

Ограничения дистантной сочетаемости согласных в русском языке

Почему в русском языке практически нет слов, которые включали бы, например, комбинации согласных **б** и **п**, разделенных одним из гласных? Звуки [б] и [п] в русском языке – очень частые звуки, настолько частые, что мы вправе были бы ожидать, что в наиболее полных словарях русского языка мы найдем сотни слов, включающих указанные комбинации. Но в стотысячном словаре мы находим лишь три подходящих слова, а именно *бантист*, *биплан* и *биполярный*. Практически нет в словаре и слов, включающих комбинации **мVп** (ни с твердыми, ни с мягкими согласными). Единственные два слова – это *мопс* и *мопед* (такие же «русские», как и приведенные три слова с комбинациями **бVп**). Такого рода ограничения на дистантную сочетаемость согласных, разделенных гласными, не случайны. Как было показано в *Pozdniakov–Segeer 2007*, они практически универсальны в языках мира и предполагают (в наиболее общей формулировке) запрет на использование в структурах CVC двух согласных одного локального ряда. Это ограничение системно проявляется и в русском языке (*Поздняков 2007*). В этой статье мне бы хотелось более детально рассмотреть данные русского языка. Прежде, чем перейти к конкретным данным, изложу кратко основные выводы, полученные в двух предыдущих публикациях на эту тему.

В *Pozdniakov–Segeer 2007* были проанализированы данные по 30-ти языкам 15-ти генетических общностей, а именно: атлантические, иджо, манде, убангийские, банту, сара-бонго-багирми, австронезийские, чадские, индоевропейские, алтайские, ностратический, баскский, кечуа, камиларой, пиджины (на английской основе). Во всех без исключения рассмотренных языках было выявлено резкое статистическое ограничение на сочетаемость согласных общего места образования в структурах CVC.

В приводимом ниже примере, иллюстрирующем это универсальное свойство (1), приняты следующие условные обозначения:

- Р обозначает все лабиальные согласные, К – велярные, Т – дентальные, С – палатальные.
- Строки представляют данные по первому согласному в комбинациях CVC, а столбцы – по второму согласному.
- Минус обозначает существенное (более 20%) отрицательное отклонение от ожидаемой частоты, рассчитанной в соответствии с частотами звуков в определенных позициях¹, плюс – существенное превышение количества комбинации по отношению к «норме». Два минуса (или два плюса) обозначают наиболее существенные отклонения от ожидаемого количества комбинаций (более 30%).
- N – количество комбинаций CVC в словаре :

¹ Техника вычисления ожидаемой частоты (в дальнейшем «нормы») подробно рассмотрена в *Pozdniakov–Segeer 2007*. Здесь продемонстрируем ее на конкретном примере. Нас будет интересовать комбинация **мVв**. По словарю А.А.Зализняка выделяется 291630 комбинаций CVC. Из них 14596 включает [м] в качестве первого согласного (что составляет 5% от общей суммы). 24653 комбинации включают [в] в качестве второго согласного (8,5%). Логично предположить, что, если эти согласные не влияют друг на друга, то 8,5% от 14596-ти слов будут включать одновременно [м] в качестве первого согласного и [в] в качестве второго. То есть $291630 * 0,05 * 0,085 = 1239$ комбинаций. В словаре же их оказывается 421, то есть в три раза меньше, чем мы бы ожидали. По используемой нами формуле, отклонение фактической частоты от ожидаемой составит $(421 - 1239) / 1239 = - 0,66$, или – 66%. Такое отрицательное отклонение безусловно относится к разряду значимых и обозначается на наших схемах двумя минусами.

(1) Фула (N = 672)				
	Р	К	Т	С
Р	---	---	+	
К	---	---		++
Т	++	++	-	-
С	+	++		---

*Индоевропейский (N=3085)				
	Р	К	Т	С
Р	---		+	
К		---	+	
Т	++	++	---	
С	+	+		-

Монгольский (N = 66407)				
	Р	К	Т	С
Р	---	-	+	++
К		---	++	
Т		++	---	
С	+			-

В (1) приводятся результаты подсчетов по трем очень разным языкам. Принципиально различаются и характер анализируемой лексики. Для подсчетов по фула использовался словарь корневых морфем, в котором только 672 последовательности CVC; количество комбинаций в словаре лексики монгольского языка – в 100 раз больше; в случае же с индоевропейским мы вообще имеем дело с реконструированными корнями, а не с реально зафиксированной лексикой. Тем не менее, распределение отклонений поразительно совпадает. Во все трех языках прослеживается недвусмысленное ограничение на сочетание согласных одного места образования (диагональ на схеме, выделенная серым цветом).

