве примера комбинаторных ограничений обратимся к букве К. Табл. 1.8 показывает, что справа от этой буквы неслучайно часто появляется И и, напротив, не встречается буква Ы¹. Теперь становится ясным, что парадигмы /17/ ТЯЖЕЛ-ый и /20/ ВЫСОК-ий² различаются ровно этим одним орфографическим ограничением. Две парадигмы можно объединить в одну.

Набор суффиксов парадигмы /19/ БЛАГОРОДНЕ-йший целиком входит в набор парадигмы /18/ ХОРОШ-ий. Анализ 39 слов с суффиксом *-йш-* покажет, что 35 из них присоединяются к основам все той же парадигмы /17/ (*благороднейший, богатейший, важнейший, вернейший, главнейший, глупейший, длиннейший, добрейший, добродушнейший, злейший, любезнейший, любимейший, малейший, милейший, нежнейший, новейший, огромнейший, отличнейший, первейший, покорнейший, полнейший, почтеннейший, превосходнейший, преданнейший, прекраснейший, прелестнейший, светлейший, сильнейший, скорейший, старейший, счастливейший, ужаснейший, умнейший, честнейший, чистейший*), слово *святейший* производно от *святой* (особого варианта парадигмы /17/), а три слова (*дальнейший, искреннейший, позднейший*) окажутся в составе парадигмы /18/. Сходным образом ведет себя суффикс *-айш* (*ближайший, величайший, высочайший, глубочайший, мельчайший, низжайший, строжайший, тончайший*) с тем отличием, однако, что сопровождается изменением конечной буквы производящей основы по правилу $k > ч, з > ж$ (и более редкого $з > ж$).

К парадигме /18/ ХОРОШИЙ присоединятся наборы глагольных суффиксов *-вшая, -вшие, -вший, -вших* (около двухсот примеров), *-ющие, -ющий, -ющим, -ющих* (103 примера), а в парадигме /17/ окажется набор глагольных суффиксов *-емое, -емые, -емый, -емых* (*делаемый, желаемый, занимаемый, именуемый, называемый, обвиняемый, обожаемый, ожидаемый, описываемый, подаваемый, получаемый, предлагаемый, предполагаемый, прерываемый, преследуемый, сдерживаемый, скрываемый, требуемый, уважаемый*). К этим 19 примерам присоединятся композиты (*многоуважаемый*) и слова с префиксом *не-* (*недосягаемый, неминуемый, необитаемый, неподражаемый, нескончаемый, неумолкаемый*).

Оставаясь в рамках микроинтервала, попробуем систематизировать морфологические результаты. Ограничим поле наблюдения двумя порогами частоты: пусть снизу это будет частота 5, а сверху — 5000. Отбросив самые редкие слова, мы застрахуемся от вторжения нерегулярных вариантов, многих имен собственных или простых опечаток (ср. *столпимши, столплением, столпления, столповиск, столповое, столповщины... яшшенка, яшшенки, яшшенку,*

¹ Точно такое же ограничение распространяется на буквы Г и Х.

² Номера парадигм из списка п. 2.2.4 будут даваться в косых скобках.

яшешку-глазунью, яй, яйва, яйве, яйве-то, яйле, яйлу). Отбрасывание самых частых слов будет полезным при попытках суммирования каких-то морфологических показателей, где функциональные слова могли бы статистически подавить регулярные явления, наблюдаемые в массе не самых частых слов.

