

The Typological constructicon database

Nikita Muravyev

University of Hamburg
nikita.muraviev@gmail.com

Nikita Gordeev

NRU HSE (Moscow)
permanentdaylight@mail.ru

Илья Макаrchук

NRU HSE (Moscow)
ilya.makarchuk@gmail.com

Maria Kukushkina

NRU HSE (Moscow)
erinqa.hecil@gmail.com

Anton Buzanov

NRU HSE (Moscow)
anton.buzanov.00@gmail.com

Abstract

The report introduces a new resource: the Typological Constructicon database. This resource contains an inventory of constructions of selected semantic fields in a number of languages of different areal and genetic affiliation. The constructions are labeled according to a number of semantic and morphosyntactic parameters and provided with a detailed description and illustrative examples.

Keywords: constructicon, typology, constructions grammar, carative constructions

DOI: 10.28995/2075-7182-2023-22-1091-1096

База данных «Типологический конструктикон»

Никита Муравьев

Университет Гамбурга
nikita.muraviev@gmail.com

Никита Гордеев

НИУ ВШЭ (Москва)
permanentdaylight@mail.ru

Илья Макаrchук

НИУ ВШЭ (Москва)
ilya.makarchuk@gmail.com

Мария Кукушкина

НИУ ВШЭ (Москва)
erinqa.hecil@gmail.com

Антон Бузанов

Университет Гамбурга
anton.buzanov.00@gmail.com

Аннотация

В докладе представляется новый ресурс: база данных «Типологический конструктикон». Данный ресурс содержит инвентарь конструкций избранных семантических полей в ряде языков различной ареальной и генетической принадлежности. Конструкции размечены по ряду семантических и морфосинтаксических параметров и снабжены подробным описанием и иллюстрирующими примерами.

Ключевые слова: конструктикон, типология, грамматика конструкций, каритивные конструкции

1 Введение

В центре внимания лингвистической типологии всегда были преимущественно грамматические единицы языка, которые сопоставлялись между языками прежде всего, с опорой на описательные грамматики. В последние два десятилетия с развитием полевой и корпусной лингвистики

подобные исследования стали включать в себя первичные данные опросов носителей и текстов. Также активно развивается лексическая типология, в центре внимания которой находятся лексические значения и способы их выразить. Существует, однако, огромный пласт единиц, называемых конструкциями, которые изучаются в отдельных языках, но практически не попадают в сферу интересов типологов. Проект «Типологический конструктикон» ставит своей целью привнести типологическую перспективу в изучение конструкций. Разрабатываемая на материале ряда языков типологическая база данных представляет собой аннотированный и иллюстрированный примерами инвентарь конструкций, который позволяет изучать морфологические, синтаксические и семантические взаимосвязи между конструкциями как внутри языка, так и в межязыковом сопоставлении.

2 Грамматика конструкций

Конструкции являются предметом изучения одного из важнейших и наиболее активно развивающихся когнитивно-функциональных направлений лингвистики, Грамматики конструкций (Fillmore et al. 1988, Goldberg 1995). Конструкцией в рамках этого направления считается языковая единица, состоящая из формы и соответствующего ей значения. Грамматика конструкций как модель языка представляет собой структурированный инвентарь конструкций, охватывающий все уровни языка от морфологии до дискурса. Грамматика конструкций существует в русле когнитивно-функциональных узусно-ориентированных подходов к языку, поэтому активно использует данные корпусов. Вследствие своего промежуточного положения между лексикой и грамматикой конструкции довольно скудно представлены в существующих словарях и грамматических описаниях. Поэтому в последнее время для отдельных языков активно разрабатываются специальные лексикографические ресурсы, конструктиконы, ср. к примеру Русский конструктикон (Janda et al. 2018).

