

**ПАРАМЕТРИЧЕСКАЯ СТРАТИФИКАЦИЯ
ИСЛАНДСКОЙ ЛЕКСИКИ**

Д. В. Багина, О. М. Воеводская

Воронежский государственный университет

Поступила в редакцию 13.06.2011 г.

Аннотация. В данной работе проводится количественный анализ лексики исландско-русского словаря по четырем параметрам: функциональному, синтагматическому, парадигматическому и эпидигматическому. Целью исследования является выделение лексико-семантического ядра исландского языка. Исследование основано на теоретической концепции параметрического анализа, разработанного В.Т. Титовым.

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

Abstract. In the article the authors analyse the vocabulary of the Icelandic-Russian dictionary according to the functional, syntagmatic, paradigmatic and epidigmatic parameters aiming to single out the lexico-semantic nuclear of the Icelandic language. The investigation is based on the theoretical conception of parametric analysis worked out by V. Titov.

Keywords: quantitative lexicology, length of words, syntagmatic activity of words, synonymic set, polysemy, the Icelandic language.

ВВЕДЕНИЕ

Одной из основных целей любого исследования является последовательная минимизация информации об изучаемых объектах, чтобы, в конечном счете, из огромного исходного материала получить компактную таблицу, содержащую в себе самую важную информацию об этих объектах.

Примером такого исследования является количественный анализ лексики национальных языков. Статистические исследования различных словарей дают 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 четырём параметрам – функциональному (длина слов), синтагматическому (количество фразеосочетаний), парадигматическому (максимальный размер синонимического ряда, в который входит данное слово) и эпидигматическому (количество значений), можно получить представление о лексико-семантическом ядре языка и стратификации составляющей его лексики по системному (параметрическому) весу.

Материалом для данного исследования послужил корпус исландской лексики размером в 33406 слов, который был получен в результате превращения электронной версии словаря-источника [1] в базу данных.

Исследование является частью работы по исследованию лексикологии германских языков, которая проводится на кафедре Теоретической и прикладной лингвистики Воронежского госуниверситета.

1. ФУНКЦИОНАЛЬНАЯ СТРАТИФИКАЦИЯ ИСЛАНДСКОЙ ЛЕКСИКИ

Длина слова может быть измерена в буквах, но 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 исландского языка [2: 206–210] было подсчитано общее количество звуков и вычислен коэффициент совершенства орфографии (КоСОГ), который оказался равным $\approx 0,97$. Это значение показало, что расхождение в длинах слов, выраженных буквами и звуками, составляет около 3 %, что можно трактовать как архаичность произношения исландского языка.

Наиболее активными в функциональном отношении оказались 2 полнозначные омонимичные лексемы *á*: существительное ‘река’ и глагол ‘останавливаться в пути, делать привал’, остальные 8 – неполнозначные лексемы (предлоги, междометия, наречия). Поскольку мы изучаем лексико-семантический уровень языка, неполнозначные слова находятся за пределами нашего исследования. Кроме служебных слов, из дальнейшего рассмотрения были исключены местоимения, числительные и наречия.

В табл. 1 представлено распределение слов по длине, измеряемой в количестве звуков в слове.

После удаления слов, не являющихся существительными, прилагательными и глаголами, получено функциональное ядро исландской лексики размером в 2909 слов. Самым короткими из них являются две уже упоминавшиеся однобуквенные омонимичные лексемы *á*: существительное ‘река’ и глагол ‘останавливаться в пути’, самым длинным – существительное длиной в 22 буквы *nordurheimskautsbaugur* ‘северный полярный круг’.

Графическое соотношение между длиной слова и количеством слов данной длины представлено на рисунке 1.

Таблица 1

Распределение слов по длине в звуках

Звуков	Слов	Накопл	Вес	Звуков	Слов	Накопл	Вес
1	2	2	0,99994	12	1553	30440	0,06385
2	85	87	0,99732	13	900	31340	0,03617
3	695	782	0,97595	14	574	31914	0,01851
4	2127	2909	0,91054	15	295	32209	0,00944
5	3214	6123	0,81169	16	155	32364	0,00467
6	3251	9374	0,71171	17	93	32457	0,00181
7	3895	13269	0,59192	18	31	32488	0,00086
8	4457	17726	0,45485	19	18	32506	0,00031
9	4748	22474	0,30883	20	7	32513	0,00009
10	3892	26366	0,18914	21	2	32515	0,00003
11	2521	28887	0,11161	22	1	32516	0,00000

