

Москва, 14–17 июня 2023 г.

## **Diachronicon: a new resource for the study of Russian constructions in a microdiachronic perspective**

**Evgeniya Budennaya**  
HSE University  
Institute of Linguistics,  
jane.sdrv@gmail.com

**Maxim Bazhukov**  
HSE University  
mbazhukov@hse.ru

**Liubov Barkova**  
HSE University  
labarkova@hse.ru

**Darya Kharlamova**  
HSE University  
dskharlamova@edu.hse.ru

**Akhmed Dugrichilov**  
HSE University  
akhmed1797@yandex.ru

**Tatiana Reznikova**  
HSE University  
tanja.reznikova@gmail.com

**Anastasia Yakovleva**  
HSE University  
avyakovleva@hse.ru

**Kristina Litvintseva**  
HSE University  
tinalitvina@gmail.com

**Anastasia Andreeva**  
HSE University  
aaandreeva\_15@edu.hse.ru

### **Abstract**

The article is devoted to the linguistic characteristics of the database "Diachronicon" and describes the features of the diachronic markup of Russian language constructions, as well as tags specially designed for searching through a diachronic database. A special comment field used in the database is separately justified. In addition, the computer interface of the "Diachronicon" is presented and described.

The developed resource provides extensive opportunities for systematic study of not only specific constructions, but also general mechanisms of idiomatization and grammaticalization. The database allows the researcher to simultaneously compare several separate plots, search through a list of constructions and their characteristics in diachrony, track the history of syntactic and semantic changes and limitations of compatibility of different constructions.

**Keywords:** database; construction; Russian language; diachrony

**DOI:** 10.28995/2075-7182-2023-22-1041-1051

## **Диакроникон: новый ресурс для изучения русских конструкций в микродиакронической перспективе**

**Евгения Буденная**  
НИУ ВШЭ  
Институт языкознания РАН  
jane.sdrv@gmail.com

**Максим Бажуков**  
НИУ ВШЭ  
mbazhukov@hse.ru

**Любовь Баркова**  
НИУ ВШЭ  
labarkova@hse.ru

**Дарья Харламова**  
НИУ ВШЭ  
dskharlamova@edu.hse.ru

**Ахмед Дугричилов**  
НИУ ВШЭ  
akhmed1797@yandex.ru

**Татьяна Резникова**  
НИУ ВШЭ  
tanja.reznikova@gmail.com

**Анастасия Яковлева**  
НИУ ВШЭ  
avyakovleva@hse.ru

**Кристина Литвинцева**  
НИУ ВШЭ  
tinalitvina@gmail.com

**Анастасия Андреева**  
НИУ ВШЭ  
aaandreeva\_15@edu.hse.ru

### Аннотация

Статья посвящена лингвистическим характеристикам базы данных «Диахроникон». В статье описаны особенности диахронической разметки конструкций русского языка, а также теги, специально разработанные для поиска по диахронической базе данных. Отдельно обосновывается особое поле комментария, используемое в базе. Кроме того, представлен и описан компьютерный интерфейс «Диахроникона».

Разработанный ресурс предоставляет обширные возможности для системного изучения не только конкретных конструкций, но и общих механизмов идиоматизации и грамматикализации. База данных позволяет исследователю одновременно сравнивать несколько отдельных сюжетов, проводить поиск по списку конструкций и их характеристик в диахронии, отслеживать историю синтаксических и семантических изменений и ограничений сочетаемости разных конструкций.

**Ключевые слова:** база данных; конструкции; русский язык; диахрония

## 1 Введение

«Диахроникон» представляет собой проектируемую базу данных конструкций русского языка в микродиахронической перспективе (XVIII–XXI вв.), с подробным систематизированным описанием синтаксических и семантических изменений, происшедших с каждой из конструкций за исследуемый период. Под *конструкциями*, в рамках проекта, подразумеваются идиоматизированные единицы на пересечении лексики и грамматики, состоящие из неизменяемой (якорной) части и слотов, которые могут заполняться различными лексемами [Fillmore et al., 1988; Croft, 2001; Рахилина, 2010; Эндерсен и др., 2020]: N-Gen.PL Cop (*хоть пруд пруди, не дай бог* (NP-Dat) VP-Pfv.Inf, NP-Dat *не улыбается* VP/NP и др. По своей сути, «Диахроникон» углубляет ранее разработанный ресурс «Русский конструктик»<sup>1</sup> [Janda et al. 2018; Эндерсен и др. 2020] в область исторических изменений и является первой базой данных для системного диахронического изучения фразеологизированных конструкций русского языка.

