

Компьютерная лингвистика и интеллектуальные технологии:
по материалам международной конференции «Диалог 2020»

Москва, 17–20 июня 2020 г.

РУССКИЙ КОНСТРУКТИКОН: НОВЫЙ ЛИНГВИСТИЧЕСКИЙ РЕСУРС, ЕГО УСТРОЙСТВО И СПЕЦИФИКА

Эндресен А. А. (anna.endresen@uit.no)

Университет Тромсё — Норвежский арктический университет

Жукова В. А. (valentina.zh96@gmail.com)

Национальный исследовательский университет
«Высшая Школы Экономики»

Мордашова Д. Д. (mordashova.d@yandex.ru)

Институт языкознания РАН, МГУ им. М. В. Ломоносова

Рахилина Е. В. (rakhilina@gmail.com),

Ляшевская О. Н. (olesar@yandex.ru)

Национальный исследовательский университет
«Высшая Школы Экономики»;

Институт русского языка им. В. В. Виноградова РАН

DOI: NN.NNNNN/ANNNNNNNANNNNNNN-N

THE RUSSIAN CONSTRUCTICON: A NEW LINGUISTIC RESOURCE, ITS DESIGN AND KEY CHARACTERISTICS

Endresen A. A. (anna.endresen@uit.no)

Universitetet i Tromsø: Norges arktiske universitet

Zhukova V. A. (valentina.zh96@gmail.com)

National Research University Higher School of Economics

Mordashova D. D. (mordashova.d@yandex.ru)

Institute of Linguistics (Russian Academy of Sciences),
Lomonosov Moscow State University

Rakhilina E. V. (rakhilina@gmail.com),

Lyashevskaya O. N. (olesar@yandex.ru)

National Research University Higher School of Economics;
Vinogradov Institute for Russian language (Russian Academy
of Sciences)

We present a new open-access electronic resource named the Russian Constructicon that offers a searchable database of Russian constructions accompanied by descriptions of their properties and illustrated with corpus examples. The project was carried out over the period 2016–2020 and at present contains an inventory of over 2200 multi-word constructions of Contemporary Standard Russian. We prioritize “partially schematic” constructions that lie between the two extremes of fully compositional syntactic sequences on the one hand and fully idiomatic (phraseological) expressions on the other hand. Constructions of this type are difficult to account for in terms of either lexicon or grammar alone, and are often underrepresented in reference works of Russian. A typical construction in our database contains a fixed part (anchor words) and an open slot that can be filled with a restricted set of lexemes. In this paper we first focus on key characteristics of this resource that make it different from existing constructicons of other languages. Second, we describe how the new interface will be designed and how it will serve the needs of both linguists and L2 learners of Russian. In particular, we discuss various search possibilities relevant for different users and those parameters that are available for specifying the retrieval output. An example of an entry is given to show how the information about each construction is structured and presented. Third, we provide an overview of our multi-level semantic classification of constructions. We argue that our system of semantic and syntactic tags subdivides our items into meaningful classes and smaller groups and eventually facilitates the identification of constructional families and clusters. This methodology works well in turning the initial list of constructions as unrelated units into a structured network and makes it possible to refine and expand the collected inventory of constructions in a systematic way.

Keywords: Russian, constructions, semantics, syntax, language technology, electronic database, language pedagogy, typology

1. Введение

Русский Конструктикон представляет собой новый электронный лингвистический ресурс по конструкциям современного русского языка — базу данных, которая на сегодняшний день содержит более 2200 единиц. Примерами таких конструкций являются **NP-Nom так и не VP** (*Он так и не женился*)¹, **NP-Nom возьми и VP-Imper** (*А он возьми и купи новую машину!*)², **NP-Gen бы** (*Чаю бы!*) и др. Каждая конструкция сопровождается толкованием, семантическими, синтаксическими и стилистическими пометами, а также примерами из Национального корпуса русского языка (НКРЯ; подробнее о структуре словарной статьи см. [Janda et al. 2018]). Ресурс является бесплатным и находится в открытом доступе³.

Настоящая статья посвящена лингвистическим характеристикам ресурса. В **разделе 2** мы рассмотрим отличительные особенности Русского Конструктикона, сопоставив его с существующими конструктиконами других языков. В **разделе 3** мы кратко опишем его новый интерфейс, над которым идет работа, после чего в **разделе 4** представим принципы семантической разметки конструкций. В заключении будет сформулировано значение проекта как для русистики, так и для *конструктикографии* в целом.

¹ В Русском Конструктиконе принята система обозначений, с помощью которых записываются конструкции в базе: **NP-Nom** = именная группа с существительным в именительном падеже.

