

ПАРАМЕТРИЧЕСКИЕ ХАРАКТЕРИСТИКИ АНГЛИЙСКОЙ ЛЕКСИКИ

О. М. Воеводская

Воронежский государственный университет

Поступила в редакцию 4 апреля 2012 г.

Аннотация: *статья посвящена количественному анализу лексики англо-русского словаря по четырем параметрам: функциональному, синтагматическому, парадигматическому и эпидигматическому. Целью работы является выделение лексико-семантического ядра английского языка. Исследование основано на теоретической концепции параметрического анализа, предложенного В. Т. Титовым.*

Ключевые слова: *количественная лексикология, длина слов, синтагматическая активность слов, синонимический ряд, многозначность, английский язык.*

Abstract: *the article deals with the quantitative analysis of the vocabulary of the English-Russian dictionary according to the functional, syntagmatic, paradigmatic and epidigmatic parameters. The aim of the investigation is to single out the lexico-semantic nucleus of the English language. The work is based on the theoretical conception of parametric analysis suggested by V. Titov.*

Key words: *quantitative lexicology, length of words, syntagmatic activity of words, synonymic set, polysemy, the English language.*

Английский язык входит в западногерманскую подгруппу германских языков. Это государственный язык Великобритании, США, Австралии, Новой Зеландии, один из двух государственных языков Ирландии, Канады, один из государственных языков ЮАР. В ряде бывших английских колоний он сохраняется как единственный государственный (Гамбия, Гана, Замбия и др.) или как один из государственных (Индия, Пакистан, Ботсвана, Бруней и др.). Общая численность лиц, для которых он является родным, составляет примерно 400 млн, что в 2,5 раза превышает число говорящих на всех других германских языках вместе взятых.

Статистические методы исследования словарей национальных языков являются одним из действенных инструментов познания лексико-семантической системы языка. Параметрический анализ лексики, предложенный В. Т. Титовым в его монографиях «Общая количественная лексикология романских языков» и «Частная количественная лексикология романских языков» [1, 2], дает возможность познать лексико-семантическую систему языка, минимизировав ее, но не произвольно, а в неразрывной связи с качественными параметрами лексико-семантического, морфологического и синтаксического уровней. Анализируя язык по четырем параметрам – функциональному (длина слов), синтагматическому (число фразеосочетаний и иллюстративных примеров с данным словом в словарной статье), парадигматическому (максимальный размер синонимического ряда, в который входит данное слово) и эпидигматическому

(число значений), можно получить представление о лексико-семантическом ядре языка и стратификации составляющей его лексики по системному (параметрическому) весу.

Материалом для данного исследования послужил корпус английской лексики размером в 53 581 слово, который был получен в результате превращения электронной версии словаря-источника [3] в базу данных.

Функциональная стратификация лексики английского языка

Показателем функциональной активности слова является его длина в буквах. Со времен Джорджа Ципфа известно, что частотность слов обратно пропорциональна их длине: чем короче слово, тем (при прочих равных условиях) чаще оно употребляется, и наоборот [4].

Однако более достоверным показателем являются данные о длине слов в звуках, так как именно звуковая форма является первичной реальностью языка. После обработки имеющихся данных по правилам чтения английского языка [5, с. 10–13] и с учетом приведенной в словаре транскрипции было подсчитано общее число букв и звуков и вычислен коэффициент совершенства орфографии, который оказался равным $\approx 0,905$. Это значение показывает, что расхождение в длинах слов, выраженных разными способами, составляет около 9 %.

Поскольку мы изучаем лексико-семантический уровень языка, то из дальнейшего рассмотрения исключим местоимения, числительные, наречия и служебные слова.

Распределение лексики английского языка по длине слов в звуках представлено в табл. 1.

Самыми короткими оказались следующие однозвучные слова английского языка: глагол *err* 'ошибаться, заблуждаться'; существительные – *are* 'ар' (мера земельной площади 100 кв. м; слово заимствовано из французского языка), *or* 'золотой или желтый цвет' (геральд.), а также три пары омонимов – существительное *awe* 'благоговейный страх, трепет' и глагол 'внушать страх, благоговение'; существительное *oar* 'весло' и глагол 'грести'; существительное *ore* 'руда' и прилагательное 'рудный'.

