

Компьютерная лингвистика и интеллектуальные технологии:
по материалам международной конференции «Диалог 2017»

Москва, 31 мая — 3 июня 2017

РУССКИЕ ПРЕДИКАТИВЫ В ЗЕРКАЛЕ ЭКСПЕРИМЕНТА И КОРПУСНОЙ ГРАММАТИКИ¹

Циммерлинг А. В. (fagraey64@hotmail.com)

Московский педагогический государственный
университет / Институт языкознания РАН, Москва, Россия

В статье описывается двойной эксперимент по определению объема и структуры активного словаря предикативов дативно-предикативной конструкции (ДПС). Вначале представляется онтология ситуаций, покрываемых данной конструкцией в русском языке, и тестовый словарь, состоящий из 422 стимулов, разбитых на 15 тематических классов. Словарь тестируется на группе из 18 испытуемых и составляется выборка предикативов на основе рейтинга одобрения Socio. Тот же набор из 422 стимулов тестируется на НКРЯ и составляется вторая выборка на основе показателя *m*, отражающего число клауз ДПС в контексте мне Z-vo. Объем и структура словаря ДПС определяются с помощью соотношения *m* и Socio. В ходе эксперимента строится модель конструкции ДПС, с выделением ее словарного ядра {Lex_k} и грамматического расширения {Gram_k}, предсказывающая, что носители языка могут использовать одинаковые принципы семантической селекции и добавления элементов в {Gram_k} при несовпадающем инвентаре лексических единиц.

Ключевые слова: грамматика конструкций, словарь, предикативы, дативно-предикативные структуры, идиолект, лексическое расширение, лексическая редукция, тематический класс

RUSSIAN PREDICATIVES IN THE PERSPECTIVE OF SOCIOLINGUISTIC EXPERIMENT AND CORPUS GRAMMAR

Zimmerling A. V. (fagraey64@hotmail.com)

Moscow State University of Education / Institute of Linguistics,
Russian Academy of Sciences, Moscow, Russia

¹ Работа выполнена при поддержке гранта РНФ 16-18-00203а «Структура значения и ее отображение в системе лексических и функциональных категорий русского языка», реализуемого в МПГУ. Я благодарю анонимных рецензентов конференции «Диалог», а также А. Ч. Пиперски и Л. Л. Иомдина за ценные замечания. Ответственность за все недочеты лежит на авторе.

This paper aims at measuring the active vocabulary of Russian predicatives licensing dative-predicative-structures (DPS) and introduces the results of the sociolinguistic experiment paired with the corpus study for the same set of stimuli. The sociolinguistic experiment aimed at measuring the active DPS vocabulary in the idiolects of 18 native speakers of Russian and produced a sample of 422 stimuli ranked according to their acceptance rate. The same set of stimuli was tested on data from the Russian National Corpus, which produced the second sample based on the frequency ranks of DPS clauses in the diagnostic context *мне Z-во*, with the priority dative argument placed contiguously to the predicative at the distance <-1:1> in the same clause. The correlation of two ranked samples predicts the volume of the DPS vocabulary in an average idiolect of Russian and helps to establish the proportion between the lexical nucleus of the DPS construction and its extension. The analysis verifies the hypothesis that native speakers of Russian apply to the same underlying principles of semantic selection for DPS elements while using non-identical sets of lexical elements. The paper introduces a formal model of lexical extension with the entire DPS being grouped onto 15 thematic classes. The results prove that each thematic class includes high-, middle- and low-frequency DPS elements. Low-frequency DPS predicatives are modelled after high-frequency elements from the same thematic class.

Key words: construction grammar, lexicon, predicatives, dative-predicative-structures, idiolect, lexical extension, lexical reduction, thematic class

1. Конструкции ДПС

Дативно-предикативная-структура (ДПС), предполагающая маркировку субъекта несогласуемого неглагольного предикатива дат. п., — одна из наиболее известных черт русской грамматики [Щерба 2008; Поспелов 1955]. Вместе с тем, точно не установлено, какие лексемы обслуживают конструкцию ДПС. Теоретические соображения свидетельствуют о том, что валентность на дат. п. лица — характеристика семантического класса предикатов состояния в русском языке [Циммерлинг 1998; 1999; Бонч-Осмоловская 2009], но условия его пополнения неясны [Циммерлинг 2010; Say 2013].

2. Русские предикативы в теоретической и корпусной лингвистике

Доступного справочного ресурса, с помощью которого можно проверить состав класса предикативов ДПС, нет. Лакуна частично восполняется синтаксическим словарем системы «Этап 3» [Этап-3; Apresjan et alii 2003], который насчитывает около 800 единиц, для которых на основе регулярных или окказиональных употреблений типа *мне было не очень эротично*, *Х-у было не эргономично делать Z* и т. п. постулируется валентность на дат. п.

лица. Словари данного типа ориентированы на парсинг, поэтому учет окказионализмов, подтверждающих, что ЭРГОНОМИЧНО, ЭРОТИЧНО имеют валентность на дат. п. лица в речи отдельных носителей, необходим. Но это не доказывает совместимости подобных употреблений друг с другом в речи одного носителя русского языка. В «Грамматике 1980» [Шведова 1982: 151] ИГ в дат. п. лица при предикативе трактуется не как приоритетный актант, но как т. н. детерминант. Предикативы ДПС изучались на материале НКРЯ, ср. [Маркова 2015], но прогресс в этом направлении сдерживается тем, что «с предикативами на -о само составление ... списка и определение того, насколько тот или иной предикат имеет валентность на дативный актант, является непростой исследовательской задачей» [Бонч-Осмоловская 2015: 83]. Еще хуже изучены неадеквативные предикативы (X-y) НЕ ПО СЕБЕ, ТАК И НАДО, ПО СИЛАМ и т. п. [Циммерлинг 2016: 362]. Такие выражения обычно не снабжаются пометой «предикатив» и не распознаются в качестве единых элементов.

