

УТОЧНЕНИЕ КРИТЕРИЕВ ФУНКЦИОНАЛЬНОЙ ПАРАМЕТРИЗАЦИИ ЛЕКСИКИ (НА ПРИМЕРЕ МАЛОГО АКАДЕМИЧЕСКОГО СЛОВАРЯ РУССКОГО ЯЗЫКА)

Ю. А. Суворова

Воронежский государственный университет

Поступила в редакцию 10 сентября 2014 г.

Аннотация: статья посвящена уточнению критериев функциональной параметризации лексики в рамках параметрического анализа лексики, предложенного В. Т. Титовым, на конкретном материале Малого академического словаря русского языка.

Ключевые слова: параметрический анализ лексики, функциональный параметр, критерии вхождения слов в функциональное ядро лексики, Малый академический словарь русского языка.

Abstract: the paper specifies criteria necessary for functional parametrisation of vocabulary within a framework of parametric lexical analysis developed by V. T. Titov. The research is based on the data from Concise Academic Dictionary of Russian Language.

Key words: parametric analysis of vocabulary, functional parameter, criteria for inclusion of words in a functional lexical core, Concise Academic Dictionary of Russian Language.

Функциональный параметр, характеризующий активность слова в речи, является одной из четырех составляющих *параметрического анализа лексики* (ПАЛ) в версии В. Т. Титова [1; 2]. Одной из важнейших характеристик словаря является длина слов, поскольку частота слов связана обратной зависимостью с их длиной [1, с. 10–11].

Цель данной статьи – на конкретном материале показать, каким образом могут быть уточнены и конкретизированы применяющиеся в ПАЛ приемы выделения функционального ядра лексики.

Для получения функционального ядра русской лексики мы используем базу данных МАС-2 размером 83 596 словарных статей, полученную из электронной версии словаря [3].

В словарных статьях МАС-2 представлены не только слова, но и *приставки, начальные и конечные части сложных слов*, которые не могут рассматриваться наравне со словами. За вычетом 503 словарных статей на приставки, начальные и конечные части сложных слов из общего числа словарных статей словаря-источника мы получаем 83 093. Таково число *слов* (полнозначных и служебных) в словарных статьях МАС-2.

В позиции леммы иногда даются не слова, а словосочетания: *в наклон, до зарезу, за глаза, то бишь, то есть* и т.п. Такие словосочетания исключались из дальнейшего рассмотрения, поскольку одним словом они не являются, а считать из порознь – удваивать частоту составляющих их слов: *в, до, за, то, бишь, есть, наклон* и др.

Измерение длины слова в буквах для новых информационных технологий является самым удоб-

ным способом. На рис. 1 показано распределение русской лексики, представленной в МАС-2, по длине в буквах.

Как видно из графика, наибольшее количество слов приходится на слова длиной в 9 букв. По количеству слов рядом находятся 8- и 10-буквенные слова.

Наиболее достоверные результаты дают данные о длине слова в звуках, поскольку первичной является звуковая форма слова. Для получения данных о длине слов в звуках необходимо осуществить практическую транскрипцию русских слов, задача которой – сделать количество букв равным количеству звуков в слове; например, так: *пять > пят*. Описание принципов и алгоритма реализации этой практической транскрипции могло бы составить предмет отдельной статьи. Распределение лексики МАС-2 по длине в звуках представлено на рис. 2.

Из рис. 2 следует, что наибольшее количество слов приходится на слова в 9 звуков. Максимально приближены к ним слова в 8 звуков. Разница в их количестве составляет всего лишь 65 слов.

После получения результатов о длине слова в буквах и звуках был высчитан КоСОГ (коэффициент совершенства орфографии), предложенный В. Т. Титовым [1]. Сумма звуков словаря была разделена на сумму букв словаря. Таким образом, КоСОГ оказался равен 0,957. Это означает, что на каждую тысячу употребленных русских букв в среднем приходится 957 звуков. Следовательно, расхождение в длинах слов, выраженных разными способами, чуть более 4 %. Тем не менее даже при таком незначительном расхождении за единицу подсчета корректнее признать звук. Как следует из рис. 3, максимум распределения по длине в звуках и буквах совпадает и равен 9.