² Данная конструкция тесно связана с конструкцией **взять и VP** (*он взял и увез меня в деревню*) и рассматривается как ее подтип в статье [Stojnova 2007]. В Русском Конструктиконе это две разные конструкции. Такой подход позволяет объяснить различие употреблений типа *А ты возьми и купи машину!* (оба императива в императивном значении, семантика побуждения к действию, согласование предиката с подлежащим, конструкция **взять и VP**) и употреблений типа *А мы возьми и купи машину!* (псевдоимперативные формы, императив в неимперативном значении, семантика неожиданно совершенного действия, отсутствие согласования подлежащего и предиката по лицу и числу, конструкция **NP-Nom возьми и VP-Imper**).

³ Новый сайт Русского Конструктикона находится в разработке, его адрес будет указан на интернет-странице проекта (<https://site.uit.no/russian-constructicon/>). Более ранняя версия интерфейса доступна по адресу <https://spraakbanken.gu.se/karp/#?mode=konstruktikon-rus> и устроена по образцу шведского конструктикона. Работа над созданием Русского Конструктикона ведется более четырех лет (2016–2020) в рамках совместного проекта сотрудничества Университета Тромсе — Норвежского Арктического Университета (UiT) и Национального исследовательского университета «Высшая Школы Экономики» (НИУ ВШЭ) в Москве. В разработке ресурса участвуют Л. А. Янда, Т. Нессет, А. А. Эндресен (группа CLEAR, UiT); Е. В. Рахилина, О. Н. Ляшевская, В. А. Жукова (Школа лингвистики НИУ ВШЭ); Д. Д. Мордашова (МГУ им. М. В. Ломоносова, Институт языкознания РАН); Ф. Тайерс (Университет Индианы, США). Программное обеспечение и создание нового интерфейса осуществляет Р. Баст (Отдел цифровых исследований, UiT), ранее работа велась А. Г. Клезович (Школа лингвистики НИУ ВШЭ) и Ф. О. Сизовым (РГГУ). В проекте также принимают участие студенты НИУ ВШЭ и UiT.

2. С NP-Ins не сравнить: Русский Конструктикон в ряду других конструктиконов

В рамках проекта *конструкция* понимается, в соответствии с Грамматикой Конструкций [Fillmore et al. 1988]; [Croft 2001]; [Goldberg 2006]; [Rakhilina 2010], как центральная единица языковой структуры. В последние годы стал активно употребляться термин *конструктикон*, обозначающий, с одной стороны, систему конструкций определенного языка, а с другой стороны, детальное описание этой системы.

Это второе, прикладное, понимание конструктикона было впервые введено Ч. Филлмором [Fillmore 2008]; [Fillmore et al. 2012], который инициировал разработку конструктикона для английского языка⁴ в дополнение к ресурсу FrameNet. Идея Филлмора довольно быстро была подхвачена и нашла применение в отношении других языков. В настоящий момент идет работа по созданию конструктиконов для шведского⁵, японского⁶, бразильского португальского⁷, немецкого⁸ и русского. Разработка этих ресурсов вылилась в появление нового направления, получившего название *конструктикография* и объединяющего теоретические основы грамматики конструкций и методы лексикографии [Lyngfelt et al. 2018a]. Выявление типологически сходных и лингвоспецифичных конструкций в разных языках входит в задачи многоязычной конструктикографии [Boas et al. 2019].

2.1. Объем

Объем Русского Конструктикона пока значительно превышает размеры других аналогичных ресурсов. На сегодняшний день, база данных Русского Конструктикона включает более 2200 конструкций, из которых первые 600 предварительно описаны. Ведется активная работа по пополнению базы новыми конструкциями. Для сравнения, шведский конструктикон, самый крупный из имеющихся, содержит описание около 400 конструкций, а конструктикон Беркли для английского языка включает 73 конструкции. Впрочем, такое расхождение в количественных оценках в определенной степени зависит от уровня обобщения (абстракции) в записи конструкций, который частично определяется целевой аудиторией (см. [раздел 2.2](#)), отбором материала (см. [раздел 2.3](#)) и связью конструктиконов с платформой FrameNet (см. [раздел 2.4](#)).

⁴ The Berkeley FrameNet Constructicon: <http://www1.icsi.berkeley.edu/~hsato/cxn00/21colorTag/index.html>.

⁵ The Swedish Constructicon: <https://spraakbanken.gu.se/konstruktikon>.

⁶ The Japanese FrameNet Constructicon, доступ через японский FrameNet: <http://jfn.st.hc.keio.ac.jp>.

⁷ The Brazilian Portuguese Constructicon, или the FN-Br Constructicon: <http://webtool.framenetbr.ufjf.br/index.php/webtool/report/cxn/main>.

⁸ The German Constructicon (GCon): <https://gsw.phil.hhu.de/constructicon/>.

2.2. Полифункциональность

Целевой аудиторией Русского Конструктикона являются исследователи языка (русисты, типологи), студенты и преподаватели русского языка как иностранного (РКИ), а также специалисты по автоматической обработке текста.

