О СПОСОБАХ ЛЕКСИКОГРАФИЧЕСКОГО ОПИСАНИЯ РУССКОЙ СПОНТАННОЙ РЕЧИ (КОРПУСНОЕ ИССЛЕДОВАНИЕ)

Осьмак Н. А. (natalia.osmak@yandex.ru)

Санкт-Петербургский государственный университет
Санкт-Петербург, Россия

В докладе представлены результаты анализа возможных способов лексикографического описания спонтанной речи в целях создания словаря русской бытовой речи. Материалом для анализа стал блок «Один речевой день» Звукового корпуса русского языка. Работа над словарем русской бытовой разговорной речи может осуществляться через описание функционирования наиболее частотных лексем/словоформ или через лексикографический анализ отдельных лексико-семантических групп. В статье последовательно рассматриваются оба способа, а также анализируются их достоинства и недостатки.

Ключевые слова: номинация женщины, номинация мужчины, корпусное исследование, проблемы лексикографии, спонтанная речь

METHODS OF LEXICOGRAPHIC DESCRIPTION OF RUSSIAN SPONTANEOUS SPEECH (A CORPORA RESEARCH)

Os'mak N. A. (natalia.osmak@yandex.ru)

Saint Petersburg State University,
7–9, Universitetskaya nab.,St.Petersburg, 199034

The paper presents the results of an analysis of possible methods of lexicographical description of Russian spontaneous speech aimed at the creation of a General Russian Conversational Dictionary. The analysis is based on the material of the Corpus of Spoken Russian "One Speech Day", which presents unique linguistic material, enabling fundamental research in many aspects: study of real spontaneous speech, phonetics and grammar of spoken language, psycholinguistics, communication studies, etc. The General Russian Conversational Dictionary can be created through the description of most frequent units or through the analysis of lexical semantic groups. Advantages and disadvantages of the two methods are analyzed. Tentative lexicographical description of a lexical semantic group representing nominations of humans are given.

Key words: nomination of humans, , corpus study, lexicography, spontaneous speech

В настоящее время, наряду с существующей академической лексикографией, ставящей перед собой в качестве цели «отражение только тех фактов лексики и семантики, которые характерны для стандартного современного общелитературного, общеразговорного употребления как в его письменной, так и в устной форме; ничего индивидуального, ничего окказионального» (Герд 1990: 29–30), стала появляться традиция «речевой лексикографии», основные принципы которой до конца еще не разработаны. Так, все острее ощущается необходимость создания специального словаря, представляющего реальную речевую ситуацию, те лексические и грамматические единицы, которые уже давно фиксируются в различных корпусах и базах данных, однако в большинстве словарей до сих пор не находят отражения. Данная задача, в свою очередь, требует обращения непосредственно к устной спонтанной речи, поскольку именно она является наиболее полным и ярким отражением состояния речи на настоящий момент.

Такую возможность предоставляет Звуковой корпус русского языка (ЗКРЯ), особенно та его часть (блок «Один речевой день» — ОРД), которая фиксирует повседневную бытовую речь[1] носителей русского языка в ее наиболее естественной

[1] В настоящий момент материал ОРД представляет собой около 300 тыс. словоупотреблений в расшифровках, около 400 часов звучания, записанных у 50 информантов и 600 коммуникантов.

форме. Методика 24-часовой записи, используемая при формировании данного блока ЗКРЯ, обеспечивает почти полную свободу говорящего от диктофона и от тех или иных коммуникативных заданий (чтение, пересказ, описание изображения, рассказ на заданную тему), которые неизбежно усложняют его задачу и обеспечивают «на выходе» хоть и вполне спонтанную, но все же экспериментальную речь (см. подробнее о Звуковом корпусе: Богданова и др. 2008; Степанова 2008). В пользу обращения к ЗКРЯ говорит и то, что именно корпусный подход «отражает точку зрения многих лингвистов на то, каким должно быть адекватное "эмпирическое" изучение языка» (Meyer 2004: 141), а также позволяет выявить новые тенденции, новые языковые явления, отражающиеся, прежде всего, в индивидуальной речи конкретных носителей языка.

Материал Звукового корпуса может быть использован для различных аспектов лексикографического описания спонтанной речи, в том числе для создания, например, словарей контекстных экспрессем русской разговорной речи, дискурсивных единиц и редуцированных форм (см.: Богданова, Осьмак 2011). Однако наиболее перспективной представляется работа над ***Словарем русской бытовой разговорной речи***, поскольку ни один из существующих словарей не дает возможности в полной мере проследить особенности употребления лексических единиц в современной спонтанной речи носителей языка.

Представляется, что работа по составлению такого словаря может осуществляться различными способами. Во-первых, через описание функционирования наиболее частотных лексем/словоформ, а во-вторых, через лексикографический анализ отдельных лексико-семантических групп. Рассмотрим каждый из этих способов подробнее.

Стоит отметить, что целью настоящей статьи является не проспект нового словаря и не полномасштабный лексический анализ тех или иных единиц или групп единиц, а обсуждение различных способов лексикографического описания русской спонтанной речи.