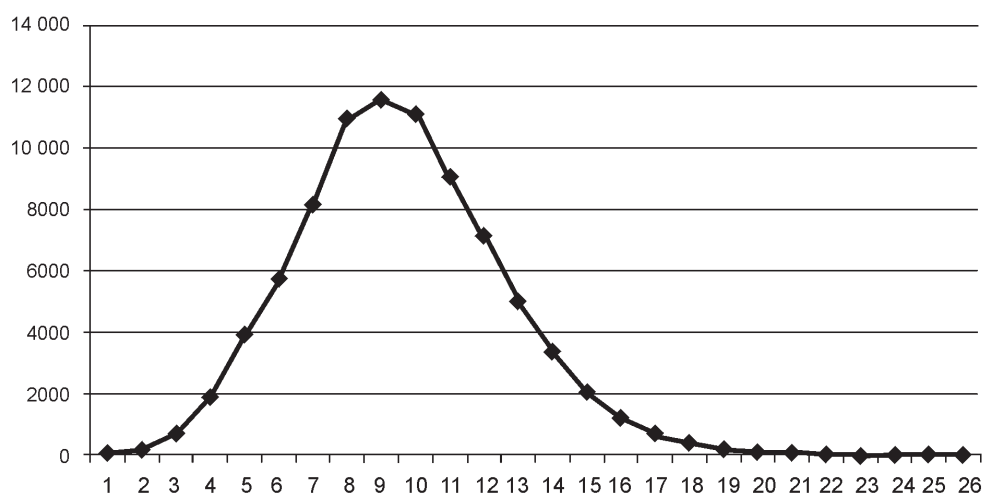


Рис. 1. Зависимость количества слов русской лексики от их длины в буквах по данным МАС-2

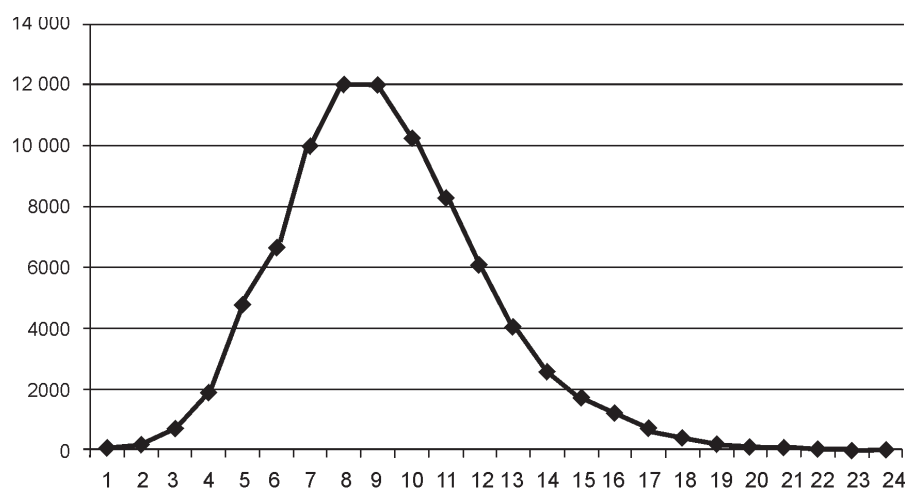


Рис. 2. Зависимость количества слов русской лексики от их длины в звуках по данным МАС-2

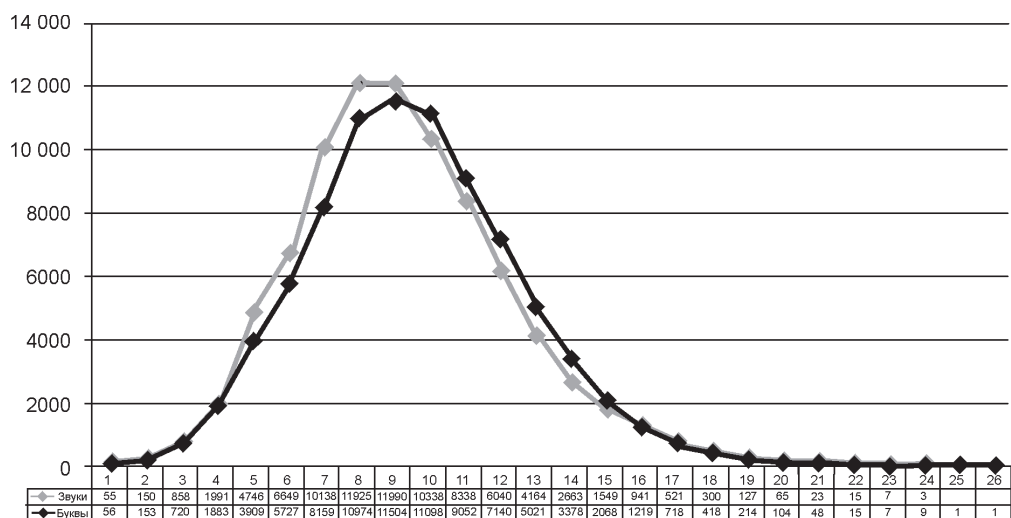


Рис. 3. Распределение длин русских слов в буквах и звуках по данным МАС-2

Функциональное ядро русской лексики по данным МАС-2 составят 1063 слова из 1–3 звуков.

В функциональном ядре МАС-2 оказалось 55 самых коротких слов, состоящих **из одного звука**.