2.1. Семантический инвариант конструкции ДПС

Если семантический инвариант конструкции ДПС существует, вероятным кандидатом является значение внутреннего состояния, соотнесенного с одушевленным субъектом в конкретной референтной ситуации² [Циммерлинг 2014: 52].

- (i) валентность на дат. п. лица отражает способность предикатива ДПС обозначать актуальные положения, соотнесенные с референтным одушевленным субъектом в течение некоторого отрезка времени.

Гипотеза (i) является модификацией тезиса Л. В. Щербы и Н. С. Поспелова [Поспелов 1955: 65] о том, что конструкция ДПС выражает в русском языке значение состояния. Она не требует представлять все предикативы ДПС как метафоры физического состояния или аффекта, ср. критику в [Селиверстова 1982: 131]. Конструкция ДПС покрывает некоторый спектр лексических значений, которые можно представить в виде онтологии ситуаций.

2.2. Семантическая селекция основ признаков слов

Употребление конструкции ДПС в русском языке поддерживается существованием большого числа признаков слов, которые в данной конструкции употребляться не могут, причем этот признак не скоррелирован с невозможностью предикативного употребления в структуре без согласования. Этот факт впервые отмечен в [Циммерлинг 1999: 224]. Предикаты типа *желтый* всегда обозначают свойства предмета или лица и не имеют сентенциальных

² О других дативных схемах русского языка, где семантический субъект может быть неодушевленным, в том числе, модальной конструкции *пирог у надо остыть*, см. [Циммерлинг 2010; 2016: 358].

валентностей. Предикаты типа *солнечный* формируют нульместную схему, ср. *Сегодня солнечно, здесь пыльно*, но не конструкцию ДПС: **мне солнечно, *нам пыльно*. Признак «быть СОЛНЕЧНО/ПЫЛЬНО» осмысливается как свойство внешней ситуации, одинаково воспринимаемое всеми потенциальными наблюдателями, и не может быть представлен как внутреннее состояние конкретного одушевленного субъекта в некоторый момент времени. Аналогичным образом, оценочные предикаты типа *аморальный*, которые допускают употребления *Z-во, что P, Z-во делать P*, отсеиваются конструкцией ДПС. Хотя квалификация поступка *U* как *аморального алогичного* может быть личным мнением говорящего, навязываемая языком концептуализация заставляет считать, что любой *X* оценит *U* как аморальный или алогичный. Наконец, предикаты, обозначающие активные свойства одушевленного субъекта, ср. *злой, гневный, бешеный*, в отличие, например, от древнеисландского или болгарского [Циммерлинг 2010], тоже отсеиваются конструкцией ДПС.

В [Циммерлинг 1998: 73; 2010: 553] производящие основы прилагательных, от которых образуются только предикаты постоянного свойства, отнесены к типу т. н. актантно-поляризованных. Производящие основы другого типа порождают предикативы ДПС, ср. *X-у грустно*, но параллельно используются и для порождения предикатов свойства: *грустный X*. Данные основы определяются как амбивалентные. Основы третьего типа, которые определяются как ситуативно-поляризованные, используются только для порождения предикативов ДПС, а коррелятивные формы прилагательного утрачены: *X-у стыдно, известно, *стыдный X, *совестный X*. Данную классификацию можно расширить, учтя предикативы неадъективной морфологии, что показано в табл. 1.

Таблица 1. Типы производящих основ предикативов ДПС в русском языке

	Тип I: актантно-поляризованные основы	Тип II: амбивалентные основы	Тип III: ситуативно-поляризованные основы
А. Адъективные основы	<i>желт-; пыльн-; аморальн-, алогичн-; гневн-, зло-; стыдлив-;</i>	<i>нужн-, грустн-;</i>	<i>стыдн-; совестн-; можн-;</i>
В. Неадъективные основы	(...)	<i>по силам, по душе, некстати</i>	<i>жаль, пора, надо, нельзя, наплевать; невдомек; все равно; нечего, некогда</i>

2.3. Пополнение класса ДПС

Основной динамической тенденцией системы, представленной в табл. 1, является пополнение типов II и III. Главным ресурсом служит переход $A1 \rightarrow A2$. Если в 1800–1850 гг. основной корпус НКРЯ фиксирует ок. 200 предикативов ДПС, к 1900 г. их уже более 300 [Циммерлинг, Астахова 2017], а в современном языке — около 500. Переход $A1 \rightarrow A2$ иллюстрируется распространившимися

в XX в. употреблениями *X-у слабо́, жирно, лениво* и т. п., которые характерны для речи определенных поколений [Циммерлинг 2010: 554]. Переход A2 → A3, т. е. утрата производящей основы атрибутивной формы, редок, ср. *стыдно, совестно, боязно* при **??стыдный, *??совестный, *боязный* [Циммерлинг 1998: 73]. Переход B2 → B3 более обычен. В XIX в. *невдомек, невтерпеж, невоготу* употреблялись не только в ДПС *X-у было Z-во*, но и в дативно-номинативной схеме *Y-и X-у были Z-ы*, что в современной норме не допускается. Под рубрику B2 → B3 можно также подвести случаи размежевания предикатива и его субстантивного источника, отражающиеся в смене модели согласования: *X-у была (крайняя) неохота делать P → X-у было очень неохота делать P* и т. п.³

2.4. Модель лексического расширения конструкции ДПС

Гипотезу (i) можно проверить для ситуации, когда лексическое заполнение класса предикативов ДПС у разных носителей языка неодинаково, например, при синхронном варьировании или при диахроническом сдвиге. Построим модель расширения класса ДПС, при которой котором принцип тернарного распределения основ, иллюстрированный выше на рис. 1, т. е. семантическая избирательность конструкции ДПС, не нарушается.