После введения этих двух порогов общее число разных графических слов (первоначально 458 тысяч) уменьшилось почти вдвое. Введем еще одно ограничение. Установим цифру 3 как минимальное число суффиксов, необходимое для попадания основы в пул возможных участников регулярной парадигматики. В нашем уменьшенном вдвое поле наблюдения встретилось слово *кабы* с частотой 956, но ни потенциальная основа *каб-*, ни основа *кабы* не имеет других продолжений-суффиксов из числа парадигм, открытых на микроинтервале. Основа *кавалькад-* имеет два продолжения *-а* и *-ы*, оба они входят в парадигмы /17/ и /27/, но мы пока исключаем их из рассмотрения (слова *кавалькаде* с частотой 4, *кавалькадой* и *кавалькаду* с частотой 1 заставили бы нас выбрать парадигму /27/ ПРИЧИН-а, но эти слова не преодолели частотный порог). Слово *кавардак* с частотой 8 не имеет продолжений. Но основа *кавказ* имеет 4 продолжения: *-θ*, *-а*, *-ец*, *-ом*, три из которых заставляют отнести эту основу в парадигму /21/ ОБРАЗ, что хорошо соответствует ожиданиям человека, знающего русский язык. Наконец основа *кавказск-* имеет 6 продолжений: *-ие*, *-ий*, *-их*, *-ого*, *-ой*, *-ом*, все они указывают на главную адъективную парадигму. К этой же парадигме отнесем основу *кажд-* с 14 продолжениями: *-ая*, *-о*, *-ого*, *-ое*, *-ой*, *-ом*, *-ому*, *-ою*, *-у*, *-ую*, *-ые*, *-ый*, *-ым*, *-ых*. Тринадцать из них входит в набор парадигмы /17/, но появился и новый кандидат в суффиксы — продолжение *-у*. Если бы в рамках данной парадигмы *-у* появлялось многократно, пришлось бы расширить парадигму. Можно ожидать, что и перечень парадигм подвергнется пересмотру: объединятся парадигмы /5/ и /6/, /9/ и /10/, /22/ и /23/, /25/ и /26/; некоторые парадигмы исчезнут, появятся новые.

Теперь число основ, удовлетворяющих нашим требованиям, сократилось до 13 тысяч. Дадим эскиз будущего списка в новой нотации, оставляя в косых скобках старые названия и номера парадигм.

А	/Глагольные парадигмы/
Аа	/1/, /3/
Аб	/5/, /6/
Ас	/9/, /10/
Ад	/12/
В	/Субстантивные парадигмы/
Ва	/21/, /22/
Вб	/25/, /26/
Вс	/35/, /36/
Вд	/37-40/
Ве	/27/, /30/, /31/
Вф	/28/
С	/ Адъективные парадигмы /
Са	/17/, /20/
Сб	/18/, /19/

4.1.0.1. Адъективные парадигмы

Начнем наш обзор с самого ясного класса — С.

Парадигма Са (более 2500 основ) содержит самый большой набор суффиксов: *-а*, *-ая*, *-ее*, *-ей*, *-о*, *-ого*, *-ое*, *-ой*, *-ом*, *-ому*, *-ою*, *-ую*, *-ы*, *-ые*, *-ый*, *-ым*, *-ыми*, *-ых*. К этим восемнадцати суффиксам теперь добавился еще один *-θ* (нулевой суффикс), наблюдаемый почти у двухсот основ. Назовем самые частые формы:

белокур	17	глух	43	женат	273	ленив	40
богат	301	гол	19	жесток	54	лют	20
бодр	47	горазд	46	жив	725	мал	131
быстр	19	горд	147	задумчив	88	мертв	61
вежлив	57	готов	1690	застенчив	36	мил	207
велик	271	груб	84	здоров	505	милосерд	22
весел	331	густ	19	здрав	22	милостив	197
виноват	1489	далек	54	зелен	22	млад	50
вспыльчив	18	дик	34	знаком	459	молод	359
всяк	218	доверчив	19	красив	117	наг	25
высок	91	един	61	ласков	121	находчив	18

невелик	32	одинок	36	справедлив	59	тяжел	73
невредим	38	подл	30	стар	310	угрюм	54
недалек	23	помешан	37	строг	109	учен	24
недвижим	22	прав (?)	982 ¹	стыдлив	19	учтив	26
нездоров	139	приветлив	26	суров	40	колост	28
незнаком	41	проклят	54	счастлив	715	храбр	23
некрасив	19	пьян	567	сыт	136	худ	85
необразован	20	расчетлив	18	таков	534	цел	114
необходим	79	ревнив	33	тверд	77	честолюбив	18
неправ	45	самолюбив	60	терпелив	20	чист	104
несчастлив	108	свят	93	тих	91	чужд	33
нетерпелив	19	скуп	65	толст	40	широк	41
обязан	504	слаб	222	трезв	36	эдак	307
одинаков	18	слеп	38	туг	20	этак	1115

В тех случаях, когда при отбрасывании положительных суффиксов исход основы содержит сочетание согласных букв, перед последней буквой появляется гласная. Так возникают закамouflированные нулевые формы: *долот* 26 (долгий), *полон* 201 (полный), *смешон* 119 (смешной). Особенно характерны они при основах, кончающихся буквой -к.