Объем «Диахроникона» значительно превышает аналогичные проекты для других языков, где внимание в первую очередь уделяется синхронному состоянию: так, если синхронные системы конструкций (*конструктики*, в терминологии Ч. Филлмора [Fillmore, 2008]), в последние годы активно изучаются и разрабатываются<sup>2</sup>, то диахронические изменения исследовались более фрагментарно и в первую очередь только для определенного вида конструкций (см., например, работы [Падучева, 2018; Сай, 2010, 2014] для русского языка). С этой точки зрения, «Диахроникон», основанный на данных Национального корпуса русского языка (НКРЯ)<sup>3</sup>, Генеральном интернет-корпусе русского языка (ГИКРЯ, использовался как дополнительный источник в случае недостаточной выборки по данным НКРЯ)<sup>4</sup> и сервисе Google Ngram Viewer<sup>5</sup> (использовался как агрегатор текстов, а не как корпус; к этому источнику обращались только в случае недостаточной выборки, с дальнейшей верификацией даты вхождения), представляет собой самый полный ресурс, позволяющий одновременно сравнивать множество отдельных сюжетов и проводить поиск по заданному списку конструкций и их характеристик в динамике, а также отслеживать историю синтаксических и семантических изменений, которые происходили с конструкцией.

Заделом для базы данных «Диахроникон» послужили корпусные исследования студентов Школы лингвистики НИУ ВШЭ по диахронии идиом, выполненные в рамках курса «Академическое письмо» в 2015–2022 гг. [Klezovich, Golosov, 2018; Клезович, Голосов, Фесенко, 2018; Добрушина, Ахапкина, 2019] (в работах использовались выборки из 300 примеров). На данный момент с помощью коллектива преподавателей и студентов Школы филологических наук и Школы лингвистики НИУ ВШЭ, ведется активная работа по разработке и редактированию системы разметки конструкций, а также созданию компьютерной оболочки базы данных. Объем размеченных единиц составляет более 100 конструкций. Все конструкции были взяты из собрания [Баранов, Добровольский, 2016] и содержат единицы, в той или иной мере обозначающие количество, время, отрицание и дейксис. Критерием отбора конструкций была презентативность их

<sup>1</sup> Проект разработан исследователями Школы лингвистики НИУ ВШЭ (Москва) совместно с Арктическим университетом Тромсё (Норвегия) и расположен по адресу <https://constructicon.ruscorpora.ru/>

<sup>2</sup> См., например, ресурсы <https://spraakbanken.gu.se/konstruktikon> для шведского языка, <http://jfn.st.hc.keio.ac.jp> для японского, <http://webtool.framenetbr.ufjf.br/index.php/webtool/report/cxn/main> для бразильского португальского, <https://gsw.phil.hhu.de/constructicon/> для немецкого и уже упомянутый «Русский конструктик» для русского языка.

<sup>3</sup> <https://ruscorpora.ru/>

<sup>4</sup> <http://www.webcorpora.ru/>

<sup>5</sup> [books.google.com/ngrams](https://books.google.com/ngrams)

вариативности и различных синтаксических и семантических изменений. При разметке проводился ручной отбор и верификация примеров с целью устранить ошибки распознавания и, в некоторых случаях, отделить примеры с нужной конструкцией от примеров с похожим на неё сочетанием слов.

Настоящая статья посвящена лингвистическим характеристикам базы данных «Диахроникон». В разделе 2 будут описаны особенности диахронической разметки конструкций. Раздел 3 будет посвящен компьютерной оболочке базы данных. Заключительный раздел 4 суммирует состояние проекта на текущий момент и обозначает его дальнейшие перспективы.

## 2 Диахроническая разметка

В ходе работы над диахронической разметкой команда проекта учитывала необходимость ее дальнейшей интеграции в систему поиска: вводимые пометы позволят выявлять закономерности в семантической и синтаксической эволюции конструкций, обнаруживать регулярно воспроизводимые модели их развития, в частности, пути конструкциализации свободных сочетаний и де-конструкциализации устойчивых фразем.

### 2.1 Цели и задачи разметки

При разработке разметки основной упор делался на ее мнемоничность и интерпретируемость для потенциального пользователя.

Строки разметки для одной конструкции соответствовали этапам её развития; *этапом* в базе «Диахроникон» называется состояние конструкции в определённый период, являющееся результатом некоторого изменения.