Целям лингвистического описания русских конструкций служит разветвленная система семантических тегов, ориентированная на терминологию, принятую в типологической литературе («универсальный грамматический набор» значений, см. [Plungian 2011: 65]). Студентам и преподавателям РКИ адресован перевод толкований на другие языки, а также пометы сложности конструкций, соответствующие уровню владения языком. Разметка синтаксических структур и якорных элементов конструкции ориентирована на задачи в том числе автоматического выявления конструкций в текстах.

Среди других проектов только шведский конструктикон [Lyngfelt et al. 2018b] рассматривает в качестве своей целевой аудитории не только профессиональных лингвистов, но и студентов, изучающих шведский как иностранный. Остальные ресурсы на данный момент преследуют скорее общетеоретические цели: описание конструкций и ограничений на их элементы, объединение конструкций в семьи. Заметим, что эти задачи решаются и в Русском Конструктиконе.

2.3. Отбор материала

При отборе материала предпочтение отдается неоднословным синтаксическим конструкциям, которые лежат на пересечении лексики и грамматики и являются «частично идиоматичными» (*partially schematic*, см. [Ehrlemark et al. 2016], ср. также принятое в русскоязычной традиции понятие «конструкции малого синтаксиса» [Apresjan, Iomdin 2010]). С одной стороны, такие конструкции слишком некомпозициональны, чтобы быть понятными неносителям русского языка. С другой стороны, они не настолько идиоматичны, чтобы считаться фразеологизмами. Как правило, такие конструкции имеют неизменяемую (якорную) часть и слоты, которые могут заполняться различными лексемами. Примером может служить конструкция (не) время VP-Inf (Самое время вершить большие дела), где якорной частью является слово *время*, допускающее перед собой отрицание, а слот заполняется глагольной группой в инфинитиве.

Наряду с лексикализованными конструкциями, конструктикон Беркли для английского языка содержит описание конструкций более высокого уровня абстракции в плане синтаксиса, такие как: *Absolute_clause* (конструкция с нефинитной клаузой), *Coordination* (сочинительная конструкция), *Infinitival_relative_modal* (инфинитивная конструкция, модифицирующая именную группу) и др. Конструктикону Беркли близки аналогичные ресурсы для японского, немецкого и бразильского португальского языков.

Шведский конструктикон так же, как и русский, фокусируется в основном на частично идиоматичных конструкциях, однако в перспективе предполагает покрытие всего лексико-грамматического континуума конструкций, которые будут иерархически организованы по принципу наследования признаков,

от более абстрактных синтаксических до более частных лексических. Это связано с общим представлением языка как разветвленной системы конструкций [Lyngfelt 2018: 6], разделяемым создателями всех конструктиконов. В Русском Конструктиконе идея системного представления конструкций реализуется в выявлении *семей* конструкций и объединении их в более крупные группы — *гнезда* [Endresen et al. 2019].

2.4. Связь с другими семантическими ресурсами

Конструктикон Филлмора и его аналоги «выросли» из платформы FrameNet, представляющей классификацию семантических фреймов в связи с их лексическим и синтаксическим выражением в том или ином языке [Fillmore et al. 2003]. FrameNet имеет жесткую структуру, устанавливая однозначное соответствие между формой и значением у каждого из аргументов предиката, а фреймовая семантика определяет роль отдельных элементов в общем фрейме. В этих рамках разрабатывать описание менее прозрачных неоднозначных выражений крайне затруднительно. Конструктикон же позволяет описывать единицы, имеющие сложную внутреннюю и внешнюю структуру в отношении как семантики, так и синтаксиса, и видится как дополнение к FrameNet [Ohara 2018: 143].

Многие конструктиконы дают отсылку к связанному с конструкцией фрейму (*evoked frame*), если таковой имеется (ср. классификацию конструкций в соответствии с тем, провоцируют ли они возникновение того или иного фрейма в сознании говорящего, предложенную в [Ohara 2018]). Так, например, конструкция *be_recip* (e. g. *She is good friends with her mother*) в конструктиконе Беркли адресует к фрейму *Reciprocity* (Отношение взаимности), см. [Lee-Goldman, Petruck 2018: 32]. В конечном счете, уникальная метка FrameNet ID позволяет связать узлы Конструктикона с другими семантическими ресурсами и онтологиями (WordNet, PropBank, SemLink+ и др. [Palmer et al. 2014]), что открывает перспективы для его использования в IT-технологиях.

Заметим, что связывание через фреймы не является единственно возможным рецептом: FrameNet не раз критиковали за неполноту фреймовой сети, отсутствие единого принципа выделения фреймов в разных частях ресурса (см. [Baker, Ruppenhofer 2002]; [Shi, Mihalcea 2005]). Так, шведский Конструктикон ориентируется не на Swedish FrameNet++, а на SALDO — систему, представляющую лексикографические ресурсы как иерархически организованную сеть.