близок	118	короток	36	ловок	79	резок	37
гадок	63	крепок	71	мягок	32	робок	36
дерзок	48	кроток	40	неловок	32	сладок	20
жалок	77	легок	54	низок	47	тонок	23
						узок	23

Если основа заканчивается суффиксом -н, в нулевой форме перед этим суффиксом появляется -е⁻² (на -е- заменяется и мягкий знак в последовательности -льн). В нашем материале встретилось 280 форм с -ен. Приведем важнейшие:

беден	145	добродетелен	10	нежен	72	почтителен	40
безмолвен	10	добродушен	15	неизвестен	36	прекрасен	47
безобразен	11	доволен	908	ненавистен	25	прелестен	12
беспокоен	24	должен	4821	неопытен	17	привычен	11
беспощаден	10	дружен	45	неосторожен	19	приличен	18
бессилен	20	дурен	81	неподвижен	40	причастен	11
благодарен	319	жаден	20	непонятен	10	приятен	66
благоразумен	12	желчен	12	неприятен	55	простодушен	14
благороден	79	забавен	17	неравнодушен	32	противен	116
блуден	193	заметен	30	несносен	15	равен	19
блуден	565	замечателен	38	несогласен	14	равнодушен	109
важен	67	занимателен	11	несчастен	56	радостен	14
великодушен	41	звучен	10	ничтожен	41	развязен	19
верен	124	злобен	14	нужен	681	раздражителен	29
виден	185	знатен	22	опасен	48	разумен	16
виновен	82	известен	231	опрятен	12	решителен	13
властен	43	изящен	10	опытен	16	ровен	76
внимателен	39	интересен	53	оригинален	14	свободен	184
возможен	30	искусен	14	осторожен	47	серьезен	56
волен	71	исправен	17	остроумен	13	силен	215
вреден	27	красен	42	отвратителен	17	скромен	65
выгоден	17	любезен	120	откровенен	82	скрытен	11
годен	52	любопытен	27	памятен	19	скучен	57
голоден	82	мнителен	30	печален	46	славен	21
грешен	61	мрачен	83	повинен	41	слышен	222
грозен	33	надобен	37	подобен	11	смирн	42
грустен	76	наивен	21	подозрителен	19	снисходителен	30
грязен	13	небрежен	16	покорен	20	согласен	687
деликатен	26	невинен	37	понятен	36	спокоен	279
деятелен	18	недоволен	289	послушен	16	способен	383

¹ Вопросительный знак при слове *прав* указывает на большую вероятность омонимии.

² Вышеприведенный пример *смешон* — единственное исключение.

странен	65	сходен	21	удачен	12	холоден	86
страшен	142	темен	15	удобен	16	чуден	18
строен	32	тесен	12	ужасен	46	черен	13
сумрачен	17	труден	22	умен	371	честен	99
						ясен	69

Теперь, когда набор суффиксов парадигмы Са окончательно определен, можно обратиться к анализу относительной частоты отдельных суффиксов. Материал пятидесяти частых основ¹ с общей частотой 227 680 словоупотреблений дает такую картину:

-0	,046	-о	,126	-ому	,017	-ый	,117
-а	,030	-ого	,073	-ою	,010	-ым	,058
-ая	,073	-ое	,066	-ую	,038	-ыми	,028
-ее	,015	-ой	,087	-ы	,020	-ых	,066
-ей	,004	-ом	,028	-ые	,088		

В табл. 4.1 даются абсолютные частоты сочетаний суффиксов парадигмы с некоторыми характерными основами².

Таблица 4.1

	бел-	глуп-	готов-	жив-	мил-	молод-	нов-	полн-	русск-	свят-
0	26	465	1690	725	208	359	15	201	-	93
а	87	212	707	334	205	271	25	200	6	19
ая	379	311	60	285	1105	1129	772	372	487	256
ее	40	155	2	122	80	-	9	54	-	18
ей	17	27	-	80	36	1	-	4	-	13
о	34	765	494	771	418	33	137	1606	-	152
ого	331	109	46	360	176	1303	1262	312	958	291
ое	267	146	55	350	194	234	1148	617	260	136
ой	552	148	35	624	211	4191	912	471	872	536
ом	621	31	35	43	42	124	497	484	256	49
ому	69	40	6	44	56	346	290	38	220	63
ою	89	30	11	45	31	63	109	153	33	36
ую	242	107	31	182	106	338	873	357	273	132
ы	55	146	544	196	74	178	35	142	9	17
ые	682	219	89	234	258	1344	1323	192	902	220
ый	564	349	178	8	1790	-	1712	482	1158	84
ым	304	134	78	214	147	594	646	422	398	166
ыми	376	57	22	103	36	213	269	87	189	98
ых	440	104	47	398	95	1190	967	62	962	252
	5175	3555	4130	5118	5267	11911	11001	6256	6985	2631

