

## ПАРАМЕТРИЧЕСКИЙ АНАЛИЗ ХОРВАТСКО-РУССКОГО СЛОВАРЯ

И. А. Меркулова

*Воронежский государственный университет*

Поступила в редакцию 12 ноября 2014 г.

**Аннотация:** статья посвящена анализу лексики хорватского языка, представленной в двуязычном словаре, с помощью методов количественной лингвистики. На материале словаря выявляются значимые элементы лексико-семантической системы: наиболее употребительные, многозначные, синтагматически активные и синонимически богатые лексемы. Анализ завершается выделением лексико-семантического ядра хорватского языка.

**Ключевые слова:** хорватский язык, лексика, словарь, параметрический анализ, лексико-семантическое ядро.

**Abstract:** the paper analyses vocabulary of the Croatian language. The quantitative analysis is based on the Croatian-Russian dictionary. The analysis reveals significant elements of the lexical system of the Croatian, i.e. the lexemes that are the most frequent and polysemous. These words are also used in various collocation and have many synonyms. The analysis presents the core vocabulary of the Croatian language.

**Key words:** Croatian language, vocabulary, dictionary, parametric analysis, lexico-semantic core.

Хорватский язык – один из языков южных славян, являющийся официальным для Хорватии, Боснии и Герцеговины. Как отмечают авторы коллективного международного труда «Славянская лексикография», история хорватской лексикографии иная, нежели история лексикографии других южнославянских языков, а «это доказывает, что существование хорватского литературного языка не связано с существованием самостоятельного хорватского государства, т.е. он жил столетиями, задолго до провозглашения Хорватии государством» [1, с. 336].

В последнее время хорватский язык все чаще стал появляться под своим собственным, а не двойным (сербско-хорватский, сербо-хорватский, хорватско-сербский и т.п.) названием. Соответственно, стали выходить и словари, в которых хорватский язык представлен в соотношении с каким-либо другим славянским или неславянским языком. Хотя в настоящее время все еще ощущается большая потребность в словарях, в которых хорватский язык был бы представлен в соотношении с каким-либо другим славянским языком [2]. Нас интересует, в первую очередь, соотношение хорватский-русский. Истории известны рукописный словарь *Dizionario italiano-slavo-moscovitico* (1751) уроженца Дубровника Ивана Мария Матияшевича, а также Краткий хорватско-русский словарь (1894) А. Суворова. Среди изданий, вышедших в наше время, можно назвать Русско-хорватский словарь с грамматикой Радослава Ф. Полянца и Серафимы М. Мадатовой-Полянец (последнее издание 2002 г.) и двухтомный хорватско-русский

словарь Матии Даутовича (Загреб, 2002). В России вышли Хорватско-русский словарь (1999, 2001, 2003) и Новый хорватско-русский словарь (2007) А. Р. Багдасарова.

Источником нашего исследования явился хорватско-русский словарь [3].

Метод исследования – параметрический анализ лексики, подробно описанный в работах проф. В. Т. Титова [4; 5].

Первый параметр характеристики лексики – длина слова, которую можно оценить в буквах. Статистика приводится в табл. 1.

Т а б л и ц а 1

*Функциональная стратификация хорватской лексики*

Количество букв	Количество слов	Накопление	Ф-вес
1	2	3	4
2	8	8	0,99932
3	222	230	0,98039
4	473	703	0,94005
5	1034	1737	0,85188
6	1467	3204	0,72678
7	1704	4908	0,58148
8	1903	6811	0,41920
9	1690	8501	0,27509
10	1250	9751	0,16850
11	791	10 542	0,10105
12	561	11 103	0,05321
13	325	11 428	0,02550
14	156	11 584	0,01219
15	76	11 660	0,00571
16	31	11 691	0,00307

О к о н ч а н и е т а б л. 1

1	2	3	4
17	17	11 708	0,00162
18	10	11 718	0,00077
19	6	11 724	0,00026
20	1	11 725	0,00017
21	1	11 726	0,00009
22	1	11 727	0,00000

Как видим, при выделении функционального ядра можно опираться на слова длиной от 2 до 5 букв.