К однозвучным словам относятся: 5 приставок (*в...*, *о...*, *с...*, *у...*, *а...*); 7 предлогов (ϵ^2 , ϵ^2 , ϵ^2 , κ^2 , σ^2 , ς , υ^4); 6 междометий (a^4 , y^2 , y^3 , ϵ^2 , u^4 , σ^3); 6 частиц (*ль*, a^3 , $\bar{\sigma}^2$, ж^2 , u^3 , $-\varsigma$); 5 союзов (a^2 , a^2 , a^2 , a^2 , u^2) и 26 названий букв (a^1 , $\bar{\sigma}^1$, $\bar{\sigma}^1$, ς , δ , ж^1 , ς , u^1 , $\bar{\imath}$, κ^1 , $л$, $м$, $н$, σ^1 , $п$, $р$, ς^1 , $т$, y^1 , ϕ , $х$, ψ , $ч$, $ш$, ψ , ϵ^1).

Приставки (*в...*, *о...*, *с...*, *у...*, *а...*) исключаются из дальнейшего рассмотрения, поскольку это морфемы, входящие в словообразовательную систему русского языка, а не слова, включенные в лексико-семантическую систему.

Служебные слова (предлоги, частицы, союзы) относятся скорее к синтаксису (т.е. внутренней технике языка), чем к лексико-семантической системе языка, описывающей внеязыковую действительность.

Междометия – единицы пред-речи и предъ-языка: они не могут быть названы предложениями, не являются членами предложения. Они внеположны (а исторически – предшествуют) членораздельному человеческому языку и, таким образом, не входят в лексико-семантическую систему языка.

Следуя принципам параметрического анализа лексики, в исследовании лексико-семантического ядра русского языка мы должны учитывать только самостоятельные части речи: существительные, прилагательные и глаголы.

Названия букв словарь трактует как несклоняемые существительные среднего рода. Казалось бы, нет оснований отказывать им во включении в функциональное ядро лексики. Но, во-первых, только *а*, *и*, *о*, *у*, *э* действительно однозвучны. Остальные содержат отсылку к двузвучным именам, например, «Д нескл., ср. см. дэ». Во-вторых, даже *а*, *и*, *о*, *у*, *э* являются «названиями N-й буквы русского алфавита». А НАЗВАНИЕ в МАС-2 толкуется как: «1) словесное обозначение предмета, явления, понятия и т. п.; <...> 3) имя...». По требованиям ПАЛ к ядру лексико-семантической системы не относятся *онимы* (имена собственные), составляющие особую периферийную подсистему словаря. В данном случае мы имеем дело с разновидностью имени собственного: с индивидуальным названием буквы. Ср. «ИМЯ СОБСТВЕННОЕ Существительное, являющееся личным, индивидуальным названием кого-, чего-л. (ср. имя нарицательное)» [4]. Ядро же лексико-семантической системы языка составляют имена нарицательные – «обобщенные названия однородных предметов и понятий (например: *брат*, *озеро*, *страна*, *победа*)» [там же]. Поскольку имена букв явля-

ются не «обобщенными названиями однородных предметов», а «индивидуальными названиями», т.е. онимами, они исключаются из дальнейшего рассмотрения.

Таким образом, среди однозвучных слов не оказалось лемм, относящихся к лексико-семантическому ядру лексики.

Слов, состоящих **из 2 звуков**, в МАС-2 150. С удалением слов, не прошедших частеречный фильтр, множество двузвучных слов сократилось до 19: *ад*, *аз*, *ар*, *ас*, *ил*, *ля*, *ми*, *ом*, *ось*, *па*, *пи*, *ре*, *си*, *уж*¹, *ум*, *ус*, *фа*, *щи*, *эль*, среди которых исключительно имена существительные.

Одним из критериев принадлежности к лексико-семантическому ядру лексики является **немаркированность слова** – отсутствие стилистической, территориальной, профессиональной или любой другой социальной ограниченности слова: его **стилистическая нейтральность, общеизвестность и общенародность**.

В соответствии с приведенным выше значением слово *ад* относится к религиозной терминологии и тем самым тематически и стилистически маркировано. Но это значение не единственное. Два других не имеют признаков маркированности какого бы то ни было рода:

«АД – 2. перен. Обстановка, условия, пребывание в которых мучительно, невыносимо. *Жизнь Александры Михайловны и Зины обратилась в беспросветный ад. Они не знали, как стать, как сесть, чтоб не рассердить Андрея Ивановича* (Вересаев В. В. Два конца). *В манеже стоял ад. От раскормленных за зиму коней валил пар. Все неслись кто как умел, стучаясь о барьеры и наезжая друг на друга* (Игнатъев А. А. Пятьдесят лет в строю).

3. обычно какой или где. О нравственных страданиях, душевных муках, испытываемых кем-л. *Целый ад был у него [Калиновича] в душе* (Писемский А. Ф. Тысяча душ). *Ах, это писательское ремесло! Это не только мука, но целый душевный ад* (Салтыков-Щедрин М. Е. Мелочи жизни)».

В связи с этим в методику параметрического анализа лексики следует внести уточнение: слово исключается из рассмотрения лишь в том случае, если маркированы **все** его значения или если маркировано его **единственное** значение.