На первый взгляд, обращение к наиболее частотным речевым единицам кажется самым логичным способом начать лексикографическое описание русской устной спонтанной речи. Это подразумевает выявление основных проблем такого описания, а также конкретных тенденций лексических изменений. Более подробный анализ «верхушки» частного словника, полученного в результате обработки записей корпуса «Один речевой день», позволяет судить о том, насколько данный способ применим в целях составления словаря бытовой разговорной речи.

В таблице ниже представлены 150 словоформ этой «верхушки», ранжированных по частотности.

Таблица 1. Наиболее частотные словоформы русского языка (по материалам корпуса «Один речевой день»)

Ранг	Слово	Кол-во	Ранг	Слово	Кол-во	Ранг	Слово	Кол-во
1	я	5 823	51	за	544	100,5	для	294
2	в	5 598	52	здесь	539	102	от	281
3	не	5 480	53,5	будет	538	103	сегодня	273
4	вот	5 195	53,5	очень	538	104	тогда	267
5	ну	5 146	55	потом	537	105	нам	264
6	да	5 096	56	когда	530	106	ж	261
7	а	4 867	57	тебе	519	107	быть	255
8	что	4 460	58,5	такой	512	108,5	три	254
9	и	4 213	58,5	тут	512	108,5	пока	254
10	это	3 655	60	говорит	506	110	почему	250
11	то	3 386	61	или	499	111	них	248
12	там	3 267	62	его	493	112	была	245
13	у	3 103	63	может	485	113	думаю	244
14	с	3 087	64	значит	484	114	такие	242
15	э	2 894	65	чего	483	115	него	241
16	на	2 736	66	хорошо	475	116	какие	240
17	так	2 671	67	ой	474	117	сколько	239
18	как	2 240	68	знаешь	468	118	ему	234
19	ты	1 798	69	тебя	458	110	во	227
20	всё	1 754	70	только	452	120	этого	225
21	нет	1 655	71,5	давай	440	121,5	пять	216
22	он	1 611	71,5	говорю	440	121,5	понятно	216
23	мне	1 284	73	потому	438	123	типа	212
24	угу	1 248	74	можно	429	124	слушай	211
25	она	1 212	75	кто	406	125	нормально	210
26	есть	1 157	76	ничего	405	126	куда	208
27	меня	1 118	77	ага	395	127	эта	207
28	по	1 103	78	этот	394	128	общем	204
29	они	1 069	79	короче	386	129	время	202
30	мы	1 066	80	конечно	377	130	наверное	201
31	сейчас	1 058	81,5	такая	364	131,5	могу	196
32	еще	1 011	81,5	раз	364	131,5	туда	196
33	уже	956	83	до	353	133	много	192
34	но	935	84	о	350	134,5	ей	191
35	бы	927	85	вас	349	134,5	х…й	191
36	надо	923	86	такое	340	136	самое	190
37	же	893	87,5	какой	335	137,5	понимаешь	189
38	просто	864	87,5	их	335	137,5	нужно	189
39	м	782	89	чтобы	332	139	эту	186
40	вообще	763	90	даже	330	140	этом	184
41	к	701	91	блин	325	141	чем	182
42	знаю	693	92	эти	323	142	какая	181

Ранг	Слово	Кол-во	Ранг	Слово	Кол-во	Ранг	Слово	Кол-во
43	нас	682	93	два	318	143	сюда	180
44	если	661	94,5	её	313	144	двадцать	179
45	вы	639	94,5	из	313	144	со	179
46	тоже	629	96	был	306	144	хочу	179
47	все	609	97	один	302	147	вчера	176
48	бл..ь	599	98	вам	301	148	моему	173
49	где	588	99	ли	300	149	больше	170
50	было	562	100,5	ладно	294	150	четыре	169

Рассмотрим подробнее первые десять позиций этого частотного словника. Как видно из таблицы, самым частотным оказалось местоимение *я* (5 823 употреблений). Второе место занимает предлог *в* (5 598 употреблений). Далее следует отрицательная частица *не* (5 480 употреблений), при этом в некоторых ситуациях *не* является редуцированной формой слова-предложения *нет* и иногда употребляется в составе конструкций ***не**-а* и ***не-не-не***, например:

- ***не*** */ ну она мне сказала / что да вот / она вот случайно;*
- *а* ***не-не-не*** */ нам нужно завернуть направо / у меня уже..;*
- ***не-не-не*** *вот я уже поняла;*
- *ты читала ? #* ***не-а****.*

Кроме того, зачастую частица *не* употребляется в составе хезитационных конструкций, таких как ***не*** *это самое, (я)* ***не*** *знаю, (я)* ***(не)*** *думаю (что)* и под., которые стоит рассматривать не только в качестве единого целого при лексикографическом описании, но и в составе совсем другого словаря — например, дискурсивных единиц, или хезитационных конструкций (см.: Богданова 2011а, б; 2012; Хан 2012).