Оказалось также, что ограничения распространяются не только на согласные одного места образования, но и системно затрагивают некоторые комбинации согласных разных рядов, а именно: в большинстве языков мира наблюдаются ограничения на сочетания лабиальных и велярных, с одной стороны, а также дентальных и палатальных, с другой². Таким образом, наиболее общая закономерность состоит в следующем: «периферийные» («низкие») согласные, с одной стороны, и «центральные» («высокие») согласные, с другой, в комбинациях CVC сочетаются ограниченно, а иногда и не сочетаются вовсе.

Еще один важный вывод из наших статистических наблюдений состоит в том, что языки мира системно показывают не только одинаковые «плохие» клетки в комбинациях CVC, но и универсальные «хорошие» комбинации. Оказывается, что существует практически универсальная иерархия не только запретов, но и предпочтений, то есть бывают слова, которые фонетически «объективно» хорошие, бывают слова менее хорошие, бывают слова плохие, а бывают совсем плохие. Прочитирую (Pozdniakov–Seegerer 2007) итоговую таблицу по 15-ти группам языков (2).

(2)	++	+	-	---	=норм.нет		Итого
К-К			1	14			-15
Р-Р			1	14			-15
Т-Т			6	4	5		-10
С-С		1	3	4	3	4	-6
Р-К	1		4	2	8		-5
Т-С		2	5	2	4	2	-5
С-Т		1	6		8		-5
К-Р	1		3	1	9	1	-3
Итого	6		70		37	6	
Р-С	4	2	1		8		5
К-С	2	3			8	2	5
С-К	5	4	1		3	2	8
С-Р	2	6			5	2	8
К-Т	1	7			7		8
Р-Т	3	8			4		11
Т-Р	6	6			1	2	12
Т-К	10	3			2		13
Итого	72		2		38	8	

² Это хорошо видно в (1) на примере фула

В во всех 15-ти группах языков выявляется запрет на сочетания велярных с велярными, а также лабиальных – с лабиальными. Напротив, резко повышенной частотой в 13-ти группах из 15-ти характеризуются комбинации дентального и велярного – это «хорошие» слова. Английские слова *dog* ‘собака’ и *cat* ‘кошка’ – «очень хорошие» *bug* ‘клоп’ – «нехорошее слово», а *toad* ‘жаба’ – «совсем плохое слово» (интерпретацию выявленной иерархии см. в *Pozdniakov–Segerer 2007*).

Естественно, в каждом конкретном языке указанные общие тенденции проявляются по-разному. Проиллюстрирую интересные аспекты проблемы на материале современного русского языка³.

В качестве материала для статистических подсчетов использовалась, в частности, электронная версия Грамматического словаря А.А.Зализняка, полученная нам от С.А.Старостина и включающая 97328 слов. Эти слова, как отмечалось выше, дают 291630 комбинаций CVC (в среднем 3 интересующих нас комбинации в одном слове)⁴. Приведем фрагмент общей таблицы, в которой указано фактическое число конкретных комбинаций, а также указано, сколько комбинаций мы бы ожидали встретить в соответствии с частотами согласных в позиции C1 и C2 (3)

(3)	Фактическое количество					Ожидаемое количество					Отклонение (в %)				
	п	б	ф	в	м	п	б	ф	в	м	п	б	ф	в	м
п	349	205	45	826	425	606	430	137	1661	664	-41	-53	-67	-50	-37
Б	64	157	32	480	127	271	192	61	744	297	-76	-21	-49	-36	-58
Ф	6	113	15	56	56	99	71	22	273	109	-94	+55	-35	-80	-50
В	397	236	14	1238	275	949	674	214	2603	1041	-58	-66	-94	-52	-74
м	96	95	65	421	135	452	321	102	1239	496	-78	-71	-37	-66	-73
СУМ					5928					13729					-57%