На рисунке хорошо видно, что максимум слов приходится на длину в 8 букв.

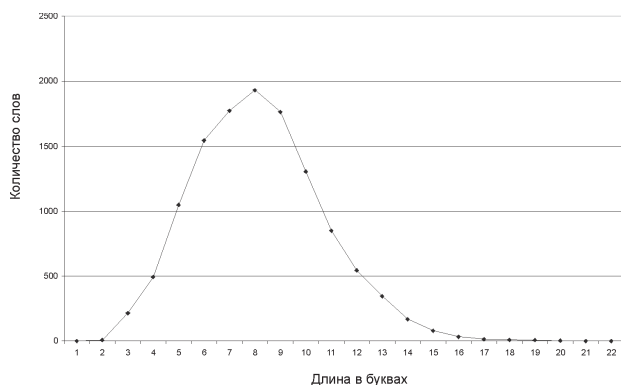


Рисунок. Распределение слов хорватского языка по длине (в буквах)

Приведем примеры слов хорватского языка, имеющих наибольший функциональный вес. Это двухбуквенные — *ep* ‘эпос’, *la, fa* ‘музыкальные ноты’, *os* ‘ось’, *uš* ‘вошь’, местоимения *mi* ‘мы’ и *vi* ‘вы’. Далее следуют 215 слов длиной в 3 буквы, среди которых *bič* ‘кнул, хлыст’, *bik* ‘бык’, *bob I* ‘боб’, *bob II* ‘боб, бобслей’, *Bog* ‘бог’, *bok* ‘бок’, *bol* ‘боль’, *brk* ‘ус’, *brz* ‘быстрый, скорый’, *crn* ‘черный’, *crv* ‘червяк’, *čar* ‘прелесть, очарование’, *čas* ‘мгновение, минута’, *ček* ‘чек’, *čep* ‘пробка’, *čip* ‘чип’, *dah* ‘дыхание’, *dan* ‘день’, *dar* ‘дар, подарок’, *dim* ‘дым’, *dio* ‘часть’, *dom* ‘дом, палатка’, *dug I* ‘долг’, *dug II* ‘длинный’, *gaj* ‘роща’, *gol* ‘голый’, *grk* ‘горький’, *hod* ‘ход’, *ime* ‘имя’, *iće* ‘пища, еда’, *ići* ‘идти, ходить’, *jad* ‘горе, беда’, *jak* ‘сильный, мощный’, *jal* ‘зависть’, *kor* ‘хор’, *kos I* ‘дрозд’, *kos II* ‘косой’, *krv* ‘кровь’, *kus* ‘кусочек’, *laž* ‘ложь’, *led* ‘лёд’, *let* ‘полет’, *lud* ‘сумасшедший’, *miš* ‘мышь’, *niz I* ‘ряд’, *nos* ‘нос’, *nov* ‘новый’, *oči* ‘глаза’, *osa* ‘оса’, *pas I* ‘пояс’, *pas II* ‘собака’, *peč* ‘печь’, *pod* ‘пол’, *put I* ‘тропа’, *rad* ‘работа, труд’, *raj* ‘рай’, *rat* ‘война’, *raž* ‘рожь’, *san* ‘сон’, *sat* ‘час’, *sit* ‘сытый’, *živ* ‘серый’, *tih* ‘тихий’, *tlo* ‘почва’, *tup* ‘тупой’, *živ* ‘живой’, *zlo* ‘зло’, *zub* ‘зуб’, *žut* ‘желтый’ и некоторые другие.

Второй параметр характеристики лексики — синтагматическая активность, которая косвенно оцени-

вается по количеству фразеосочетаний со словом. Мы не стали ограничиваться только ромбовидной частью словарной дефиниции, представленной очень бедно, а учли все примеры употребления лексемы. Результаты приводятся в табл. 2.

