

СЕМАНТИЧЕСКИЕ РОЛИ И СЕТЬ КОНСТРУКЦИЙ В СИСТЕМЕ FRAMEBANK¹

Кашкин Е. В. (egorkashkin@rambler.ru)

Московский государственный университет
им. М. В. Ломоносова, Москва, Россия

Ляшевская О. Н. (olesar@gmail.com)

Национальный исследовательский университет «Высшая
школа экономики», Москва, Россия

Словарь русских глагольных конструкций — часть системы FrameBank, которая постоянно пополняется по данным Национального корпуса русского языка. Семантическая разметка глагольных конструкций включает а) определение значения глагола и семантических ролей (экспликаций) участников ситуации, б) формулировку семантических ограничений на заполнение валентностей, в) установление отношений между конструкциями одного глагола и между конструкциями разных глаголов в графе конструкций.

Инventарь семантических ролей устроен иерархически, что позволяет масштабировать его размеры от десятка проторолей до множества частных определений. Инвентарь «базового уровня», описанный в работе, коррелирует с семантической классификацией глагольной лексики в НКРЯ.

В статье также формулируются принципы построения графа конструкций, отражающего как семантические переходы на множестве глагольной лексики, так и наследование / сдвиги в морфосинтаксическом оформлении конструкций.

Обсуждаются возможности практического использования FrameBank в задаче semantic role labeling, а также теоретические вопросы соотношения семантических классов глаголов, семантических ролей и семантических ограничений на заполнение валентностей.

Ключевые слова: конструкции, семантические роли, полисемия, семантические переходы, лексикография, корпусная лингвистика.

SEMANTIC ROLES AND CONSTRUCTION NET IN RUSSIAN FRAMEBANK

Kashkin E. V. (egorkashkin@rambler.ru)

Lomonosov Moscow State University, Moscow, Russia

Lyashevskaya O. N. (olesar@gmail.com)

National Research University "Higher School of Economics",
Moscow, Russia

¹ Работа выполнена в рамках программы фундаментальных исследований Президента РАН «Корпусная лингвистика».

This paper reports on a research project in progress which involves a dictionary of Russian lexical constructions and a corpus tagged with FrameNet-like annotation scheme. Russian FrameBank, originally conceived as an analogue of Berkeley FrameNet, takes into account some recent approaches adopted in Construction Grammar and Russian lexical semantics, as well as certain features of Russian lexical system and grammar.

We focus on the semantic annotation of constructions in FrameBank. First, the article describes the inventory of semantic roles used in FrameBank which correlates with the semantic classification of verbs and other predicates. Semantic roles form a hierarchy: all the 88 roles are classified into six clusters (those of Agent, Patient, Experiencer, Instrument, Addressee, Circumstances), which are further subdivided into some smaller groups. The hierarchical organization makes the inventory of semantic roles more flexible for use in theoretical research and computational applications (such as automatic semantic role labeling). We also show that many examples are annotated in a more appropriate way by introducing syncretic semantic roles (e. g. Instrument-Place or Result-Manner).

Second, we touch upon an ongoing project on the systematization of semantic shifts in verbal lexemes (metaphor, metonymy, and rebranding, which is argued to be a special type of a semantic shift, see, for example, [Rakhilina et al. 2010a]) and the corresponding changes in argument structure constructions (including changes of a morpho-syntactic pattern, omission of a participant which belongs to a known class, etc.). The labels for the shifts are provided, along with examples of their realization. Lexical constructions are defined on constant (lexicalized) slots, mainly verbs and other predicates in a particular meaning. Frames are thus seen as the signifié side of constructional clusters formed by synonymous predicates, aspectual pairs, etc. Since it is not uncommon for polysemous lexemes that the formal facade of constructions is inherited from sense to sense, we claim that the frame nets cannot be routed without taking into account sense relations in polysemous predicates.

The final discussion deals with the relation between semantic classes of verbs, semantic roles, and lexical/semantic constraints on the classes of participants as provided by FrameBank data.

Key words: constructions, semantic roles, polysemy, semantic shifts, lexicography, corpus linguistics.

1. Русский FrameBank

Статья посвящена описанию двух разработанных классификаций: инвентаря семантических ролей и переходов между конструкциями одного глагола, которые составляют теоретическую основу семантической разметки системы FrameBank (www.framebank.ru). Прежде чем перейти к основной части изложения, охарактеризуем кратко общее устройство и назначение системы (см. также [Ляшевская, Кузнецова 2009, Lyashevskaya 2010, Lyashevskaya 2012]).

FrameBank — общедоступный онлайн-ресурс, объединяющий словарь лексических конструкций русского языка и размеченный корпус их реализаций в текстах НКРЯ. Конструкции включают предикатно-аргументные структуры глаголов, существительных, прилагательных, наречий и предикативов, а также идиомы, в которых часть элементов фиксированы, а часть представляют собой переменные (т. н. конструкции «малого синтаксиса»).

FrameBank относится к семейству FrameNet-ориентированных ресурсов (см. https://framenet.icsi.berkeley.edu/fndrupal/framenets_in_other_languages)², однако и по идеологии, и по формату несколько отличается от их родоначальника, системы FrameNet, разработанной в Беркли. Если FrameNet строится вокруг фреймов — типовых ситуаций с известным набором участников и расписанными ролями, то русский FrameBank — вокруг конструкций конкретных лексем. Каждое значение лексемы соответствует своему фрейму, а близкие значения ряда лексем, описывающие типовую ситуацию, входят в общий фрейм. Во FrameNet средства выражения участников в тексте приписываются сразу всем предикатам, обозначающим данный фрейм, — во FrameBank подразумевается, что конструкция каждого предиката имеет индивидуальные особенности (ограничения на заполнение элементов, отличия в значении), даже если предикаты относятся к одному фрейму. В основе FrameNet лежит теория Frame Semantics, разработанная Ч. Филлмором в середине 1970-х гг., и идея, что сеть фреймов универсальна во всех языках, а русский FrameBank ориентирован в большей мере на исследования лексических конструкций в духе Грамматики Конструкций того же Ч. Филлмора, А. Голдберг и др. и Московской семантической школы (Ю. Д. Апресян, Е. В. Падучева и др.).