Для каждого этапа размечались:

1. Индекс предыдущего этапа (*former change*), позволяющий устанавливать отношения преемственности между этапами;
2. Новый вид той части конструкции, которая подверглась изменению (*part of construction changed*). Это позволило отмечать, какие изменения затрагивают якорь конструкции, какие – её изменяемую часть, а какие – синтаксис и семантику всей конструкции в целом;
3. Тип (*type of change*) и подтип изменения (*subtype of change*), благодаря которому возник данный этап, а также уровень языка (*level*), к которому относится это изменение (синтаксис или семантика, *synt vs. sem*);
4. Теги для обозначения морфосинтаксических и семантических категорий (*morphosyntags* и *semantags*), характеризующие конструкцию на данном этапе ее развития;
5. Пример первого и последнего вхождения конструкции в НКРЯ; дата первого и последнего вхождений тем не менее условна, НКРЯ постоянно пополняется новыми примерами и работа строится в соответствии с тем, чтобы при обновлении базы употреблений разметка не теряла своей релевантности.
6. Текстовый комментарий, более детально описывающий суть конкретного изменения.

Пример разметки можно увидеть в разделе 2.4.

### 2.2 Типы и подтипы изменений

Для начального этапа, соответствующего первому зафиксированному употреблению конструкции в НКРЯ, в качестве "типа изменений" вносилась помета "*source*", т.е. источник. Он может быть как композиционным, так и идиоматическим (если идиома, ставшая якорем конструкции, впервые появляется в корпусе именно в некомпозиционном значении).

В рамках проекта каждый новый этап развития конструкции характеризовался изменением на синтаксическом или семантическом уровне. Следует отметить, что в некоторых случаях, по

данным НКРЯ, и синтаксическое, и семантическое изменения фиксировались фактически одновременно, в одном и том же примере<sup>6</sup>.

Для каждого семантического и синтаксического изменения выделялись тип и подтип. Список типов является закрытым, а список подтипов – открытым.

Типы синтаксических изменений размечались в зависимости от того, в какой части конструкции произошло изменение. Так, оно могло произойти:

1. В ядре конструкции (*change in anchor*).

Подтипы этого изменения описывают усиление или ослабление синтаксического и семантического единства якоря, а также изменения в составе якоря.

Так, при добавлении компонента (*adding a component*) в якорь конструкции “как не бывало” получает возможность добавляться усилительная частица *и*:

(1) *Потом вдруг как и не бывало литературных обозрений!* [В. Г. Белинский. Русская литература в 1842 году (1843)]<sup>7</sup>

2. Во внутреннем синтаксисе конструкции (*change in inner syntax*).

Под внутренним синтаксисом конструкции подразумеваются те возможные наборы составляющих, которые зависят от наиболее вложенной составляющей якоря конструкции. В большинстве случаев они представляют собой слоты в изменяемой части конструкции. Подтипы этого изменения описывают появление и исчезновение зависимых от якоря или изменения в наборе синтаксических типов существующих зависимых. Так, при подтипе *change in government: clause* слот в конструкции с якорем “без понятия” со значением информации раньше мог заполняться только предложной группой, а теперь заполняется и финитной клаузой:

(2) *И вновь нас зомбируют, только уже в другом направлении, хотя без понятия как это влияет на верующих* [ГИКРЯ]

3. Во внешнем синтаксисе конструкции (*change in outer syntax*).

Под внешним синтаксисом конструкции подразумеваются те возможные наборы составляющих, в которые может быть вложена наименее вложенная часть якоря конструкции.

Подтипы этого изменения описывают изменения в полярности конструкции и в наборе синтаксических ролей, в которых она может выступать.

Так, при изменении поляризованности (*change in polarity*) конструкция “без царя в голове”, которая раньше была положительно поляризованной, стала употребляться и в отрицательных предложениях:

(3) *Нужны люди, а я человек бывалый, опытный и не без царя в голове, чего еще?* [М. Е. Салтыков-Щедрин. Пестрые письма (1884-1886)]

Типы семантических изменений выделялись исходя из более широкого набора факторов. Так, были выделены:

1. Новые идиоматичные употребления конструкций (*new idiomatic use*).

Подтипы этого изменения описывают семантические переходы: метафора, метонимия, ребрендинг, а также расширение и сужение значения.

<sup>6</sup> Строго говоря, даже в этом случае два изменения происходят последовательно, так как одно из них является причиной другого (например, ребрендинг является результатом расширения семантической сочетаемости (Rakhilina, et al. 2010)). Тем не менее, на данный момент в разметке отражены только те факты о развитии конструкций, которые фиксируются в выдаче НКРЯ - интерпретация же этих фактов пока полностью предоставляется пользователю. Поэтому пока что мы размечаем несколько изменений, впервые зафиксированных в одном примере, как произошедшие одновременно. Установление подлинного порядка произошедших в таких случаях изменений остается задачей пока что не решенной, но входящей в перспективы развития проекта.

<sup>7</sup> Этот и все дальнейшие примеры взяты из НКРЯ.