Если не пользоваться этим принципом, то и слово *корень* можно отнести к маркированной лексике: «КОРЕНЬ – 1. Подземная часть растения, *посредством которой оно укрепляется в почве и получает из земли воду с растворенными в ней минеральными веществами*. <...> 8. устар. и разг. Род, семья; начало поколения. 9. *грамм.* Основная часть слова (без приставок и суффиксов), которая не поддается дальнейшему разложению на значащие составные части.

10. *мат.* Число, дающее данное число при возведении его в определенную степень» [МАС-2].

Первое значение в выделенной курсивом части содержит научную (энциклопедическую) информацию о корне, полученную биологией, и фактически является биологическим термином. Ср. определение в «Советском энциклопедическом словаре»: «КОРЕНЬ – один из основных органов растений, служащий для прикрепления к почве, поглощения воды, минеральных соединений, а также для выделения некоторых продуктов обмена» [5, с. 637]. Значения же 8, 9 и 10 маркированы в МАС-2 самым явным образом.

В соответствии с вышеизложенным мы не исключаем слово *ад* из дальнейшего рассмотрения.

Названия букв: *бэ, вэ, гэ, дэ, е, ё*, а также устаревшее название буквы «А» – *аз* исключаются из рассмотрения по изложенным выше причинам.

Уже по этим же основаниям можно было бы исключить и слова, называющие звуки – *до, ре, ми, си, фа* и др., которые, кроме того, относятся к музыкальной терминологии и маркированы профессионально.

Профессионально маркированы научные термины, обозначающие единицы измерения: *ар, ом, пи* и др.

Профессионально маркированы также слова *ас* (авиация – характеристика летчика) и *па* («танцевальный шаг» – балетное искусство), связанные с профессиональной деятельностью людей.

Слово *эль* несет инокультурный и иноязычный маркер в своем определении: «Светлое английское пиво, густое и крепкое, изготавливаемое из ячменного солода».

В итоге у нас остаются: *ад, ил, ось, ужэ¹, ум, ус, ици*.

Интересная лексикографическая проблема связана со словом *ус*: оно содержит отсылку: «см. *усы*». Вся информация, которая, как правило, дается при форме Им. п. ед. ч., в данном случае дается при форме Им. п. мн. ч. Леммой в словаре избирается форма множественного числа.

Попробуем понять причину, побудившую авторов словаря принять такое решение.

В НКРЯ имеется 11 649 вхождений слова *усы*. Мы сделали 10 выборок и получили среднее отношение частот форм *усы*: *ус* по данным НКРЯ (табл. 1).

Затем по данным НКРЯ мы исследовали соотношение форм Им. п. (В. п.) мн. ч. : Им. п. (В. п.) ед. ч. в других словах русского языка, обозначающих парные части тела: *рука, нога, глаз, плечо*. Результат представлен в табл. 2.

Результат получился неожиданный. Если в паре *усы* : *ус* форма множественного числа встречается чаще в 3,8 раза, чем форма единственного числа, то в паре *руки*: *рука* форма *руки* встречается в 4,03 чаще, чем форма *рука*, а в паре *ноги* : *нога* форма *ноги* – в 14,22 раза чаще, чем форма *нога*.

Если основанием для выбора в качестве леммы формы множественного числа является относительная частота употребления словарной формы в тексте, то лексикографы явно непоследовательны: коль скоро отступать от общего принципа, ссылаясь на частоту формы, то для лемм РУКИ и НОГИ есть больше оснований, чем для леммы УСЫ.

Но, может быть, лексикографы ориентировались на суммарную частоту форм единственного и множественного числа? Проверим и эту гипотезу.

Из НКРЯ было взято 2797 словоупотреблений слова УС-УСЫ, распределенных по 10 выборкам. Соответствующие данные представлены в табл. 3.

Как видим из табл. 3, среднее отношение форм единственного числа к формам множественного числа (мн. : ед. ч.) превышает 5,5 раза.

Сравнение этого значения с аналогичными значениями слов *рука, нога, глаз, плечо* представлено в табл. 4.

Как видим, в слове *усы* формы единственного числа почти в 6 раз превосходят формы множественного числа, тогда как даже в слове *нога* это соотношение не достигает 3. У слова *глаз* – чуть больше 2, а у слов *рука* и *плечо* не достигает 1,5.

Можно предположить, что основанием для избрания в качестве леммы формы **множественного** числа именительного падежа существительного является превышение **суммарной частоты** форм единственного числа формами множественного числа **не менее чем в 3 раза**. При превышении этой величины в 5 и более раз избрание в качестве леммы формы Им. п. мн. ч. практически неизбежно.

Таким образом, мы обнаружили верифицируемое основание для отступления от принципа выбора в качестве леммы существительного – формы Им. п. ед. ч.