FrameBank документирует:

- русскую лексическую систему, структуру русских лексико-семантических групп и полисемии (фреймы не универсальны);
- парадигматические отношения между значениями многозначных слов — в том, как они отражаются в системе связанных с этими значениями лексических конструкций;
- лексико-семантические ограничения на слоты конструкций;
- грамматические особенности русского языка (порядок слов, падежи, согласование и т. п.).

Ядро системы FrameBank составляют 2200 частотных русских глаголов и ассоциированных с ними конструкций и корпусных примеров. Словарь русских глагольных конструкций представляет каждую конструкцию как шаблон, в котором указаны а) морфосинтаксические характеристики элементов конструкции; б) синтаксический ранг участника; в) экспликация (роль) участника; г) семантические ограничения на заполнение слота конструкции.

Например, для глагола *резать* выделено около 30 шаблонов конструкций (обозначаемых для наглядности ярлыками-примерами), среди которых есть кластер шаблонов *Продащица режет сыр, Она режет хлеб на тонкие куски, Василий начал резать хлеб длинными ломтями, Портной режет ткань большими ножницами*, реализующих с помощью различных моделей управления значение

² Помимо версий FrameNet для японского, китайского, бразильского португальского, шведского, испанского и немецкого языков, клонирующих полностью или частично его структуру и формат, к семейству FrameNet-ориентированных ресурсов можно отнести современные корпусно-ориентированные словари валентностей, соответствующие компоненты WordNet-ов, а также корпуса с глубокой разметкой, отражающей аргументную структуру, кореференцию, дискурсивные стратегии и мн. др. (ср., например, тектограмматику в Prague Dependency Treebank, тестовые корпуса для машинного перевода с Comreno-разметкой ABVYU и др.).

глагола 'разделять объект на части давлением острого инструмента', кластер шаблонов *Старик резал четки из кипариса* и *Он резал деревянные ложки простым ножом*, соответствующих значению 'изготавливать что-л. посредством резьбы', а также кластер *В боку режет* и *У него в желудке резало* (см. рис. 1), где глагол *резать* описывает определенный тип болевого ощущения.

Схема: в + Sloc Vimpers y + Sgen

Пример: У него в желудке резало.

Буква	Вершина	Экспликация	Ранг	Сем. ограничения
Z	в + Sloc	часть субъекта физиологического ощущения	Периферия	часть тела
-	Vimpers	-	Предикат	-
T	y + Sgen	посessor	Периферия	одушевленный

Рис. 1. Шаблон одной из конструкций глагола *резать* в значении физиологического ощущения

– Вот он, один из этих шукарей из Варьете, – поспышался грозный голос над онемевшим бухгалтером. И тут же Василия Степановича арестовали.

Имя	Вершина	Группа	Заполнение вершины	Заполнение группы	Экспликация	Ранг	Семантические ограничения
X	Snom	NPnom	неопределенно-личная конструкция		агенс	Субъект	лицо / организация
Реализация	-	-			-	Не выражен	-
-	арестовать	арестовать			-	Предикат	-
Реализация	арестовать	N/A	арестовали	N/A	-	Предикат	-
Y	Sacc	NPacc	стандартный	стандартный	пациенс	Объект	лицо
Реализация	Sacc	NPacc	Василия	Василия Степановича	пациенс	Объект	лицо
№	Тип	Группа					
1	сирконстант	тут же					

Рис. 2. Разметка примера на шаблон конструкции «Полиция арестовала преступника». Каждому элементу конструкции отведено две строки: в верхней отражается информация из словаря, а также вариант реализации (пассивная конструкция, неопределенно-личная и т. п.), в нижней — разметка реализации в данном примере.

В корпусной части ресурса FrameBank представлено около 100 примеров из НКРЯ на каждый глагол (в дальнейшем выборки планируется расширить). Разметчик должен соотнести пример с конструкцией (если нужный шаблон в системе отсутствует, разметчик пополняет словарь конструкций). Затем определяется вариант реализации конструкции (включая нестандартные, например, при деепричастии, пассивном причастии и т.п.), в примере определяются группы, соответствующие элементам конструкции (а также синтаксические и семантические вершины групп), несовпадения в морфосинтаксическом оформлении и ограничениях на лексико-семантическую сочетаемость. Отдельно размечаются сирконстанты и модальные слова, которых шаблон, естественно, не охватывает (см. рис. 2).

Два компонента системы FrameNet не нашли пока отражения во FrameBank. Граф фреймов (frame grapher³) видится как над надстройкой над графом конструкций, о котором речь пойдет ниже, а сами фреймы — как генерализация индивидуальных значений лексических конструкций. Корпус со сплошной полнотекстовой framenet-разметкой (full-text annotation) в русском ресурсе также пока отсутствует, и эта задача представляется приоритетным направлением дальнейшего развития системы.

Создаваемый ресурс выполняет как лексикографические задачи (например, предъявляет пользователю список конструкций того или иного глагола, выдает все глаголы с выражаемой при них ролью Инструмента), так и служит инструментом для углубленного исследования русских конструкций с использованием корректно размеченных и не содержащих поискового «шума» примеров (о теоретической актуальности последней задачи ср. в частности [Рахилина 2010]).

FrameBank может быть использован и при решении прикладных задач, как размеченный вручную стандарт для машинного обучения. Наиболее тесно с нашим ресурсом связана задача автоматического распознавания семантических ролей (semantic role labeling (SRL), см. [Gildea, Jurafsky 2000, Màrquez et al. 2008, Кузнецов (в печати)]). Эта задача складывается из а) идентификации и определения ролей семантических аргументов предиката в тексте (при том, что его фрейм известен) и б) определения фрейма (значения) предиката, отсутствующего в словаре.

Решение пользовательских и компьютерных задач чувствительно к классификации семантических ролей и самих фреймов и конструкций. Ожидания пользователей об инвентаре этих единиц и круге явлений, которые они охватывают, могут быть разными. Точно так же успех задачи SRL зависит от дробности ролей и успешного «вытягивания» похожих конструкций через сеть фреймов/конструкций. Далее в статье описывается начатый в 2012 г. проект по системной семантической разметке глагольных конструкций, которая строится на иерархическом связывании конструкций и ролей. Именно этот принцип должен обеспечить гибкость в приспособлении к разным задачам. В разделе 2

³ Во FrameNet граф строится от фреймов самого частного уровня (например, фрейм, кодируемый глаголом to shop и существительным shopping) к промежуточным уровням (например, купля-продажа в перспективе продавца (продажа) VS в перспективе покупателя (купля)) и далее к еще более абстрактным (например, фрейм, общий по отношению к разнообразным вариантам купли-продажи, фрейм посессивного отношения, транзитивный фрейм и т.п.).