Так, при подтипе *widening* (расширение) конструкция *не дай бог* ранее содержала интенсивность и негативную оценку; а на новом этапе – только интенсивность:

(4) — *Я, кажись, ни за что бы не выдержал, потому я до лошадей, не дай бог, какой охотник! Ежели увижу где хорошего коня, у меня душа мрет, а у самого так руки и чешутся украсть его.* [Ф. Ф. Тютчев. На скалах и долинах Дагестана (1903)]

2. Изменения семантической сочетаемости конструкции (*change in semantic compatibility*). Подтипы этого изменения описывают изменения в наборе семантических типов зависимых конструкции или вершин конструкции. Так, при подтипе *extension: new type of NP dependent* (расширение множества имён-зависимых, с которыми может сочетаться выражение) в конструкции *NP-Gen не напасешься* в качестве аргумента используется слово, не обозначающее множество, хотя ранее это было невозможно:

(5) *На всех проходящих и Маслова не напасешься.* [И. Т. Кокорев. Савушка (1847)]

3. Изменения в прагматике конструкции (*change in pragmatics*). Подтипы этого изменения описывают прагматикализацию и депрагматикализацию дискурсивных формул, а также изменения в иллокутивной цели употребления конструкции. Так, при подтипе *depragmaticalization of a routine* конструкция *ни-ни* ранее употреблялась как дискурсивная формула, а теперь может быть частью предложения:

(6) — *Нет, брат! играть с тобой еще можно, но позволять тебе карты сдавать — ни-ни!* [М. Е. Салтыков-Щедрин. Помпадур и помпадурши (1863-1874)]

### 2.3 Теги

Особо следует сказать об используемой в разметке системе тегов: она была введена с целью найти в базе данных этапы, имеющие отношение к тем или иным морфологическим, синтаксическим и семантическим явлениям. В связи с этим при разметке использовались два набора тегов: морфосинтаксические и семантические.

Инвентарь **морфосинтаксических** тегов включает в себя следующие подгруппы:

1. Частеречные теги (Noun, Adj, Verb и т.д.).
2. Теги для грамматических категорий глаголов (например, Infinitive, Past Tense, Subjunctive и т.д.), имен и местоимений (Animacy, Singular, Dative и т.д.).
3. Теги для синтаксических категорий составляющих (например, Sentential Actant, Argument, Discontinuity – разрывность и т.д.).
4. Теги для поляризованных единиц (NPI, PPI).

Большая часть **семантических** тегов заимствована из Конструктикона (см. описание в [Endersen et al. 2020]); среди них – тег *Maleficiary*, обозначающий участника, которому действие приносит вред. Использование данного тега в диахронической разметке можно увидеть в разделе 2.4.

Однако некоторые семантические теги были разработаны для нужд данного проекта. Среди них выделяются следующие группы:

1. Теги для явлений, связанных с квантификацией (Large Quantity, Small Quantity, Mass Noun).
2. Теги, обозначающие стиль (Obscenity).
3. Теги для режимов дискурса (Dialogue, Narrative).
4. Теги для семантического класса всей конструкции или её элемента (в частности, тег *Negation*, обозначающий отрицание какой-либо пропозиции в семантике конструкции).

## 2.4 Пример разметки: *NP-Dat не улыбаться NP*<sup>8</sup>

change	former change	part of construction changed	level	type of change	subtype of change	semantags	morphosyntags
1		NP-Dat не улыбаться NP		source	compositional source	Negation	
2	1	NP	sem	change in semantic compatibility	extension: new type of NP dependent	Negation, Maleficiary	
3	1	entire construction	sem	new idiomatic use	rebranding	Negation, Maleficiary	Noun
4	2, 3	не Adv улыбаться	synt	change in anchor	deidiomatization	Negation, Maleficiary	Discontinuity
5	4	не улыбаться CI	synt	change in inner syntax	change in government: clause	Negation, Maleficiary	Verb, Infinitive
6	5	NP-Dat улыбаться NP	synt	change in outer syntax	change in polarity	Maleficiary	NPI

Таблица 1: Фрагмент разметки выражения *NP-Dat не улыбаться NP*

Описанную выше разметку можно проследить на примере выражения *не улыбается* (см. Таблицу 1). За свою историю оно претерпело несколько существенных изменений: изначально мы видели, что конструкция была композиционной (этап 1) и имела несколько актантов: того, кто улыбается, и того, кому улыбаются.

(7) *Да ведь и Иван Самойлыч молодой человек, а он уж и хил, и желт, и согнут, да и актриса ему не улыбается...* [М. Е. Салтыков-Щедрин. Невинные рассказы / Запутанное дело (1848-1863)]

Затем сочетаемость расширилась (этап 2): зафиксированы употребления, в котором NP обозначает кого-либо, кто не может улыбаться в буквальном смысле этого слова<sup>9</sup>.