Т а б л и ц а 1

Соотношение частот форм *усы* : *ус* по данным НКРЯ

Формы	1	2	3	4	5	6	7	8	9	10	Среднее значение
<i>усы</i>	115	142	125	170	163	130	87	126	77	69	
<i>ус</i>	20	29	28	45	44	36	26	40	25	33	
Отношение	5,75	4,90	4,46	3,78	3,70	3,61	3,35	3,15	3,08	2,09	3,79

Таблица 2

Соотношение частот форм Им. п. мн. ч. : Им. п. ед. ч. в словах, обозначающих парные части тела по данным НКРЯ

Лемма	Мн. : Ед.
нога	14,22
рука	4,03
усы	3,79
глаз	2,00
плечо	1,29

Таблица 3

Распределение частот слова *усы* в 10 выборках по данным НКРЯ

РангЧ	Число	Гр. зн.	Словоформы	1	2	3	4	5	6	7	8	9	10	Сумма
1	мн.	Им. п. /В. п. мн. ч.	усы	115	142	131	172	163	127	88	77	71	126	1212
2	мн.	Т. п. мн. ч.	усами	73	82	70	59	83	52	58	29	44	71	621
4	мн.	Р. п. мн. ч.	усов	25	37	42	25	36	28	38	49	18	18	316
5	мн.	П. п. мн. ч.	усах	9	9	15	12	10	4	10	5	4	9	87
6	мн.	Д. п. мн. ч.	усам	9	8	9	7	8	3	6	5	3	5	63
		ВСЕ	мн. ч.	232	285	275	285	306	219	202	169	143	238	2354
3	ед.	Им. п. /В. п. ед. ч.	ус	20	29	36	45	44	28	27	25	33	40	327
7	ед.	П. п. ед. ч.	усе		9	8	2	7	5	7	3		6	47
8	ед.	Р. п. ед. ч.	уса	3	1	2	1	3	9	5	1		8	33
9	ед.	Т. п. ед. ч.	усом	2	4		2	1	2	5	2		10	28
10	ед.	Д. п. ед. ч.	усу					1			6		1	8
		ВСЕ	ед. ч.	25	43	46	50	56	44	44	37	33	65	443
Ср. зн.	5,52		мн. ч. : ед. ч.	9,28	6,63	5,98	5,70	5,46	4,98	4,59	4,57	4,33	3,66	

Таблица 4

Соотношение относительных частот форм множественного и единственного числа в словах, обозначающих парные части тела по данным НКРЯ

Леммы	мн. ч.	ед. ч.	мн. ч. : ед. ч.
усы	0,84	0,16	5,52
нога	0,74	0,26	2,91
глаз	0,68	0,32	2,14
рука	0,60	0,40	1,48
плечо	0,59	0,41	1,45

Важно отметить, что слова, леммы которых состоят из двух звуков, – *ил*, *ось*, *уж*¹, *ум*, *ус*, *щи*, раньше были длиннее. Так, например, рус. *ил* происходит от прасл. **jь/ь*, которое восходит к и.-е. **īlu-* [ЭССЯ-8:221]. «Праслав. **osъ* продолжает и.-е. **ak'si* – 'ось'» [ЭССЯ-36:83]. Рус. *уж*¹ восходит к прасл. **zъ* родственно др.-прussk. *angis* «змея», лит. *angis* «ядовитая змея», «уж» [Фасмер-4:150-151]. Рус. *ум* родственно лит. *auituo* «разум», вост.-лит. *aitmenis* «память» [Фасмер-4:161]. Рус. *ус* «родственно др.-прussk. *wanso*» [Фасмер-4:170]. А русское слово *щи* имеет две этимологии: от **сътъ*, мн. *съти*, др.-рус. *съто* «пропитание» (Соболевский) или от **съчь* : *сокъ*

(Якобсон) [Фасмер-4:506]. В любом случае длина этого слова, как и остальных, была больше: *Щ*- восходит к **сът*- или **сък*-.

Из этого мы можем сделать вывод, что слова длиной в два звука в русском языке представляют собой фонетически *вырожденный* случай, обусловленный ходом фонетических процессов. Этот вывод, как и немногочисленность двузвучных слов, свидетельствует о **нетипичности** двузвучных слов для русского языка.

В функциональное ядро русской лексики по данным МАС-2 вошло 858 слов, состоящих **из трех звуков**. Пропустив это множество через частеречный

фильтр, мы получили 547 трехзвучных слов, среди которых также оказалась лексика стилистически и тематически маркированная.

По причинам, описанным выше, из рассмотрения исключаются 9 слов, обозначающих устаревшие названия букв (*ер, ерь, хер, юс¹, ять*) и буквы латинского алфавита (*зет, икс*), название согласного звука *йот*, восходящего к названию греческой буквы «йота», и название 5-го звука музыкальной гаммы *соль²*.

Слова *бог* и *рай* относятся к религиозной терминологии. Соответственно, они стилистически и тематически маркированы.