Для Русского Конструктикона связывание с другими ресурсами — задача на перспективу. Аналогом FrameNet для русского является русский ФреймБанк⁹ [Lyashevskaya, Kashkin 2015a], [Lyashevskaya, Kashkin 2015b]. Однако, последний покрывает меньше лексических зон (фокус прежде всего на глаголах, а не на именных и адвербиальных предикатах), а уровень представления семантики менее абстрактен: так, например, в сеть связываются конструкции

⁹ Доступ к ресурсу ФреймБанк: <https://github.com/olesar/framebank>.

индивидуальных глаголов, а не ассоциированные с ними обобщенные фреймы. Вместе с тем, материал ФреймБанка и Конструктикона пересекается. С одной стороны, во ФреймБанке документируются все конструкции предиката, засвидетельствованные в текстах, в том числе и такие идиоматизированные, как «стоило NP-Dat V-Inf, как/и Cl». С другой стороны, в ограниченном объеме Русский Конструктикон все же включает конструкции «нестандартного» глагольного управления, и в этом отношении ФреймБанк является одним из источников для пополнения списков конструкций Конструктикона. Наконец, Конструктикон и ФреймБанк объединяют общие принципы выделения и классификации семантических ролей (меток, приписываемых элементам-слотам конструкции).

Несмотря на то, что прямой адресации между этими ресурсами сейчас нет, в будущем предполагается интегрировать информацию о тех фреймах, которые описаны во Фреймбанке, в документацию конструкций в Русском Конструктиконе. В более широком ключе, представляется перспективным ориентироваться, в рамках идеологии linked data, на инициативы по созданию мультязычных прикладных *Net-, *Bank- и *Link- ресурсов, а также исследовательских типологических баз данных по лексике, фразеологии и конструкциям.

С учетом названных отличительных особенностей Русского Конструктикона, прежде всего его полифункциональности, объема и содержания, разработан новый интерфейс, который ориентирован на решение и общетеоретических, и прикладных задач и подойдет пользователям разного типа. О нем речь пойдет в следующем **разделе 3**.

3. NP-Nom как на ладони: устройство ресурса и возможности поиска

На новом сайте Русского Конструктикона пользователям будут предложены различные возможности поиска конструкций. Для каждой конструкции на сайте будет отдельная словарная статья, включающая в себя несколько полей¹⁰, рассчитанных как на студентов РКИ, так и на лингвистов. Структура словарной статьи показана в **таблице 1**.

Поле **Имя конструкции/Name** содержит обобщенную запись конструкции, включающую как якорные элементы, так и слоты с ограничениями на их лексическое заполнение. Поле **Иллюстрация/Illustration** дает короткий репрезентативный пример употребления конструкции. Затем следует поле **Толкование/Definition**, в котором раскрывается семантика конструкции

¹⁰ Названия всех полей на сайте даны на английском, тогда как содержание полей дается на русском, за исключением принятых сокращений (типа NP для именной группы, Gen для генитива и пр.) и переводов толкований. Для удобства читателя мы приводим русские эквиваленты английских обозначений.

и указываются её отличия от синонимических вариантов¹¹. В поле **Перевод толкования/Translation of Definition** предлагается перевод толкования на английский и норвежский. Далее представлены: **Структура конструкции/Structure** в терминах универсальных зависимостей (UD)¹², **Якорные слова/Anchor words** (фиксированная часть конструкции), **Частотные заполнители слотов/Common fillers**, а также **Примеры/Examples** использования конструкции из НКРЯ. Поле **Уровень/CEFR Level** соотносит конструкцию с уровнем владения русским языком, которому она соответствует. Отдельное поле содержит **Семантические теги/Semantic tags**, которые обобщенно отражают семантику конструкции (подробнее см. **раздел 4**). Затем следуют поля, где указывается **Синтаксическая функция/Syntactic function** конструкции целиком и **Синтаксическая структура якорной части/Syntactic structure of the anchor**. В поле **Стилистический тег/Usage label** дается информация о регистре, в котором конструкция уместна. Наконец, поле **Комментарий/Comment** содержит дополнительную информацию о конструкции: ссылки на синонимичные и антонимичные конструкции, представленные в конструктиконе, как опция могут приводиться ссылки на научные исследования данной конструкции.

На стартовой странице нового сайта (**Browse**) пользователь сможет посмотреть полный список конструкций, по которым можно осуществлять поиск. Вводя в поисковую строку якорные слова или словоформы, можно будет искать конструкции по точным элементам. Например, если ввести в строку слово *ещё*, пользователю будет предложен список конструкций, включающих эту лексему.

Сайт будет содержать несколько вкладок.

Вкладка **Дневная порция/Daily Dose** будет выдавать случайную подборку из пяти конструкций, соответствующих уровню владения языком, задаваемому в меню. Такая обозримая порция конструкций может использоваться студентами РКИ для самостоятельной работы.