Примем за точку отсчета употребление I_0^n , где имеется n предикативов ДПС. Содержательно I_0^n можно интерпретировать как идиолект-эталон, либо как обобщенное употребление некоторого периода, восстанавливаемое по корпусу. Множество $\{n\}$ должно состоять из непустых подмножеств $\{Lex\}$ и $\{Gram\}$, куда входят, соответственно, элементы, воспроизводимые как словарные единицы со своей моделью управления, и элементы, принадлежность которых к классу ДПС, т. е. наличие у них диагностической валентности на дат. п. лица, устанавливается на уровне речи, в соответствии с некоторыми правилами. Пусть I_1^{n+m} — идиолект, в котором имеется на m больше предикативов ДПС, чем в эталоне I_0^n . По условию, хотя бы часть предикативов множества $\{Lex\}$, элементов которого наследуются из словаря, должна быть общими в I_0^n и I_1^{n+m} . Расширение класса предикативов ДПС происходит за счет пополнения $\{Gram\}$. Возможны два сценария.

А. $\{Gram_0\}$ и $\{Gram_1\}$ имеют непустое пересечение, и сохраняются признаковые основы типа I, которым не соответствуют никакие элементы в $\{Gram_0\} \cup \{Gram_1\}$ ⁴.

В этом случае семантическая избирательность конструкции сохранится.

³ Аналогично — при *охота, жаль, пора, грех, досуг, недосуг, лень*. О квазипосессивных конструкциях с дат. п. и предикатным именем в истории русского языка см. [Июмдин & Июмдин 2014: 221].

⁴ Данное условие является дефиниционным. Основы типа I являются источником пополнения класса предикативов ДПС. Если при расширении $I_0^n \rightarrow I_1^{n+m}$ не останется основ типа I, принцип семантической селекции основ перестанет действовать.

В. $\{Gram_0\}$ и $\{Gram_1\}$ имеют пустое пересечение, и сохраняются признаковые основы типа I, которым не соответствуют никакие элементы в $\{Gram_0\} \cup \{Gram_1\}$.

Дополнительное распределение элементов $\{Gram_0\}$ и $\{Gram_1\}$ совместимо с семантической избирательностью, если конструкция ДПС ориентирована не на конкретные лексемы, а на ряды значений. В этом случае тот факт, что, например, для выражения значения «X-у безразлично, P или не P» в I_0 и I_1 используются разные единицы, не нарушит принципов функционирования конструкции.

Для основного сценария можно уточнить прогноз о пополнении класса ДПС:

- (ii) В ситуации расширения класса предикативов ДПС $I_0^n \rightarrow I_1^{n+m}$ главным источником пополнения $\{Gram_0\} \rightarrow \{Gram_1\}$ является переход основ признаковых слов из типа I (актантно-поляризованные) в тип II (амбивалентные): $A1 \rightarrow A2, A2 \rightarrow B2$.

2.5. Семантическая избыточность рядов предикативов ДПС

В [Циммерлинг 2000: 199] утверждается, что предикативы ДПС в русском языке образуют ограниченно открытые ряды синонимов. Уточним эту гипотезу применительно к ситуации варьирования. Пусть в идиолекте I_1 для выражения значения M в конструкции ДПС используется лексема $a_1 \in \{a\}$, а в идиолекте I_2 — лексема $a_2 \in \{a\}$, и существует идиолект I_k , где для выражения данного значения используются a_1 и a_2 . Тогда в I_k имеет место семантическая избыточность, которая устранена в I_1 и I_2 .

3. Словник социолингвистического эксперимента

Для проверки гипотез (i) и (ii) был составлен словник предикативов ДПС, использующий принцип разбиения на тем. классы. Словник тестировался на группе носителей русского языка от 20 до 63 лет, уроженцев Москвы или Санкт-Петербурга, с высшим образованием. Всего было выделено 15 тем. классов. В основной список вошло 342 стимула. 80 стимулов дополнительного списка добавлялись в те же 15 классов, т. е. представляли собой их расширение.

3.1. Структура словника

Словник эксперимента представляет синтезированный идиолект. В тем. классы включались и предикативы ДПС, допустимые в литературном языке, X-у *лень*, *охота*, и субстандартные формы, ср. X-у *лениво*, *влом* и т. п. Словник соответствует модели идиолекта I_k с семантической избыточностью, где для выражения значения M^a в конструкции ДПС используется ряд лексем $a_1, a_2 \dots a_k \in \{a\}$. Каждому тем. классу соответствует набор значений $M^a, M^b \dots M^i$.

Тем самым, используется представление о том, что конструкции ДПС в русском языке соответствует некоторый закрытый и строго конечный набор значений, при том, что в идиолектах этим значениям могут соответствовать нетождественные ряды лексем. Одобрение испытуемым элемента доп. списка может интерпретироваться либо как добавление конкретной лексемы к ряду, закреплённому за уже имеющимся значением, либо как добавление нового значения к набору значений, присущих тем. классу, либо как отступление от принципов функционирования всей конструкции ДПС в эталонном идиолекте.