Однако не будем спешить с выводами. У слова *рай* МАС-2 выделяет 5 значений: «РАЙ – 1. Во многих религиях: место, где души умерших праведников ведут блаженное существование. 2. Сад, где, по библейскому сказанию, жили Адам и Ева до грехопадения. 3. перен. (часто в сочетании с прил. «земной»). Красивое место, доставляющее удовольствие, наслаждение. – *Да, место точно прекрасное, – сказал Беттельгейм, – надо еще осмотреть залив Мельвиль да один пункт на северной стороне – это рай* (Гончаров И. А. Фрегат «Паллада»); – *Ты видала когда-нибудь море? А виноградники? Идешь, а вокруг тебя рай* земной: *груши зреют, апельсины на ветках... виноград висит вот такими кистями!* (Телешов Н. Д. Сухая беда); *На новый год мы ходили на елку и в кукольный театр. Для меня тогда он казался каким-то раем* (Барто А. Л. Найти человека). 4. Очень хорошие условия жизни. [Оброшенов:] *Ваша мать хозяйка была отличная, дома у меня рай* был (Островский А. Н. Шутники); – *Да, – подумал Лукашин, – вот она, жизнь человеческая. Будет, будет сытно в Пекашине, обязательно будет, может, рай* даже *будет* (Абрамов Ф. А. Пути-перепутья). 5. обычно какой или где. Блаженство, радость, испытываемые от чего-л. [Председатель:] *Меня когда-то она считала чистым, гордым, вольным – И знала рай* в объятиях моих (Пушкин А. С. Пир во время чумы); *Тынянов хотел изобразить в искусстве горький рай* творчества, он хотел написать о Пушкине (Шкловский В. Б. Ю. Тынянов – писатель и ученый).

Как и слово *ад*, слово *рай*, в соответствии со внесенным нами в ПАЛ дополнением, не может быть исключено из дальнейшего рассмотрения.

Сложнее обстоит дело со словом *бог*, также относящимся к религиозной сфере, но имеющим в МАС-2 всего одно значение: «По религиозному представлению: верховное существо, сотворившее мир и управляющее им, или (при многобожии) одно из таких существ» [МАС-2].

Формально есть все основания для исключения слова *бог* из дальнейшего рассмотрения. И все же слово *бог* вряд ли можно исключить из числа обще-

известных слов. Во-первых, поскольку даже неверующие определяются по их отношению к Богу: *безбожники* или *а-теисты* (ср. греч. *θεός* 'бог'). Трудно представить себе человека, не знающего слова *Бог*. Во-вторых, при комплексном анализе словаря нельзя абстрагироваться от богатства синтагматики (МАС-2 дает за ромбом 10 фразеологизмов со словом *бог*) и парадигматики этого слова. В-третьих, словарная статья МАС-2, по-видимому, из идеологических соображений советского времени содержит неполную информацию о количестве и качестве значений слова *бог* в русском языке. В этом легко убедиться хотя бы на следующем примере:

*А в углу мы «богов» не повесим,
И не будет лампадка тлеть.*

Твардовский А. Т. Новая изба (1925)

В данном случае, несомненно, имеется в виду третье значение слова *бог*, выделенное БАС-3: «Материализованный образ такого существа в искусстве» [6, с. 81]. У Твардовского речь идет о православных иконах: «живописных изображениях – обычно на доске – Иисуса Христа, Богоматери, святых, сцен из Священного Писания и т.п., которым приписывается священное значение, которые являются предметом поклонения; образ *П* (в православии и католицизме)» [7].

НКРЯ дает и такой пример: «*Комбат меня притянул. Я был комсомольский бог средних размеров*»; «*Ну, чего творить будем по факту хищения?*» [Александр Терехов. Мемуары срочной службы (1991)] [НКРЯ: <http://search.ruscorpora.ru/>]. Этот пример можно подвести под определение 4-го значения слова *бог* в БАС-3: «4. перен. О человеке могущественном, обладающем властью над людьми (обычно в функции сказуемого) // О человеке необычайной творческой одаренности; гений <> Бог чего-л., какой-л. *Бог музыки, кулинарный бог*» [6, с. 81]. Ср. употребление гендерного эквивалента: *И никаких богов в помине, лишь только дело – бой кругом. Но комсомольская богиня, ах, это братцы, о другом* (Окуджава Б. Песенка о комсомольской богине).

Кроме этого, БАС-3 дает и еще одно немаркированное значение: «5. перен. О предмете поклонения, восхищения; кумир (обычно в функции сказуемого). *Нет у меня другого бога, кроме животицы* (Каверин В. Перед зеркалом)» [там же].

В-пятых, даже в «Частотном словаре русского языка» Л. Н. Засориной [8], созданном в советскую эпоху на корпусе текстов советской эпохи¹ слово *бог*

¹ «Словари Э. А. Штейнфельд (1963), Л. Н. Засориной (1977), Л. Леннгрена (1993) и др. были созданы на основе относительно небольших коллекций текстов (400 тысяч – 1 миллион слов) и в большой степени отражают специфику русского языка советского периода: частоты слов *товарищ* и *партия* в них сопоставимы со служебными словами, а слово *расческа* отсутствует» [9].