Среди наиболее длинных слов английского языка можно назвать существительные *amplitude modulation* 'амплитудная модуляция' и *variety entertainment* 'варьете' (20 звуков), *proletarianization* 'пролетаризация' (21 звук), *delegated legislation* 'право министров издавать приказы, имеющие силу законов' (22 звука); пять девятнадцатизвучных слов – *arteriosclerosis* 'атеросклероз' (мед.), *valetudinarianism* 'болезненность, мнительность'; *incomprehensibility* 'непонятность, непостижимость'; *progenitress* 'прародительница, основательница рода'; *radiospectroscopy* 'радиоспектроскопия, техника панорамного приема (электромагнитной энергии)'. Большинство из них представляют собой специальную терминологическую лексику.

Что касается самых длинных слов, то, как известно, в английском языке не существует жестких правил, регламентирующих написание сложных слов. В одном случае имеет место написание их через дефис, в другом – слитное написание, в третьем – написание в два слова. Здесь проявляется закон языковой эволюции в формулировке Т. М. Николаевой: «Язык стремится к передаче всё большего количества ин-

формации в единицу времени» [6, с. 30]. Эта тенденция осуществляется посредством компрессии и суперсегментации. С данной точки зрения, наиболее длинные слова английского языка – это словосочетания, находящиеся на пути к универбации. Потребность в ней возникает тогда, когда частая употребительность словосочетаний приходит в противоречие с их длиной.

Графическое соотношение между длиной слов в звуках и числом слов данной длины представлено на рис. 1.

Как видно из графика, наибольшее число слов английского словаря приходится на семь звуков. Средняя длина английского слова – 7,2 звуков.

На рис. 2 показано распределение слов в соответствии с их удельным весом.

В ядро необходимо включить все слова длиной до трех звуков включительно, следовательно, функциональное ядро будет насчитывать 3119 слов.

Синтагматическая стратификация английской лексики

Оценка синтагматического веса английской лексики производится путем подсчета фразеосочетаний в ромбовой части словарной статьи, а также иллюстративных примеров, в которых встречается данное слово. В анализируемом словаре таких слов оказалось 10 358, и объединяют они 28 012 фразеологизмов и иллюстративных примеров.

Самое большое число фразеосочетаний и иллюстративных примеров зафиксировано у глагола *go* 'идти, ходить; быть в движении; передвигаться (в пространстве или во времени)' – 83.

«Верхушку» синтагматического ядра составили в основном глаголы: *put* 'класть, положить; поставить'

Т а б л и ц а 1

Распределение полнозначной лексики английского языка по длине слов в звуках

Звуков	Слов	Накопл.	Ф-вес	Звуков	Слов	Накопл.	Ф-вес
1	9	9	0,99983	12	1494	50 441	0,03234
2	267	276	0,99471	13	854	51 295	0,01596
3	2843	3119	0,94017	14	416	51 711	0,00798
4	5450	8569	0,83561	15	214	51 925	0,00388
5	6564	15 133	0,70969	16	122	52 047	0,00153
6	7044	22 177	0,57456	17	53	52100	0,00052
7	7518	29 695	0,43033	18	16	52 116	0,00021
8	7071	36 766	0,29468	19	7	52 123	0,00008
9	5620	42 386	0,18687	20	2	52 125	0,00004
10	4016	46 402	0,10983	21	1	52 126	0,00002
11	2545	48 947	0,06100	22	1	52 127	0,00000