Вкладка **Продвинутый поиск/Advanced Search** будет предоставлять разнообразные возможности поиска, предусмотренные для лингвистов. Здесь можно будет задать грамматические и синтаксические параметры, по которым будет осуществляться запрос. В разделе **Морфология/Morphology** пользователь сможет задать поиск конструкции по таким грамматическим категориям, как **Падеж/Case**, **Род/Gender**, **Число/Number**, **Вид/Aspect**, **Наклонение/Mood**, **Время/Tense** и др.

¹¹ Толкования конструкций составляются таким образом, чтобы они были содержательны, но компактны и лаконичны. Мы по возможности избегаем узких лингвистических терминов, так чтобы и студенты РКИ без специальной лингвистической подготовки тоже могли эффективно пользоваться этим ресурсом. При составлении толкования мы опираемся на данные НКРЯ, сопоставляем конструкции внутри *семей* и *гнезд*, а также стараемся учитывать опубликованные результаты исследований конструкций русского языка, прежде всего работы Л. Л. Иомдина, Д. О. Добровольского, А. Н. Баранова, С. С. Сая, В. Ю. Апресян, Н. Р. Добрушиной, А. Б. Летучего, Е. Л. Вилинбаховой, Н. М. Стойновой и др. Ссылки на ученные работы планируется приводить в поле Комментария, а также привести на сайте проекта по адресу <https://site.uit.no/russian-constructicon/sources/>.

¹² Подробнее о принципах анализа в терминах UD см. <https://universaldependencies.org>.

Таблица 1. Структура словарной статьи в Русском Конструктиконе

Имя конструкции/ Name	какой/какое (там) VP-Inf.Imp! ¹³
Иллюстрация/ Illustration	<i>Какое там гулять! У нас нет времени!</i>
Толкование/ Definition	Конструкция обозначает, что говорящий отклоняет [действие] _{Action} , названное в предыдущем контексте, потому что считает это действие неуместным или несвоевременным. Обычно конструкция используется в ситуации диалога и выражает реакцию говорящего на предшествующую реплику собеседника. Конструкция также встречается при передаче внутренней речи говорящего, которая конструируется как диалог говорящего с самим собой.
Перевод толкования/ Translation of Definition	This construction means that the speaker rejects the suggested action named in the preceding context, because he/she finds this action inappropriate or poorly timed. Usually, this construction is used in a dialogue and expresses the speaker's reaction to the preceding remark of his/her conversation partner. The construction can also occur when presenting the inner monologue of the speaker that is constructed as a conversation in which the speaker is talking to him- or herself.
Структура конструкции/ Structure	[[advmod какой [advmod там]] VP-Inf.Imp!] [[advmod какое [advmod там]] VP-Inf.Imp!]
Якорные слова/ Anchor words	какой, какое, там
Частотные заполнители слотов /Common fillers	спать, работать, гулять, отдыхать
Примеры/Examples	— <i>Ладно, пойду [спать]_{Action}!</i> — <i>Какой [спать]_{Action}, Марина Михайловна? — возмутилась Ира. — Посидите с нами, хоть чайку попейте.</i> — <i>Вам суп [наливать]_{Action}?</i> <i>Я посмотрела на суп. Какое там [наливать]_{Action}! Серый, непрозрачный, он не вызывал у меня аппетита. — Нет, благодарю, ответила я.</i> <i>И вот народ галдит. Я там присел, попытался [помедитировать]_{Action} как-то. Какое там [медитировать]_{Action}!</i> <i>Ничего подобного там и близко невозможно — весь этот галдеж и тусовка.</i>

¹³ Через косую черту указываются равноправные варианты элементов конструкции, а в скобках даются факультативные элементы.

Имя конструкции/ Name	какой/какое (там) VP-Inf.Imp! ¹³
Уровень/CEFR Level	C1
Семантические теги/Semantic tags	Reaction to the previous discourse: Disagreement
Синтаксическая функция/Syntactic function	Clause
Синтаксическая структура якорной части/Syntactic structure of the anchor	Adj (APRO); Adv (PART)
Стилистический тег/Usage label	Colloquial
Комментарий/ Comment	см. конструкции какой/какое (там) VP-Past! <i>Какой там поел!</i> и какой/какое (там) Adv/Adj! <i>Какое там быстро!</i>