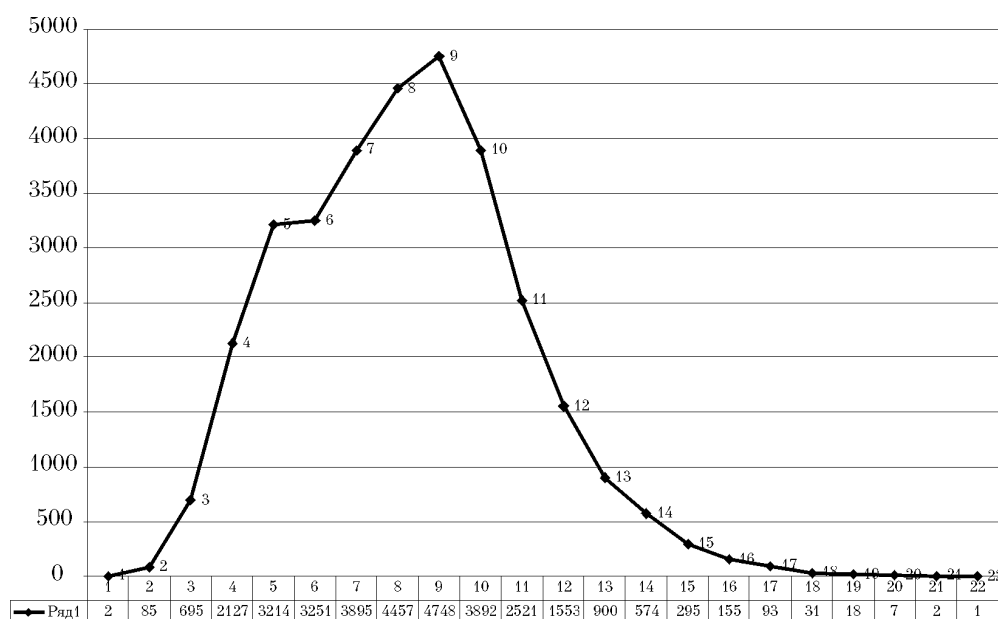


Рис. 1. Распределение слов исландского языка по длине

2. СИНТАГМАТИЧЕСКАЯ СТРАТИФИКАЦИЯ ИСЛАНДСКОЙ ЛЕКСИКИ

Оценка синтагматического веса исландской лексики производится путем подсчета фразеосочетаний в ромбовой части словарной статьи. В анализируемом словаре таких слов оказалось 3793, и объединяют они 7718 фразеологизмов. Самое большое количество фразеосочетаний оказалось у глаголов: *láta* 'класть; ставить' (38) и *bera* 'носить, нести' (34).

Несколько ниже оказались показатели у существительных – *hönd* 'рука' (30) и *vegur* 'дорога, путь' (28). Прилагательное с наибольшим количеством фразеосочетаний – *ýungur* 'грузный, тяжелый' (23). Таким образом, самой синтагматически активной оказалась глагольная лексика.

Поранговая динамика синтагматического веса исландских слов представлена в следующей таблице:

В данном случае к ядру будут относиться все слова, имеющие при себе два и более фразеосочетаний, а к периферии – не имеющие их или имеющие по одному фразеосочетанию. Итоговый размер синтагматического ядра исландской лексики – 1194 слова.

3. ПАРАДИГМАТИЧЕСКАЯ СТРАТИФИКАЦИЯ ИСЛАНДСКОЙ ЛЕКСИКИ

После анализа синонимических рядов исландской лексики были получены следующие результаты:

Таблица 3

Показатели парадигматической активности лексики исландского языка

Синонимов	Рядов	Накопл. рядов	П-вес
14	2	2	0,99994
12	1	3	0,99992
11	2	5	0,99986
10	1	6	0,99983
9	2	8	0,99978
8	3	11	0,99969
7	14	25	0,99930
6	18	43	0,99880
5	72	115	0,99680
4	162	277	0,99228
3	550	827	0,97697
2	2759	3586	0,90012
1	32316	35902	0,00000

Таблица 2

Показатели синтагматической активности лексики исландского языка

Кол-во ФС:	Слов:	Накопл:	Вес:	Кол-во ФС:	Слов:	Накопл:	Вес:
38	1	1	0,99997	14	7	52	0,99840
34	1	2	0,99994	13	3	55	0,99831
30	2	4	0,99988	12	13	68	0,99791
28	1	5	0,99985	11	14	82	0,99748
27	1	6	0,99982	10	11	93	0,99714
26	2	8	0,99975	9	29	122	0,99625
25	2	10	0,99969	8	26	148	0,99545
23	5	15	0,99954	7	31	179	0,99450
22	1	16	0,99951	6	56	235	0,99277
21	3	19	0,99942	5	77	312	0,99040
20	4	23	0,99929	4	125	437	0,98656
19	2	25	0,99923	3	206	643	0,98023
18	4	29	0,99911	2	551	1194	0,96328
17	5	34	0,99895	1	2599	3793	0,88335
16	6	40	0,99877	0	28723	32516	0,00000
15	5	45	0,99862				

Поскольку ближе всего к 1000 множество из 827 рядов, то в ядро войдут все слова, имеющие 3 и более синонимов, т.е. 2978 слов.

Наибольшее количество синонимов в ряду (14) зафиксировано у лексем со значениями 'шуметь' (*brauka, busla, dunka, fjanda, gnauða, kliða, nauða, skrattast, skurka, svarra, penjasig, prasa, pysja, ymjá*) и 'щедрый' (*artarlegur, gjafmildur, gjöfull, greiðasamur, greiðugur, höfðinglegur, mildur, ör 3, örlótlegur, örlótur, örlyndur, styrlatur, stýrmenskufullur, stýrtækur*).