Далее в «верхушке» частотного словника следуют частицы (тоже по преимуществу хезитативы) *вот, ну, да* (5 195, 5 146 и 5 096 употреблений соответственно), уже давно привлекающие внимание лингвистов. Обратимся лишь к некоторым из последних работ. Так, например, П. Бокале (Бокале 2011а, б) говорит о преимущественной инициальной и финальной позиции слова *вот*, где оно служит для установления контакта либо, соответственно, маркирует завершение планируемого отрезка речи. Ю. В. Дараган выделяет следующие функции данного слова: 1) концептуализация; 2) контроль; 3) оценка; 4) апелляция (Дараган 2011). Особо хотелось бы отметить работу А. И. Кислощук, в которой выделены такие функции дискурсивного слова *вот* в устной спонтанной речи (на материале сбалансированного блока ЗКРЯ), как указательная и усилительная частицы (как правило, в составе различных конструкций, не зафиксированных в существующих словарях), поиск слова или иного продолжения речи, заполнитель паузы хезитации и дискурсивный маркер (Кислощук 2011).

Частица *ну*, являясь текстовым коннектором, часто употребляется в начале фрагмента текста, что, по мнению О. А. Лаптевой, является «маркированной функцией» (Лаптева 1999: 492). Кроме того, *ну* может использоваться для уточнения сказанного или для побуждения, может стимулировать продолжение прерванного высказывания, выражать отказ от первоначального намерения говорить или выступать в противительно-выделительном значении, заменяя при этом союз *но* (Bocale 2011б).

Частица *да* используется в ОРД в значениях, фигурирующих в толковых словарях:

- ***да*** *(:) нашли //* — выражение утверждения, согласия;

- *а я был один раз на () маслобойке вот на Украине //* ***да ?*** *в Донецке () куда я ездил // мы возили семечки в обмен на масло такое знаешь () ну ароматное //* — выражение вопроса при желании получить подтверждение чему-либо;

- ***да*** *ладно ! *П чем помогала //* — придание высказыванию большей выразительности.

Кроме этого, *да* встречается в качестве заполнения пауз хезитации, как самостоятельно, так и в составе хезитационных конструкций *да вот*, *ну да*:

- *как приводим (...) к единой шкале /* ***да*** */ к стандартизированным баллам / мы открываем с вами страницу / триста восемнадцать*;

- *дальше / вот пока с этим понятно /* ***да*** */ как в стены переводить ? #* ***ну да*** *// # вам понятно ? @ вам всем там / понятно ?*;

- *у нашего испытуемого /* ***да вот*** */ *П мы смотрим средний /* ***да*** *... *П пять баллов / пять стенов / средний уровень (...) невротизации / то есть / по сути дела что мы можем *Н как его охарактеризовать*.

Текстовый коннектор *а*, служащий для оформления вопроса, сцепления реплик и введения новой темы является, по мнению П. Бокале, одним из наиболее часто встречающихся, особенно в стартовой позиции (Бокале 2011б):

- ***а*** *интересно ...*;

- *так /* ***а*** *теперь куда ?*

- ***а*** *зажигалку дала ?*

- ***а*** *теперь куда / ты видишь ?*

- *четыре ... / **а** четыре вот здесь.*

Финальное *а,* сигнализирующее о просьбе реакции, отклика со стороны собеседника, встречается реже:

- *почему панели такие узкие / **а** ?*

- *ну мне кажется / мне надо больше # какой нам в коридор лучше / **а?***

Восьмое и девятое места в «верхушке» частотного словника ОРД занимают служебные слова *что* и *и* (4 460 и 4 213 употреблений соответственно). Наконец, замыкает десятку местоимение *это* (3 655 употреблений), которое также входит в ряд таких хезитационных конструкций, как *не это самое*, *это самое* (см. Богданова 2011б; 2012), да и само может быть хезитативом.

Таким образом, беглый взгляд на «частотную верхушку» ОРД дает понять, что лексикографическое описание по этому пути приведет к описанию не лексического состава речи, а, скорее, только дискурсивных или служебных слов.

Первый полноценный глагол в частотном словнике ОРД — *знаю* (693 употреблений; 42 место) При этом ни одного полноценного существительного в списке 150 самых частых слов не обнаружилось вообще. На 48 месте оказалось слово *бл…ь* (599), сопоставимы с ним по употребительности следующие единицы:

  - *Блин* — 325 (91 место);

  - *Типа* — 212 **(**123 место);

  - *Время* — 202 (129 место).

При этом *блин* употребляется в ОРД на 100 % в роли не существительного, а междометия (просторечного, грубого, бранного):

- *нам надо было **блин** ! где Мойка ? как проще вернуться ? здесь ? туда ? или с этой стороны ?;*

- *вот дело в том что знаешь / это в общем короче да / и муж / вот этой подруги взвился говорит / ты вообще там идиотка **блин** / и я это всё видел.*

Частеречная принадлежность слова *типа* может существенно различаться в зависимости от контекста. Данная лексема может выступать в качестве и частицы, и союза, и существительного, и даже для заполнения паузы хезитации (Богданова 2010а: 6–8).