(8) *Им не улыбается мысль, что лучше быть первым в деревне, нежели вторым в Риме; им не приходит в голову даже то совершенно естественное предположение, что, сделавшись участником столичного движения, они не только не внесут никакой новой струи, но сами утонут в департаментском соре.* [М. Е. Салтыков-Щедрин. Письма о провинции (1868-1870)]

Этап 1 стал причиной ребрендинга (этап 3): наблюдаем вхождение, где выражение стало идиомой и субъект начал обозначать что-либо неприятное для объекта:

(9) *Министр финансов хотел бы созвать совещание с участием гг. Гирса и Абазы, и перспектива необходимости вернуться в город до своего ближайшего доклада, т. е. до 4 июня, вовсе не улыбается моему дорогому начальнику.* [В. Н. Ламсдорф. Дневник (1891)]<sup>10</sup>

Далее, в НКРЯ фиксируется вариант, где между частями выражения теперь может ставиться модификатор (этап 4, развивающийся из двух предыдущих этапов - см. таблицу 1).

<sup>8</sup> В разметке использован инфинитив глагола “улыбаться”, чтобы отразить вариативность форм глагола, которые могут занимать данную позицию.

<sup>9</sup> Возможное влияние метафорического употребления “улыбаться” в значении абстрактных существей (например, удача мне (не) улыбается). Тем не менее, изменения, происходящие с конструкцией “NP-Dat не улыбаться NP” (и отмеченные в Таблице 1) не влияют на вариант с метафорическим употреблением.

<sup>10</sup> Данное вхождение не является первым примером употребления конструкции в НКРЯ; первый пример этапа 3 совпадает с первым примером этапа 2.

(10) *Хотя мне эта комбинация не особенно улыбается, но, вероятно, придется прибегнуть к ней.* [В. О. Ковалевский, С. В. Ковалевская. Письмо А. О. Ковалевскому (1874)]

Подобный тип изменений, связанных с тем, что якорь получает возможность изменяться там, где раньше он был неизменяемым, то есть теряет одно или несколько свойств идиомы<sup>11</sup>, мы назвали *деидиоматизацией*.

На следующем этапе у конструкции фиксируется возможность присоединять предикаты (этап 5):

(11) [Телемахов Прокопий Евсеевич (муж, профессор)] *Мне вовсе не улыбается, Валентин Николаевич, быть вестником несчастья и врываться в вашу... идиллию.* [Л. Н. Андреев. Профессор Сторицын. Драма в четырех действиях (1912)]

Последний этап заключается в изменении якоря конструкции (этап 6): фиксируются варианты, где *не* перестает быть постоянной частью якоря и может заменяться на другие лицензоры NPI:

(12) *Это мне не улыбается — да и вряд ли кому вообще улыбается.* [Г. С. Эфрон. Дневники. Т. 1. 1941 (1941)]

Общую схему микродиакронического развития конструкции *не улыбается* можно видеть на Рис. 1.

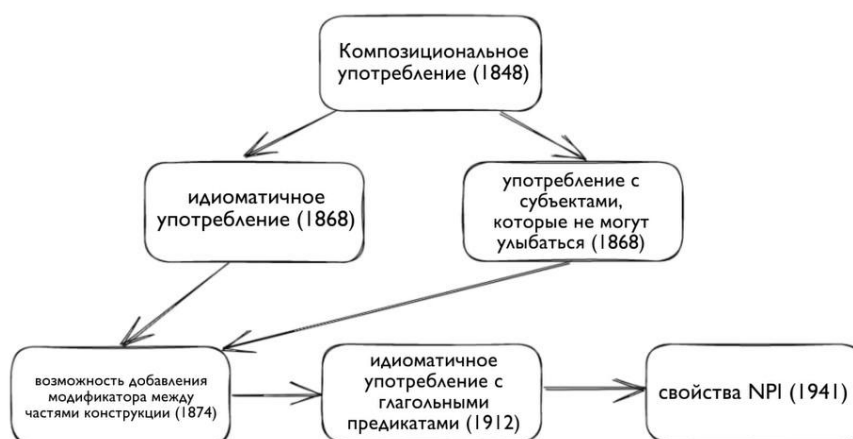


Рис. 1: Этапы микродиакронического развития конструкции *не улыбается* по данным НКРЯ

### 3 База данных с веб-интерфейсом

В рамках проекта предполагается, что база данных «Диакроникон» будет доступна онлайн. На данный момент для неё разрабатываются реляционная архитектура (в тестовой версии она представляет собой базу данных SQLite) и веб-сервер, который предоставит возможность просмотра данных в удобном виде и поиска на основе описанной выше разметки. Сервер создан при помощи фреймворка Flask для языка Python. Ниже будет кратко описана архитектура базы данных, а также реализованные и планируемые способы поиска и представления данных.