3 Типологический конструктикон

Типологический конструктикон представляет собой прежде всего научно-исследовательский ресурс. Он задуман как инструмент для типологов и исследователей конструкций, позволяющий на примере избранных семантических полей строить типологические и теоретические обобщения о структуре и функционировании языка, на материале, который до сих пор относился к периферии как лексической, так и грамматической семантики. Второе его предназначение – справочно-переводной ресурс для изучающих иностранные языки, так как содержит переводные эквиваленты конструкций с подробным описанием и примерами. Однако в отличие от существующих монолингвальных конструктиконов его основной задачей является не максимально полный охват существующих конструкций, а сбор данных по отдельным семантическим полям, наиболее насыщенным конструкциями.

Реляционная база данных PostgreSQL устроена таким образом, чтобы обеспечить представление морфологических и семантических особенностей данных. Между конструкциями из разных языков связи устанавливаются вручную. Мы предполагаем, что тщательная разметка способна обеспечить поиск похожих конструкций в разных языках. Кроме того, наша разметка позволяет искать конструкции, похожие по как по разному количеству параметров, так и по разным их комбинациям. Таким образом, в группе похожих конструкций могут оказаться не только синонимичные конструкции, но и конструкции устроенные одинаково с точки зрения морфосинтаксиса.

Интерфейс Типологического конструктикона включает возможности простого и расширенного поиска конструкций, а также общего обзора всех конструкций в рамках заданного семантического поля. Кроме того, сайт содержит общую информацию о проекте, его авторах, представленных семантических полях, а также инструкцию по пользованию ресурсом, релевантную библиографию и ссылки на другие ресурсы по конструкциям.

3.1 Простой поиск

Простой поиск представляет собой строку поискового запроса, в которую можно ввести конструкцию целиком или ключевое слово. В выдаче показывается список содержащихся в базе

данных конструкций, упорядоченных по релевантности, а также список наиболее релевантных аналогов найденных конструкций в языках, отличных от языка запроса. Гиперссылки в списке ведут на страницы соответствующих конструкций с подробной информацией о каждой, см. далее раздел 3.3.

3.2 Расширенный поиск

Расширенный поиск представляет более детальный поиск с применением различных фильтров, аналогично тому, как это устроено в корпусах (ср., к примеру, лексико-грамматический поиск в НКРЯ). Аннотация каждой конструкции включает в себя следующие параметры.

- запись конструкции в графике оригинала
- латинская транслитерация
- глоссирование
- якорные элементы и их значения
- информация о языке конструкции (семья, группа, язык)
- семантическое поле конструкции
- семантический класс и подкласс конструкции
- короткое описание значения
- примеры
- и др.

3.3 Обзор конструкции

При выборе конкретной конструкции из списка выдачи открывается страница со справочной информацией и примерами к данной конструкции. Набор характеристик соответствует параметрам расширенного поиска.

3.4 Обзор семантического поля

Отдельно от поиска на сайте Типологического конструктиконе имеется страница “Обзор”, содержащая данные о конструкциях в форме таблицы. При открытии данной страницы пользователю предлагается выбрать семантическое поле, класс и подкласс, а также языки, семьи и группы с интересующими конструкциями. В выдаче, помимо заданной пользователем информации о языке и семантике, предлагается список релевантных конструкций с переводом, глоссами, валентностной рамкой и информацией о деривационном источнике якорных слов. Страница обзора также доступна по гиперссылке с названием конкретного семантического поля в разделе “Семантическое поле” при описании конструкции.

4 Каритивные конструкции в конструктиконе

Разберем в качестве примера семантическое поле каритивов и их представление в Типологическом конструктиконе. Каритив представляет собой грамматическое значение отсутствия в ситуации ожидаемого участника, ‘без X’. Морфологически каритивная семантика обычно выражается при помощи специализированных предлогов/послелогов, как *без* в русском, *sans* во французском или аффиксов, как *-løs* в норвежском, *-li* в хантыйском. Каритивам посвящен ряд конкретно-языковых и типологических работ, ср. (Оскольская и др. 2020) и обзор литературы в статье. Помимо своей грамматической функции кодирования глагольной валентности на отсутствующего участника, каритивные показатели также входят в состав семантически нерегулярных конструкций, в которых исходное каритивное значение взаимодействует со значениями других компонентов, образуя новое значение. Такие конструкции относятся к целому ряду семантических фреймов, в терминологии Ч. Филлмора: лишения, лишения, уступки и избегания, которые будут проиллюстрированы далее.