Схожая ситуация наблюдается со словоформой *время*, которая, наряду с традиционным употреблением в качестве имени существительного, встречается и в других — адвербиальных — значениях (в составе различных

конструкций), среди которых *в ближайшее время* (= совсем скоро), *время от времени* (= иногда), *все время* (= постоянно, не переставая), *на время* (= временно, на какой-то срок) и т. д. (Богданова 2010а: 9–11).

Несмотря на кажущуюся логичность этого пути лексикографического описания устной спонтанной речи — обращения к «верхушке» частотного словника как к основному материалу такого описания — от него пока приходится отказаться. Это объясняется, как уже упоминалось выше, тем, что наиболее частотные лексемы словника ОРД представляют собой, в основном, служебные слова, хезитативы и другие дискурсивные единицы, а значит, лексикографическое описание таких «слов» приведет к созданию не словаря русской разговорной бытовой речи, а, скорее, словаря служебных или дискурсивных единиц. Второе же направление — анализ отдельных лексико-семантических групп — напротив, ставит во главу угла семантику слова, а не его функции, и позволяет проследить семантические, коннотативные и стилистические изменения в употреблении тех или иных лексем. Работа в этом направлении уже позволила получить ряд интересных результатов.

Материалом для пилотного исследования стали две группы русских слов сравнительно небольшого объема — единицы, обозначающие лиц женского и мужского пола. Критерием отбора материала для анализа было наличие у слова наиболее общего значения[2]:

1) общие наименования (*женщина*, *баба*, *мужчина, мужик, дядя*);
2) возрастные наименования (*девочка*, *бабушка, юноша, старик*);
3) наименования с учетом семейного статуса — замужняя/женатый (*женщина — девушка*, *девка, мужчина — парень*)[3].

Первым этапом исследования стало создание корпуса конкретного материала — методом сплошной выборки из компьютерной версии «Малого академического словаря» (далее — МАС)[4]. В итоговый список вошли 40 лексем для лиц женского пола и 38 — для лиц мужского пола.

Следующим шагом анализа стало обращение к корпусу ОРД. После расшифровки записей в программе e-Kar (электронная картотека) был составлен конкорданс всех интересующих нас единиц. Затем данные были экспортированы в программу eLex, позволяющую составлять словарные статьи, подключая контексты ОРД, а также значения слов по МАС. Интерфейс данной программы выглядит следующим образом (см. рисунок).

---

[2] За пределами исследования были, соответственно, оставлены многие слова, имеющие схожую семантику, такие, например, как термины родства, наименования лица с учетом внешности, социального положения, особенностей характера и т. п. Представляется, что все эти группы слов, отражающих лингвокультурологические особенности определенной эпохи, требуют отдельного изучения.

[3] В данной работе не рассматривались также слова, которые могут быть использованы для обозначения лица как мужского, так и женского пола (ребенок, ребята, ребятишки, дитя и пр.), а также все единицы — обращения к лицу.

[4] Евгеньева А. П. (1999). Выбор данного словаря объясняется тем, что в дальнейшем в исследовании использовалась компьютерная программа eLex, предназначенная для составления словарных статей на основе значений слов именно по МАС.

Многие единицы из исходного списка в ОРД попросту не встретились, и материал для дальнейшего лексикографического анализа сократился до 16 лексем в первом случае и 12 во втором.

Результаты исследования показали, что при попытке лексикографического описания слов, обозначающих лицо мужского или женского пола, наблюдаются следующие явления:

1) совпадение значений МАС и значений, выделяемых из контекстов ОРД (*мужчина, мальчишка, женщина, девушка*);

2) появление новых лексических единиц данного типа, не зафиксированых в МАСе (*дедун, мужичок, мужчинка, бабец, бабуська, бабулька, молодой человек*);

3) особые случаи сочетаемости лексем (*дед, дядя, тетя*);

4) отсутствие значения по ОРД в МАСе (*мальчик, девочка, девчонка*).