обсуждается классификация семантических ролей, а разделе 3 — принципы построения графа конструкций, отражающего внутрилексемные и межлексемные семантические связи целевых глаголов.

2. Инвентарь семантических ролей

Инвентарь семантических ролей для русского языка может иметь разный состав и объем (ср. в частности [Апресян 1974/1995: 125–126], [Апресян и др. 2010: 370–377], [Падучева 2004: 587–588], а также обзор различных подходов и теоретических проблем в [Fillmore 1968, 1977, 1982], [Dowty 1991], [Лютикова и др. 2006: 17–22], [Плунгян 2011: 160–165]), что во многом определяется конкретными нуждами его использования. Следует, однако, иметь в виду, что:

- роль — это инвариант над разнообразием морфосинтаксических способов кодирования участника; так же и семантически — это генерализация функций участника в круге ситуаций, обозначаемых группой предикатов;
- роли в описании семантически близких лексем должны либо системно совпадать, либо системно различаться;
- полный инвентарь ролей должен описывать все области лексики.

Идея применения инвентаря семантических ролей к описанию больших массивов данных сама по себе не нова, ср. в частности известные проекты FrameNet, «Лексикограф», НОСС и RussNet. В упомянутых ресурсах, однако, этот инвентарь играет лишь вспомогательную роль при описании других свойств лексем и конструкций. Так, разработчики системы RussNet ориентированы в первую очередь на создание детального тезауруса русской лексики, применимого в сфере автоматической обработки естественного языка, а не на подробную классификацию и анализ семантических ролей. В современной версии системы FrameNet семантическая роль служит лишь для пояснения конфигурации участников внутри одного фрейма, и к ее названию не предъявляется никаких требований. В результате FrameNet предлагает слишком широкий и, как кажется, потенциально неограниченный набор семантических ролей, часто и вовсе заводимых ad hoc для одного узкого фрейма — ср., например, выделение отдельного фрейма AGRICULTURE, покрывающего лексические единицы *to cultivate, to farm* и *farming*, участникам которого приписываются такие роли, как *Agriculturist* (тот, кто возделывает сельскохозяйственную культуру) и *Food* (возделываемая сельскохозяйственная культура). Неудивительно, что получить классификацию семантических ролей в онлайн-версии системы невозможно, и она не входит в число официальных компонентов FrameNet.

Проект «Лексикограф» идеологически более близок нашим задачам, однако на данный момент охватывает не все значения и тематические классы глаголов с одинаковой степенью детальности: так, в версии базы от 30.10.2010, доступной сейчас онлайн, детально разработаны глаголы физического воздействия, перемещения, звука, однако отсутствуют такие глаголы, как *видеть, слышать, понимать, светиться, греметь* и мн. др. (а для включенных в базу глаголов учтены далеко не все значения и конструкции, ср. глагол *бить*, для которого в «Лексикографе»

имеется только два входа — «БИТЬ 1 (палкой по забору)» и «БИТЬ 2 (кого)». В этой связи, говорить о полном инвентаре ролей не приходится.

Наш проект можно рассматривать как масштабирование идеи «Лексикографа» на большой объем данных⁴. Была поставлена задача создания инвентаря семантических ролей, строящегося на следующих принципах:

- инвентарь должен быть иерархически организован с целью создания более гибкого инструмента поиска и кластеризации: при желании, его можно свести к 5–10 проторолям, в других случаях, он может быть расширен до нескольких десятков и даже сотен ярлыков;
- интерпретация первого и второго аргумента в большей мере зависит от семантики предиката, нежели трактовка третьих, четвертых и т. д. аргументов типа Инструмента, Траектории и т. д.;
- инвентарь коррелирует с семантической классификацией глагольной лексики⁵: в частности, это означает, что традиционные очень широко понимаемые роли Агенса и Пациенса должны в разных группах получать разные ярлыки;
- объем роли строится по принципу прототипа и периферии: например, прототипом Пациенса является участник, претерпевающий изменение под физическим воздействием контролирующего ситуацию Агенса, периферийные случаи (пациенс нефизического процесса, пациенс, не претерпевающий изменения, пациенс, который создается в результате физического действия и пр.) получают собственные ярлыки (ср. Тема, Результат) и считаются частным случаем роли Пациенса;
- предусматривается возможность сдвоенных ролей и расщепления ролей [Апресян 1974/1995].

За основу для составления списка был взят инвентарь семантических ролей, приведенный в [Апресян и др. 2010: 370–377]. Практическая работа с имеющимися в системе шаблонами конструкций потребовала, однако, внесения в этот инвентарь ряда изменений. Помимо незначительной правки технического характера (так, вместо ярлыка «Пациенс!» мы использовали более самодостаточное наименование «Подвергающаяся воздействию часть пациенса»), в список Ю. Д. Апресяна были внесены изменения в связи с тем, что ряд содержащихся в нем ролей объединяет достаточно разнородные семантические сущности. Если роль соотносилась с несколькими семантическими классами глаголов, то она разделялась нами на несколько — например, роли Экспериментера в нашей разметке соответствуют семантические роли Субъект восприятия (*видеть, слышать*), Субъект ментального состояния (*думать, понимать*), Субъект психологического состояния (*бояться, любить*), Субъект физиологического ощущения (*болеть*,

⁴ При этом мы не преследуем задачу приписать каждому глаголу толкование по некоторой заданной схеме, как это делается в «Лексикографе», а сосредотачиваемся на детальном описании конструкций и связей между ними.

⁵ Поскольку FrameBank является «дочерним» ресурсом НКРЯ, с надстроенным слоем разметки и интегрированным словарем, он ориентирован на систему глагольных классов Основного корпуса [Kustova et al. 2009], с учетом их дополнения и расширения. Вместе с тем, сам принцип иерархического выделения ролей может быть связан с любыми другими лексическими классификациями.