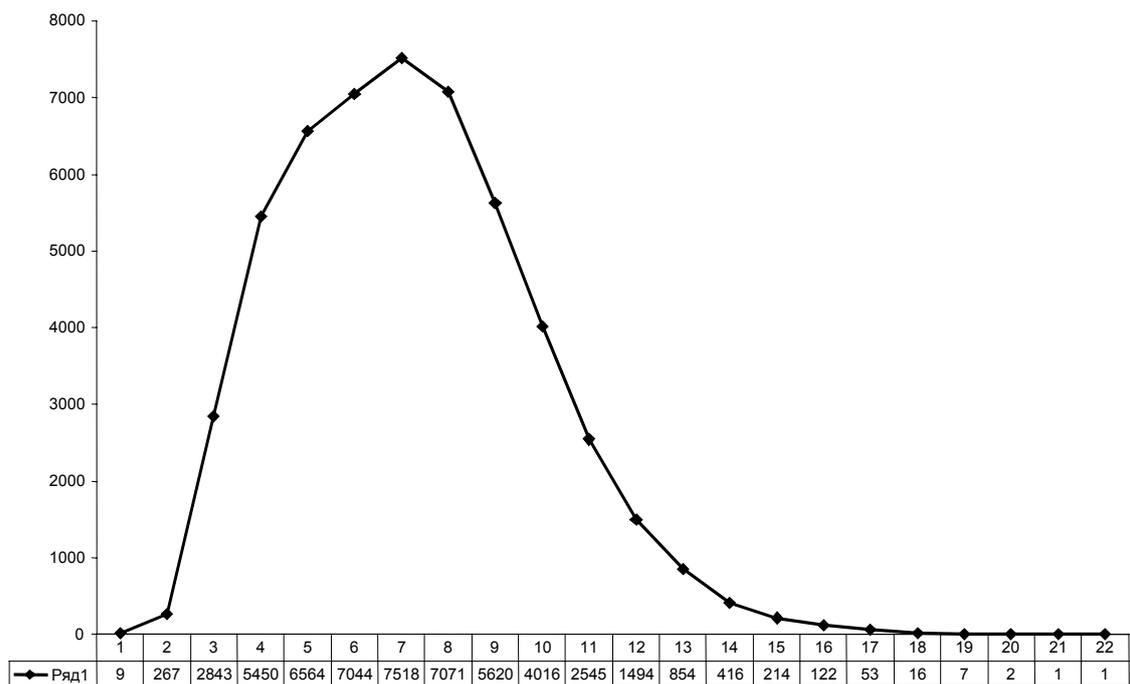


Рис. 1. Распределение лексики английского языка по длине в звуках

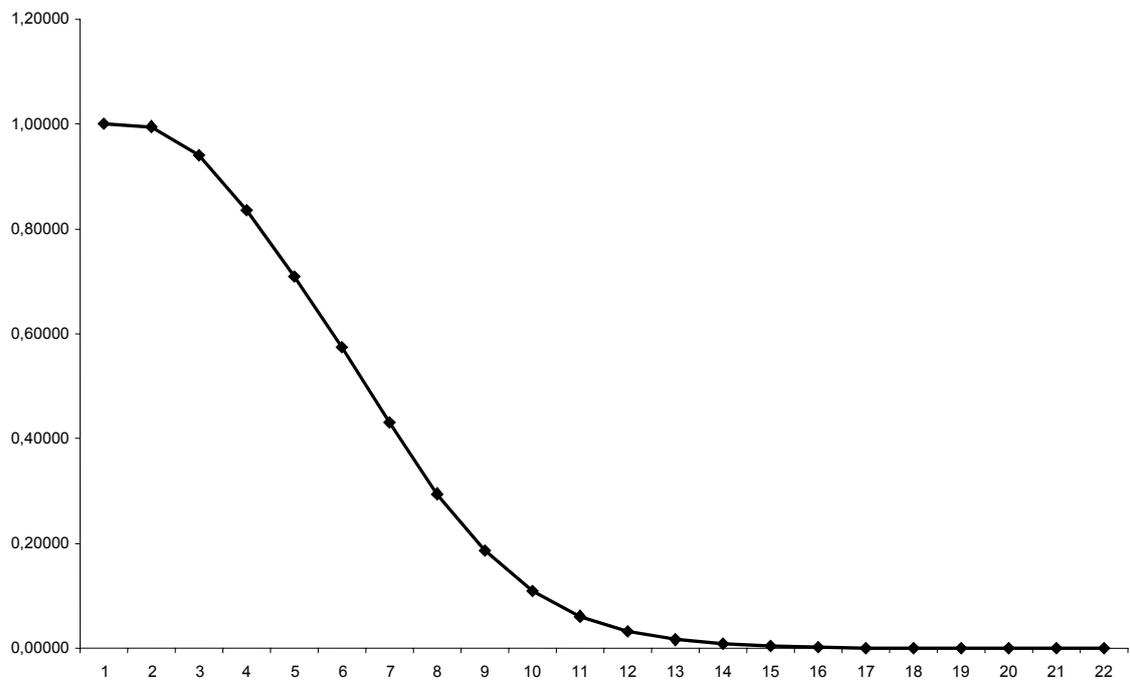


Рис. 2. Распределение лексики английского языка по функциональному весу

– 63, *get* ‘получать; доставать, добывать’ – 62, *take* ‘брать’ – 60, *fall* ‘падать, спадать, понижаться’ – 51, *come* ‘приходить, подходить’ – 50, *stand* ‘стоять’ – 49. Самые высокие показатели синтагматической активности у существительного *time* ‘время’ – 63, у прилагательного *open* ‘открытый’ – 31.

Поранговая динамика синтагматического веса английских слов представлена в табл. 2.

В синтагматическое ядро будут включены все слова с числом фразеосочетаний 3 и более, т.е. 2885 слов.

Парадигматическая стратификация английской лексики

В результате анализа синонимических рядов английской лексики были получены следующие результаты (табл. 3).

В ядро войдут все слова, имеющие 3 и более синонимов, т.е. 1513 синонимических рядов и 5537 слов.

Верхушку парадигматической «пирамиды» составили синонимические ряды, включающие в себя следующее число членов: ‘деньги’ – 18 (*money, cash, crap, dough, jack, kail, necessary, pocket, rhino, rubbish, scads, shekel, shell, shiner, siller, splosh, spondulicks, sugar*), ‘пьяный’ – 16 (*boozy, canned, cockeyed, drunk, high, jagged, loaded, malty, pickled, pissed, pixilated, plastered, screwy, soused, squiffed, stewed*), ‘случайный’ – 14 (*accidental, aleatory, casual, chance, coincidental, episodic, fortuitous, haphazard, happy-go-lucky, lucky, odd, promiscuous, scratch, stray*), ‘сильный