Таблица 2. Тематические классы предикативов ДПС в словнике. Прописными буквами выделены предикативы неадеквативной морфологии

	Тем. класс	Примеры	Число стимулов осн. списка	Число стимулов доп. списка
1	Физические состояния	<i>Х-у холодно, плохо₁, НЕ ПО СЕБЕ</i>	27	10
2	Модальности	<i>Х-у НАДО, неприлично, ПОРА</i>	44	5
3	Эмоциональные состояния	<i>Х-у грустно, НЕ ПО ДУШЕ.</i>	57	24
4	Моральные оценки	<i>Х-у стыдно, ГРЕХ, ТАК И НАДО</i>	16	0
5	Удобство исполнения	<i>Х-у удобно, сподручно, ПО ПУТИ</i>	8	0
6	Уместность/неуместность	<i>Х-у уместно, В САМЫЙ РАЗ</i>	13	2
7	Внутренняя потребность	<i>Х-у НЕВМОГОТУ, ПОТРЕБНО.</i>	7	2
8	(не) соответствие задаче	<i>Х-у ВПОРУ, жирно, слабó, ЛЕНЬ</i>	11	0
9	Трудность выполнения	<i>Х-у сложно, трудно, нелегко</i>	10	4
10	(Не) желание выполнять	<i>Х-у ОХОТА, НЕОХОТА, ВЛОМ</i>	9	1
11	Общая оценка	<i>Х-у хорошо, плохо₂, НЕ ПО СЕБЕ</i>	41	7
12	(Не) релевантность	<i>Х-у важно, НАПЛЕВАТЬ</i>	16	4
13	Эффективность	<i>Х-у вредно, полезно, БЕЗ ТОЛКУ</i>	6	1
14	Сенсорные и интеллектуальные реакции	<i>Х-у видно, слышно, ясно, известно, любопытно</i>	25	12
15	Параметризуемый признак	<i>Х-у темно, великó, ПО ПОЯС.</i>	52	8
ВСЕГО:			342	80

3.2. Тестовое задание

Испытуемые должны были оценить 422 стимула в плане употребления в одном из четырех контекстов: 1) *Мне было очень приятно.* 2) *Мне было очень приятно сделать Р.* 3) *Мне приятно, что Р.* 4) *Мне это очень приятно* (где слово *это* не отсылает к существительному ср. р.). Инструкция предписывала ответ «ДА» в случае, если испытуемый может сам употребить данный стимул хотя бы в одном из 4 перечисленных диагностических контекстов, и «НЕТ» во всех остальных случаях, включая знакомство с соответствующими употреблениями в речи других людей. Разрешалось менять формы лица, местоимения на существительные, порядок слов и лексическое наполнение.

4. Результаты эксперимента

В таблице 3 представлены данные об употреблении предикативов ДПС в идиолектах I_1 — I_{18} . Возраст указан на момент эксперимента (март 2016 г.). В нижней строчке указан процент от общего числа стимулов осн. списка (342) в идиолекте каждого информанта.

Таблица 3. Заполнение тем. классов предикативов ДПС в идиолектах испытуемых. Зеленым выделены идиолекты с лексическим расширением, красным — идиолекты с лексической редукцией

		ИДИОЛЕКТЫ ИНФОРМАНТОВ																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
		F25	M51	F37	F61	M63	F61	F32	F59	F35	M25	M27	F56	F25	F34	F28	F22	M55	M20
Осн.		190	266	205	282	268	288	210	169+7	260	244	292	210	205	236	275	139	306	313
Доп.		5	0	9	14	9	20	6	7	12	10	25	6	10	6	41	1	27	38
		55,5%	77,7%	59,9%	82,4%	78,4%	84,2%	61,2%	49,4%	76%	71,3%	85,4%	61,4%	59,9%	69%	80,4%	40,6%	89,5%	91,5%

4.1. Грамматика идиолектов

В 16 идиолектах из 18 объем словаря ДПС уместился в диапазон от 49,3% (I_8) до 89,5% (I_{17}). Стоит исключить anomalно низкий результат I_{16} (40,5%, 139 предикативов), и anomalно высокий результат I_{18} (91,5%, 313 предикативов). В I_{16} и I_8 большинство стимулов представляют {Lex}, т. е. носители данных идиолектов не пользуются или мало пользуются правилами расширения {Gram}. У I_{18} налицо проблемы с семантической селекцией, он забраковал всего 29 стимулов осн. списка и 42 элемента доп. списка. Тем самым, I_{18} с трудом вычлняет основы класса I. У I_{15} anomalно высок процент одобренных элементов доп. списка (41 из 80), но процент элементов осн. списка (80,4%, 275 стимулов) находится в пределах основного диапазона. Тем самым, I_{15} — информант со стандартной для русского языка грамматикой ДПС, но смещенной границей между элементами осн. и доп. списка.

4.1.1. Редукционные идиолекты и идиолекты с расширением

Чтобы интерпретировать синхронное варьирование словаря ДПС как случай расширения/редукции при общих принципах селекции стимулов, нужно выбрать в качестве эталона I_0 значение в нормальном диапазоне. Если взять значение нижней медианы выборки $I^m = I_{10} = 71,3\%$, редукционными будут I_{16} , I_8 и $I_1, I_3, I_{13}, I_7, I_{12}, I_{14}$.

4.2. Ранжирование тестовых стимулов по рейтингу Socio

18 положительных ответов из 18 соответствуют высшему рангу 1, а 0 положительных ответов из 18 — низшему рангу 19. 48 предикативов (14% осн.

списка) имеют ранг 1, что можно расценить как подтверждение того, что они воспроизводятся как словарные единицы и принадлежат $\{Lex_k\}$. 4 элемента получили 0 баллов, т. е. они отсутствуют в активном словаре ДПС у всех информантов. Далее количество баллов называется 'рейтингом Socio'.

4.2.1. Наборы стимулов

Из 80 стимулов доп. списка 68 (85%) относятся к 6 низшим рангам ($Socio < 6$), всего 7 предикативов (*X-у обязательно, некомфортно, чуждо, симпатично, дешево, твердо, неплохо*) получило более 50% положительных ответов ($Socio \geq 9$). Напротив, в осн. списке более 50% положительных ответов получило 245 стимулов (71,6%). Тем самым, разделение стимулов на осн. и доп. списки подтвердилось, а исключения касаются небольшого числа единиц, где оценки приемлемости в эталоне I_0 и усредненном идиолекте выборки I_k разошлись.