Т а б л и ц а 2

## Синтагматическая стратификация хорватской лексики

ФС	Количество слов	Накопление	С-вес
10	11	11	0,99906
9	4	15	0,99872
8	15	30	0,99744
7	18	48	0,99591
6	32	80	0,95541
5	68	148	0,98738
4	82	230	0,98039
3	185	415	0,96462
2	373	788	0,93282
1	1006	1794	0,84705
0	9935	11729	0,00000

Очевидно, что синтагматическое ядро будут составлять все слова, имеющие при себе хотя бы одно сочетание. Приведем примеры слов, имеющих наибольший синтагматический вес.

**Kuća** ‘дом’: *drvena* ~ деревянный дом; *javna* ~ публичный дом; *kamena* ~ каменный дом; *nakladna* ~ издательский дом; *prizemna* ~ одноэтажный дом; *robna* ~ универмаг; *stambena* ~ жилой дом; *trgovačka* ~ торговый дом; *višekatna* ~ многоэтажный дом; *zidana* ~ кирпичный дом; **kosjfa** ‘волосы’: *bojiti* ~ красить волосы; *bujna* ~ пышные волосы; *čma* ~ черные волосы; *ispadanje* ~ выпадение волос; *kovrčava* ~ кудрявые, курчавые волосы; *plava* ~ светлые волосы; *počešljati* ~ причесать волосы; *prati* ~ мыть волосы; *sijeda* ~ седые волосы; *suha* ~ сухие волосы; *valovita* ~ волнистые волосы; **rečenica** ‘предложение’: *glavna* ~ главное предложение; *izjavna* ~ повествовательное предложение; *jednostavna* ~ простое предложение; *nezavisno složena* ~ сложносочиненное предложение; *niječna* ~ отрицательное предложение; *pogodbena* ~ условное предложение; *poticajna* ~ побудительное предложение; *složena* ~ сложное предложение; *upitna* ~ вопросительное предложение; *usklična* ~ восклицательное предложение; **meso** ‘мясо’: *kuhano* ~ отварное мясо; *meso na ražnju* ~ мясо на вертеле; *meso žaropekuroštilju* ~ мясо на гриле (барбекю); *miješano* ~ мясное ассорти (горячее блюдо); *mljeveno* ~ рубленое мясо; *pečeno* ~ запеченное [жареное] мясо; *prijano* ~ тушеное мясо; *pohano* ~ мясо в панировке; *suho* ~ копченое мясо; *topovsko* ~ фраз. пушечное мясо; **stil** ‘стиль’: *administrativni* ~ административный стиль; *funkcionalni* ~ функцио-