Например, можно будет найти все конструкции с именными группами в дательном падеже. Пользователю можно будет выбрать как один пункт из списка, так и комбинацию грамматических признаков. В разделе **Синтаксис/Syntax** параметрами поиска будут теги, характеризующие синтаксическую функцию конструкции в целом и состав якорных элементов конструкции. Здесь пользователь сможет найти, к примеру, все биклаузальные конструкции или все конструкции с производными предлогами. Кроме того, поиск будет возможен и по таким структурным особенностям, как инверсионный порядок слов, наличие отрицания и редупликация лексических элементов конструкции. В разделе **Семантика конструкции/Semantics** можно будет задать поиск конструкций по семантическим тегам и их подтипам. Так, можно будет найти все конструкции, обозначающие причину события (тег Cause) или условие его совершения (тег Condition), и проч. В разделе **Семантические роли/Semantic roles** можно будет задать поиск по тегам Агенса, Пациенса, Адресата и др., которыми размечены толкования конструкций и примеры употребления (см. **таблицу 1**). В разделе **Уровень/CEFR** можно будет ограничить выдачу конструкций только тем уровнем, который интересует пользователя. Уровни будут варьироваться от A1 до C2 по общеевропейской шкале оценки уровня владения языком. Пользователь сможет выбрать один или несколько уровней из списка. Кроме того, при оформлении запроса все указанные параметры можно будет комбинировать.

Также на сайте будет вкладка с инструкцией для пользователей (**Instructions**) и вкладка с информацией о ресурсе и разработчиках (**About**).

4. NP-Nom что надо: семантическая разметка конструкций

Для семантической разметки конструкций в Русском Конструктиконе используется двухуровневая система семантических помет (тегов), которая включает 53 общих тега и 87 уточняющих тегов-подтипов. Так, например, общий тег Фаза действия имеет подтипы Инхоатив, Континуатив, Терминатив и Кунктатив согласно [Plungian 1999]. Общий тег Сравнение имеет подтипы Сходство, Идентичность, Несходство, Контраст и Имитация, в соответствии с [Treis 2018]. В **таблице 2** приведен список из 25 «крупных» общих тегов, каждым из которых в Русском Конструктиконе размечено более 40 конструкций. В правом столбце таблицы приведено количество конструкций для каждого тега.

Таблица 2. Общие семантические теги, описывающие более 40 конструкций в базе Русского Конструктикона

Семантический тег	Кол-во
Интенсификация/Intensity	280
Оценка/Assessment	223
Отношение/Attitude	209
Сравнение/Comparison	167
Реакция на предшествующий дискурс/Reaction to the previous discourse	149
Темпоральность/Time	139
Категоризация объектов/Categorization	139
Образ действия/Manner	134
Измерение количества/Measure	120
Временное состояние/Temporary state	109
Фаза действия/Phase of action	105
Причина/Cause	96
Пространство/Spatial expressions	87
Условие/Condition	83
Побуждение к действию/Causation	82
Степень уверенности/Degree of certainty	77
Уступка/Concession	69
Временная референция/Timeline	59

¹⁴ Тег Миративность/Mirative широко употребляется в типологической литературе для обозначения лингвистической категории, выражающей удивление говорящего неожиданно установленным фактом [DeLancey 1997]; [Aikhenvald 2012]. Примерами миративных конструкций являются NP-Nom (как) давай VP-Inf (он как давай петь), XP-таки, (а Cl)! (Пришёл-таки, а мы уже и не ждали!) и др.

¹⁵ Термин Прохибитив/Prohibitive описывает лингвистические средства, выражающие запрет на совершение действия: нечего VP-Inf (Нечего баловать детей). Следуя [Rakhilina 2013], в класс прохибитивных конструкций мы также включаем континуативные прохибитивы, выражающие побуждение прервать некоторую ситуацию: хорош VP-Inf! (Эй, наверху! Хорош прыгать!), полно Pron-Dat VP-Inf (полно тебе плакать!), по-VP-Past и хватит (погуляли и хватит) и др.

Семантический тег	Кол-во
Миративность/Mirative ¹⁴	57
Степень точности/Degree of accuracy	52
Аддитив/Additive	50
Прохибитив/Prohibitive ¹⁵	49
Цель/Purpose	49
Глагольная множественность/Pluractionality	47
Эмфаза/Emphasis	41

Те общие семантические теги, которые не вошли в **таблицу 2**, описывают меньшие группы конструкций: Опасение/Apprehension, Угроза/Threat, Просьба/Request, Результат/Result, Дистрибутив/Distributive, Каритив/Caritive; Конатив/Conative, Кванторы/Quantification, Реализованная возможность/Actuality, Calculation/Счет, Подмножество/Subset, Инструмент/Instrument и др.

Многие конструкции имеют более одного общего семантического тега. Так, 933 конструкции (42%) имеют по два общих тега, а 144 конструкции имеют по 3 общих тега. Например, конструкция (уж/ну) Pron-Pron, a Cl (уж кому-кому, а ему можно доверять) имеет теги Реакция на предшествующий дискурс, Степень уверенности и Исключительность, а конструкция Noun-Nom на Noun-Loc (ошибка на ошибке) имеет набор тегов Измерение количества, Отношение и Интенсификация.

Система семантических тегов используется для выявления, анализа и системного пополнения семей и гнезд конструкций (см. [Endresen et al. 2019]). По семантическим тегам можно осуществлять поиск конструкций и их групп и сравнивать этот материал с конструкциями других языков. Кроме того, в процессе работы исследуется, насколько широко конструкции русского языка покрывают набор значений, которые в других языках могут выражаться грамматически.