4.2.2. Дополнительные стимулы

В I_0 перечисленные 7 предикативов не входят в $\{Lex_0\}$ и отсеиваются из $\{Gram_0\}$, так их основы в I_0 относятся к типу I (актантно-поляризованные). В I_k данные 7 предикативов либо входят в $\{Lex_k\}$, либо помещаются в $\{Gram_k\}$ благодаря переходу в тип амбивалентных ($A1 \rightarrow B1$) по принципу (ii).

4.2.3. Основные стимулы

74 стимула, т. е. 21,7% осн. списка, получили меньше 50% положительных ответов. 64 стимула, т. е. 18,7% осн. списка относятся к диапазону $9 \leq Socio \leq 12$. Их естественно отнести к $\{Gram_k\}$. Логично принять, что стимул не воспроизводится как словарная единица, если он получает треть и более отрицательных ответов.

4.3. Интерпретация данных

Эксперимент подтвердил, что класс предикативов ДПС в речи испытуемых не наследуется из словаря целиком, а достраивается по правилам грамматики. Правила селекции у большинства информантов оказались сходными при значительном варьировании лексического наполнения класса предикативов ДПС. Модель расширения $\{Gram_k\}$ применима к большинству идиолектов. Нижняя медиана выборки $I^m = 71,3\%$ соответствует уровню в 244 стимула осн. списка. Почти идентичный показатель (245) дает рейтинг Socio: этому числу соответствует та часть стимулов осн. списка, которую одобрили более 50% испытуемых, т. е. элементы с $Socio \geq 9$.

5. Корпусная грамматика предикативов ДПС

Второй этап исследования был связан с тестированием 422 стимулов на основном корпусе НКРЯ⁵.

⁵ 23803881 предложение, дата обращения 05.03.2017.

5.1. Аппроксимативная мнe-мера

Использовалась выборка употреблений предикатива ДПС с 1 л. ед. ч. личного мест. на расстоянии $1 < -1: 1 >$, т. е. контексты *мне Z-во ~ Z-во мне*. Обоснование такого решения состоит в том, что большинство предикативов ДПС — автореферентные слова, ориентированные на употребление в 1 л. ед. ч. наст.вр. [Циммерлинг 2014]. Использовалась функция точного поиска, с последующей очисткой выборки аннотатором. Реализации ДПС и все прочие вхождения элемента в корпус трактовались как омонимы.

5.2. Выборка по мнe-мере и общее число вхождений

Между числом клауз с ДПС в выборке по *мнe-мере* (m), и общим числом вхождений элемента (с учетом омонимии) в НКРЯ (N), нет явной корреляции. При больших значениях n коэффициент m/N зависит не столько от абсолютной величины m , сколько от продуктивности употреблений, не связанных с ДПС.

5.3. Частотные классы предикативов

Выборка делалась для тех же 342 стимулов осн. и 80 стимулов доп. списка. Кроме того, проверялись 56 стимулов, которые не вошли в анкету: данная группа указана в таблице 4 в строке «компенсированный список». Стимулы были распределены по 11 частотным классам: для класса I, $m > 1000$, в классе XI, $m = 0$, т. е. выборка дает 0 клауз ДПС.

Таблица 4. Распределение предикативов ДПС по частотным классам в выборке НКРЯ по *мнe-мере*

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
	> 1000	400–1000	200–399	100–199	50–99	20–49	10–19	5–9	2–4	1	0
Основ.	7	17	17	18	24	25	37	33	50	35	79
Доп.	0	0	0	0	0	1	0	3	9	9	58
Компенс.	1	0	0	2	3	6	1	9	7	4	23
	8	17	17	20	27	32	38	45	66	48	160

5.4. Соотношение ранжированных выборок

422 стимулов анкеты, ранжированных по Socio, были поделены на 7 групп в диапазонах: 16–18 баллов, 13–15 баллов, 9–12 баллов, 6–8 баллов, 3–5 баллов, 1–2 балла, 0 баллов. Соотношение этой выборки с распределением предикативов ДПС по частотным классам, показывает таблица 5.

Таблица 5. Соотношение ранжированных выборок эксперимента и выборки НКРЯ по *мне-мере* для предикативов ДПС (по осн. и доп. спискам)

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI		
	> 1000	400–1000	200–399	100–199	50–99	20–49	10–19	5–9	2–4	1	0		
16–18 баллов (137)	7	16	14	13	19	15	17	9	12	7	8	137	
13–15 баллов (70)	0	0	2	3	3	6	14	8	17	1	17	70	
9–12 баллов (70)	0	1	1	1	1	3	2	11	12	17	21	70	
6–8 баллов (42)	0	0	0	1	1	1	1	4	5	10	19	42	
3–5 баллов (44)	0	0	0	0	0	0	2	3	9	3	27	44	
1–2 балла (39)	0	0	0	0	0	1	1	0	3	5	29	39	
0 баллов (20)	0	0	0	0	0	0	0	1	1	1	17	20	
422	7	17	17	18	24	26	37	36	59	44	137		
	Высокочастотные предикативы (146)							Среднечастотные предикативы (95)		Низкочастотные предикативы (181)			

Высокочастотные предикативы почти всего являются высокорейтинговыми, в первых 6 частотных классах ($m \geq 20$) всего 4 предикатива с $Socio < 9,3$ из которых являются архаизмами. Обратное не всегда верно: в частотном классе XI ($m = 0$) есть предикативы всех групп, хотя низкорейтинговые ($Socio < 9$) преобладают (68, 14%).