При полной независимости вероятностей суффиксов и основ мы ожидали бы, что слово *молодой* должно появиться $(11\ 911 \times 0,117) \approx 1393$ раза, между тем оно не встретилось ни разу; математическое ожидание формы *молодой* = 1036, реально же оно появилось 4191 раз. Униформное распределение суффиксов дало бы математическое ожидание 738 у формы *полно*, реально же ее частота составила 1606.

Для оценки неслучайности отклонений реальных частот от математического ожидания снова воспользуемся нашей основной формулой (т. 1, с. 31)

$$S = (x - m - 1) / \sqrt{m}.$$

Приняв в качестве порога $S = 2$, заменим превышения над математическим ожиданием знаком плюс. Тогда из табл. 4.1 получим табл. 4.2.

¹ В эту группу входят: *нов-*, *стар-*, *добр-*, *цел-*, *молод-*, *русск-*, *велик-*, *бел-*, *мил-*, *разн-*, *полн-*, *черн-*, *бедн-*, *знаком-*, *высок-*, *готов-*, *стран-*, *собствен-*, *умн-*, *весел-*, *красн-*, *длинн-*, *сильн-*, *особенн-*, *счастлив-*, *прост-*, *родн-*, *темн-*, *жив-*, *больн-*, *подобн-*, *прекрасн-*, *богат-*, *свят-*, *глуп-*, *чист-*, *огромн-*, *пуст-*, *честн-*, *главн-*, *красив-*, *тяжел-*, *тих-*, *пьян-*, *довольн-*, *широк-*, *важн-*, *бледн-*, *благородн-*, *несчастн-*. За пределами группы оказалось 8 самых частых основ (*котор-*, *так-*, *друг-*, *как-*, *сам-*, *перв-*, *всяк-*, *кажд-*), выходящих за границы нашего поля наблюдения.

² Напомним, что орфографические *-и* после *к-*, *г-* и *х-* заменены на *-ы*.

Таблица 4.2

	бел-	глуп-	готов-	жив-	мил-	молод-	нов-	полн-	русск-	свят-
0		+	+	+						
а		+	+	+	+					
ая		+			+	+				+
ее		+		+						
ей		+		+						
о		+		+				+		
ого						+	+		+	+
ое							+	+		
ой	+					+			+	+
ом	+						+	+	+	
ому						+	+		+	+
ою	+							+		
ую							+	+	+	
ы		+	+	+						
ые	+					+	+		+	
ый					+		+		+	
ым										
ыми	+									+
ых	+					+	+		+	+

Таблицы такого типа могут дать информацию о внутренней неоднородности данных — как об основах, так и о суффиксах. Основы *глуп-* и *жив-* имеют общие плюсы в шести строках, что говорит об их сходстве. С другой стороны, суффиксы *-0*, *-а* и *-ы* имеют общие плюсы в трех столбцах (*глуп-*, *готов-* и *жив-*), что также говорит об их дистрибутивном (а может быть, и о семантическом) сходстве.

Заменив частоты на плюсы, мы превратили количественные данные в качественные и можем воспользоваться коэффициентом связи, введенным Дж. Юлом еще в 1900 г. [Upton, Cook 2014: 435]:

$$Q = (ad - bc)/(ad + bc).$$

Применим эту формулу к случаю суффиксов *-0* и *-а* в таблице, аналогичной табл. 4.2, но охватывающей весь список 50 частых основ. Оба суффикса появляются как плюсы у 9 основ (*бледн-*, *больн-*, *богат-*, *глуп-*, *готов-*, *довольн-*, *жив-*, *счастлив-*, *умн-*). Две основы дают плюс только для суффикса *-0* (*знаком*, *пьян*), четыре основы имеют плюс только у суффикса *-а* (*велика*, *добра*, *мила*, *чиста*). Эти данные представлены в четырехклеточной таблице:

	-0	не-0	Итого	
-а		9 (a)	2 (b)	11
не-а		4 (c)	35 (d)	39
Итого	13		37	50

Отсюда $Q = (9 \times 35 - 2 \times 4) / (9 \times 35 + 2 \times 4) = 307/323 = 0,95$ — необычайно высокий показатель дистрибутивного сходства. Для пары суффиксов *-0* и *-ы* $Q = 0,99$, для *-а* и *-ы* $Q = 0,75$. Все это дает нам некоторый кластер из трех суффиксов в рамках общей адъективной парадигмы. Содержательная оценка этого кластера сразу бы определила — речь идет о кратких прилагательных; быть может, минимальный интервал обнаружит здесь что-нибудь средствами ДСА. Как бы то ни было, существование кластера есть доказанный факт. Ниже будут показаны некоторые черты внутренней структуры субстантивных парадигм, открываемые на микроинтервале.