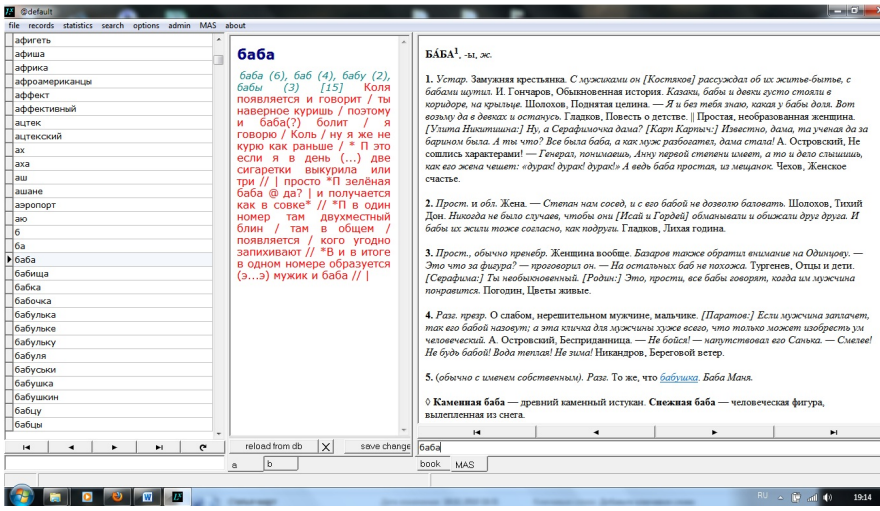

Рис. 1. Окно программы eLex, позволяющей сопоставлять данные ОРД и МАС

Ряду лексем из данных групп уже посвящены отдельные исследования. Так, например, было доказано, что в современной устной русской речи параллельно существуют две противоположные коннотации слова *девка*, ранее не зафиксированные в словарях: негативная ('девица или легкого поведения, или малообразованная, из нижних слоев общества или неблагополучной семьи') и положительная ('хорошая, искренняя, прямая девушка') (см. Осьмак 2011). Новую коннотацию в разговорной спонтанной речи получила и лексема *старушка*. Традиционно толковые словари определяют *старушку* как женщину, достигшую старости. Однако в ОРД встречается употребление данной лексемы в качестве обращения-обзывания:

- ***старушка*** *// сама ты* ***старушка*** *// поэтому ...* ***старушка****(?) / да(?)*

Анализ лексем *девочка/девчонка* и *мальчик* показал, что у них появились новые значения, которые можно охарактеризовать следующим образом:

1) *девочки/девчонки* — ‘лицо женского пола одного возраста с говорящим или младше’; характерно преимущественно для женской речи, а также для речи определенных возрастных групп мужчин (до 25 и после 60 лет);
2) *мальчик* — а) ‘лицо мужского пола одного возраста с говорящим или младше’, характерно только для женской речи; б) ‘лицо мужского пола младше говорящего’, характерно для мужской речи (Осьмак 2012).

Случаи особой сочетаемости лексем *дядя* и *тетя* (по преимуществу с именем собственным) привели к появлению таких новых фразеологизмов, как *тетя Мотя* (см. о подобных сочетаниях слов *дядя* и *тетя*: *Богданова* 2010б, в, 2011в).

Отдельный интерес вызывают лексемы, выявленные в ОРД, но не представленные в МАС. В основном это типичные диминутивы от *бабуля*, *бабуся, мужик* и *мужчина*: *бабулька*, *бабуська*, *мужичок* и *мужчинка*, ср.:

- *я говорю конечно / если она с ребёнком не умеет разговаривать / что это / **бабулька**-то* ?

- *там сидят **бабуськи** такие / на остановке // у них глаза такие по полтиннику // ну представляешь / они сидят / ждут автобуса;*

- *есть тут жилые дома вообще ? (э-э) дак (э) никто не знает // а потом () какой-то вот **мужичок** мне указал / что вот вам туда надо;*

- *ну как всегда у неё // достаёт классный телефон / тут ни с того ни с сего / начинает мне говорить / вот этот **мужчинка** / я уже не раз от неё такое слышала / она вот () рассказывала // да?*

Из примеров видно, что ничего принципиально нового в значениях этих слов, по сравнению с исходными, ОРД не дает. Различия могут быть, по-видимому, только стилистические или коннотативные.

Исключение представляет лексема *мужчинка*. Результаты блиц-опроса носителей языка (обращения к их ментальному лексикону) показали, что в настоящее время под *мужчинкой* понимается ‘недомужчина’, т. е. мужчина, по каким-то характеристикам недотягивающий до настоящего мужчины. Он может быть социально неуспешным, внешне несимпатичным или инфантильным, маленького роста, плюгавеньким. Однако нельзя сказать, что отношение носителей языка к *мужчинке* исключительно пренебрежительное. Часто данное слово употребляют в ироническом значении, даже с какой-то долей симпатии. В то же время все чаще встречается употребление лексемы *мужчинка* для обозначения представителей сексуальных меньшинств. Кроме этого, большинство носителей языка отмечают характерность употребления данного слова в женской речи. Богатейший материал

для исследования значения этой лексемы предлагает и Интернет, в котором встречаются целые эссе-размышления о том, кто же такой *мужчинка*. Например:

- ***Мужчинка*** *— женское слово, редкий в русском языке пример однополого употребления. Это слово используют женщины на войне полов. Его феминистские корни не очевидны. Скорее, это взгляд сверху, лексикон высокомерия. Кто это там идет? Чего-то я не вижу… —* ***Мужчинка.*** *Плюгавенький такой, безденежный этакий. Что с него взять? Он ни на что не годится. С ним в ресторан пойти стыдно, но можно, хотя зачем? В кровать — ни за что. У него все маленькое, никудышное, жиденькое, нефотогеничное. Голый* ***мужчинка*** *— это, девки, обхохочешься!*

- *Бывают такие кадры, которых иначе и не назовёшь. Хиленькие, беспомощные, ведро воды на 2–3 метра не пронесут, не то чтоб девушку, маменькины сыночки — шагу в сторону без разрешения не сделают, того и гляди маму под окошком будут звать, такого на ручки разве что взять и пожалеть. И такие довольно часто встречаются, между прочим. Как их иначе назвать? — хлюпик? лучше* ***мужчинка****. И вроде обидно (должно быть для нормального мужчины) и без прямых оскорблений.*

- *Слово замечательное! Есть «мужчина» — это тот, за кем как за каменной стеной, добытчик и защитник. А есть «****мужчинка****» — такое нечто, мужеского пола, маменькин сынок до старости…*

Все эти примеры подтверждают необходимость более углубленного изучения данной лексемы.