В ресурсе представлены на данный момент преимущественно крупные языки Евразии: русский, английский, норвежский, французский, испанский языки, также в ближайшей перспективе будут добавлены немецкий, арабский, иврит и японский. В то же время, немаловажным представляется включение в базу также материала малых языков, находящихся под угрозой исчезновения. В настоящий момент в базу включены данные севернохантыйского языка, в ближайшее

время планируется сбор данных по северносаамскому языку. Материал планируется в дальнейшем расширять на другие языки как в Евразии, так и за её пределами. Источником материала служат полевые исследования авторов, онлайн опросники на платформе Google Forms и данные текстовых корпусов.

4.1 Краткий обзор каритивных конструкций

В семантическом поле каритивов можно выделить по меньшей мере четыре основных фрейма, вокруг которых сосредоточены встречающиеся в обсуждаемых языках конструкции: лишение, лишенность, уступка и избегание. Семантический фрейм **лишения** описывает динамическую ситуацию утраты сущности или свойства и состоит из двух участников: исходный обладатель утрачиваемой сущности и утрачиваемое обладаемое. Лишение может выражаться переходным глаголом, в котором обладатель заполняет позицию субъекта, а обладаемое — прямого объекта, как *X* потерял *Y*-асс в русском языке (1). Альтернативный способ — при помощи непереходной конструкции с предлогом, как *X s'est retrouvé Y* во французском языке в (2).

(1) *Петя потерял ключи.*

(2) *Pierre s' est re-trouvé sans argent.* Французский
 Пьер REFL быть.PRS REP-найти.PST.ПТСР без деньги
 'Пьер оказался без денег.'

Семантический фрейм **лишенности** описывает временное состояние отсутствия ожидаемой или необходимой сущности, как норвежская конструкция *X sitte uten Y* в (3).

(3) *Det var kjedelig å sitte uten å gjøre noe.* Норвежский
 это быть.PST скучно INF сидеть без INF делать что.INDEF
 'Было скучно сидеть без дела.'

Семантический фрейм **уступки** описывает деятельность, как в (4) из испанского языка, или существование, как в (5) из севернохантыйского, в отсутствие некоторой сущности или дополнительной ситуации, которая оценивается говорящим как ожидаемая или необходимая.

(4) *Pedro se las arregla sin electricidad.* Испанский
 Педро REFL ЗРЛ.О управлять.NPST без электричество
 'Педро справляется без электричества.'

(5) *pet 'a-jen atəl pa xətəl oləm-li pak-əs.* Севернохантыйский
 Петя-POSS.2SG ночь и день сон-CAR терпеть-PST
 'Петя продержался сутки без сна.'

Семантический фрейм **избегания** описывает либо целенаправленное достижение отсутствия нежелательной с точки зрения говорящего ситуации, как в английском примере (6), либо деятельность или существование в отсутствие данной ситуации, как в русском примере (7).

(6) *Peter avoided talking to his father.* Английский
 'Питер избежал разговора с отцом.'

(7) *Мероприятие прошло без эксцессов.*

Уже по приведенным выше примерам можно судить о широком разнообразии каритивных конструкций. В разных языках можно встретить переходные (1, 6) и непереходные конструкции. Среди непереходных можно выделить активные (3, 5) и рефлексивные (2, 4) конструкции, конструкции с каритивным предлогом (2, 3, 4, 7) и морфологическим каритивным показателем (5).