Обратимся теперь к неоднородности основ.

Статистически значимое сокращение форм с суффиксом *-ый* (как в случаях *живый* и *святой*) или их полное отсутствие не может быть объяснено в рамках ДСА, ведь ударение почти не находит отражения в русской орфографии. Подобное поведение обнаружено у 146 основ (или 6 %), ср. *босой*, *восьмой*, *второй*, *иной*, *кривой*, *крутой*, *лихой*, *молодой*, *немой*, *плохой*, *простой*, *прямой*, *пустой*, *рябой*, *святой*, *седой*, *седьмой*, *скупой*, *сухой*, *тугой*, *тупой*, *холостой*, *шестой* и т. д. В основах с суффиксом *-ов-* подобные случаи составляют до 40 % (*боковой*, *вековой*, *верховой*, *вестовой*, *годовой*, *городовой*, *громовой*, *деловой*, *долговой*, *круговой*, *мастеровой*, *меховой*, *мировой*, *моровой*, *мостовой*, *носовой*, *передовой*, *полковой*, *половой*, *родовой*, *роковой*, *рядовой*, *становой*, *трудовой*, *часовой* и др.). В словах с суффиксом *-н-* доля их меньше 10 %, но представлены они очень широко (*головной*, *грудной*, *дверной*, *двойной*, *земной*, *крепостной*, *лесной*, *мясной*, *ночной*, *основной*, *отставной*, *переводной*, *перекладной*, *пивной*, *портной*, *потайной*, *резной*, *речной*, *ржаной*, *родной*, *ручной*, *смешной*, *стальной*, *стенной*, *степной*, *тройной*, *хмельной*, *цветной*, *цепной*, *четвертной*, *шерстяной* и т. п.). Встречаются они и среди основ

с суффиксом *-ск-* (4 %), ср. *воровской, городской, донской, людской, мастерской, мирской, морской, мужской, слободской, тверской, ямской*.

Рассмотренная парадигма *Са* включает огромное число основ с суффиксами. Самый частый суффикс *-н-* образуется по преимуществу адъективы от субстантивов (более 80 %). При этом происходит изменение конечной буквы производящей основы:

-к > -ч,
-ц > -ч,
-г > -ж,
-х > -ш.

Назовем эту схему основным чередованием согласных (она будет неоднократно повторяться у суффиксов именных парадигм). Примерами суффикса *-н-* могут служить:

брачн-	(брак)	ручн-	(рука)	бумажн-	(бумага)
вечн-	(век)	сказочн-	(сказка)	денежн-	(денег)
восточн-	(восток)	табачн-	(табак)	дорожн-	(дорога)
горячечн-	(горячка)	фабричн-	(фабрика)	дружн-	(друг)
желудочн-	(желудок)	цветочн-	(цветок)	жемчужн-	(жемчуг)
загадочн-	(загадка)	циничн-	(циник)	каторжн-	(каторга)
звучн-	(звук)	чахоточн-	(чахотка)	книжн-	(книга)
копеечн-	(копейка)	шуточн-	(шутка)	ночлежн-	(ночлег)
крошечн-	(крошка)	ярмарочн-	(ярмарка)	острожн-	(острог)
кулачн-	(кулак)			отважн-	(отвага)
лихорадочн-	(лихорадка)	больничн-	(больница)	пирожн-	(пирог)
молочн-	(молоко)	горничн-	(горница)	присяжн-	(присяга)
научн-	(наука)	единичн-	(единица)	радужн-	(радуга)
оброчн-	(оброк)	заграничн-	(граница)	снежн-	(снег)
отрывочн-	(отрывок)	конечн-	(конец)	судорожн-	(судорога)
ошибочн-	(ошибка)	мельничн-	(мельница)	тревожн-	(тревога)
политичн-	(политика)	месячн-	(месяц)	южн-	(юг)
порочн-	(порок)	пшеничн-	(пшеница)		
порядочн-	(порядок)	сердечн-	(сердце)	воздушн-	(воздух)
праздничн-	(праздник)	солнечн-	(солнце)	грешн-	(грех)
практичн-	(практика)	столичн-	(столица)	душн-	(дух)
публичн-	(публика)	уличн-	(улица)	потешн-	(потеха)
пушечн-	(пушка)			смешн-	(смех)
речн-	(река)			страшн-	(страх)
				успешн-	(успех)