нальный стиль; *gotički* ~ готический стиль; *književnoumjetnički* ~ литературно-художественный стиль; *novinski* ~ газетный стиль; *pjesnički* ~ поэтический стиль; *poslovni* ~ деловой стиль; *publicistički* ~ публицистический стиль; *razgovorni* ~ разговорный стиль; *to je njezin* ~ это ее стиль; *stvar* 'вещь': *stvar je krenula* дело пошло; *stvar je u tome da...* дело в том, что...; *nazvati ~i pravimmenom* называть вещи своими именами; *prijedimo na* ~ перейдем к делу; *osobne -i* личные вещи; *propala* ~ пропавшее [гибкое] дело; *staviti ~i u prtljajnik* ~ положить вещи в багажник; *školske ~i* школьные вещи; *što je na ~i?* в чем дело?; *to je ~ navike* это дело привычки; *stanje* 'состояние, положение': *bračno* ~ семейное положение; *čvrsto* ~ твердое состояние; *financijsko* ~ финансовое положение; *imovno* ~ имущественное положение; *izvanredno* ~ чрезвычайное положение; *opsadno* ~ осадное положение; *pravno* ~ правовое положение; *ratno* ~ военное положение; *tekuće* ~ жидкое состояние; *teško* ~ тяжелое состояние; *pravo* 'право': *biračko* ~ избирательное право; *državno* ~ государственное право; *imovinsko* ~ имущественное право; *građansko* ~ гражданское право; *kazneno* ~ уголовное право; *međunarodno* ~ международное право; *nasljedno* ~ наследственное право; *obiteljsko* ~ семейное право; *pomorsko* ~ морское право; *povijesno* ~ историческое право; *rjecnik* 'словарь': *častotni* ~ частотный словарь; *dvojezični* ~ двуязычный словарь; *džepni* ~ карманный словарь; *etimološki* ~ этимологический словарь; *frazološki* ~ фразеологический словарь; *jednojezični* ~ одноязычный словарь; *mali* ~ краткий словарь; *nazivoslovni* ~ терминологический словарь; *normativni* ~ нормативный словарь; *objasmbeni* ~ объяснительный словарь; *crn* 'черный': *crna boja* ~ черная краска; *crna kava* ~ черный кофе; *crna Marica* ~ черный ворон; *crne misli* ~ черные мысли; *crne oči* ~ черные глаза; *crni dani* ~ черные дни; *crni humor* ~ черный юмор; *crno pivo* ~ темное пиво; *crno tržište* ~ черный рынок; *crno zlato* ~ черное золото; *električni* 'электрический': *električna mreža* ~ электросеть; *električna peć* ~ электропечь; *električna stolica* ~ электрический стул; *električna struja* ~ электрический ток; *električni krug* ~ электроцепь; *električni napon* ~ электрическое напряжение; *električni otpor* ~ электрическое сопротивление; *električni stroj* ~ электромашинка; *električni vlak* ~ электропоезд; *električno brojilo* ~ электрический счетчик.

Результаты обработки словаря по эпидигматическому параметру представлены в табл. 3.

Из таблицы видно, что в эпидигматическое ядро придется включать все слова, имеющие при себе минимум два значения. Заметим, что такая ситуация является типичной для двуязычных словарей малого объема.

Т а б л и ц а 3

Эпидигматическая стратификация хорватской лексики

Значения	Количество слов	Накопление	Э-вес
5	2	2	0,99983
4	30	32	0,99727
3	179	211	0,98201
2	1130	1341	0,88568
1	10 389	11 730	0

По пять значений имеет видовая пара *odbiti-odbijati* 'отбить-отбивать': 1) отбить; 2) отразить; 3) отказать в чем-либо, отказаться от чего-либо; 4) вычистить; 5) отучить.

Далее следуют сразу тридцать слов с четырьмя значениями: *zametnuti* 'затеять, начать'; *zakuhati* 'заварить'; *prijava* 'прописка'; *list* 'лист'; *tjelesni* 'физический, телесный'; *obljubljivati* 'любить'; *izdati* 'выдать, издать'; *poziv* 'звонок'; *provaljivati* 'взламывать'; *rampa* 'платформа'; *pogonom* 'привод, передача'; *provaliti* 'взломать'; *jabučica* 'яблочко'; *krunica* 'коронка'; *podignuti* 'поднять'; *podizati* 'поднимать'; *gajba* 'клетка'; *kocka* 'куб'; *zakuhavati* 'заваривать'; *broj* 'число'; *oklop* 'броня'; *obljubiti* 'полюбить'; *zaleđe* 'тыл'; *donijeti* 'принести'; *ploča* 'плита'; *izdavati* 'выдавать'; *provala* 'взлом'; *donositi* 'приносить'; *pokretač* 'двигатель'; *patron* 'патрон, покровитель'.

И наконец, последний параметр – парадигматический. Он определяется количеством синонимов у той или иной лексемы.

Данные по парадигматическому параметру были получены методом позиционного анализа дефиниций [6]. Результаты представлены в табл. 4.