удар’ – 12 (*biff, clip, haymaker, lick, pelt, plunk, slog, smite, swash, thwack, wallop, winder*), ‘бродяга’ – 11 (*casual, drifter, hobo, landloper, loafer, prowler, rogue, stiff, stroller, tramp, vagabond*), ‘глупый’ – 11 (*anserine, apish, balmy, balsamic, calvish, cloddish, half-baked, inept, nitwitted, sappy, spoony*), ‘проститутка’ – 11 (*bawd, drab, moll, nightwalker, prostitute, scarletwhore, scarletwoman, slut, streetwalker, strumpet, trollop*), ‘грубый’ – 10 (*bearish, churlish, crass, crude, curmudgeonly, earthy, glaring,*

Т а б л и ц а 2

Показатели синтагматической активности лексики английского языка

ФС	Слов	Накопл.	С-вес	ФС	Слов	Накопл.	С-вес
83	1	1	0,99998	25	6	65	0,99875
63	2	3	0,99994	24	10	75	0,99856
62	1	4	0,99992	23	8	83	0,99841
60	1	5	0,99990	22	9	92	0,99824
51	1	6	0,99988	21	6	98	0,99812
50	1	7	0,99987	20	9	107	0,99795
49	1	8	0,99985	19	9	116	0,99777
48	3	11	0,99979	18	19	135	0,99741
47	1	12	0,99977	17	30	165	0,99683
44	1	13	0,99975	16	20	185	0,99645
43	1	14	0,99973	15	32	217	0,99584
42	1	15	0,99971	14	32	249	0,99522
41	2	17	0,99967	13	44	293	0,99438
40	1	18	0,99965	12	48	341	0,99346
39	1	19	0,99964	11	45	386	0,99260
38	1	20	0,99962	10	65	451	0,99135
37	1	21	0,99960	9	97	548	0,98949
35	4	25	0,99952	8	110	658	0,98738
33	3	28	0,99946	7	151	809	0,98448
32	4	32	0,99939	6	215	1024	0,98036
31	3	35	0,99933	5	312	1336	0,97437
30	3	38	0,99927	4	525	1861	0,96430
29	6	44	0,99916	3	1024	2885	0,94465
28	3	47	0,99910	2	1903	4788	0,90815
27	5	52	0,99900	1	5570	10 358	0,80129
26	7	59	0,99887	0	41 769	52 127	0,00000