Для диахронических изменений в конструкциях в базу данных интегрирована таблица изменений, с разметкой описанной в Разделе 2. Помимо этого, в базе данных представлена разметка современного этапа **конструкции**, где вместо типов изменения представлены параметры «современное значение» (*contemporary meaning*) и «синтаксическая функция якоря конструкции» (*synt\_function\_of\_anchor*). Кроме того, здесь добавлена схема (глоссирование) якорной части

<sup>11</sup> Например, в якоря конструкции порядок слов из фиксированного становится свободным; лексический состав якоря при этом обычно остаётся неизменным.

конструкции, и альтернативные формулы для современной конструкции (*variation*). Предусмотрены указатели для связи с Русским Конструктиконом.

Кроме того, отдельно размечаются синтаксические и семантические **ограничения** на конструкцию в целом и на ее элементы, которые данным этапе имеют вид текстовых комментариев. При этом сохранены метаданные и **общая информация** о конструкции: в частности, в базе данных представлена выборка, на основе которой получена разметка, и указаны разметчики конструкции.

Все эти данные подробно отражены на страницах отдельных конструкций, см. Рис. 2.

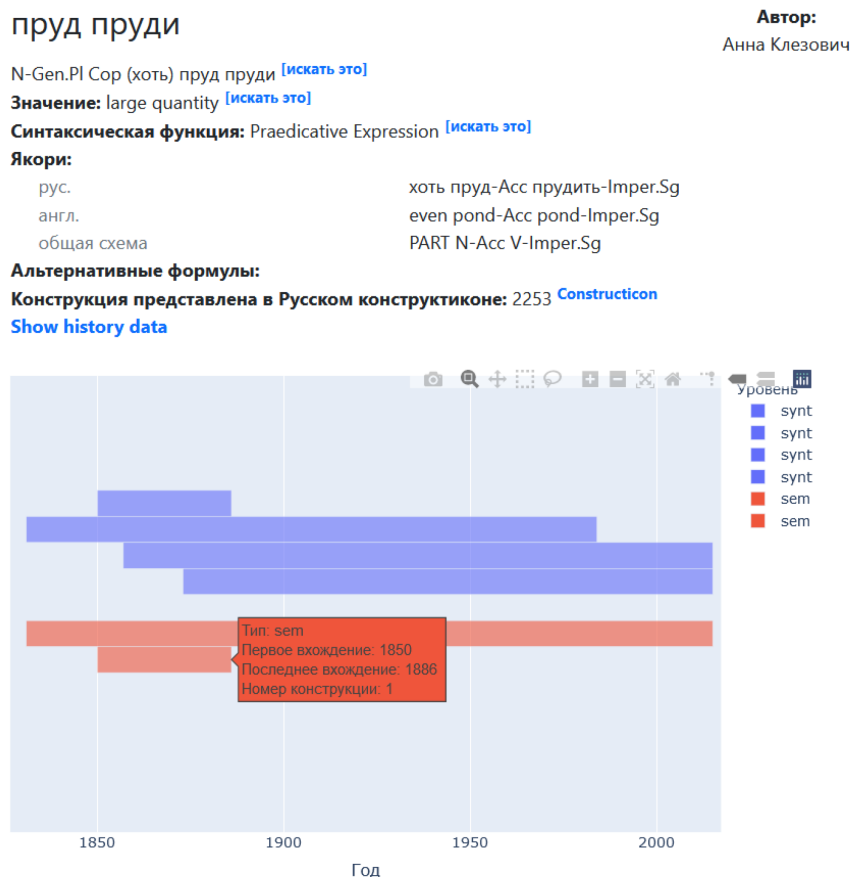


Рис. 2: Страница конструкции *пруд пруди*

На Рис.2 представлена базовая информация о конструкции, включающая большую часть перечисленных выше параметров: значение, якорь и альтернативные формулы. Здесь также доступен линейный график периодов в истории конструкции, а в нижней части страницы изменения отражены в более подробном табличном виде (Рис. 3). Наконец, приведена информация о разметчиках конструкции и выборке.



стадия	уровень	тип изменения	первое	последнее	первое вхождение	последнее вхождение	комментарий
			вхождение (дата)	вхождение (дата)			
V N-Acc (NP-Ins)	synt	source	1850	1886	Видеть она, что враги хотя́ть прудить ровь и стрѣляютъ по стѣнамъ. [А. В. Дружинин. Письма иногороднего подписчика о русской журналистике (1850)]	<...> убирали хлеб и сено (на монастырских запашках), прудили пруды, сады оплетали, на невод ходили, хлебы пекли, пиво варили, лен пряли». [М. Ф. Владимирский-Буданов. Обзор истории русского права (1886)]	Фразеологизм происходит от сочетания форм глагола «прудить» (ставить плотину поперёк водоёма) и существительного «пруд». Материал, из которого ставится запруда, выражался творительным падежом.
physical action	sem	source	1850	1886	Видеть она, что враги хотя́ть прудить ровь и стрѣляютъ по стѣнамъ. [А. В. Дружинин. Письма иногороднего подписчика о русской журналистике (1850)]	<...> убирали хлеб и сено (на монастырских запашках), прудили пруды, сады оплетали, на невод ходили, хлебы пекли, пиво варили, лен пряли». [М. Ф. Владимирский-Буданов.	