5. Во всех отношениях AP: вместо заключения

Конструкции представляют собой то звено, где происходит пересечение лексики и грамматики. При этом уже ведется работа по пополнению базы морфологическими конструкциями, и таким образом учитывается роль словообразовательных средств: ср. конструкцию по-XP-ee(ей) (смотреть на жизнь повеселее) для выражения семантики аттенуативного сравнения [Boguslavsky, Iomdin 2009], пере-V все NP-Acc.Plur (перемыть все тарелки в доме) с семантикой дистрибутива [Tatevosov 2013], и пр.

В целом, обширный материал по конструкциям русского языка, собранный и описанный в Конструктиконе, дает возможность уточнить наше понимание устройства русской грамматической системы. В ходе проекта выверена детальная и многоуровневая семантико-синтаксическая разметка конструкций, на основе которой разработана эффективная методика выявления семей и гнезд конструкций (позволившая, в частности, значительно и системно расширить инвентарь конструкций в базе).

Ближайшей задачей Русского Конструктикона является дальнейшая систематизация семантико-синтаксической иерархии, которая позволит структурировать толкования конструкций, а в дальнейшем добиться большей адекватности в выборе их переводных эквивалентов — для английского, норвежского и других языков. Опыт работы над Русским Конструктикonom может помочь и в создании аналогичных ресурсов для других языков.

References

1. *Aikhenvald, A.* (2012), The essence of mirativity, *Linguistic Typology*, Vol. 16, pp. 435–485.
2. *Apresjan, Ju. D., Iomdin, L. L.* (2010), Small syntax constructions [Konstruktsii malogo sintaksisa]. In Apresjan Ju. D., Boguslavsky I. M., Iomdin L. L., Sannikov V. Z. (eds.). *Theoretical problems of the Russian syntax: Interaction of grammar and lexicon [Teoreticheskie problemy russkogo sintaksisa: Vzaimodeistvie grammatiki i slovary]*, Moscow, pp. 59–280.
3. *Baker, C., and Ruppenhofer, J.* (2002), Framenet's frames versus Levin's verb classes, *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*, Berkeley (CA): BLS, pp. 27–38.
4. *Boas, H. C., Lyngfelt, B., Torrent, T. T.* (2019), Framing constructicography, *Lexicographica*, Vol. 35(1), pp. 41–85.
5. *Boguslavsky, I., Iomdin, L.* (2009), Semantics of Attenuated Comparatives in Russian, *Proceedings of the 4th International Conference on Meaning-Text Theory. Observatoire de linguistique Sens-Texte (OLST)*, pp. 65–77.
6. *Croft, W.* (2001), *Radical Construction Grammar*, Oxford University Press, Oxford.
7. *DeLancey, S.* (1997), Mirativity: The grammatical marking of unexpected information. *Linguistic Typology*, Vol. 1, pp. 33–52.
8. *Ehrlemaek, A., Johansson, R., Lyngfelt, B.* (2016), Retrieving Occurrences of Grammatical Constructions, *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, Osaka, Japan, pp. 815–824.
9. *Endresen, A., Klezovich, A., Lyashevskaya, O., Mordashova, D., Nordrum, M., Rakhilina, E., Tyers, F., Zhukova, V.* (2019), Building a Constructicon for Russian: How to identify families of constructions, A talk presented at the 15th International Cognitive Linguistics Conference (ICLC-15), Kwansei Gakuin University, Nishinomiya, Japan, August 2019.
10. *Fillmore, C. J.* (2008), Border Conflicts: FrameNet Meets Construction Grammar, Bernal, E., DeCesaris, J. (Eds.), *Proceedings of the XIII EURALEX International Congress*, Barcelona, Universitat Pompeu Fabra, pp. 49–68.
11. *Fillmore, Ch. J., Kay, P., O'Connor, M. C.* (1988), Regularity and idiomaticity in grammatical constructions: The case of *let alone*, *Language*, Vol. 64(3), pp. 501–538.
12. *Fillmore, C. J., Johnson, C. R., Petruck, M. R.* (2003), Background to FrameNet, *International Journal of Lexicography*, Vol. 16.3, pp. 235–250.