колоть в боку) и Субъект физиологической реакции (*смеяться, тошнить*). Роль Агенса была сохранена для ядерных агентивных контекстов, но в дополнение к ней в список были включены экспликации Говорящий, Субъект поведения (*лечься, медлить*), Субъект социального отношения (*дружить, помириться*), Субъект перемещения (последняя экспликация используется для всех, не только агентивных, одноместных глаголов перемещения, коррелируя тем самым с их выделением в особый класс; агентивность глагола в этом случае однозначно устанавливается по одушевленности субъекта).

В результате для разметки шаблонов конструкций был использован список из 88 экспликаций, классифицированный по принципу семантической близости на несколько групп: блок Агенса, блок Пациенса; блок Экспериенцера; блоки Инструмента и Адресата, блок обстоятельственных характеристик. Внутри блоков можно выделить группу посессивных ролей, группы ролей Места, Времени, Параметров, Признаков, Причины и Цели; группа Источников и Ресурсов объединяет роли из блока Агенса и Места; при максимальном сжатии инвентаря роли группы Экспериенцера можно распределить между агентивными и пациентивными ролями. Семантические роли и их блоки образуют единый граф (см. рис. 3)⁶, что позволяет выбирать между разными уровнями дробности поиска, релевантными для конкретной задачи (например, найти как все шаблоны конструкций, в которых реализуются семантические роли из Блока Агенса, так и все шаблоны конструкций, где есть участник с ролью Говорящего).

Для целого ряда шаблонов конструкций оказалось невозможным приписать участнику ровно одну роль, поскольку имело место сочетание семантики двух различных ролей. В этих случаях в разметку вводились двойные семантические роли (ср. здесь [Апресян 1974/1995: 140] об отдельных примерах синкретичного выражения валентностей). Так, в контексте *обрабатывать детали на станке* речь идет об инструменте совершения действия, но одновременно этот инструмент имеет локативные свойства, поэтому в данном случае использовалась экспликация Инструмент-Место. Конструкция *Пехотинцы строились клином* описывает результат (то, что получилось в результате построения) и одновременно способ совершения действия, и в этом и подобных случаях в разметку вводилась двойная экспликация Результат-Способ. В конструкции *«Вот это фокус», — удивился он* участник-лицо получил двоякую роль Говорящий-Субъект психологического состояния. Очевидно, что двоякие роли присутствуют в конструкциях, где либо участник размечен морфосинтаксически нестандартно (ср. *на станке*), либо предикат относится к нескольким лексическим классам (ср. *удивился* — эмоциональное психологическое состояние и говорение).

⁶ Иерархические отношения между семантическими ролями обозначены на графе сплошными линиями. Пунктирные линии соответствуют семантическим связям между ролями, не связанными непосредственным иерархическим отношением. Стрелки при ролях блока Экспериенцера показывают семантическую близость этих ролей к блоку Агенса или к блоку Пациенса. О-блок объединяет в себе шесть групп обстоятельственных ролей, которые традиционно не сводят к одной гиперроли. Пространственное расположение ярлыков (например, сверху vs. справа) относительно ярлыков ролей верхнего уровня в иерархии не несет какой-либо смысловой нагрузки.

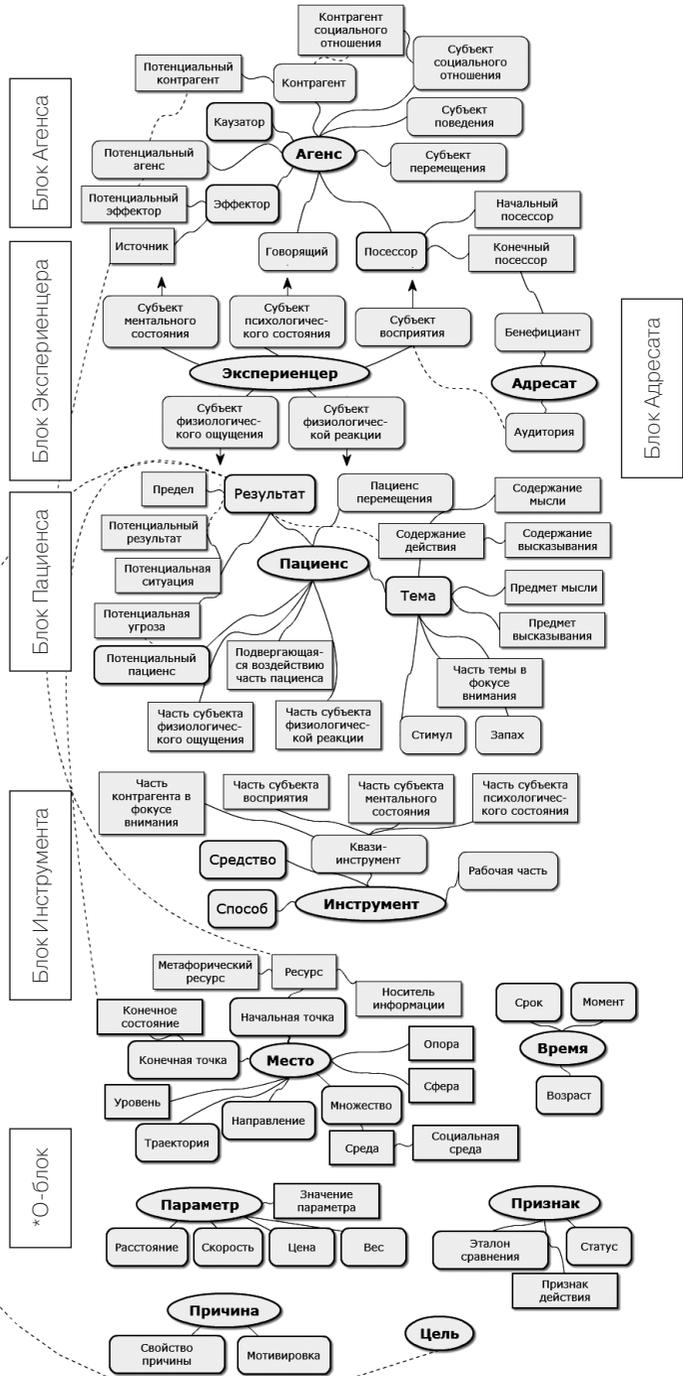


Рис. 3. Иерархия семантических ролей системы «ФреймБанк»

Следует оговорить, что FrameBank предусматривает и более дробное представление ролей участников: например, для глаголов *служить* и *спасти* экспликации в стиле FrameNet «тот, кому служат» и «тот, кого спасают» будут более точными ярлыками, нежели Контрагент и Пациент — однако, с одной стороны, такие индивидуальные ярлыки будут редко востребованы пользователями, а с другой стороны, они могут быть порождены автоматически по определенной схеме. В этой связи, основной рабочий статус в системе получает инвентарь из 88 базовых ролей⁷.