Рис. 3: Фрагмент таблицы с изменениями на странице конструкции

Общая информация	Якорь	Изменения конструкции	Ограничения
id	Синт. функция якоря	Формула в этот период	Элемент
Формула	Схема якоря	Этап в истории конструкции	Выберите тип ограничения
Значение	Якорь (рус.)	Синтаксическое	Ограничение
<input type="checkbox"/> Есть в конструктиконе	Якорь (англ.)	Подтипы изменений	
id (Конструктикон)		Длительность это (≤, ≥ или =)	
		Длительность периода	
		Первое вхождение в таком виде	
		Последнее вхождение в таком виде	

Рис. 4: Поддерживаемые поля для поиска

По большинству указанных выше параметров возможен поиск (Рис. 4). Поля «значение» и «синт. функция якоря», «подтип изменения» допускают поиск по одному из предусмотренных в разметке значений. Поле «подтип изменения» позволяет находить конструкции по диахронической разметке, описанной в разделе 2.2. В совокупности с другими полями это учитывает большую часть вариативности конструкций.

Возможен поиск и по формуле. Он учитывает конструкцию, а также ее альтернативные формулы, варианты, выявленные при разметке, и допускает использование символа-джокера \*. Ниже приведен пример запроса (12). При поиске по формуле с именной группой в любом падеже, что выражено символом \*, на выходе выдаются конструкции с отдельными «подрезультатами» (данная особенность связана с тем, что часть конструкций имеют альтернативные формулы, которые тоже содержат именную группу).

(12) [Запрос] Формула: пр\*  
[Фрагмент выдачи]:

Результатов по искомым параметрам: 21 (конструкций — 14)

1. NP Cop не что (иное) как NP:
  - 1 NP Cop не что (иное) как NP
  - 2 NP Cop не кто (иной) как NP

Наконец, раздел “изменения конструкции” позволяет найти все изменения с определёнными характеристиками: формулой заданного вида, типа изменений (синтаксических vs. семантических), начавшихся или закончившихся в определённом году, либо длившихся заданный период времени.

Веб-интерфейс пока поддерживает не всю описанную в предыдущем разделе разметку. Тем не менее он уже сейчас позволяет отвечать на разные вопросы. Так, интересны могут быть «стабильные» этапы, на которых конструкции задерживаются долго без изменений. Найти такие этапы можно удобным запросом в (13), его результат представлен на Рис.5.

(13) [Запрос] Длительность периода это: Минимум ( $\geq$ )  
(&)Длительность периода: 150

Результатов по искомым параметрам: 24 (конструкций — 16)

- 1 **N-Gen.PI Cop (хоть) пруд пруди**  
1 1831 — 1984  
**synt** : adding concessive particle      *NP-Ins хоть N-Acc V-Imper*
- 2 **в точности PronDem**  
2 1716 — 2019  
**synt** : *source*      *VP в точности*
- 3 **Prep N-Dat.Sg не по адресу**  
3 1841 — 2021  
**sem** : metaphor      *negative assessment*
- 4 **NP-Gen не\*густо**  
4 1795 — 2017  
**sem** : *source*      *entire construction*
- 5 1830 — 2018  
**sem** : metaphor      *entire construction*

Рис. 5: Конструкции с этапами изменений длиной более 150 лет, найденные через веб-интерфейс

Так, видно, например, что для конструкций *в точности*, *пруд пруди* долго фиксируется употребление в исходном синтаксически или семантически виде (*source*).

Логичным расширением функционала видится возможность поиска по нескольким изменениям и обобщающих поисков. Так, можно найти конструкции, с которыми произошло меньше 2 изменений или конструкции, с которыми произошло больше 4 семантических изменений за период, определяемый конкретными годами или своей длительностью. Возможность выделить период на основе конкретных годов важна, чтобы выявлять года, когда среди разных конструкций в целом изменений происходит много и года, когда наоборот, конструкции устойчивы. Такие вопросы могут задаваться как лингвистами, так и филологами. Эти функции активно разрабатываются.

#### 4 Заключение

На данный момент в рамках проекта “Диакроникон” разработана детальная и многоуровневая разметка диахронического развития конструкций, позволяющая обобщить и систематизировать изменения их синтаксиса и семантики, а также создан веб-интерфейс для поиска конструкций в базе данных и наглядного отображения изменений и свойств конструкций.