Как правило, производная основа имеет меньшую частоту, чем производящая основа. Обратное соотношение в парах *мрачн-* (*мрак*), *скучн-* (*скука*), *должн-* (*долг*), *нежн-* (*нега*) заставляет предполагать ослабление семантической зависимости производной основы от производящей. В некоторых случаях основа с исходом на *-чн-* не укладывается в основную схему чередований, и тогда можно предположить существование более протяженного суффикса *-ичн-* (в парах *будничн-* — *будни*, *вторичн-* — *второй*, *симпатичн-* — *симпатия*, *энергичн-* — *энергия*). Сходная трактовка напрашивается при анализе основ *зажиточн-* и *распивочн-*.

Можно считать доказанным существование отглагольного суффикса *-тельн-*, подтверждаемого внушительным списком:

блистательн-	изумительн-	обстоятельн-	отрицательн-
взыскательн-	исключительн-	обходительн-	очаровательн-
влиятельн-	ласкательн-	обязательн-	ощутительн-
внимательн-	медлительн-	оглушительн-	повелительн-
внушительн-	мнительн-	одобрительн-	подозрительн-
возмутительн-	мстительн-	окончательн-	позволительн-
вопросительн-	мучительн-	омерзительн-	положительн-
выразительн-	назидательн-	оправдательн-	поразительн-
действительн-	наставительн-	оскорбительн-	последовательн-
заклучительн-	настоятельн-	ослепительн-	поучительн-
замечательн-	нюхательн-	основательн-	почтительн-
занимательн-	обаятельн-	осязательн-	предварительн-
заразительн-	обличительн-	отвратительн-	представительн-
значительн-	обольстительн-	отличительн-	предупредительн-

предусмотрительн-	решительн-	стеснительн-	удовлетворительн-
презрительн-	ругательн-	стремительн-	указательн-
привлекательн-	снисходительн-	томительн-	умилительн-
пригласительн-	соблазнительн-	требовательн-	уморительн-
приготовительн-	созерцательн-	трогательн-	унизительн-
признательн-	сознательн-	убедительн-	успокоительн-
продолжительн-	сомнительн-	уважительн-	утвердительн-
проницательн-	сострадательн-	увеселительн-	утешительн-
раздражительн-	спасительн-	увлекательн-	утомительн-
рассудительн-	старательн-	удивительн-	чувствительн-

Популярность суффикса *-н-* способствует его экспансии, в частности, в сферу европейских заимствований, ср. *гениальн-, колоссальн-, музыкальн-, натуральн-, национальн-, нормальн-, официальн-, патриархальн-, радикальн-, рациональн-, сентиментальн-, социальн-, танцевальн-, театральн-, центральн-*. Загадочным кажется пространство *-альн-* на русские основы: *буквальн-, погребальн-, поминальн-, прощальн-*.

Около 300 основ рассматриваемой парадигмы содержат суффикс *-ск-*, присоединяемый к субстантивным основам (суффиксация соответствует схеме В --> Са). Этот суффикс обычно вызывает основное чередование согласных (ср., однако, *петербургский, оренбургский*). Присоединение его к производящим основам на *-л-* сопровождается появлением *-ь-*, (*генеральский, дьявольский, сельский, тульский, уральский*). Напротив, всякий след «мягкости» может исчезнуть после *-н-* и *-р-*: *деревенский, конский, богатыйрский, зверский, морской, сибирский, ухарский, царский*. Перед данным суффиксом появляется *-е-* после конечных *-ч-* или *-ш-* производящей основы (независимо от причины появления этих букв): *всяческ-, греческ-, девическ-, иноческ-, купческ-, отческ-, творческ-, человеческ-, языческ-; монашеск-, юношеск-*. После *-ж-* твердое правило установить не удастся, ср. *дружеск-, княжеск-, пажеск-, супружеск-*, но *волжск-, калужск-, коллежск-, мужск-, парижск-*.