5.4.1. Высокочастотные предикативы

Емкость словаря ДПС в среднем идиолекте находится на уровне 244–245 стимулов. Такое количество единиц соответствует суммарному объему частотных классов I–VIII ($m \geq 5$) и части класса IX ($m \geq 2$). Примем, что большая часть этих элементов принадлежит $\{Lex_k\}$. Тогда понижение рейтинга Socio интерпретируется по модели редукции. Нижнюю границу $\{Lex_k\}$ стоит ограничить классом VII ($m \geq 10$), после него в выборке возрастает доля предикативов с $Socio < 13$. Тем самым, в группе высокорейтинговых/высокочастотных оказывается 145 стимулов.

5.4.2. Среднечастотные предикативы

К среднечастотным относятся стимулы классов VIII и IX, где $2 \leq m \leq 9$. В данной зоне 84 элемента, они могут относиться либо к $\{Lex_k\}$, либо к $\{Gram_k\}$. Ключ дает разбиение употреблений на два периода: 1760–1974 гг. vs. 1975–2016 гг. Знание вектора развития, значений Socio и m/N дает достаточные сведения для проверки гипотез о том, реализует ли употребление элемента сценарий редукции или расширения.

5.4.3. Низкочастотные предикативы

К данной группе относятся 113 элементов классов X ($m = 1$) и XI ($m = 0$). 33 из них имеют высокий рейтинг Socio (от Socio = 13 до Socio = 18). Низкий показатель m в данном случае отражает то, что эти предикативы описывают онтологически ограниченные ситуации. Информанты легко восстанавливают

контекст, где уместны высказывания X -у *малб* <о платье>, X -у *горько* <во рту>, <в ручье> X -у *было по щиколотку*, но соответствующие контексты редко фиксируются корпусом.

5.5. Тематические классы

Соотношение частотных и тем. классов показано в таблице 6.

Таблица 6. Распределение 342 предикативов осн. списка по частотным классам НКРЯ в выборке по мне-мере

Частотные классы			Тематические классы														
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Высокочастотные предикативы	I	>1000		3	1					1			1		1		
	II	400–999	1	5	5	1						2			2	1	
	III	200–399		1	5	2				1	1	1	2		2	2	
	IV	100–199	3	3	3	1				1		3			3	1	
	V	50–99	5	4	2		1			1	2	2	1	1		1	4
	VI	20–49	1	1	6	2			1	1				3	2	5	3
	VII	10–19	2	5	10	2	3	1	2			1	2	2	2	1	4
			12	22	32	8	4	1	3	2	5	4	9	9	4	15	15
Среднечастотные	VIII	5–9	4	4	3	2	1	0	2	3	2	0	7	1	0	2	3
	IX	2–4	5	7	9	3	0	4	0	2	1	2	8	0	1	2	6
			9	11	12	5	1	4	2	5	3	2	15	1	1	4	9
Низкочастотные	X	1	2	2	5	3	1	3	0	0	1	0	2	2	0	3	10
	XI	0	4	9	8	0	2	5	2	4	1	3	15	4	1	3	18
			6	11	13	3	3	8	2	4	2	3	17	6	1	6	28
Емкость класса			27	44	57	16	8	13	7	11	10	9	41	16	6	25	52

145 высокочастотных предикативов ДПС ($m > 9$) интерпретируются как ядро словаря ДПС, т. е. элементы $\{Lex_k\}$. 113 низкочастотных предикативов ДПС ($m \leq 1$), интерпретируются как расширения $\{Gram_k\}$, либо как элементы, исключенные из $\{Lex_k\}$ по сценарию лексической редукции. 84 среднечастотных предикатива ДПС ($1 < m < 10$) образуют промежуточную зону между ядром и расширением конструкции.

6. Выводы и перспективы

Проведенный двойной эксперимент позволил установить объем класса предикативов ДПС в идиолектах русского языка и смоделировать усредненный идиолект I_k с подмножествами $\{Lex_k\}$ (словарь) и $\{Gram_k\}$ (грамматика) по данным НКРЯ. Гипотеза о том, что одна часть предикативов ДПС принадлежит общерусскому словарю $\{Lex_k\}$ и является общей для всех носителей, а другая

часть предикативов ДПС порождается информантами в речи по правилам семантической селекции и добавляется в $\{Gram_i\}$ в их идиолектах, подтвердилось. Средний объем класса ДПС в идиолекте носителя современного русского языка составил 244 стимулов осн. списка (71,3%), что почти совпало с числом стимулов осн. списка, одобренных более 50% испытуемых — 245 стимулов, 71,6%.

Целью проведенного эксперимента было тестирование словника конструкции ДПС, а не изучение варьирования в разных социальных и возрастных группах. В то же время, для оценки устойчивости результатов полезно уточнить требования, предъявляемые к испытуемым и к корпусам текстов. От информанта требуется способность строить проверочный контекст для каждого стимула. Заранее неизвестно, обладают ли все испытуемые способностью к семантической селекции стимулов ДПС (в проведенном эксперименте такая способность была подтверждена для 16 из 18 человек), но для проверки гипотезы о лексическом расширении $\{Gram_k\}$ нужно, чтобы большинство испытуемых группы такой способностью обладали. Для создания ранжированной корпусной выборки предикативов ДПС подходит любой корпус, который отражает все 15 тем. классов словника (т.е. полную онтологию конструкции ДПС) и включает предикативы, зафиксированные в разные периоды существования языка.