Разработанный ресурс предоставляет обширные возможности для системного изучения не только конкретных конструкций, но и общих механизмов идиоматизации и грамматикализации. База данных позволяет исследователю одновременно сравнивать несколько отдельных сюжетов, проводить поиск по списку конструкций и их характеристик в диахронии, отслеживать историю синтаксических и семантических изменений и ограничений сочетаемости разных конструкций. Ближайшей задачей проекта является пополнение инвентаря конструкций в базе и уточнение разметки имеющихся конструкций, а также дальнейшее совершенствование методов разметки и классификации диахронических изменений. Что касается развития веб-интерфейса, планируется разработка функции поиска по нескольким изменениям, которая позволит исследователям запрашивать в базе конструкции, с которыми произошло определенное количество изменений за конкретный период времени. Такое расширение функционала сделает ресурс особенно ценным инструментом для системного изучения языковых изменений.

Опыт работы над Диахрониконом может быть полезен также для создания аналогичных ресурсов для других языков и изучения механизмов конструкционализации в типологической перспективе.

## Acknowledgements

Статья подготовлена по результатам проекта «База данных по историческим изменениям русских конструкций «Diachronicon»» при поддержке Фонда академического развития ФГН НИУ «Высшая школа экономики» в 2022- 2024 году.

## References

- [1] Croft W. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press, 2001.
- [2] Fillmore C. J. *Border Conflicts: FrameNet Meets Construction Grammar*. // *Proceedings of the XIII EURALEX International Congress 2008*. — Barcelona, Spain, 2018. — P. 49–68.
- [3] Fillmore Ch. J., Kay P., O'Connor M. C. *Regularity and idiomatization in grammatical constructions: The case of let alone* // *Language* 1988. — Vol. 64(3). — P. 501–538.
- [4] Janda L. A., Lyashevskaya O., Nessel T., Rakhilina E., Tyers F. M. *A construction for Russian: Filling in the gaps*. // *Constructicography: Constructicon development across languages*. Amsterdam: John Benjamins, 2018. — P. 165–181.
- [5] Klezovich A.G., Golosov F.V. (2018a), *The Database “Phrasal Diachronicon”: constructional change in Russian expressions with quantitative semantics* // *Proceedings of Constructional semantics: Cognitive, functional and typological approaches 2018*. — Helsinki, Finland, 2018. — Poster session.
- [6] Баранов А. Н., Добровольский Д. О. *Основы фразеологии (краткий курс)*. — Москва: Флинта Наука, 2016.
- [7] Добрушина Е. Р., Ахапкина Я. Э. *Пограничный русский язык: как рождаются экспрессивные кванторные выражения*. — СПб.: Алетейя, 2019.
- [8] Клезович А. Г., Голосов Ф. В., Фесенко В. П. *База исторической фразеологии (Phrasal Diachronicon): отражение изменений в функционировании лексических единиц с семантикой количества*. // *Симпозиум по Русскому Конструктивному в филиале НИУ ВШЭ в Вороново, 17-23 марта 2023*. — М.: 2018.
- [9] Падучева Е. В. *Из наблюдений над языком Л. Толстого (к вопросу о малых диахронических сдвигах)*. // *Вопросы языкознания*. — 2018 — Вып. 5. — С. 49–63.
- [10] Рахилина Е. В. (ред.) *Лингвистика конструкций*. — М.: Азбуковник, 2010.
- [11] Рахилина Е. В., Резникова Т. И., Карпова О. С. *Семантические переходы в атрибутивных конструкциях: метафора, метонимия и ребрендинг*. — М.: Азбуковник, 2010. — С. 396–455.
- [12] Сай С. С. *Динамика развития обстоятельств времени со значением предшествования на интервал в русском языке*. // *Избыточность в грамматическом строе языка*. [Acta Linguistica Petropolitana. Труды Института лингвистических исследований РАН. — 2010. — Вып. 6(2)]. / Воейкова М. Д. (ред.). — СПб.: Наука, 2010. — С. 131–183.
- [13] Сай С.С. *Лексические механизмы грамматического дрейфа: конструкции с дативным субъектом в русском языке XVIII – XXI веков* // *Acta Linguistica Petropolitana. Труды Института лингвистических исследований РАН*. / Казанский Н. Н. (отв. ред.). СПб.: Наука, 2014. — Т. 10(3). — С. 568–610.
- [14] Эндерсен А., Жукова В., Мордашова Д., Рахилина Е., Ляшевская О. *Русский конструктивон: новый лингвистический ресурс, его устройство и специфика*. // *Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог-2020»*. — М.: РГГУ, 2020. — Т. 19. — С. 226–241.