Т а б л и ц а 4

Парадигматическая стратификация хорватской лексики

Количество слов	Количество рядов	Накопление	П-вес
7	2	2	0,99983
6	1	3	0,99975
5	4	7	0,99941
4	15	22	0,99816
3	79	101	0,99155
2	925	1026	0,91413
1	10 922	11 948	0,00000

Лидером парадигматики по материалам словаря стали существительные, объединенные значениями *'девушка'*: *cura, djevojka, gerla, gospođica, puca, treba, udavača* и *'кусоч, ломоть'*: *komad, kus, zalogaji, kocka, peča, režanj, feta*. Далее идет семема *'разделить'*: *razdvojiti, razdijeliti, rastaviti, kalati, lamati, razbijati* с синонимическим рядом из шести синонимов. Затем

семемы 'бить', 'болтать', 'еда', 'умереть' с рядами в пять синонимов.

В парадигматическое ядро включаются все лексеммы, образующие синонимические ряды.

В результате слияния частных ядер, полученных по каждому из четырех параметров, мы получаем

четыре множества слов, из которых наибольший интерес представляют два множества: слова, вошедшие во все четыре ядра, и слова, вошедшие в три частных параметрических ядра. Общее количество таких слов для хорватского языка – 345, причем малое ядро – 53 слова (оно приводится в табл. 5).

Таблица 5

Малое ядро хорватской лексики по двуязычному словарю

Слово	Значение	Д	П	С	Ф	Σ
1	2	3	4	5	6	7
glas	голос	0,98201	0,91413	0,98738	0,94005	3,82357
gust	густой	0,98201	0,91413	0,98738	0,94005	3,82357
kola	повозка	0,98201	0,91413	0,98738	0,94005	3,82357
broj	число	0,99727	0,91413	0,96462	0,94005	3,81607
boja	цвет, окраска	0,88568	0,99155	0,99744	0,94005	3,81472
ploča	плита	0,99727	0,99155	0,96462	0,85188	3,80532
imati	иметь	0,98201	0,99816	0,96462	0,85188	3,79667
list	лист	0,99727	0,91413	0,93282	0,94005	3,78427
sat	час	0,88568	0,91413	0,99744	0,98039	3,77764
ljut	злой, сердитый	0,98201	0,91413	0,93282	0,94005	3,76901
poći	пойти	0,98201	0,91413	0,93282	0,94005	3,76901
vući	волочить, тащить	0,98201	0,91413	0,93282	0,94005	3,76901
zbor	корпус	0,98201	0,91413	0,93282	0,94005	3,76901
komad	кусок	0,98201	0,99983	0,93282	0,85188	3,76654
led	лед	0,88568	0,91413	0,98039	0,98039	3,76059
udar	удар	0,88568	0,91413	0,98738	0,94005	3,72724
čep	пробка	0,98201	0,91413	0,84705	0,98039	3,72358
dah	дыхание	0,98201	0,91413	0,84705	0,98039	3,72358
čast	честь	0,88568	0,91413	0,98039	0,94005	3,72025
sлом	провал, неудача, крах	0,88568	0,91413	0,98039	0,94005	3,72025
ići	идти, ходить	0,88568	0,91413	0,93282	0,98039	3,71302
kraj	край	0,88568	0,91413	0,96462	0,94005	3,70448
kocka	куб	0,99727	0,91413	0,93282	0,85188	3,69610
poziv	звонок (телефонный)	0,99727	0,91413	0,93282	0,85188	3,69610
kapa	головной убор	0,88568	0,91413	0,93282	0,94005	3,67268
tvrd	твердый, жесткий	0,88568	0,91413	0,93282	0,94005	3,67268
kotač	колесо	0,98201	0,99155	0,84705	0,85188	3,67249
jelo	еда, пища	0,88568	0,99941	0,84705	0,94005	3,67219
tući	бить, ударять	0,88568	0,99941	0,84705	0,94005	3,67219
brdo	гора (небольшая)	0,88568	0,99155	0,84705	0,94005	3,66433
krema	крем	0,88568	0,91413	0,99591	0,85188	3,64760
izbor	выбор, отбор	0,88568	0,91413	0,98738	0,85188	3,63907
tanak	тонкий	0,88568	0,91413	0,98039	0,85188	3,63208
dom	дом	0,88568	0,91413	0,84705	0,98039	3,62725
muš	муж	0,88568	0,91413	0,84705	0,98039	3,62725
izlaz	выход	0,88568	0,91413	0,96462	0,85188	3,61631
junak	герой	0,88568	0,91413	0,96462	0,85188	3,61631
naići	появиться	0,88568	0,91413	0,96462	0,85188	3,61631