Таким образом, становится очевидным, что работа с отдельными лексико-грамматическими группами слов дает богатейший материал для анализа семантических, коннотативных и даже стилистических изменений в функционировании слова в речи. В большинстве случаев для составления полноценной словарной статьи требуется дополнительное обращение к ментальному лексикону носителей языка через серию экспериментов в форме опроса — например, электронного — большого числа носителей русского языка, однако несомненно, что подобный лексикографический анализ следует провести в отношении и других лексико-грамматических групп, что и позволит, в конечном счете, составить ***словарь русской бытовой речи.***

Литература

1. *Богданова Н. В.* (2010а) Частотный словник ОРД — взгляд с «высоты птичьего полета». Материалы XXXIX Международной филологической конференции. Выпуск 23. Полевая лингвистика. Интегральное моделирование звуковой формы естественных языков. СПб., 3–13.
2. *Богданова Н. В.* (2010б) Кто Вы, тетя Мотя? Мир русского слова, 4. СПб., 53–62.
3. *Богданова Н. В.* (2010в). Тети Моти, дяди Вади и проблемы перевода и преподавания русского языка. HUMANIORA: LINGUA RUSSICA. Труды по русской и славянской филологии. Лингвистика XIII. Развитие и вариативность языка в современном мире, I . Тарту, 9–26.
4. *Богданова Н. В.* (2011а) Об одном типе хезитационной вставки в русской устной спонтанной речи. Языковые категории и единицы: синтагматический аспект. Материалы девятой международной конференции. Владимир, 50–56.
5. *Богданова Н. В.* (2011б) Это самое: грамматические формы и функционирование в русской спонтанной речи. Материалы XL Международной филологической конференции. Выпуск 24. Полевая лингвистика. Интегральное моделирование звуковой формы естественных языков. СПб., 18–47.
6. *Богданова Н. В.* (2011в) Тетя Мотя в зеркале русской фразеологии. “WORT — TEXT — ZEIT”. Die slawische Phrazeologie in onomasiologischer, linguokultureller und phrazeographischer Sicht. Greifswalder Beiträge zur Slawistik X . Greifswald, 23–24.
7. *Богданова Н. В.* (2012) Про не это самое. Слово. Словарь. Словесность: Литературный язык вчера и сегодня (к 300-летию со дня рождения М. В. Ломоносова). Материалы ежегодной Всероссийской научной конференции. СПб., (в печати).
8. *Богданова Н. В., Бродт И. С., Куканова В. В., Павлова О. В., Сапунова Е. М., Филиппова Н. С.* (2008) О «корпусе» живой речи: принципы формирования и возможности описания. Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной международной конференции «Диалог 2008». М., 57–61.
9. *Богданова Н. В., Осьмак Н. А.* (2011) О некоторых лексических «открытиях» на материале русской спонтанной речи (корпусное исследование). Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции «Диалог 2011». М., 110–124.
10. *Бокале П.* (2011а) Дискурсивные слова в разговорной речи // russian.slavica.org/article9375.html.
11. *Бокале П.* (2011б) Роль текстовых коннекторов в молодежной речи // russian.slavica.org/article514.html.
12. *Герд А. С.* (1990) Новый академический словарь — исходные позиции и ориентиры. Актуальные проблемы разработки нового академического словаря русского языка. Л., 29–30.

13. *Дараган Ю. В.* (2011) Функции слов-«паразитов» в русской спонтанной речи. Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Междугородней Конференции «Диалог 2000». // http://www.dialog-21.ru/materials/archive.asp?id=6260&vol=6077&y=2000
14. *Евгеньева А. П.* (1999) Словарь русского языка. В 4-х т. М.
15. *Кислощук А. И.* (2011) Опыт систематизации функций слова «вот» в спонтанном монологе. Актуальные проблемы современной науки. Материалы международной научной сессии «XIII Невские чтения». СПб., 157–164.
16. *Лаптева О. А.* (1999) Живая русская речь с телеэкрана. М., 520с.
17. *Осьмак Н. А.* (2011а) Из истории лексикографического описания русской спонтанной речи (далеко ли до Даля?). Материалы XL международной филологической конференции. Выпуск 24. Полевая лингвистика. Интегральное моделирование звуковой формы естественных языков. СПб., 160–182.
18. *Осьмак Н. А.* (2011б) Новые коннотации старых слов (о лексикографическом описании русской спонтанной речи). Образовательные технологии в виртуальном лингвокоммуникативном пространстве. IV Международная виртуальная конференция по русистике, литературе и культуре. Вермонт, 166–170.
19. *Осьмак Н. А.* (2012) Новые значения старых слов (корпусное исследование на материале повседневной русской речи). (в печати).
20. *Степанова С. Б., Асиновский А. С., Богданова Н. В., Русакова М. В., Шерстинова Т. Ю.* (2008) Звуковой корпус русского языка повседневного общения «Один речевой день»: Концепция и состояние формирования. Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной международной конференции «Диалог-2008». М., 488–494.
21. *Хан Н. А.* (2012) Конструкция (я) не знаю в русской спонтанной речи: соотношение разных функциональных типов. Мир русского слова. СПб., (в печати).
22. *Meyer C. F.* (2004) English Corpus Linguistics: An Introduction. Cambridge.