3. Граф глагольных конструкций

3.1. Полисемия глагольной лексики и ее подача в системе FrameBank

Разработка семантической разметки системы FrameBank предполагает также системный анализ полисемии глагольной лексики и создание поискового инструмента при исследованиях в этой области. На решение этой задачи и нацелена ведущаяся на данном этапе работа. Речь идет не только об анализе отдельных семантических полей (ср. [Кустова 2004] о глаголах физического воздействия и лексемах с экспериенциальной семантикой, [Падучева 2004] о фазовых и бытийных глаголах, глаголах восприятия, эмоций, звучания, речи, перемещения, и мн. др.), но и о создании такого ресурса, который представлял бы информацию о системных закономерностях семантических переходов глагольной лексики — причем о таких закономерностях, которые формулируются в терминах не только наиболее базовых противопоставлений вроде «конкретный предмет» vs. «абстрактное имя», но и более детальной классификации семантических ролей и семантических ограничений.

В нашей работе мы опираемся на теоретический опыт школы Е. В. Падучевой в осмыслении глагольной полисемии (см. [Падучева 2004] и др.), а также на опыт проекта Базы данных по многозначным качественным прилагательным и наречиям русского языка ([Рахилина и др. 2009], [Карпова и др. 2010], [Карпова и др. 2011]), призванного решить аналогичную задачу для признаковой лексики. Естественно, набор используемых нами ярлыков для семантических переходов не является копией аналогичного инвентаря из базы данных прилагательных и наречий — в первую очередь в силу значительно большей вариативности конструкций у глагольной лексики по сравнению с признаковой лексикой.

При разметке системы FrameBank для каждого из глаголов строится семантическая сеть, иллюстрирующая направления и типы переходов между всеми шаблонами его конструкций. Прежде всего, выделяются внутрифреймовые

⁷ Это число не является абсолютным и, безусловно, со временем будет меняться. В частности, не исключено, что потребуются расширение инвентаря при разметке конструкций имен прилагательных и существительных.

и межфреймовые связи. Связи между конструкциями, относящимися к одному значению глагола и, соответственно, к одному фрейму, маркируют введение нового участника, замену участника при переносе фокуса внимания, мену коммуникативного статуса и морфосинтаксического оформления и т. п. явления. Межфреймовые связи соединяют конструкции, относящиеся к разным значениям глагола.

Кроме того, различаются связи, маркирующие семантический переход, и формальные связи, маркирующие изменение или наследование формального паттерна конструкции. Примером семантического перехода является метафора — например, при переходе от конструкции *Гром гремит* к конструкции *Директор гремел, обличая пороки*. Вместе с тем, для конструкции *Директор гремел...* устанавливается отношение точного формального наследования (Spom V) как с конструкцией *Гром гремит*, так и с конструкцией *Друзья говорили до утра*, обозначающей прототипическую речевую ситуацию.

3.2. Типы переходов между конструкциями одного глагола

Ниже приводится пилотная классификация типов переходов между конструкциями одного глагола, выделенных по результатам обработки 10% глагольной лексики в системе «Фреймбанк»; разметка связей между конструкциями разных глаголов еще предстоит. Классификация включает различные комбинации изменения/сохранения плана выражения и плана содержания конструкций.

- A1. Мена морфосинтаксического оформления участника: *занес этот факт в протокол* → *занес в протокол, что судья обрывает его*; в т. ч. в зависимости от типа лексического заполнения элемента: *он занес ногу за порог* → *через плетень* → *на ступеньку брички*;
- A2. Мена статуса участников, диатетический сдвиг: *протираю стол от пыли* → *протираю пыль на столе*;
- A3. Невыражение участника, относящегося к известному классу: *наши следы занесло снегом* → *наши следы занесло*
- A4. Невыражение участника, дейктически или ситуативно известного: *он выписал все адреса из справочника* → *он выписал все адреса*
- A5. Невыражение участника, неопределенного (неважного) в ситуации: *птицы летят на юг* → *летят птицы*
- A6. Добавление участника: *вода собирается* — *вода собиралась каплями*; частный случай добавления — гибрид двух конструкций: *рыбку занесло из реки, рыбку занесло в протоку* → *рыбку занесло из реки в солёный океан*
- A7. Мена участников (перенос фокуса с одного участника на другого): *он выписал все адреса из справочника* → *он выписал все адреса в тетрадь*
- Б. Добавление неядерного участника ситуации: в производную конструкцию эксплицитно добавляется участник, не предусмотренный прототипом фрейма: *Птица летит* → *Птица летит за кормом*. *Вахтер выписал пропуск* → *Вахтер выписал мне пропуск*.