Литература

1. *Апресян Ю. Д., Богуславский И. М., Иомдин Л. Л., Санников В. З.* (2010) Теоретические проблемы русского синтаксиса: Взаимодействие лексики и грамматики / отв. ред. Ю. Д. Апресян. М.: 408 с.
2. *Бонч-Осмоловская А. А.* (2009) Дативные субъектные конструкции с предикативами на — о/-е // *Корпусные исследования по русской грамматике*. Ред.-сост. К. Л. Киселева, В. А. Плунгян, Е. В. Рахилина, С. Г. Татевосов. М. С. 157–183.
3. *Бонч-Осмоловская А. А.* (2015) Квантитативные методы в диахронических корпусных исследованиях // *Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог-2015»*, вып. 14 (21). М. С. 80–94.
4. *Иомдин Б. Л., Иомдин Л. Л.* (2014) Валентности русских предикатных существительных и микросинтаксические конструкции // *Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог-2014»*, вып. 13 (20). М. С. 219–231.
5. *Маркова В. Р.* (2015) Предикативы состояния с финалью -о в русском языке // *Типология морфосинтаксических параметров. Материалы международной конференции «Типология морфосинтаксических параметров 2015»* / Под ред. Е. А. Лютиковой, А. В. Циммерлинга, М. Б. Коношенко. Вып. 2. М. С. 266–278.
6. *Поспелов Н. С.* (1955) В защиту категории состояния // *Вопросы языкознания*. 1955. № 2. С. 55–65.

7. Селиверстова О. Н. (1982) Второй вариант классификационной сетки и описание Некоторых предикатных типов русского языка // Семантические типы предикатов / Отв. ред. О. Н. Селиверстова. М. С. 88–157.
8. Циммерлинг А. В. (1998) История одной полемики // Язык и речевая деятельность. 1998. № 1. С. 63–87.
9. Циммерлинг А. В. (1999) Субъект состояния и субъект оценки // Логический анализ естественного языка. Образ человека в культуре и языке. П. ред. Н. Д. Арутюновой, И. Б. Левонтиной. М. С. 221–228.
10. Циммерлинг А. В. (2000) Этические концепты и семантические поля // Н. Д. Арутюнова, Т. Е. Янко, Н. К. Рябцева (ред.). Логический анализ языка. Языки этики. М.. С. 190–199.
11. Циммерлинг А. В. (2010) Именные предикативы и дативные предложения в европейских языках // Компьютерная лингвистика и интеллектуальные технологии. Вып. 9 (16). Материалы международной конференции «Диалог 2010». М. С. 549–558.
12. Циммерлинг А. В. (2014) Предикаты состояния и фактивность // Вопросы филологии. 2014. № 3–4. С. 52–60.
13. Циммерлинг А. В. (2016) Предикативы параметрического признака в русском языке // Труды ИРЯ РАН. Вып. 10. М. С. 358–369.
14. Циммерлинг А. В., Астахова Т. С. (2017) Корпусная грамматика и язык русских писателей XIX в. М., 2017. В печати.
15. Щерба Л. В. (2008). О частях речи в русском языке // Щерба Л. В. Языковая система и речевая деятельность. Изд. 4-е. М. С. 77–100.
16. Шведова Н. Ю. (1982). Распространение предложения в целом (детерминация) // Русская грамматика, том 2, М. С. 149–180.
17. Apresjan Ju., Boguslavsky I., Iomdin L. et al. (2003) ETAP-3 Linguistic Processor: a Full-Fledged NLP Implementation of the MTT // First International Conference on Meaning-Text Theory (MTT'2003). June 16–18, 2003. Paris, 2003. pp. 279–288. (Соавторы: Ю. Д. Апресян, И. М. Богуславский, А. В. Лазурский, В. З. Санников, В. Г. Сизов, Л. Л. Цинман)
18. Say S. (2013). On the nature of dative arguments in Russian constructions with “predicatives” // Current studies in Slavic linguistics. Amsterdam, Philadelphia, 2013. pp. 225–246.
19. Лингвистический процессор ЭТАП-3: <http://proling.iitp.ru/ru/etap3>

References

1. Apresjan Ju., Boguslavsky I., Iomdin L. et al. (2003) ETAP-3 Linguistic Processor: a Full-Fledged NLP Implementation of the MTT // First International Conference on Meaning-Text Theory (MTT'2003). June 16–18, 2003. Paris: Ecole Normale Supérieure, 2003, pp. 279–288. (Soavtory: Ju. D. Apresjan, I. M. Boguslavskij, A. V. Lazurskij, V. Z. Sannikov, V. G. Sizov, L. L. Cinman)
2. Say S. (2013). On the nature of dative arguments in Russian constructions with “predicatives” // Current studies in Slavic linguistics. Amsterdam, Philadelphia, 2013, pp. 225–246.