Т а б л и ц а 3

Показатели парадигматической активности лексики английского языка

Синонимов	Рядов	Накопл. рядов	П-вес
18	1	1	0,99999
16	1	2	0,99998
14	1	3	0,99996
12	1	4	0,99995
11	3	7	0,99992
10	4	11	0,99987
9	5	16	0,99981
8	11	27	0,99968
7	31	58	0,99932
6	34	92	0,99891
5	119	211	0,99751
4	349	560	0,99339
3	953	1513	0,98214
2	4911	6424	0,92418
1	78 308	84 732	0,00000

horse, rough, unrefined), ‘простак’ – 10 (*boob, bullcalf, coot, flat-foot, greener, innocent, jay, lamb, lummoх, simpleton*), ‘скупой’ – 10 (*cheeseparing, churlish, close, covetous, curmudgeonly, hardfisted, niggard, parsimonious, penurious, scrimpy*), ‘тюрьма’ – 10 (*cage, choky, college, gaol, jail, jug, mill, pound, prison, stone-jug*), ‘болтовня’ – 9 (*babble, cackle, chatter, chit-chat, gammon, gossip, jive, prattle, wishwash*), ‘напыщенный’ – 9 (*bombastic, declamatory, high, highfaluting, highblown, mouth-filling, mouthy, pompous, tumid*), ‘обманывать’ – 9 (*beguile, defraud, duff, gammon, hocus, mump, short-change, stick, victimize*), ‘проклятие’ – 9 (*anathema, bane, cuss, damn, execration, imprecation, malediction, perdition, tarnation*), ‘хитрый’ – 9 (*artful, astute, fly, foxy, furtive, leery, loopy, slick, slim*).

Воспользуемся методом аксиологического анализа, предложенного Е. В. Долбиловой и А. А. Кретовым [7], с целью выяснения того, какова система духовных ценностей, объективно присущих английской ментальности.

Парадигматической доминантой английской лексики, по данным словаря-источника, является семема ‘деньги’, отмеченная в синонимическом ряду, включающем 18 членов. Особое внимание англичан к деньгам не вызывает никаких вопросов, так как именно Англия стала родиной первой промышленной революции и центром крупнейшей колониальной империи. Дух предпринимательства всегда был

свойствен английскому народу, который Наполеон пренебрежительно называл «нацией лавочников».

К доминанте непосредственно примыкают семемы ‘обманывать’, ‘хитрый’ и ‘скупой’, поскольку операции купли-продажи могут включать в себя надувательство и стремление получить дополнительную выгоду.

Следующий семантический кластер образуют семемы ‘бродяга’, ‘проститутка’ и ‘пьяный’, объединенный общим компонентом «нарушение нормы/закона». Если рассматривать эти семемы как маркированные члены оппозиции, можно предположить, что в качестве оппозиции к ‘бродяге’ будет семема ‘привязанность к дому’ (вспомним известную английскую поговорку «мой дом – моя крепость»), хотя, наряду с этим, близость моря будила в англичанах страсть к приключениям и путешествиям, породила моряков, торговцев и авантюристов. Оппозицией к семеме ‘проститутка’ может послужить ориентация на семейные ценности. Оппозиция к ‘пьяный’ – трезвость.

Логическим продолжением цепочки ‘нарушение нормы’ может служить семема ‘грубый’. Англичанам приписывается традиционно строгое следование этикету и нормам поведения, иногда доходящее до абсурда.