Приведенные в таблице (3) цифры поразительны. Наблюдается совершенно очевидная и системная тенденция к ограничению комбинаций CVC с участием двух губных. – очевидная, потому что таких комбинаций почти на 60% меньше, чем мы бы ожидали; системная, потому что огромные отрицательные отклонения отменены практически в каждой клетке таблицы. В комбинации **фVп** отрицательное отклонение составляет –94% (6 слов вместо ожидаемых 99-ти. При этом 5 из них имеют структуру CVVC, а не CVC: *эфюп*, *~-ка*, *~-ский*; *орфоэтия*, *орфоэтический*. Если бы мы принимали во внимание только слова структуры CVC, то в нашем списке оказалось бы единственное слово *амфюпод* (имеющее, к тому же, морфемную границу). То есть одно слово вместо сотни! Заметим попутно, что ни одного слова структуры **fvр** нет и во французском языке, несмотря на то, что частота [f] во французском существенно выше, чем в русском. В английском языке имеется единственно слово *фор* ‘глупец’ с яркой экспрессивной окраской⁵. Итак, в трех «больших» языках мы имеем дело практически с полным запретом определенной комбинации, и, насколько мне известно, этот запрет никогда не привлекал внимания лингвистов.

Отметим, что словарь А.А.Зализняка, несмотря на его полноту, не является оптимальным источником для исследования интересующей нас проблемы. И, прежде всего, по двум основным причинам:

³ Предварительные результаты анализа были опубликованы в *Поздняков 2007*. Публикация доступна на <http://pozdniakov.free.fr>

⁴ Точнее, эти комбинации характеризуют две структуры: CVC и CVVC. Так, к комбинации **бVл** были отнесены как *бал*, так и *баул*. Напротив, *мак* и *маяк* были отнесены к разным комбинациям – **мVк** и **мVй** / **йVк** соответственно. Слова структуры CVVC, как сейчас ясно, следовало бы исключить из рассмотрения.

⁵ Со второй половины XVII века также ‘денди; сноб’.

1. Если в внутри сочетания CVC проходит морфемная граница, то такие сочетания допускаются. Как отмечалось в начале статьи, в русском языке очень мало комбинаций структуры **бVп**. Напротив, комбинаций структуры **пVб** в словаре оказывается достаточно много. Однако практически все они представляют собой сочетания префикса **по-** с корнем (*побег, побывать, по-братски* и т.п.). Если же исключить комбинации с префиксом **по-**, то в словаре останется лишь *публика, наблизити и наб* (столь же «исконно русские», как *мопс, мопед, баптист, биплан и биполярность*⁶). Это означает, для нашей целей больше подходит словарь корней русского языка, в котором фактор морфемной границы не проявляется.
2. Если от какого-либо корня образовано очень много слов, то это, конечно же, влияет на статистические результаты. Так, в русском языке повышенная частота комбинации **рVд** в значительной степени определяется тем, что корень **род** имеет очень высокую продуктивность – от него образовано как минимум 110 слов (*Кузнецова 1986*). Единственное положительное отклонение в таблице 3 отмечено для комбинации **фVб**. Легко убедиться, что оно объясняется очень большим количеством в словаре различных *фобий* и *фабрик*.

К этому обстоятельству мы еще вернемся, а пока рассмотрим распределение интересующих нас комбинаций в русском языке (4а, 4б), полученное по материалам морфемного словаря (*Кузнецова 1986*), в котором отсутствуют указанные «помехи»⁷.

(4а) Русский язык (5 локальных рядов)

	Р	К	Т	С	С
Р	--		+		
К		--	+	-	-
Т	++	++	--	+	+
С	+	-	+	--	--
С				--	-

(4б) Русский язык (4 локальных ряда)

	Р	К	Т	С
Р	--		+	
К		--	+	-
Т	++	++	--	+
С				--

Комментарий: В (4а) согласные сгруппированы по 5-ти рядам: 1) Р – **п, б, ф, в, м**; 2) К – **к, г, х**; 3) Т – **т, д, л, р, н**; 4) С – **с, з, ц**; 5) С – **ч, ж, ш, щ, й**. В (4б) объединены данные по рядам S и C.