1	2	3	4	5	6	7
volja	воля	0,88568	0,91413	0,96462	0,85188	3,61631
čemer	попынь	0,98201	0,91413	0,84705	0,85188	3,59507
čudo	чудо	0,88568	0,91413	0,84705	0,94005	3,58691
jeza	дрожь, озноб	0,88568	0,91413	0,84705	0,94005	3,58691
utok	устье	0,88568	0,91413	0,84705	0,94005	3,58691
leden	ледяной, леденящий	0,88568	0,91413	0,93282	0,85188	3,58451
miran	мирный	0,88568	0,91413	0,93282	0,85188	3,58451
popis	список	0,88568	0,91413	0,93282	0,85188	3,58451
zгода	удобный случай	0,88568	0,91413	0,93282	0,85188	3,58451
svući	снять	0,88568	0,99155	0,84705	0,85188	3,57616
kugla	шар	0,88568	0,91413	0,84705	0,85188	3,49874
matun	кирпич	0,88568	0,91413	0,84705	0,85188	3,49874
molba	просьба	0,88568	0,91413	0,84705	0,85188	3,49874
staza	тропинка	0,88568	0,91413	0,84705	0,85188	3,49874
ureći	ужалить, укусить	0,88568	0,91413	0,84705	0,85188	3,49874

Доминантой хорватской лексики стали сразу 3 слова, имеющие одинаковый параметрический вес: *glas* – голос, *gust* – густой и *kola* – повозка.

Таким образом, мы получили возможность статистически описать словарь хорватского языка. Более детальное рассмотрение содержательного наполнения каждого из представленных параметров даст возможность сопоставления лексики хорватского и других языков с учетом системных связей и характеристик.

#### ЛИТЕРАТУРА

1. *Тафра Б.* Хорватская лексикография / Б. Тафра, Ж. Финк-Арсовски // Славянская лексикография : международная коллективная монография / отв. ред. М. И. Чернышева. – М. : Азбуковник, 2013. – С. 310–340.

*Воронежский государственный университет*  
*Меркулова И. А., кандидат филологических наук,*  
*доцент кафедры теоретической и прикладной лингвистики*  
*E-mail: igell@yandex.ru*  
*Тел.: 8 (473) 220-41-49*

2. *Тафра Б.* Хорватско-славянские лексикографические контакты / Б. Тафра // *Studia z Filologii Polskiej i Słowiańskiej.* Issue no. 42. – Warszawa, 2007. – S. 225–245.

3. *Багдасаров А. Р.* Хорватско-русский словарь / А. Р. Багдасаров. – М. : Астрель-АСТ : Транзиткнига, 2003. – 192 с.

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

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

6. *Кретов А. А.* Позиционный анализ парадигматики романско-русских словарей / А. А. Кретов, В. Т. Титов // *Проблемы компьютерной лингвистики : сб. науч. трудов.* – Воронеж : Издат.-полиграф. центр Воронеж. гос. ун-та, 2008. – Вып. 3. – С. 266–282.

*Voronezh State University*  
*Merkulova I. A., Candidate of Philology, Associate*  
*Professor of the Theoretical and Applied Linguistics Department*  
*E-mail: igell@yandex.ru*  
*Tel.: 8 (473) 220-41-49*