Важной с аксиологической точки зрения является семема ‘болтовня’, которая также может рассматриваться как некое нарушение нормы. Следовательно, молчание – качество, высоко ценимое англичанами.

Что касается семем ‘глупый’ и ‘простак’, то, если рассматривать их как маркированные члены оппозиции, можно сделать вывод о том, что англичане ценят ум и предприимчивость. Если же обе эти семемы включить в кластер, который объединяет семемы ‘обманывать’ и ‘хитрый’, то их можно рассматривать как объекты, на которые направлены действия по одурачиванию не слишком сообразительных и смелых людей.

И наконец завершает кластер аномальных характеристик семема ‘случайный’. В качестве маркированного члена оппозиции она может свидетельствовать о том, что англичане стараются избегать всего случайного, и для этого тщательно планируют свою жизнь и следуют строгому распорядку. Любая случайность воспринимается ими как неприятность или досадная помеха и надолго остается в памяти.

Семемы ‘тюрьма’ и ‘проклятие’ могут быть объединены в кластер ‘наказание’, в частности, за те нарушения нормы и закона, о которых говорилось выше.

Семема ‘напыщенный’ ассоциируется с приписываемым англичанам снобизмом. Снобизм – поведенческая модель или стиль жизни, заключающийся

в подражании аристократическим манерам, вкусам, поведению высшего общества; претензия человека на принадлежность к высшим социальным слоям.

Семема 'сильный удар' может быть отнесена скорее не к стихийной драке и потасовке, а к драке, регламентируемой определенными правилами, например, боксу, родоначальниками которого англичане и являются. Принципы спортивной этики стали для англичан основами подобающего поведения, морали, мерилom порядочности. В боксе одним из основополагающих правил является умение держать удар, сохранять самообладание. Держать удар, т.е. стойко переносить тяготы и испытания – вот, что англичане считают главным достоинством человеческого характера.

Таким образом, язык демонстрирует, что умеренность, сдержанность, стремление следовать правилам, законопослушность – ключевые характеристики английского народа. Английская ментальность ориентирована в первую очередь на человека, а также на следование установленным в обществе и обязательным для всех его членов нормам и правилам поведения.

Эпидигматическая стратификация английской лексики

Стандартный учет эпидигматического потенциала английских слов, основанный на данных о многозначности, позволил получить следующий результат (табл. 4).

Таким образом, в английском языке насчитывается 20 796 многозначных слов. В ядро необходимо

включить все слова с числом значений от 4 и выше, следовательно, эпидигматическое ядро будет составлять 4499 слов.

Максимальное число значений отмечено у глагола *run* 'бежать, бегать' – 34. Следует отметить, что наиболее многозначной является лексика с «динамичной» семантикой: *pass* 'двигаться вперед; проходить, проезжать' (31), *draw* 'тащить, волочить' (27), *run* 'бег, пробег' (27), *go* 'идти, ходить; быть в движении' (27), *play* 'играть, резвиться, забавляться' (27). Ее несколько уравнивают глаголы с более «статичной» семантикой: *set* 'ставить, класть, помещать' (30) и *take* 'брать'.

Большое число значений отмечено также у существительных, называющих части тела человека и растения: *head* 'голова' (27) и *stock* 'главный ствол (дерева)' (28).

Параметрическая стратификация английской лексики

Итак, проведя анализ английской лексики по четырем параметрам, мы получили: 1) функциональное ядро размером в 3119 слов, 2) синтагматическое ядро размером в 2885 слов, 3) парадигматическое ядро размером в 5537 слов, 4) эпидигматическое ядро размером в 4499 слов. Сложение этих множеств позволило получить корпус английской лексики в 16 040 слов.