Более 60 основ кончаются на последовательность *-ическ-*, лишь одна из них (*девическ-*) целиком состоит из русских морфем; все остальные — результат адаптации европейских прилагательных (во французском на *-ique*): *гармоническ-, географическ-, героическ-, гимназическ-, готическ-, демократическ-, деспотическ-, дипломатическ-, ироническ-, истерическ-, историческ-, католическ-, классическ-, комическ-, критическ-, лирическ-, логическ-, магическ-, математическ-, меланхолическ-, механическ-, мистическ-, нервическ-, патетическ-, патриотическ-, педагогическ-, периодическ-, политическ-, поэтическ-, практическ-, психическ-, психологическ-, романическ-, романтическ-, саркастическ-, сатирическ-, симпатическ-, систематическ-, социалистическ-, статистическ-, студенческ-, сценическ-, теоретическ-, трагическ-, тропическ-, фанатическ-, характеристическ-, фактическ-, фантастическ-, физиологическ-, физическ-, фотографическ-, химическ-, эгоистическ-, экономическ-, эксцентрическ-, электрическ-, энергическ-, эстетическ-, юмористическ-, юридическ-*.

Основы на *-ск-* и *-ическ-* обладают рядом дистрибутивных ограничений — они не сочетаются с *-л-*, *-а-*, *-о-*. С другой стороны, формы с суффиксом *-и-*, по-видимому, ассоциированы не с *-л-* и *-а-*, но функционально аналогичны формам с конечным *-о* остальных основ нашей парадигмы. Так это, или не так, покажет минимальный интервал. Ср. *воровски, гармонически, детски, иронически, мастерски, плутовски, предательски, старчески, фактически, физически, философски, человечески, юношески* (всего 55 форм).

Из остальных основообразующих суффиксов упомянем:

- ист- (30 основ): *болотистый, волнистый, глинистый, голосистый, душистый, жилистый, золотистый, извилистый, каменистый, лесистый, лучистый, морщинистый, мускулистый, мясистый, неказистый, объемистый, окладистый, осанистый, отрывистый, порывистый, прерывистый, приземистый, пушистый, размашистый, речистый, росистый, серебристый, скалистый, смолистый, тенистый;*
- лив- (51 основа): *болтлив-, боязлив-, брезглив-, визглив-, говорлив-, горделив-, добычлив-, догадлив-, дождлив-, жалостлив-, заботлив-, завистлив-, затейлив-, кокетлив-, криклив-, молчалив-, назойлив-, насмешлив-, особлив-, отчетлив-, приветлив-, прихотлив-, причудлив-, прозорлив-, пуглив-, пытлив-, расчтетлив-, слезлив-, словоохотлив-, смазлив-, сметлив-, смешлив-, сонлив-, справедлив-, стыдлив-, суетлив-, счастлив-, талантлив-, терпелив-, тороплив-, тосклив-, труслив-, удушлив-, уродлив-, услужлив-, хвастлив-, хлопотлив-, шутлив-, щекотлив-;*
- чив- (16 основ, в конечном счете, глагольного происхождения): *вкрадчив-, вспльщив-, доверчив-, задумчив-, заманчив-, заносчив-, застенчив-, навязчив-, настойчив-, находчив-, обидчив-, обманчив-, разборчив-, разговорчив-, сбивчив-, уклончив-;*
- оват- (16 основ, чаще производных от адъективных же основ): *белесоват-, беловат-, глуповат-, голубоват-, грубоват-, желтоват-, зеленоват-, красноват-, плутоват-, подследоват-, продолговат-, простоват-, сероват-, сутуловат-, угловат-, щеголеват-.*

Отглагольные адъективы будут рассмотрены в связи с морфологией глагола.

Сб («Мягкая парадигма адъективов»). Данная парадигма отличается от только что рассмотренной тем, что 1) суффиксы *-ого, -ое, -ой, -ом, -ому, -ою* заменяются на *-его, -ее, -ей, -ем, -ему, -ею*; 2) суффиксы *-ие, -ий, -им, -ими, -их* появляются не только после *ш-, ж-, ч-* и *щ-* (где буква *-ы* невозможна), но и в остальных случаях; 3) после указанных согласных сохраняются орфографические *-ая* и *-ую*, в остальных же случаях вместо них появляются *-яя* и *-юю*; 4) в первоначальном наборе суффиксов нет *-о*. Проверка всех основ на возможность включения в данную парадигму добавила несколько суффиксов: *-л, -а, -е, -и*. К парадигме несомненно относится 186 основ (с исходом на *ш-* — 9 основ, с исходом на *щ-* — 8 основ, с исходом на *ж-* — 18 основ, с исходом на *ч-* — 31 основа, с исходом на *н-* — 120 основ¹).