Представленные в таблице данные показывают системный характер отклонений. В корнях русского языка в комбинациях структуры CVC резко ограничиваются сочетания согласных одного места образования. Ограничиваются и сочетания между препалатальными (**с, з, ц**) и палатальными (**ч, ж, ш, щ, й**). Несколько меньше ожидаемой нормы встречаются сочетания, в которых первый согласный велярный, а второй – (пре)палатальный. Наиболее предпочтительными комбинациями (знак «+ +») являются комбинации, в которых первый согласный «центральный» – Т (**т, д, л, р, н**), а второй – «периферийный» – лабиальный Р (**п, б, ф, в, м**) или велярный К (**к, г, х**): **TVР, TVК**. Следующие по предпочтительности комбинации также предполагают сочетание периферийных и центральных согласных, но в них вектор направлен от периферии к центру: **PVT, KVT** (знак «+»). Комбинации между периферийными согласными разного места образования (**PVK, KVP**) не ограничиваются, но и не относятся к предпочтительным (отсутствие значимых отклонений). Все эти тенденции практически полностью отвечают универсальной иерархии «запретов» и «разрешений» в языках мира (за исключением нетипичного отрицательного отклонения в комбинации **KVC**, а также положительного отклонения в **TVC**).

Отметим, что приведенная таблица (4), позволяя выявить наиболее общие закономерности, скрывает интересные нюансы, которые характеризуют каждую из ячеек. Так, например, в целом

⁶ Заметим, что в последних двух словах очевидна морфемная граница.

⁷ В словаре А.И.Кузнецовой 2936 корней, включающих 3750 комбинаций CVC.

велярные практически не сочетаются с велярными в комбинациях CVC, но это неверно для комбинаций **xVx**. В словаре А.А.Зализняка 66 таких комбинаций, а их должно быть только 35, то есть почти в 2 раза меньше. В *Поздняков 2007* было показано, что слова с этой комбинацией а) ярко экспрессивны б) часто имеют неясную этимологию. Об экспрессивности русской фонемы /x/ написано очень много. Слова типа *хохотать, хихикать, хрустеть, хранить*, многократно рассматривались в контексте дискуссии об источниках происхождения [x] в словах, в которых этот звук не выводится из индоевропейского *s. Часто слова типа *соха* рассматриваются как аргумент в пользу реконструкции глухих придыхательных в индоевропейском ([x] выводится в таких словах из *k^h). Сбои в других словах с неясным [x] обычно объясняются экспрессивным характером этого звука в русском языке. Однако, по нашим статистическим данным, повышенной частотностью характеризуется в словаре не сам звук, а именно комбинация **xVx**.

Приведу еще один интересный факт. Как было отмечено, сочетания дентальных в комбинациях CVC ограничено, как и сочетания всех других согласных общего места образования. Однако это касается только двух дентальных смычных (**т/дVт/д**), либо двух дентальных спирантов (**р/л/нVр/л/н**), но не сочетаний смычных со спирантами, которые не только допускаются, но и имеют повышенную частоту в словаре. Слова *прореха, стрелять, нанотехнология, аврал, баррель, мораль, доллар, надутый, дотуда, велярный* – плохие, а *родня, ладный, подарок, Виноградов, лето, динозавр, доля, дентальный* – хорошие. В морфемном словаре лишь 17 корней, включающих комбинацию **рVл** (например, *мораль*), при том, что мы бы ожидали найти 64 таких комбинации. В результате, двойной повтор «запретного» сочетания в *без руля и без ветрил* придает лермонтовской строке дополнительную экспрессивную окраску, что возможно и способствовало формированию фразеологизма, закрепившегося в языке⁸. В данном случае, мы имеем дело с ограничением сочетаемости не по месту, а по способу образования согласных. И здесь статистические данные дает неожиданные результаты. Рассмотрим один из таких результатов подробнее. Он касается распределения глухих и звонких согласных в комбинациях CVC.

Мы бы ожидали встретить 6 корней **гVп**, а в словаре есть единственный корень **гипс**. Естественно, там могли бы оказаться и гепард с гиппопотамом, и гипноз с гепатитом, и гипюр с гопаком (а также гоппер-фидер с флогопитом), но это был бы уже большой словарь, в котором, как отмечалось, почти 300000 комбинаций CVC, а не 3750. Соответственно, другим было бы и ожидаемое количество слов с данной комбинацией. В лексическом словаре 165 комбинаций **гVп**, среди которых много слов вроде *благопристойный, богопротивный, бумагопрядильня, книгопечатник, многопалубный, налогоплательщик, снегопад*, то есть сочетаний с морфемной границей. Тем не менее, и в большом словаре мы ожидали бы найти существенно больше слов с этой комбинацией (226), чем имеется. В морфемный же словарь единственный корень **гипс** попал только потому, что для него выделяется 4 производных слова (нижний порог, принятый в *Кузнецова 1986 : 16* для включения заимствований в морфемный словарь). Итак, мы имеем дело с фонотактическим ограничением, причем оно не имеет отношения к описанной выше тенденции избегать согласных одного ряда в комбинациях CVC – перед нами сочетание велярного и губного⁹. Можно ли установить правило, которое описывало бы данное ограничение?