В результате слияния частнопараметрических ядер (далее – ЧПЯ) получаем четыре множества: 1) малое параметрическое ядро английского словаря

Т а б л и ц а 4

Показатели эпидигматической активности лексики английского языка

Значен.	Слов	Накопл.	Д-вес	Значен.	Слов	Накопл.	Д-вес
34	1	1	0,99998	15	13	61	0,99883
31	1	2	0,99996	14	22	83	0,99841
30	1	3	0,99994	13	24	107	0,99795
28	1	4	0,99992	12	48	155	0,99703
27	4	8	0,99985	11	58	213	0,99591
26	1	9	0,99983	10	82	295	0,99434
25	1	10	0,99981	9	147	442	0,99152
24	2	12	0,99977	8	168	610	0,9883
23	1	13	0,99975	7	307	917	0,98241
21	3	16	0,99969	6	537	1454	0,97211
20	4	20	0,99962	5	972	2426	0,95346
19	5	25	0,99952	4	2073	4499	0,91369
18	5	30	0,99942	3	4728	9227	0,82299
17	6	36	0,99931	2	11 569	20 796	0,60105
16	12	48	0,99908	1	31 331	52 127	0

(лексика, вошедшая во все четыре ЧПЯ – 31 слово), 2) большое параметрическое ядро (лексика, вошедшая в три ЧПЯ – 515 слов), 3) периферия большого параметрического ядра (лексика, вошедшая в два ЧПЯ – 1966 слов), 4) «мусор, отходы» (лексика, вошедшая в одно ЧПЯ).

Слово с максимальным значением параметрического веса будет считаться доминантой лексико-семантической системы. В английском языке это слово – *head* ‘голова’, отражающее не только осознание человеком самого себя, но и подчеркивающее значимость той части человека, которая отвечает за мышление.

Не считая доминанты, в первую «десятку» вошли лексемы, которые семантически объединяются вокруг нее: существительные *man* ‘человек’, *neck* ‘шея’ (часть тела между головой и туловищем), прилагательные, обозначающие различные ментальные характеристики и качества человека: *keen* ‘острый’ (в основном сочетается со словами, обозначающими ум, чувства, ощущения), *mad* ‘сумасшедший, безумный’, глагол говорения *tell* ‘рассказывать’, предполагающий сообщение какой-либо информации собеседнику.

Отдельную подгруппу представляют прилагательные, выражающие признаки предметного мира: *loose* ‘свободный, несвязанный’, *rough* ‘грубый’, *thin* ‘тонкий’.

Воронежский государственный университет

Воеводская О. М., кандидат филологических наук, доцент кафедры теоретической и прикладной лингвистики

E-mail: oxavoev@mail.ru

Тел.: 8-960-110-41-78

Лексемы *force* ‘сила’ и *arm* ‘оружие’ в сочетании с доминантой ‘голова’ отражают то самое могущество, которое и обеспечило Великобритании статус великой морской державы, мастерской мира, родины великих ученых и изобретателей.

ЛИТЕРАТУРА

1. *Титов В. Т.* Общая квантитативная лексикология романских языков : монография / В. Т. Титов. – Воронеж : Изд-во Воронеж. гос. ун-та, 2002. – 240 с.

2. *Титов В. Т.* Частная квантитативная лексикология романских языков : монография / В. Т. Титов. – Воронеж : Изд-во Воронеж. гос. ун-та, 2004. – 552 с.

3. *Мюллер В. К.* Англо-русский словарь. – 22-е изд., стер. / В. К. Мюллер. – М. : Русский язык, 1988. – 848 с.

4. *Zipf G. K.* The Psycho-Biology of Language : an introduction to dynamic philology / G. K. Zipf. – Cambridge : Mass. MIT Press, 1965. – 336 p.

5. *Берков В. П.* Современные германские языки / В. П. Берков. – М. : Астрель АСТ, 2001. – 336 с.

6. *Николаева Т. М.* От звука к тексту / Т. М. Николаева. – М. : Язык рус. культуры, 2000. – 680 с.

7. *Долбилова Е. В.* Парадигматическая стратификация каталанской лексики и аксиологемы каталанской ментальности / Е. В. Долбилова, А. А. Кретов // Лінгвістичні студії. – Донецьк, 2010. – Вип. 21. – С. 267–273.

Voronezh State University

Voevodskaya O. M., Candidate of Philology, Associate Professor, Department of Theoretical and Applied Linguistics

E-mail: oxavoev@mail.ru

Tel.: 8-960-110-41-78