занимает 242-ю позицию, входя в число 250 самых употребительных слов русского языка с частотой 4,74 употреблений на миллион. По данным О. Н. Ляшевской (Институт русского языка им. В. В. Виноградова, Москва) и С. А. Шарова (Университет Лидса, Великобритания), в НКРЯ слово *бог* занимает 244-ю позицию с частотой 376,63 употреблений на миллион [НКРЯ: <http://ruscorpora.ru>].

Перед нами дилемма: отдать предпочтение МАС-2 или русскому языку. В данном случае мы выбираем русский язык, фиксируя при этом неполноту МАС-2 в отражении семантической структуры слова *бог*.

В соответствии с изложенными выше соображениями слово *бог* входит в функциональное ядро русской лексики.

Во множестве трехзвучных слов оказалось много профессионально маркированной лексики. 16 слов относится к специальной терминологии и содержит соответствующую помету «*спец.*»: *бор, бут, вар, док, зол, йод, кап, лак, леи, мол, пыж, рол, ряж, тол, туш, тяж*. Все слова данной группы используются в различных специализированных сферах человеческой деятельности.

Отдельно выделено 16 трехзвучных слов, которые описывают сферу активного и интеллектуального спорта: *аут, бек, бум, вис, йог, кий, кон, куш, мат, пас, пат, поп, сет, туз, фук, шах*.

Из множества трехзвучных слов исключаем также 44 леммы, относящихся к терминологии разных областей. Профессионально маркированы научные термины – *куб, туф, шар*; исторические термины – *лен, мыт*; медицинские термины – *буж, корь, рак, сыпь, тик, тиф, шок*; ветеринарные термины – *зоб, мыт, сан*; экономические термины – *бум, вал, пул, чек*; физические термины – *ион, фот, эрг*; технические термины – *бор, вал, пек, под, рым, тал, цеп*; морские термины – *бак², бар³, лаг, линь¹, лот¹, ост, пал², рей, риф², топ², туз², фал, фок, ют, ял*.

Термин *тон* является междисциплинарным и относится к области музыки и физики: «Характеристика звука (высота звука, определяемая частотой звуковых колебаний: музыкальный звук в противоположность шуму» [3].

Также исключается из рассмотрения лексика, относящаяся к кулинарной терминологии (*пат, фри*), горному делу (*бар, бур, орт*), пчеловодству (*пад, сот, уза*) и метеорологии (*бар*).

Слова *дол, лик¹, ночь, огонь, сень* с пометой «поэтическое» также исключаются из функционального ядра. Все данные леммы, кроме *сень*, являются вариантами стилистически нейтральных немаркированных слов, что видно из их толкований: «Дол – долина»; «Лик¹ – лицо»; «Ночь – ночь»; «Огонь – то же, что огонь (в 1 знач.)». Ср. «Сень – покров, полог, укрытие,

образуемые кронами деревьев, их ветвями, листвой» [там же].

Таким образом, перечисленные слова, являясь профессионально или стилистически маркированными, не входят в функциональное ядро русской лексики.

Кроме этого, в функциональное ядро русского языка не входит лексика, имеющая **территориальную и социальную ограниченность** и уже в силу этого маркированная.

Во-первых, это слова, имеющие помету «*обл.*»: *баз, кат, кур, кош², марь¹, марь², мга, паль, шаль², ярь¹*; во-вторых, лексика с пометой «*прост.*»: *виш, зык, кус, мат⁵, мор, мочь², пыл, ржа, сыр², топ¹, уда, уйм, чес, фря, четь, хай, хинь, юс², ярь²*. Помимо одной пометы «*обл.*» или «*прост.*» слова *кат, кур, паль, шаль², мор, пыл, ржа, уда, четь, хинь, юс²* содержат дополнительную помету «устар».

Устаревшая лексика не соответствует критерию принадлежности функциональному ядру, поскольку также является стилистически окрашенной и ограниченной в употреблении. Таким образом, слова *лот, пуд, пядь, быль, лик², мат³, меч, навь, одр, око, паж, пал¹, пря, рать, сак¹, сак², сан, тать, ток⁴, тол, тук¹, тур², тьма², узы, хаз, хам, царь, чин, шеф, шииш, шут, щит, ям*, содержащие данную помету, мы, основываясь на показаниях МАС-2, также исключаем из функционального ядра русской лексики.

Аналогично устаревшей лексике исключаются из рассмотрения леммы с пометой «**разговорное**», поскольку данная лексика также маркирована. Среди трехзвучных слов оказалось 26 разговорных: *ахи, вонь, дурь, лад, люд, мот, нюх, охи, пай², пом, поп¹, пуп, раж, рань, сап², сип², сонь, сушь, темь, чет, чих¹, чох, чушь, шип³, щец, пух²*.

Слово *рок*, наоборот, относится к высокому стилю и поэтому также является стилистически маркированным.