References

1. *Bogdanova N. V.* (2010a) Frequency Word-list of ORD from the Bird's-Eye View[Chastotnyi slovnik ORD — vzgliad s “vysoty ptich'ego poleta”]. Materialy XXXIX Mezhdunarodnoi filologicheskoi konferentsii. Vypusk 23. Polevaia lingvistika. Integral'noe modelirovanie zvukovoi formy estestvennykh iazykov (Field Linguistics. Integral Modeling of Sound Form of Natural Languages. Issue 23.: Proceedings of International XXXIX Philological Conference. 2010). St.Petersburg, 3–13.
2. *Bogdanova N. V.* (2010b) Who are you, Tetia Motia? [Kto Vy, tetia Motia?] Mir russkogo slova (World of Russian Word),4. St.Petersburg, 53–62.
3. *Bogdanova N. V.* (2010v) Teti Moti, Diadi Vadi and Problems of Translation and Teaching Russian Language [Teti Moti, diadi Vadi i problemy perevoda i prepodavaniia russkogo iazyka]. HUMANORIA: LINGUA RUSSICA. Trudy po russkoi

i slavianskoi filologii. Lingvistika XIII. Razvitie i variativnost' iazyka v sovremennom mire, I. Tartu, 9–26.

4. *Bogdanova N. V.* (2011a) One Type of Hesitation Filler in Russian Spoken Spontaneous Speech [Ob odnom tipe khezitatsionnoi vstavki v russkoi ustnoi spontannoi rechi]. Iazykovye kategorii i edinitsy: sintagmaticheskii aspekt. Materialy deviatoi mezhdunarodnoi konferentsii (Language Categories and Units: Syntagmatic Aspect. Proceedings of Ninth International Conference). Vladimir, 50–56.
5. *Bogdanova N. V.* (2011b). Eto samoe: Grammatical Forms and Functioning in Russian Spontaneous Speech [Eto samoe: grammaticheskie formy i funktsionirovanie v russkoi spontannoi rechi]. Materialy XL Mezhdunarodnoi filologicheskoi konferentsii. Vypusk 24. Polevaia lingvistika. Integral'noe modelirovanie zvukovoi formy estestvennykh iazykov (Field Linguistics. Integral Modeling of Sound Form of Natural Languages. Issue 24: Proceedings of International XL Philological Conference. 2011). St.Petersburg, 18–47.
6. *Bogdanova N. V.* (2011v) Tetia Motia in the Mirror of Russian Phraseology [Tetia Motia v zerkale russkoi frazeologii]. “WORT — TEXT — ZEIT”. Die slawische Phrazeologie in onomasiologischer, linguokultureller und phrazeographischer Sicht. Greifswalder Beiträge zur Slawistik X . Greifswald, 23–24.
7. *Bogdanova N. V.* (2012) About ne eto samoe [Pro ne eto samoe]. Slovo. Slovar'. Slovesnost': Literaturnyi iazyk vchera i segodnia (k 300-letiiu so dnia rozhdeniia M. V. Lomonosova). Materialy ezhegodnoi Vserossiiskoi nauchnoi konferentsii (Word. Dictionary. Literature: Literature Language Yesterday and Today (300 Years Anniversary of Birth of M. V. Lomonosov: Proceedings of Annual Russian Scientific Conference. 2011). St.Petersburg, (in print).
8. *Bogdanova N. V., Brodt I. S., Kukanova V. V., Pavlova O. V., Sapunova E. M., Filippova N. S.* (2008) The Corpus of Spoken Russian: Design Principles and Approaches to Data Analysis [O «korpuse» zhivoi rechi: printsipy formirovaniia i vozmozhnosti opisaniia.] Komp’iuternaia lingvistika i intellektual’nye tekhnologii. Po Materialam ezhegodnoj mezhdunarodnoi konferentsii “Dialog 2008” (Computational Linguistics and Intelligent Technologies: Proceedings of the International Annual Conference “Dialog 2008”). Moscow, 57–61.
9. *Bogdanova N. V., Os’mak N.A.* (2011) Some Lexical “Discoveries” on Material of Russian Spontaneous Speech (Corpora Research) [O nekotorykh leksicheskikh «otkrytiiakh» na materiale russkoi spontannoi rechi (korpusnoe issledovanie)]. Komp’iuternaia lingvistika i intellektual’nye tekhnologii. Po materialam ezhegodnoj mezhdunarodnoi konferentsii “Dialog 2011” (Computational Linguistics and Intelligent Technologies: Proceedings of the International Annual Conference “Dialog 2011”). Moscow, 110–124.
10. *Bocale P.* (2011a) Discourse Markers in Colloquial Speech [Diskursivnye slova v razgovornoi rechi], available at: russian.slavica.org/article9375.html.
11. *Bocale P.* (2011b) Role of Text Connectors in Youth Speech [Rol' tekstovykh konnektorov v molodezhnoi rechi], available at: russian.slavica.org/article514.html.
12. *Gerd A. S.* (1990) New General Purpose Dictionary — Original Positions and Guides [Novyi akademicheskii slovar' — iskhodnye pozitsii i orientiry]. Aktual'nye problemy razrabotki novogo akademicheskogo slovaria russkogo