4. ЭПИДИГМАТИЧЕСКАЯ СТРАТИФИКАЦИЯ ИСЛАНДСКОЙ ЛЕКСИКИ

Стандартный учет эпидигматического потенциала исландских слов, основанный на данных о многозначности, позволил получить следующие результаты:

Таблица 4

Показатели эпидигматической активности лексики исландского языка

Значений:	Слов:	Наропл:	Вес:
13	1	1	0,99997
11	2	3	0,99991
10	2	5	0,99985
9	6	11	0,99966
8	14	25	0,99923
7	20	45	0,99862
6	47	92	0,99717
5	98	190	0,99416
4	308	498	0,98468
3	1052	1550	0,95233
2	4329	5879	0,81920
1	26637	32516	0,00000

Таким образом, в исландском языке насчитывается 5879 многозначных слов. Максимальное количество значений (13) зафиксировано у глагола *standa*: 1) 'стоять', 2) 'не двигаться; застревать', 3) 'находиться, быть расположенным', 4) 'быть, пребывать в состоянии', 5) 'длиться, продолжаться', 6) 'действовать, быть в силе', 7) 'происходить, иметь место', 8) 'быть написанным (напечатанным), значиться', 9) 'котироваться; быть в цене', 10) 'дуть (о ветре)', 11) 'подобать, приличествовать', 12) 'весить', 13) 'застигать, заставить, ловить (кого-л. за чем-л.)'.

5. ПАРАМЕТРИЧЕСКАЯ СТРАТИФИКАЦИЯ ИСЛАНДСКОЙ ЛЕКСИКИ

Проведя анализ исландской лексики по четырем параметрам, мы получили: 1) функциональное ядро исландской лексики размером в 2909 слов, 2) синтагматическое ядро размером в 1194 слов, 3) парадигматическое ядро размером в 2978 слов, 4) эпидигматическое ядро размером в 1550 слов.

Сложение этих множеств позволило получить корпус исландской лексики в 8631 слово. Сложение весов по каждому параметру дает суммарный параметрический вес для каждого из слов данного множества и позволяет упорядочить слова по убыванию их параметрического веса. Ближайшее к 1000 множество слов с максимальным системным весом включает 1467 слов.

Слово с максимальным значением параметрического веса будет считаться доминантой лексико-семантической системы. Доминантой исландского словаря оказалось существительное *má* 'речь, дар речи'. В первую десятку также вошли: глагол *slá* 'бить, убивать' и существительные *far* 'судно, корабль', *syk* 'обвинение', *ryt* 'корень, причина', *byt* 'улучшение, исправление', *ráð* 'совет', *tak* 'взятие', *myt* 'печать' и *lið* 'группа, компания'. Ближайшее к к верхушке ядра прилагательное *stor* 'большой, сильный, мощный' занимает по суммарному параметрическому весу лишь 18 позицию.

ВЫВОДЫ

Используемая в работе методика параметрического анализа лексики [3] позволяет «утилизировать» большой массив словаря, а выделение лексико-семантических доминант может быть использовано при типологической характеристике лексико-семантических систем родственных языков. Успешное применение статистических методов к лексикографическим материалам свидетельствует о синтезе количественной и качественной лингвистических традиций.

СПИСОК ЛИТЕРАТУРЫ

1. Берков В. П. Исландско-русский словарь / В. П. Берков; с прилож. краткого очерка грамматики исландского языка, сост. А. Бедварссоном. – М. : Гос. изд-во иностранных и национальных словарей, 1962. – 1032 с.
2. Берков В. П. Современные германские языки / В. П. Берков. – М. : Астрель АСТ, 2001. – 336 с.

3. *Титов В. Т.* Частная квантитативная лексикология романских языков : / В. Т. Титов ; Воронеж.

гос. ун-т. – Воронеж : Изд-во Воронеж. гос. ун-та, 2004. – 552 с.

Воевудская Оксана Михайловна – кандидат физико-математических наук, доцент кафедры английской филологии, факультет романо-германской филологии, Воронежский государственный университет. Тел. 2-208-489, 8-960-110-41-78. E-mail: oxavoev@mail.ru

Voevudskaya Oxana M. – Candidate of Science (Philology), Associate Professor, the dept. of English Philology, Romance and Germanic Philology Faculty, Voronezh State University. Tel. 2-208-489, 8-960-110-41-78. E-mail: oxavoev@mail.ru

Багина Диана Владимировна – студентка 3 курса кафедры Теоретической и прикладной лингвистики, факультет романо-германской филологии, Воронежский университет. Тел. 8-950-771-17-37. E-mail: baginadiana@rambler.ru

Baghina Diana V. – 3rd year student of the dept. of Theoretical and Applied Linguistics, Romance and Germanic Philology Faculty, Voronezh State University. Tel. 8 950 771 17 37. E-mail: baginadiana@rambler.ru