- В1. Специализация значения фрейма, связанная с невыражением одного из участников: *Иван пьет чай* → *Антон пьет*; *Мы говорили о прошлом* → *Ребенок уже говорит* ('умеет говорить').
- В2. Идиоматизация значения, связанная с введением в конструкцию новых лексических констант (в частности, вместо переменных-участников): *Он опустил кулак* → *Он опустил руки* ('перестать действовать, потеряв надежду').
- Г1. Метонимия: смежный участник. Используется при замене участника на другого, смежного участника в пределах одного фрейма: *слушать музыку* → *слушать Баха*
- Г2. Метонимия: перераспределение коммуникативных акцентов между участниками ситуации (при диатетическом сдвиге, А2): *Суд слушает дело* → *В суде слушают дело*.
- Г3. Метонимия: сдвиг домена: *Я вас любил* (по МАС, 'чувствовать сердечную склонность к лицу другого пола') → *Она своих девочек очень любит* ('чувствовать глубокую привязанность к кому-л., быть преданным кому-л.').
- Г4. Метонимия: смежный класс. Используется при переходе глагола в смежный с исходным тематический класс, ср. *Вечером, сидя за чаем, Семён Семёнович со скучающим видом слушал* (глагол восприятия) *жену, которая что-то записывала на бумажке...* [И. А. Ильф, Е. П. Петров (1935)] → *Хотя мэр Москвы по стилю своего публичного поведения, безусловно, принадлежит к людям, готовым слушать* ('принимать во внимание слова, просьбы, советы', глагол ментального состояния) *москвичей*. [«Известия» (2001)] → *Нет человека, властного над ветром, умеющего удержать ветер, особенно когда этот ветер в голове. Не хотят слушать* ('подчиняться распоряжениям, следовать советам, слушаться', глагол поведения) *старших — пусть идут. Пусть хлебнут горя своей золотой ложечкой*. [М. Успенский (1995)]
- Д. Метафора. Используется при смене таксономического класса какого-либо из участников ситуации, сопровождающемся сдвигом значения глагола: *В парке борются два парня* → *На выборах борются две партии*.
- Е. Ребрендинг. Понимается нами в соответствии с [Бонч-Осмоловская и др. 2009], [Рахилина и др. 2009], [Рахилина и др. 2010а], [Рахилина и др. 2010б], [Карпова и др. 2011] как семантический переход лексемы в другой таксономический класс, основанный на механизме имплицатуры (т. е. результат семантического перехода является следствием или выводом из исходного значения), ср. переход *Солдат стреляет* → *У него стреляет в голове*, где происходит явная смена таксономического класса глагола *стрелять* (глагол физического воздействия → глагол болевого ощущения), а производное значение осмысливается как вероятный результат действия, подразумеваемого исходным значением (субъект испытывает такое ощущение, как будто в боку происходит действие стрелять).

Ж. Другие, более далекие и менее прозрачные переходы. Ср., например, выступить из толпы → выступить на совещании.

З. Грамматикализация: выветривание значения в случае, когда глагол принимает роль лексической функции: *являться, выступать (свидетелем), обратить (внимание), питать (уважение)* и т. д.

Нередко глагольные значения связаны не непосредственно, а через цепочку «посредников» в значениях других предикатов. Ср., например, конструкции *занести письмо домой* и *Войдет и занесет такую чушь...* ‘начать нести (чушь)’, которые связаны через посредство разных значений глагола *нести*; конструкцию *собраться с силами*, которая очевидно соотносится с конструкцией *собрать силы* (, чтобы встать).

4. Заключение. О соотношении семантических ролей участников, семантических ограничений, классов глаголов

Разметка семантических ролей участников конструкций и систематизация переходов между конструкциями позволит выявить закономерности системы полисемии глагольной лексики в ее связи со свойствами конструкций, в которых реализуется конкретный глагол. Так, с использованием базы можно будет выявить, с одной стороны, типы переносов, характерных для глаголов какого-либо исходного семантического класса (и свойства соответствующих конструкций), с другой стороны, типы переносов (и свойства конструкций), результатами которых являются глаголы заданного класса — ср., например, перенос в семантическую область речи из областей перемещения (*Летят птицы* → *Летит молва*), психологического состояния (*Парень волновался* → «*Не догонит!*» — *волновался парень*), физического воздействия (*Хозяйка отрезала кусок хлеба* → «*И слышать об этом не хочу!*» — *отрезала хозяйка*) и др.

Обсуждение такого рода явлений поднимает и теоретические вопросы о соотношении семантических ролей участников, семантических ограничений на заполнение валентностей, а также глагольных классов. В частности, для метафорических переносов (по определению предполагающих изменение семантических ограничений на заполнение хотя бы одной валентности) в базе обнаруживаются следующие возможности:

- Смена семантического класса глагола и семантических ролей участников, ср. *Летят птицы* (перемещение, Субъект перемещения) → *Время летит* (скорость перемещения, Время) и *Летят птицы* (перемещение, Субъект перемещения) → *Летит молва* (речь, Содержание высказывания).
- Сохранение семантического класса глагола и семантических ролей участников. Ср., например, переходы *Летят птицы* → *Конь летит, ветер свистит в ушах* и *Летят птицы* → *Не раз он летел кубарем* (во всех примерах глагол лететь относится к классу глаголов перемещения, а субъект получает семантическую роль Субъект перемещения).

- Смена семантического класса глагола при сохранении паттерна семантических ролей: *Мальчик ест хлеб* (уничтожение, Агенс + Пациент) → *Мошкы едят лошадей* (физическое воздействие, Агенс + Пациент)⁸.

Детальное исследование этих вопросов предполагается сделать возможным с использованием системы FrameBank. Перспективным направлением развития системы предполагается и создание графа фреймов (аналогичного имеющемуся во FrameNet), который бы послужил дополнением к разрабатываемому сейчас графу конструкций и содержал эмпирический материал для обсуждения связи фреймов с семантическими ролями, семантическими ограничениями и классами глаголов.

Литература

1. Апресян Ю. Д. Избранные труды, том I. Лексическая семантика. М., 1995. 1-е изд.: М., 1974.
2. Апресян Ю. Д., Богуславский И. М., Иомдин Л. Л., Санников В. З. Теоретические проблемы русского синтаксиса: взаимодействие грамматики и словаря. М., 2010.
3. Апресян Ю. Д., Палл Э. Русский глагол — венгерский глагол. Управление и сочетаемость. Будапешт, 1982.
4. Бонч-Осмоловская А. А., Рахилина Е. В., Резникова Т. И. Глаголы боли: лексическая типология и механизмы семантической деривации // Концепт боль в типологическом освещении / под ред. В. М. Брицына, Е. В. Рахилиной, Т. И. Резниковой, Г. М. Яворской. Киев, 2009. С. 8–27.
5. Карпова О. С., Резникова Т. И., Архангельский Т. А., Кюсева М. В., Рахилина Е. В., Рыжова Д. А., Тагабилева М. Г. База данных по многозначным качественным прилагательным и наречиям русского языка // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 26–30 мая 2010 г.). Вып. 9 (16). М.: РГГУ, 2010. С. 163–168.
6. Карпова О. С., Рахилина Е. В., Резникова Т. И., Рыжова Д. А. Оценочные значения ребрендингового типа в признаковой лексике (по материалам Базы данных семантических переходов в качественных прилагательных и наречиях) // Компьютерная лингвистика и интеллектуальные технологии:

⁸ В этой связи встает вопрос о регулярности соответствия между семантической классификацией глаголов и приписываемыми их аргументам семантическими ролями. Действительно, в большинстве случаев глаголам разных классов в системе приписываются разные наборы ролей. Но в ряде случаев классификация глагольной лексики может быть и более дробной (что отражается, например, в моделях метафорических сдвигов), как в рассматриваемом примере, где глаголы уничтожения по сути являются подклассом глаголов физического воздействия (ср. также глаголы перемещения, внутри которых выделяется подкласс глаголов падения, но семантической ролью субъекта в любом случае является Субъект перемещения) — однако семантические характеристики самих участников и отношения между этими участниками в таких случаях, как представляется, очень близки и могут быть сведены к одинаковым наборам ролей.

- По материалам ежегодной Международной конференции «Диалог» (Бекасово, 25–29 мая 2011 г.). Вып. 10 (17). М.: РГГУ, 2011. С. 292–304.
7. Кузнецов И. О. Автоматическое выделение глагольных актантов: теоретическая основа и актуальные подходы // НТИ. Сер. 2 (в печати).
 8. Кустова Г. И. Типы производных значений и механизмы языкового расширения. М., 2004.
 9. «Лексикограф». Электронный ресурс: <http://lexicograph.ruslang.ru>
 10. Лютикова Е. А., Татевосов С. Г., Иванов М. Ю., Пазельская А. Г., Шлуинский А. Б. Структура события и семантика глагола в карачаево-балкарском языке. М., 2006.
 11. Ляшевская О. Н., Кузнецова Ю. Л. Русский фреймнет: к задаче создания корпусного словаря конструкций // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2009» (Бекасово, 27–31 мая 2009 г.). Вып. 8 (15). М.: РГГУ, 2009. С. 306–312.
 12. Падучева Е. В. Динамические модели в семантике лексики. М., 2004.
 13. Плунгян В. А. Введение в грамматическую семантику: грамматические значения и грамматические системы языков мира. М., 2011.
 14. Рахилина Е. В., Карпова О. С., Резникова Т. И. Модели семантической деривации многозначных качественных прилагательных: метафора, метонимия и их взаимодействие // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2009» (Бекасово, 27–31 мая 2009 г.). Вып. 8 (15). М.: РГГУ, 2009. С. 420–425.
 15. Рахилина Е. В., Резникова Т. И., Карпова О. С. Семантические переходы в атрибутивных конструкциях: метафора, метонимия и ребрендинг // Лингвистика конструкций / Отв. ред. Е. В. Рахилина. М., 2010а. С. 398–455.
 16. Рахилина Е. В., Резникова Т. И., Бонч-Осмоловская А. А. Типология преобразования конструкций: предикаты боли // Лингвистика конструкций / Отв. ред. Е. В. Рахилина. М., 2010б. С. 456–540.
 17. Рахилина Е. В. (ред.) Лингвистика конструкций. М., 2010.
 18. Dowty, D. R. (1991), Thematic proto roles and argument selection, *Language* 67, pp. 547–619.
 19. Fillmore Ch. J. (1968), The Case for Case, in Bach E. and Harms (Ed.), *Universals in Linguistic Theory*. New York, pp. 1–88.
 20. Fillmore Ch. J. (1977), The case for case reopened, in Cole P., Sadock J. M. (eds.), *Grammatical Relations*, Acad. Press, New York, pp. 59–81.
 21. Fillmore Ch. J. (1982), Frame semantics, *Linguistics in the morning calm: Selected papers from the SICOL-1981*, Hanship, Seoul, pp. 111–137.
 22. *FrameNet*. Электронный ресурс: <http://framenet.icsi.berkeley.edu>
 23. Gildea D., Jurafsky D. (2000), Automatic labeling of semantic roles, *Proc. of the 38th Annual Conference of the Association for Computational Linguistics (ACL-00)*, pp. 512–520.
 24. Kustova Galina I., Olga N. Lashevskaja, Elena V. Paducheva, and Ekaterina V. Rakhilina (2009), Verb taxonomy: from theoretical lexical semantics to practice

of corpus tagging, in Lewandowska B., K. Dziwirek (eds.), *Cognitive Corpus Linguistics Studies*. Frankfurt: Peter Lang.

25. *Lyashevskaya O.* (2010), *Bank of Russian Constructions and Valencies // Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, Valletta, Malta, 17–23 May 2010. Valletta: ELRA, 2010. P. 1802–1805.
26. *RussNet: тезаурус русского языка. Электронный ресурс: <http://project.phil.spbu.ru/RussNet>*

References

1. *Apresjan Ju. D.* (1995), *Selected papers, Vol. 1, Lexical Semantics [Izbrannye trudy, tom I. Leksicheskaja semantika]*, Jazyki Russkoj Kul'tury, Vostochnaja Literatura, Moscow.
2. *Apresjan Ju. D., Boguslavskij I. M., Iomdin L. L., Sannikov V. Z.* (2010), *Theoretical issues of Russian syntax: the interrelation between grammar and vocabulary [Teoreticheskie problemy russkogo sintaksisa: vzaimodejstvie grammatiki i slovar'a]*, Jazyki slavjanskih kul'tur, Moscow.
3. *Apresjan Ju. D., Pall E.* (1982), *Russian verb — Hungarian verb. Government and combinability [Russkij glagol — vengerskij glagol. Upravlenie i sochetaemost']*, Tankyonvkiado, Budapest.
4. *Bonch-Osmolovskaja A. A., Rakhilina E. V., Reznikova T. I.* (2009), *Pain verbs: lexical typology and mechanisms of semantic derivation [Glagoly boli: leksicheskaja tipologija i mehanizmy semanticheskoi derivatsii]*, in *The concert of pain from a typological point of view [Kontsept bol' v tipologicheskom osveschenii]*, Vidavnicijj Dim Dmitra Burago, Kiev, pp. 8–27.
5. *Dowty, D. R.* (1991), *Thematic proto roles and argument selection*, *Language* 67, pp. 547–619.
6. *Fillmore Ch. J.* (1968), *The Case for Case*, in Bach E. and Harms (Ed.), *Universals in Linguistic Theory*. New York, pp. 1–88.
7. *Fillmore Ch. J.* (1977), *The case for case reopened*, in Cole P., Sadock J. M. (eds.), *Grammatical Relations*, Acad. Press, New York, pp. 59–81.
8. *Fillmore Ch. J.* (1982), *Frame semantics*, *Linguistics in the morning calm: Selected papers from the SICOL-1981*, Hanship, Seoul, pp. 111–137.
9. *FrameNet*. An online resource, available at: <http://framenet.icsi.berkeley.edu>
10. *Gildea D., Jurafsky D.* (2000), *Automatic labeling of semantic roles*, *Proc. of the 38th Annual Conference of the Association for Computational Linguistics (ACL-00)*, pp. 512–520.
11. *Karpova O. S., Reznikova T. I., Arkhangel'skij T. A., Kjuseva M. V., Rakhilina E. V., Ryzhova D. A., Tagabileva M. G.* (2010), *A database of polysemous qualitative adjectives and adverbs in Russian [Baza dannyh po mnogoznachnym kachestvennym prilagatel'nyh i narechijam russkogo jazyka]*, *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog" [Komp'juternaja lingvistika i intellektual'nye*