Как и в случае парадигмы Са, подсчитаем долю каждого суффикса на основании группы 50 самых частых основ с общей частотой 61 150².

Особой парадигмой Сс следует признать набор суффиксов *-е, -его, -ей, -ем, -ему, -ею, -и, -им, -ими, -их, -ю, -я*, присоединяемых к основам с исходом на *-ь*. Теряя конечный *-ь*, эти основы соединяются с суффиксом *-ий*. Группа включает 38 основ с общей частотой 9370: *бабь-, барань-, беличь-, божь-, вдовь-, воловь-, волчь-, вражь-, девичь-, заячь-, извозчичь-, казачь-, козь-, коровь-, кошачь-, куличь-, лись-, медвежь-, мужичь-, олень-, охотничь-, павлинь-, полковничь-, помещичь-, птичь-, разбойничь-, раскольничь-, ребячь-, рыбаць-, рыбь-, собачь-, соболь-, старушечь-, стерляжь-, телячь-, треть-, человекь-, чиновничь-*. За единственным исключением (*треть-*) все эти основы производны от субстантивных основ. К основному чередованию согласных в этих случаях присоединяется чередование *д > ж* (*медвежь-, стерляжь-*) и *т > ч* (*ребячь-, телячь-*)

Роль отдельных суффиксов в адъективных парадигмах выглядит следующим образом:

	Са	Сб-ш, ж, ч, щ	Сб-н	Сс
-о	4,6%	5,1%	...	-
-а	3,0%	4,5%	0,1%	-
-ая	7,3%	6,0%	-яя 5,4%	-я 9,4%
-ую	3,8%	3,8%	-юю 5,8%	-ю 7,4%
-о	12,6%			-
		-е 1,9%	1,9%	-
-ее	1,5%			
-ей	0,4%			
-ого	7,3%	-его 10,3%	10,0%	12,2%
-ое	6,6%	-ее 9,6%	10,3%	-е 6,3%
-ой	8,7%	-ей 5,6%	20,1%	10,5%
-ом	2,8%	-ем 2,8%	4,4%	5,1%
-ому	1,7%	-ему 2,7%	2,1%	2,0%
-ою	1,0%	-ею 0,7%	0,8%	0,5%
-ы	2,0%	-и 1,5%	0,1%	-
-ые	8,8%	-ие 10,4%	10,8%	-и 8,3%
-ый	11,7%	-ий 14,8%	12,8%	27,8%
-ым	5,8%	-им 9,0%	4,5%	4,2%
-ыми	2,8%	-ими 2,9%	1,3%	1,5%
-ых	6,6%	-их 8,3%	9,2%	4,8%

Как видим, в парадигме Сб основы с исходом на *-н* не присоединяют суффиксы *-л, -а, -и*³; ту же особенность наблюдаем в парадигме Сс. Конечно, приравнивание суффиксов в адъективных парадигмах несколько опережает ход алгоритма. Минимальный интервал должен подтвердить или опровергнуть такое приравнивание.

¹ Лишь у слова *синий* конечному *-н*- основы предшествует гласная. Во всех остальных случаях здесь находим либо сочетание гласный + *н*: *ближн-, нынешн-,* либо удвоенное *-нн-*: *ранн-, осенн-*.

² В эту группу входят *городнич-, дюж-, жгуч-, летуч-, леи-, лучи-, млади-, могуч-, неуклюж-, нехорош-, ниц-, общ-, падуч-, певч-, подьяч-, похож-, приезж-, прихож-, прохож-, проч-, рабоч-, рыж-, свеж-, стряпч-, сумасшедш-, суц-, тощ-, хорош-; ближн-, верхн-, весенн-, внутренн-, дальн-, домашн-, древн-, завтрашн-, замужен-, здешн-, ихн-, крайн-, летн-, лишн-, нижн-, осенн-, передн-, последн-, презжн-, син-, соседн-*.

³ Феноменальная доля суффикса *-ей* частично объясняется формой *крайней*, входящей в трехсловное сочетание *по крайней мере*.