3. *Apresjan Ju. D., Boguslavskij I. M., Iomdin L. L., Sannikov V. Z.* (2010), Theoretic problems of Russian syntax: interaction of lexicon and grammar, Ju. D. Apresjan (ed.) [Teoreticheskie problemy russkogo sintaksisa: Vzaimodejstvie leksiki i grammatiki / otv. red. Ju. D. Apresjan]. — Moscow, 408 p.
4. *Bonch-Osmolovskaja A. A.* (2009), Dative subject constructions with the -o/-e final [Dativnye sub#ektnye konstrukcii s predikativami na — o/-e], Corpus studies in Russian grammar, K. L. Kiseleva, V. A. Plungian, E. V. Rahilina, S. G. Tatevosov (eds.) [Korpusnye issledovanija po russkoj grammatike. Red.-sost. K. L. Kiseleva, V. A. Plungjan, E. V. Rahilina, S. G. Tatevosov]. — Moscow, pp. 157–183.
5. *Bonch-Osmolovskaja A. A.* (2015), Quantitative methods in diachronic corpus studies [Kvantitativnye metody v diahronicheskikh korpusnyh issledovanijah], Computational linguistics and intellectual technologies: Proceedings of the international conference “Dialogue 2015”, Vol. 14 (21) [Komp’juternaja lingvistika i intellektual’nye tehnologii: Trudy mezhdunarodnoj konferencii «Dialog-2015», Vyp. 14 (21)], Moscow, pp. 80–94.
6. *Iomdin B. L., Iomdin L. L.* (2014), Valencies of Russian predicate nouns and microsyntactic constructions [Valentnosti russkih predikatnyh sushhestvitel’nyh i mikrosintaksicheskie konstrukcii], Computational linguistics and intellectual technologies: Proceedings of the international conference “Dialogue-2014”, Vol. 13 (20) [Komp’juternaja lingvistika i intellektual’nye tehnologii: Trudy mezhdunarodnoj konferencii «Dialog-2014», Vyp. 13 (20)], Moscow, pp. 219–231.
7. *Markova V. R.* (2015), Predicates of State with the –o –final in Russian [Predikativy sostojanija s final’ju -o v russkom jazyke], Typology of morphosyntactic parameters: Proceedings of the international conference “Typology of morhosyntactic parameter 2015”, E. A. Lyutikova, A. V. Zimmerling, M. B. Konoshenko (eds.), Vol. 2 [Tipologija morfosintaksicheskikh parametrov. Materialy mezhdunarodnoj konferencii «Tipologija morfosintaksicheskikh parametrov 2015» / Pod red. E. A. Lyutikovej, A. V. Zimmerlinga, M. B. Konoshenko. Vyp. 2.], Moscow, pp. 266–278.
8. *Pospelov N. S.* (1955), Defending the Category of State [V zashhitu kategorii sostojanija], Issues in linguistics [Voprosy jazykoznanija], 1955. № 2, pp. 55–65.
9. *Seliverstova O. N.* (1982), A second variant of the classification of the selected predicate types in Russian [Vtoroj variant klassifikacionnoj setki i opisanie nekotoryh predikatnyh tipov russkogo jazyka] // Semantic predicate types. O. N. Seliverstova (ed.) [Semanticheskie tipy predikatov / Otv. red. O. N. Seliverstova], Moscow, pp. 88–157.
10. *Shvedova N. Ju.* (1982), Extention of the whole sentence (determination) [Rasprostranenie predlozhenija v celom (determinacija)], Russian grammar, Vol. 2 [Russkaja grammatika, tom 2], Moscow, pp. 149–180.
11. *Shcherba L. V.* (2008), On parts of speech in Russian [O chastjah rechi v russkom jazyke], Shcherba L. V. Language system and speech activity [Shcherba L. V. Jazykovaja sistema i rechevaja dejatel’nost’. Izd. 4-e.], Moscow, pp. 77–100.
12. *Zimmerling A. V.* (1998), The history of a polemic [Istorija odnoj polemiki], Language and speech activity [Jazyk i rechevaja dejatel’nost’], 1998. № 1, pp. 63–87.

13. *Zimmerling A. V.* (1999), The Subject of State and the Subject of Evaluation [Sub#ekt sostojanija i sub#ekt ocenki], Logical analysis of a natural language: the Concept of a Human Being in language and culture, N. D. Arutyunova, I. B. Levontina (eds.) [Logicheskij analiz estestvennogo jazyka. Obraz cheloveka v kul'ture i jazyke. P.red. N. D. Arutyunovoj, I. B. Levontinoj], Moscow, pp. 221–228.
14. *Zimmerling A. V.* (2000), Ethic concepts and semantic fields [Etičeskie koncepty i semantičeskie polja], Logical analysis of a natural language: the languages of ethics, N. D. Arutyunova, T. E. Yanko, N. K. Rjabceva (eds.) [N. D. Arutjunova, T. E. Yanko, N. K. Rjabceva (red.). Logicheskij analiz jazyka. Jazyki etiki]. Moscow, pp. 190–199.
15. *Zimmerling A. V.* (2010), Nominal predicatives and dative sentences in European languages [Imennye predikativy i dativnye predlozhenija v evropejskih jazykah], Computational linguistics and intellectual technologies, Vol. 9 (16): Proceedings of the international conference “Dialogue 2010” [Komp'juternaja lingvistika i intellektual'nye tehnologii. Vyp. 9 (16). Materialy mezhdunarodnoj konferencii «Dialog 2010»], Moscow, pp. 549–558.
16. *Zimmerling A. V.* (2014). Predicates of State and factivity [Predikaty sostojanija i faktivnost'], Issues in philology [Voprosy filologii], 2014, № 3–4, pp. 52–60.
17. *Zimmerling A. V.* (2016), Predicatives of parametric features in Russian [Predikativy parametričeskogo priznaka v ruskom jazyke], Proceedings of the Vinogradov Institute of Russian language [Trudy IRJa RAN], Vyp. 10, 2016, pp. 358–369.
18. *Zimmerling A. V., Astahova T. S.* (2017). Corpus grammar and language of Russian authors of the XIX century [Korpusnaja grammatika i jazyk russkih pisatelej XIX v], Moscow, 2017. In press.
19. *Shherba L. V.* (2008), On parts of speech in Russian [O chastjah reči v ruskom jazyke], Shherba L. V. Language system and speech activity [Shherba L. V. Jazykovaja sistema i rečevaja dejatel'nost'. Izd. 4-e.], Moscow, pp. 77–100.
20. *Shvedova N. Ju.* (1982), Extention of the whole sentence (determination) [Rasprostranenie predlozhenija v celom (determinacija)], Russian grammar, Vol. 2 [Russkaja grammatika, tom 2], Moscow, pp. 149–180.
21. Linguistic processor ETAP-3 [Lingvističeskij processor ETAP-3]:<http://proling.iitp.ru/ru/etap3>