Как и в случае с двузвучными словами, исключаем из рассмотрения лексику, содержащую иноязычный и инокультурный маркер: *акр, бай, дож, дон¹, кеб, лев², лей, лек, лье, маг, мим, мэр, пан, пэр, саз, сэр, фут, хан, цех, шах¹, эю, эльф, эре*. Во всех толкованиях данных 23 слов содержится указание на ту или иную страну.

Слова *вуз* «высшее учебное заведение» и *нэл* «новая экономическая политика» являются аббревиатурами. Параметрический анализ лексики необходимо дополнить явным указанием на **непринадлежность аббревиатур** ядру лексико-семантической системы.

Не включаются в функциональное ядро также **слова, не имеющие самостоятельной дефиниции**, и дефинируемые в составе фразеосочетания. Например, **ВЕС 2 м**. < > *На весу* – в висячем положении, без

опоры снизу; **ЛЯД**, м. < На кой ляд (груб. прост.) – зачем, для чего. – *Что ж ты, говорит, за муж такой, жене не можешь содержать, на кой ляд ты мне сдался* (Серафимович А. С. Ледоход); *Иди (или пошел, ну тебя) к ляду!* (груб. прост.) – употребляется для выражения презрения, пренебрежения к кому-либо; **МАТ 4** *Благим матом (кричать, орать, вопить и т. п.)* (прост.) – очень громко. – *А вон лужок, где ты играл, бывало, с ребятами; такой сердитый был: чуть что не по тебе – и закричишь благим матом* (Гончаров И. А. Обыкновенная история); **ТЕК** м. < *Дать теку* (прост.) – быстро убежать, скрыться. [*Казаки*] *постреляли только для остротки и дали теку, не оказав сопротивления* (Фурманов Д. А. Чапаев); **ТУЗ** м. < *Дать туза* (прост.) – сильно ударить кулаком. – *Всякий мальчишка-клоп – и тот норовит дать тебе мимоходом туза!* (Салтыков-Щедрин М. Е. Мелочи жизни); **ЮР** < *На (самом) юру* – 1) на открытом возвышенном месте; 2) на бойком, людном месте. *Дом господский стоял одиночкой на юру, то есть на возвышении, открытом всем ветрам, каким только вздумается подуть* (Гоголь Н. В. Мертвые души); – *Кабак же, надобно тебе сказать, стоял на самом юру, ровно среди улицы* (Успенский Н. Хорошее житье).

Названия месяцев и дней недели отсеиваются как **они́мы**.

Так, среди слов длиной в 3 звука в МАС-2 оказалось 325 слов, соответствующих сформулированным выше критериям принадлежности к функциональному ядру лексики.

Проанализировав леммы длиной в 1–3 звука и отсеяв слова, не прошедшие систему фильтров, сформулированных выше, мы получили множество из 566 однозначных слов, что не соответствует требо-

ванию ПАЛ о размере частнопараметрического ядра не менее чем в 1000 слов.

Этот результат свидетельствует о необходимости включить в функциональное ядро МАС-2 также четырехзвучные слова.

ЛИТЕРАТУРА

1. *Титов В. Т.* Общая количественная лексикология романских языков / В. Т. Титов. – Воронеж : Изд-во Воронеж. гос. ун-та, 2002. – 240 с.
2. *Титов В. Т.* Частная количественная лексикология романских языков / В. Т. Титов. – Воронеж : Изд-во Воронеж. гос. ун-та, 2004. – 552 с.
3. Словарь русского языка : в 4 т. / РАН, Ин-т лингвистич. исследований ; под ред. А. П. Евгеньевой. – 4-е изд., стер. – М. : Рус. яз. : Полиграфресурсы, 1999.
4. Большой толковый словарь русского языка / гл. ред. С. А. Кузнецов. – СПб. : Норинт, 1998. – 1536 с.
5. Советский энциклопедический словарь / научно-редакционный совет: А. М. Прохоров (пред.), М. С. Гиляров, Е. М. Жуков и др. – М. : Советская Энциклопедия, 1980. – 1632 с.
6. Большой академический словарь русского языка. Т. 2 (*Благо-Внять*). – М. ; СПб. : Наука, 2005. – 658 с.
7. Большой современный толковый словарь русского языка / Т. Ф. Ефремова. – 2006.
8. Частотный словарь русского языка / под ред. Л. Н. Засорина. – М. : Русский язык, 1977. – 936 с.
9. *Ляшевская О. Н.* Частотный словарь национального корпуса русского языка : концепция и технология создания / О. Н. Ляшевская, С. А. Шаров. – 2008. – Режим доступа: <http://www.dialog-21.ru/digests/dialog2008/materials/html/53.htm>

Воронежский государственный университет

Суворова Ю. А., преподаватель кафедры теоретической и прикладной лингвистики

E-mail: SuvorovaJ84@mail.ru

Тел.: 8-903-655-09-93

Voronezh State University

Suvorova Yu. A., Lecturer of the Theoretical and Applied Linguistics Department

E-mail: SuvorovaJ84@mail.ru

Tel.: 8-903-655-09-93