13. *Fillmore, C. J., Lee-Goldman, R., Rhodes, R., (2012), The FrameNet constructicon*, Boas, H. C., Sag, I. A. (eds.), *Sign-based construction grammar*, CSLI Publications, Stanford, CA, pp. 309–372.
14. *Goldberg, A. E. (2006), Constructions at Work: The Nature of Generalizations in Language*, Oxford University Press, Oxford.
15. *Janda, L. A., Lyashevskaya, O., Nessel, T., Rakhilina, E., Tyers, F. M. (2018), A constructicon for Russian: Filling in the Gaps*, Lyngfelt, B., Borin, L., Ohara, K., Torrent, T. T. (eds.), *Constructicography: Constructicon development across languages*, John Benjamins, Amsterdam, pp. 165–181.
16. *Lee-Goldman R., Petruck, M. (2018), The FrameNet constructicon in action*. Lyngfelt, B., Borin, L., Ohara, K., Torrent, T. T. (eds.), *Constructicography: Constructicon development across languages*, John Benjamins, Amsterdam, pp. 19–40.
17. *Lyashevskaya, O., Kashkin, E. (2015a), FrameBank: a database of Russian lexical constructions*, Khachay, M. Yu., Konstantinova, N., Panchenko, A., Ignatov, D. I., Labunets, G. V. (eds.), *Analysis of Images, Social Networks and Texts. Fourth International Conference, AIST 2015, Yekaterinburg, Russia, April 9–11, 2015, Revised Selected Papers. Communications in Computer and Information Science*, Vol. 542, Springer, pp. 337–348.
18. *Lyashevskaya, O., Kashkin, E. (2015b), Annotation of lexical constructions in Russian FrameBank [Tipy informacii o leksičeskix konstrukcijax v sisteme FrameBank]*, *Trudy Instituta russkogo jazyka imeni V. V. Vinogradova [Proceedings of Vinogradov Institute of the Russian Language]*, Vol. 6, pp. 464–555.
19. *Lyngfelt, B. (2018), Introduction: Constructicons and constructicography*, Lyngfelt, B., Borin, L., Ohara, K., Torrent, T. T. (eds.), *Constructicography: Constructicon development across languages*, John Benjamins, Amsterdam, pp. 1–18.
20. *Lyngfelt, B., Bäckström, L., Borin, L., Ehrlemark, A., Rydstedt, R. (2018b), Constructicography at work: Theory meets practice in the Swedish constructicon*. Lyngfelt et al. 2018a, pp. 41–106.
21. *Lyngfelt, B., Borin, L., Ohara, K., Torrent, T. T. (eds.) (2018a), Constructicography: Constructicon development across languages*, John Benjamins, Amsterdam.
22. *Ohara, K. (2018), Relations between frames and constructions: A proposal from the Japanese FrameNet constructicon*, Lyngfelt, B., Borin, L., Ohara, K., Torrent, T. T. (eds.), *Constructicography: Constructicon development across languages*, John Benjamins, Amsterdam, pp. 141–164.
23. *Palmer, M., Bonial, C., McCarthy, D. (2014), Semlink+: FrameNet, VerbNet and event ontologies*, *Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929–2014)*, pp. 13–17.
24. *Plungian, V. A. (1999), A typology of phasal meanings*, Abraham, W., Kulikov, L. (eds.) *Tense-aspect, transitivity, and causativity: Essays in honor of Vladimir Nedjalkov*, John Benjamins, Amsterdam.
25. *Plungian, V. A. (2011), An introduction to grammatical semantics: Grammatical meanings and grammatical systems in the languages of the world [Vvedenie v grammatičeskiju semantiku: Grammatičeskie značeniya i grammatičeskie sistemy jazykov mira]*. Russian State University for the Humanities, Moscow.

26. *Rakhilina, E. V.* (ed.) (2010), *Linguistics of constructions* [Lingvistika konstrukcij], Azbukovnik, Moscow.
27. *Rakhilina, E. V.* (2013), *Conductor, press the brakes... [Konduktor, nažmi na tormoza...]*, *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"* [Komp'juternaja lingvistika i intellektual'nye texnologii: po materialam meždunarodnoj konferencii «Dialog»], Issue 12 (19), Russian State University for the Humanities, Moscow, pp. 665–673.
28. *Shi, L., Mihalcea, R.* (2005), *Putting pieces together: Combining FrameNet, VerbNet and WordNet for robust semantic parsing*, *International conference on intelligent text processing and computational linguistics*, Springer, Berlin, Heidelberg, pp. 100–111.
29. *Stojnova, N. M.* (2007), *The construction vzjat' i sdelat' 'take and do' in Russian* [Konstrukcija vzjat' i sdelat' v russkom jazyke], *Dudčuk, F. I., Ivlieva, N. V., Podobrjaev, A. V.* (eds.), *Structures and interpretations: Studies in theoretical and applied linguistics by young researchers* [Struktury i interpretacii: raboty molodyx issledovatelej po teoretičeskoj i prikladnoj lingvistike], Moscow State University, Moscow, pp. 158–189.
30. *Tatevosov, S. G.* (2013), *Destructive notes on the distributive Aktionsart* [Destrukтивnye zametki o raspredelitel'nom sposobe dejstvija], *Russkij jazyk v naučnom osveščanii*. Vol. 2(26), pp. 30–51.
31. *Treis, Y.* (2018), *Comparative Constructions: An Introduction*, *Linguistic Discovery*, Vol. 16(1). pp. i–xxvi.