4.2 Возможные применения ресурса на примере каритивов

Типологический конструктикон может быть применен для широкого круга исследовательских вопросов. Приведем для примера некоторые из них применительно к рассмотренным выше каритивным конструкциям:

Морфологическая структура каритивных конструкций. Ресурс содержит информацию, из каких компонентов (якорных слов, слотов) состоит конструкция и какой компонент является основным, отвечающим за соответствующую семантическую функцию. Так, например, в представленных языках семантика лишения может выражаться как лексически в составе значения глагола, так и морфологически, при помощи специального показателя, который, в свою очередь, может быть как аффиксом, так и, например, предлогом.

Переходность и аргументная структура конструкций. Конструкция может быть построена по переходной (*потерять* и его аналоги) квазипереходной (*лишиться* и его аналоги) или непереходной модели. В соответствии с этим можно сравнивать между собой средства кодирования участников конструкции (падежи, предлоги).

Полисемия конструкций. Ресурс позволяет выяснить, является ли конструкция моносемичной или полисемичной и какой набор функций выражается конструкцией. Также можно установить, насколько равномерно распределены функции между конструкциями, существуют ли в языке доминантные или, напротив, специализированные конструкции.

Композициональность конструкций. Предметом сравнения в Типологическом конструктиконе может быть и то, насколько семантически прозрачной является конструкция по своему составу.

Деривационные источники конструкций. Каритивные конструкции часто построены на основе десемантизированных вспомогательных глаголов. Так, например, обходиться и *arreglarse*. Сопоставляя между собой деривационные источники конструкций по языкам, можно установить, из каких по семантике единиц чаще всего или реже всего развиваются каритивные конструкции.

Генетические и ареальные особенности. Как устроены конструкции с точки зрения структуры и семантики в близкородственных и географически близких языках? Можно ли проследить какие-то генетически или ареально обусловленные закономерности?

5 Заключение

База данных «Типологический конструктикон» представляет собой одновременно справочный ресурс по конструкциям их межъязыковым соответствиям и первый шаг в направлении типологического изучения на данный момент малоизученных многословных лексико-грамматических единиц. Подробное изучение и каталогизация столь разностороннего материала подобных единиц является перспективным направлением будущих исследований.

Благодарности

Статья подготовлена по результатам проекта «Типологический конструктикон» при поддержке фонда «Гуманитарные исследования» ФГН НИУ «Высшая школа экономики» в 2022- 2024 году. Авторы благодарят Е. В. Рахилину, В. А. Жукову, А. Б. Летучего, Т. В. Симонову, А. Рохаса, Д. Д. Мордашову и П. А. Бычкову за подробное обсуждение материала и концепции ресурса.

References¹

- [1] Fillmore C. J., Kay P., O'connor M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone //Language. pp. 501-538.
- [2] Goldberg A. E. (1992). Construction grammar: a construction grammar approach to argument structure. — University of Chicago Press
- [3] Oskol'skaja S. A., Zaika N. M., Klimenko S. B., Fedotov M. L. (2020). Opredelenie karitiva kak sravnitel'nogo ponjatija (A definition of caritive as a comparative concept) //Voprosy jazykoznanija. – 2020. – No. 3. pp. 7-25.
- [4] Janda L. A., Lyashevskaya O., Nessel T., Rakhilina E., Tyers F. M. (2018). A constructicon for Russian: Filling in the gaps. // Constructicography. John Benjamins, pp. 165-182.

References²

- [1] Fillmore C. J., Kay P., O'connor M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone //Language. pp. 501-538.
- [2] Goldberg A. E. (1992). Construction grammar: a construction grammar approach to argument structure. — University of Chicago Press
- [3] Оскольская С. А., Заика Н. М., Клименко С. Б., Федотов М. Л. (2020). Определение каритива как сравнительного понятия //Вопросы языкознания. – 2020. – №. 3. с. 7-25.
- [4] Janda L. A., Lyashevskaya O., Nessel T., Rakhilina E., Tyers F. M. (2018). A constructicon for Russian: Filling in the gaps. // Constructicography. John Benjamins, pp. 165-182.

¹ References, Scopus version

² References, РИНЦ version