iazyka (Actual Problems of Design of New General Purpose Russian Language Dictionary). Leningrad, 29–30.
13. *Daragan IU. V.* (2011) The Functions of Junk Words in Russian Spontaneous Speech [Funktsii slov-"parazitov" v russkoi spontannoi rechi]. Komp'iuternaia lingvistika i intellektual'nye tekhnologii. Po materialam ezhegodnoj mezhdunarodnoi konferentsii "Dialog 2000" (Computational Linguistics and Intelligent Technologies: Proceedings of the International Annual Conference "Dialog 2000"), available at: http://www.dialog-21.ru/materials/archive.asp?id=6260&vol=6077&y=2000.
14. *Evgen'eva A.P.* (1999) Slovar' russkogo jazyka (Russian Language Dictionary), 4 volumes. Moscow.
15. *Kisloshchuk A. I.* (2011) The Experience of Classification of Functions of "vot" in Spontaneous Monologue [Opyt sistematizatsii funktsii slova "vot" v spontannom monologe]. Aktual'nye problemy sovremennoi nauki. Materialy mezhdunarodnoi nauchnoi sessii "XIII Nevskie chteniia" (Actual Problems of Modern Science: Proceedings of International Scientific Session "XIII Nevsky Readings"). St.Petersburg, 157–164.
16. *Lapteva O. A.* (1999) Zhivaia russkaia rech' s teleekrana (Living Russian Speech from Television Screen). Moscow, 520 p.
17. *Meyer C. F.* (2004) English Corpus Linguistics: An Introduction. Cambridge.
18. *Os'mak N. A.* (2011a) From the History of Lexicographical Description of Russian Spontaneous Speech (iz it far to Dal'?) [Iz istorii leksikograficheskogo opisaniia russkoi spontannoi rechi (daleko li do Dalia?)]. Materialy XL Mezhdunarodnoi filologicheskoi konferentsii. Vypusk 24. Polevaia lingvistika. Integral'noe modelirovanie zvukovoi formy estestvennykh iazykov (Field Linguistics. Integral Modeling of Sound Form of Natural Languages. Issue 24: Proceedings of International XL Philological Conference. 2011). St.Petersburg, 160–182.
19. *Os'mak N. A.* (2011b) New Connotations of Old Words (about Lexicographical Description of Russian Spontaneous Speech) [Novye konnotatsii starykh slov (o leksikograficheskom opisanii russkoi spontannoi rechi)]. Obrazovatel'nye tekhnologii v virtual'nom lingvokommunikativnom prostranstve. IV Mezhdunarodnaia virtual'naia konferentsiia po rusistike, literature i kul'ture (Educational Technologies in Virtual Linguistic-Communicative Space: Proceedings of IV International Virtual Conference on Russian Language, Literature and Culture). Vermont, 166–170.
20. *Os'mak N. A.* (2012) New Meanings of Old Words (Corpora Research on Material of Everyday Russian Speech) [Novye znacheniia starykh slov (korpusnoe issledovanie na materiale povsednevnoi russkoi rechi)]. (in print).
21. *Stepanova S. B., Asinovskij A. S., Bogdanova N. V., Rusakova M. V., Sherstinova T. Ju.* (2008) Speech Corpus of the Russian Everyday Communication "One Speaker's Day": Basic Conception and Current [Zvukovoj korpus russkogo jazyka povsednevnogo obwenija "Odin rechevoj den'": Koncepcija i sostojanie formirovanija] Komp'iuternaia lingvistika i intellektual'nye tekhnologii. Po materialam ezhegodnoj mezhdunarodnoi konferentsii "Dialog 2008" (Computational

Linguistics and Intelligent Technologies: Proceedings of the International Annual Conference “Dialog 2008”). Moscow, 488–494.

22. *Khan N. A.* (2012) Construction (ia) ne znaiu in Russian Spontaneous Speech: Correlation between Different Functional Types [Konstruktsiia (ia) ne znaiu v russkoi spontannoi rechi: sootnoshenie raznykh funktsional'nykh tipov]. Mir russkogo slova (World of Russian Word). St.Petersburg, (in print).