- tehnologii: po materialam ezhegodnoj Mezhdunarodnoj konferentsii “Dialog”], Moscow, pp. 163–168.
12. *Karpova O. S., Rakhilina E. V., Reznikova T. I., Ryzhova D. A.* (2011), Evaluative meanings of the rebranding type in qualitative lexemes [Otsenocnyye zhachenija rebrendingovogo tipa v priznakovoj leksike], *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialog”* [Komp’juternaja lingvistika i intellektual’nye tehnologii: po materialam ezhegodnoj Mezhdunarodnoj konferentsii “Dialog”], Moscow, pp. 292–304.
 13. *Kustova G. I.* (2004), Types of figurative meanings and mechanisms of linguistic broadening [Tipy proizvodnyh zhachenij i mehanizmy jazykovogo rasshirenija], *Jazyki slavjanskoj kul’tury*, Moscow.
 14. *Kustova Galina I., Olga N. Lashevskaja, Elena V. Paducheva, and Ekaterina V. Rakhilina* (2009), Verb taxonomy: from theoretical lexical semantics to practice of corpus tagging, in *Lewandowska B., K. Dziwirek* (eds.), *Cognitive Corpus Linguistics Studies*. Frankfurt: Peter Lang.
 15. *Kuznetsov I. O.* (forthcoming), Automatic extraction of verb arguments: theoretical grounds and state of the art [Avtomaticheskoe vydelenie glagol’nyh aktantov: teoreticheskaja osnova i aktual’nye podhody], *Scientific and Technical Information. Ser. 2. Information Processes and Systems* [Nauchnaja i tehničeskaja informacija. Ser. 2. Informacionnyye processy i sistemy].
 16. *Lexicographer* [Leksikograf]. An online database, available at: <http://lexicographer.ruslang.ru>
 17. *Ljutikova E. A., Tatevosov S. G., Ivanov M. Ju., Pazel’skaja A. G., Shluinskij A. B.* (2006), Event structure and verb semantics in Karachaj-Balkar [Struktura sobytija i semantika glagola v karachaevobalkarskom jazyke], IMLI RAN, Moscow.
 18. *Lyashevskaya O.* (2010), Bank of Russian Constructions and Valencies, *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC’10)*, Valletta, pp. 1802–1805.
 19. *Lyashevskaya O. N., Kuznetsova Ju. L.* (2009), Russian FrameNet: constructing a corpus-based dictionary of constructions [Russkij Frejmnet: k zadache sozdanija korpusnogo slovarja konstruksij], *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialog”* [Komp’juternaja lingvistika i intellektual’nye tehnologii: po materialam ezhegodnoj Mezhdunarodnoj konferentsii “Dialog”], Moscow, pp. 306–312.
 20. *Màrquez L., Carreras X., Litkowski K. C., Stevenson S.* (2008), Semantic role labeling: an introduction to the special issue, *Computational Linguistics*, Vol. 34–2, pp. 145–159.
 21. *Paducheva E. V.* (2004), Dynamic patterns in lexical semantics [Dinamicheskie modeli v semantike leksiki], *Jazyki slavjanskoj kul’tury*, Moscow.
 22. *Plungian V. A.* (2011), Introduction to grammatical semantics: grammatical meanings and grammatical systems of the world’s languages [Vvedenie v grammaticheskiju semantiku: grammaticheskie znachenija i grammaticheskie sistemy jazykov mira], RSUH, Moscow.
 23. *Rakhilina E. V., Karpova O. S., Reznikova T. I.* (2009), Patterns of semantic derivation in polysemous qualitative adjectives: metaphor, metonymy, and their

- interaction [Modeli semanticheskoy derivatsii mnogoznachnyh kachestvennyh prilagatel'nyh: metafora, metonimija i ih vzaimodejstvie], Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog" [Komp'juternaja lingvistika i intellektual'nye tehnologii: po materialam ezhegodnoj Mezhdunarodnoj konferentsii "Dialog"], Moscow, pp. 420–425.
24. *Rakhilina E. V., Reznikova T. I., Karpova O. S.* (2010a), Semantic shifts in attributive constructions: metaphor, metonymy, and rebranding [Semanticheskie perehody v atributivnyh konstrukcijah: metafora, metonimija i rebrending], in *Linguistics of Constructions [Lingvistika konstrukcij]*, Azbukovnik, Moscow, pp. 398–455.
25. *Rakhilina E. V., Reznikova T. I., Bonch-Osmolovskaja A. A.* (2010b), Typology of constructional changes: pain predicates [Tipologija preobrazovanija konstrukcij: predikaty boli], in *Linguistics of Constructions [Lingvistika konstrukcij]*, Azbukovnik, Moscow, pp. 456–540.
26. *Rakhilina E. V.* (ed., 2010) *Linguistics of Constructions [Lingvistika konstrukcij]*, Azbukovnik, Moscow.
27. *RussNet*: a Russian language thesaurus. Available at: <http://project.phil.spbu.ru/RussNet>