Выскажу осторожную гипотезу и приведу подтверждающие ее аргументы. Статистические данные позволяют предположить, что, наряду с ограничением сочетаний согласных одного ряда, в современном русском языке, в комбинациях CVC, действует ограничение на сочетание согласных, противопоставленных по звонкости / глухости.

⁸ Примечательно, что О.Мандельштам, перефразируя Лермонтова, сохраняет двойную комбинацию **рVл** строки оригинала: «Научи меня, ласточка хилая, / Разучившаяся летать, / Как мне с этой воздушной могилой / Без **руля** и **крыла** совладать».

⁹ Было бы соблазнительно объяснить этот пример универсальной тенденцией избегать сочетаний периферийных согласных в рамках суперклассов (см. выше), но в русском языке для этого нет достаточных оснований.

В рамках каждого локального ряда сгруппируем согласные по признаку «глухие / звонкие» и получим суммарные характеристики¹⁰. Рассмотрим сочетаемость всех выделенных классов согласных (в комбинациях SVC), фиксируя их фактическое количество в словаре, ожидаемое количество и отрицательные отклонения, свидетельствующие об ограниченности той или иной комбинации. Все существенные отрицательные отклонения для сочетаний согласных противопоставленных по глухости / звонкости, приведены в (5).

(5) Один ряд	БВМ-ПФ	ПФ-БВМ	Г-КХ	КХ-Г	ДЛН-ТР	ТР-ДЛН	З-ЦС	ЦС-З				
Фактическое N	6	8	3	2	86	174	1	4				
Ожидаемое N	34	31	13	18	158	220	5	8				
Разные ряды	Г-ПФ	З-ПФ	З-ТР	ЖЙ-ПФ	ПФ-ЖЙ	ЦС-Г	Г-ЧШЩ	З-ЧШЩ	ЧШЩ-З	ЖЙ-КХ	КХ-ЖЙ	
Фактическое N	1	1	11	0	9	5	2	1	1	9	10	
Ожидаемое N	6	3	16	7	12	9	6	3	9	15	18	

Верхняя часть таблицы (5) обобщает ограничения на сочетаемость глухих и звонких согласных в рамках одного ряда. Нижняя часть таблицы фиксирует ограничения в рамках разных локальных рядов. Верхняя часть таблицы отражает рассмотренную выше универсальную тенденцию («серая диагональ» на приведенных выше схемах)¹¹. Но нижняя часть таблицы «длиннее», чем верхняя, а ни одно из зафиксированных в ней ограничений не объясняется рядностью согласных. А между тем, каждое из отклонений в нижней части таблицы существенно. Так, например, в морфемном словаре нет ни одного корня, в котором сочетался бы один из звонких палатальных (Ж/Й) с одним из глухих губных (П/Ф)¹², в то время, как мы бы ожидали обнаружить 7 таких корней. Прежде, чем интерпретировать эти данные, рассмотрим аналогичную таблицу для корней, в которых согласные не противопоставлены по глухости / звонкости (6).

(6) Один ряд	ПФ-ПФ	БВМ-БВМ	КХ-КХ	ТР-ТР	ДЛН-ДЛН
Фактическое N	8	16	14	119	136
Ожидаемое N	13	84	41	202	172
Разные ряды	БВМ-Г	ЦС-ЧШЩ	ЧШЩ-ЦС		
Фактическое N	24	3	11		
Ожидаемое N	31	10	19		

Верхняя часть таблицы (6) интерпретируются в рамках выделенной тенденции не сочетать согласные одного ряда. Однако легко заметить, что отрицательных отклонений здесь существенно меньше, чем в таблице (5). Нижняя же часть таблицы отличается от (5) особенно резко. При 11-ти ограничениях в (5), в (6) мы видим только 3 ограничения. В (6) не только верхняя часть «короче», чем в (5), но и нижняя часть короче, чем верхняя. В (5) выделено 19 ограничений, а (6) – только 8.

Приведенные данные можно интерпретировать следующим образом. В русском языке, как и в большинстве языков мира, действует запрет на сочетаемость согласных одного ряда в комбинациях SVC. Наряду с этим, существует ограничение, препятствующее в этой комбинации сочетанию согласных, противопоставленных по глухости / звонкости. Эти две тенденции

¹⁰ Будем ориентироваться на графическую форму слов. Слово *боб* будет интерпретироваться как пример сочетания двух звонких.

¹¹ Отметим, что здесь ограничение сочетаемости палатальных не является существенным и соответственно не отмечено в таблице.

¹² Конечно же, носитель русского языка сразу вспомнит колоритное слово, включающее комбинацию Ж-П, но в оно в словарь Кузнецовой не включено. Замечу, что фонетическая экспрессивность этого слова, возможно, и определяется нелегитимностью сочетания данных согласных.

позволяют объяснить все ограничения, за исключением трех последних. Заметим, что два из них предполагают, среди прочих, комбинации **cVш** и **шVc**, универсальный запрет на которые хорошо известен и доказывается огромным количеством скороговорок (то есть труднопроизносимых звуковых сочетаний), построенных на основе комбинаций [с] и [ш] в самых разных языках. Об этом запрете хорошо знает и русский Саша на шоссе (с сушкой) и французский Sacha (с Наташей и кошкой): «Sachez, mon cher Sasha, que Natasha n'attache pas son chat!» (подробнее см. *Поздняков 2007*).

Итак, статистические данные приводят к следующему выводу: в современном русском языке, наряду с ограничениями дистантной сочетаемости согласных одного локального ряда, в структурах CVC ограничиваются комбинации согласных, противопоставленных по звонкости.

Если для языка постулируется некое фонотактическое правило, то почему оно проявляется не всегда? Почему мы вынуждены формулировать такие правила как статистические? Приведем лишь некоторые факторы, которые влияют на статистические результаты.

- ♦ Как отмечалось, на статистическое распределение влияет «мощность» словообразовательного гнезда. Так, в русском языке существуют лексические корни, от которых образованы сотни производных слов. Каждое из них включает одну и ту же комбинацию CVC, представленную в корне, и соответственно, «вес» такой комбинации резко увеличивается. В приложении к морфемному словарю (*Кузнецова 1986: 1122*) приводится список наиболее продуктивных корней русского языка (от 50-ти слов и выше). В этом списке 222 корня, от которых образовано более 19000 слов в словаре, использованном Кузнецовой (почти половина ее словаря). Естественно, многие корни включают комбинацию CVC. К числу корней этой структуры, от которых образовано более 150-ти слов, относятся по данным источника: **ста(й)** – 370 слов, **де(й)** – 250, **прав** – 190, **кат**, **рез**, **вод** – 180, **ход**, **би(й)**, **вал**, **зна(й)** – 170, **каз**, **пис**, **клад** – 150. Очевидно, что, если мы будем сравнивать, например, комбинации **pVb** и **mVп**, но на наши результаты не может существенно повлиять продуктивность корня **прав** (190 слов) и непродуктивность корня **мопс** (0 слов). Любопытно, что комбинации CVC в этом ограниченном списке наиболее продуктивных корней в основных чертах проявляют практически все основные тенденции, отмеченные выше для русского языка в целом. Если ограничиться только высокопродуктивными корнями¹³, распределение комбинаций по локальным рядам принимает следующую конфигурацию (7).

(7)	P	K	T	S	C
P	--	--	++		
K	++	--	++	--	--
T	++	++	--		
S		++	+	-	
C	++			++	--

Эти корни показывают, в частности, что комбинации с участием согласных одного ряда крайне затруднены. Так, в соответствии с частотами согласных в списке высокочастотных корней, ожидали бы, что в нем обнаружится минимум 14 корней CVC с двумя лабиальными согласными. В реальности же, у нас есть корень **бав**, от которого образовано, по данным А.И.Кузнецовой, 80 слов, и корень **мов** (в *омовение*), являющийся редким алломорфом корня **мы / мо** (*мыть* и производные). И это все! Единственный высокопродуктивный корень с двумя велярными согласными – это **скак**, восходящий к индоевропейским экспрессивным словам на **ск-**, связанным с концептом «прыгать». Комбинаций с двумя дентальными (**т**, **д**, **н**,

¹³ Корни с продуктивностью 50 слов и выше, включая их алломорфы, содержащие комбинации CVC, что дает список, в котором 334 интересующих нас комбинации.

р, л) должно быть 38 из 334 анализируемых корней. Их оказывается в два раза меньше – только 20, причем большинство из них относятся к особому типу, получившему в русском языке широкое распространение, а именно **CCVC / CVCC**, где одним из согласных в **CC** является плавный – **р** или **л**: **гляд, клад, клон, плет, след; гран, крут, стран, стрел; дерг, держ.** В два раза реже ожидаемого представлены и комбинации с двумя палатальными (только 4 корня, при том, что в соответствии с фонемными частотами их должно быть 10: **жеч** – алломорф **жиг**; **чеш** – алломорф **чес**; **чащ и чищ** – алломорфы **част** и **чист** соответственно. Легко заметить, что появление всех четырех «корней» вызвано чисто морфонологическими причинами). Итак, продуктивные корни в интересующем нас плане не отличаются от менее продуктивных, и, как оказалось, достаточно предельно малого списка в 300 корней, чтобы продемонстрировать общую тенденцию – избегать комбинации **CVC** с согласными одного места образования. Тем не менее, оптимальным для включения данных русского языка в общую типологию является, конечно же, полный словарь корней (см. 4а, 4б).

- ♦ Во многих языках, как и в русском, оказывается важным фактор морфологической границы. Если она проходит внутри сочетания **CVC**, то комбинация согласных одного ряда допускается: *по-бежать, до-тянуть, со-звучный*. В русском языке этот фактор, конечно же, принципиально влияет на статистические результаты. К тому же, в словаре много слов с «этимологической» морфемной границей, которая уже не воспринимается носителями, например, в слове *когда* (< ***gьda** < ***goda**).
- ♦ В некоторых языках при исследовании сочетаемости согласных в комбинациях **CVC** необходимо учитывать дополнительные факторы, в частности, фактор долготы гласных. В этом плане интересные результаты дает язык волоф (8)

(8a)

C ₁ VC ₂	P	K	T	C
P	---	-	++	+
K		---	++	
T	+	++	-	
C	++	+	-	---

(8б)

C ₁ VVC ₂	P	K	T	C
P	---			
K	-	---		+
T	+		---	
C		+		---

В комбинациях с долгими гласными несовместимость согласных одного ряда абсолютно не проявляется (8б). Если же ориентироваться на комбинации только с краткими гласными, то интересующая нас универсальная тенденция проявляется более рельефно, чем в общем словаре (*Pozdniakov–Seegerer 2007*).

- ♦ Во многих языках указанное ограничение не затрагивает комбинации с одинаковыми согласными. Так, во французском языке, по подсчетам Г.Сежерера¹⁴, внутри группы губных согласных, наблюдается следующее распределение (9):

(9)	p	b	f	v
p	++	--		-
b	--	+	++	
f	--	++	---	-
v			--	++

Как показывают данные таблицы (9) во французском сочетании **pVp, bVb** и **vVv** не только допустимы, но и характеризуются повышенной частотой. К этому случаю примыкает использование техники редупликации, при которой, естественно, повышается процент комбинаций с одинаковыми согласными.

¹⁴ G. Seegerer. Неопубликованные данные, полученные в результате обработки 148266 комбинаций **CVC** в фонетической версии французского словаря.

В целом, по нашим данным, наиболее резко ограничиваются дистантные комбинации согласных одного ряда, образующие минимальные пары, такие, как **bVp, fVp, kVg, rVI** и т.п. Между членами таких пар возникает определенное «напряжение», которое часто приводит в диахронии к ассимиляции (например, ***bVp > bVb**) или же к диссимиляции (например, ***bVp > bVf**). Несмотря на то, что есть и обратные примеры (хрестоматийное изменение ***arbor > arbol** в испанском), диахронические изменения, которые мы наблюдаем, в частности, в просторечных формах ***лаборатория** и ***Фильянский вокзал**, встречаются чаще. Интересно рассмотреть это явление в плане противопоставления синтагматики и парадигматики. Насколько языки не любят «опасно близкие» согласные в синтагме (а, по нашим данным – подчеркнем еще раз – речь идет об универсальном явлении), настолько минимальные пары активно используются в парадигмах. Точнее, оптимальным условием существования минимальной пары является контекст, в котором имеется два разных знака (в парадигме – *тело / дело, коготь / ноготь, какой / такой, куда / туда* или в синтагме – *гоголь-моголь, жареный-пареный, на траве дрова, жили-были, калинка-малинка, такой-сякой, туда-сюда*), но не в пределах одного знака (такие слова как *кегли* в большинстве языков являются редкими).

Существует и множество других факторов, которые могут ограничивать действие интересующего нас универсального фонотактического правила. В заключение рассмотрим еще один такой фактор, который, возможно, представляет наибольший теоретический интерес. Рассмотренное в этой статье ограничение не связано с семантикой – это чисто фонетическое ограничение: самые разные языки, судя по нашим данным, избегают сочетаний согласных одного локального ряда в комбинациях CVC. Но именно там, где существует фонетический запрет, складывается оптимальный контекст для формирования экспрессивной лексики. Нестандартная звуковая форма слова – идеальная среда для появления иконических знаков. Именно потому, что язык в целом стремится избегать сочетаний согласных одного места образования, нарушение этого принципа модально окрашивается и часто приводит к появлению идеофонов. Так, в русском языке в три раза чаще, чем мы бы ожидали, встречается сочетание **шVш**. При этом большинство таких слов характеризуются двумя свойствами: а) они экспрессивны, б) их этимология неясна. Приведу выборку из Этимологического словаря Ф.Фасмера: *шашал / шашел* ‘моль’ – «**неясно**», *шешень* ‘рисовые лепешки’ – «**неясно**», *шишабарник* ‘растение (sp.)’ – «**неясно**», *шишать / шишить / шишкать / шишать* ‘копаться, возиться, мешкать’ – «**неясно**», *шушера* ‘хлам, старье’ / *шушульки* ‘старое тряпье’ – «**неясно**», *шушулькаться / шушкаться* ‘мешкать’ – «**неясно**», *шиши* ‘разбойник, бродяга’, ‘островерхая кладь, копна сена’, ‘шишка’, ‘кукиш’, диал. ‘черт’ – «**ЭТИМОЛОГИЯ ЗАТРУДНИТЕЛЬНА**», *шашмура / шамшура* ‘головной убор замужней женщины’ «**темное слово**», *шушун* – «**по-видимому, иноязычное**», *шашни / шашень* ‘осторожный шаг назад’ / *шашал* ‘дрянной человек’ – «**недостаточно сближение с хахаль**», *шишимола* ‘большая шишка’ – «**образование представляется необычным**» (Фасмер 1950–1958, Т.4 : 416-494). Нетрудно заметить, что в значении большинства этих слов содержится ярко выраженный модальный компонент (пейоративность): «вредное насекомое», «дрянной человек», «недостойные отношения», «осторожный шаг назад», «разбойник», «черт», «мешкать», «хлам». Иными словами, сочетание **шVш** маркирует преимущественно идеофоны, то есть экспрессивную лексику. Слова детского языка в широком смысле также относятся к экспрессивной лексике, и они занимают нишу, образовавшуюся в результате действия фонетических ограничений: *папа, мама, дядя, тетя, дед, бабушка* – все эти слова включают комбинации одинаковых согласных, запретные для «нейтральных», модально не окрашенных слов. Это означает, что выявление фонотактических запретов представляет не только самостоятельный интерес – оно может оказаться эффективным инструментом для системного исследования идеофонов.

ЛИТЕРАТУРА:

- Зализняк 1977* А.А.Зализняк. Грамматический словарь русского языка. М.: «Русский язык», 1977.
- Кузнецова 1986* А.И.Кузнецова; Т.Ф.Ефремова. Словарь морфем русского языка. М.: «Русский язык», 1986.
- Поздняков 2007* К.И.Поздняков. *О «плохих» и «хороших» словах в русском языке* // Сборник к 60-летию А.К.Байбурина. СПб: Европейский Университет, 2007, с. 78-119.
- Фасмер 1950–1958* М.Фасмер. Этимологический словарь русского языка. 3-е издание. М.: «Азбука», 1996.
- Pozdniakov–Segerer 2007* K. Pozdniakov ; G. Segerer. *Similar Place Avoidance: A Statistical Universal* / Linguistic Typology, 12, 2, 2007, p.307-348.