# Семантическая структура русских предложно-падежных конструкций

И.В. Азарова, В.П. Захаров, А.Д. Москвина

Санкт-Петербургский государственный университет

i.azarovaspbu.ru, v.zakharov@spbu.ru, moskvina.anya@gmail.com

## Аннотация

В публикации рассматривается структура синонимичных и квазисинонимичных семантических отношений для первообразных и производных предлогов русского языка в рамках определенных смысловых рубрик: передачи транзитивных локативных, темпоральных и медиативных значений в структуре предложений и словосочетаний. Приведены статистические характеристики корпусной реализации предложных конструкций, описаны регулярные типы категорий главных и зависимых слов, а также типы семантико-синтаксических конструкций реализации вариантов значения смысловой рубрики.

**Ключевые слова**: корпусная статистика, семантические рубрики, русские предложные конструкции, предложные значения, транзитив, темпоратив, медиатив

## 1. Введение

Предлоги в русском языке довольно долго оставались без пристального внимания со стороны специалистов по автоматическому анализу текста: например, они входили в списки «стоп-слов», которые предотвращали их использование в векторных моделях информационного поиска. Если искать ответ на вопрос, почему они по-прежнему игнорируются при информационном поиске, то аргументом в пользу такого подхода будет их низкий показатель «специфичности» (tf-idf): они обладают высокой частотностью практически во всех документах, в которых производится поиск. Однако для семантически ориентированного анализа текста они безусловно важны, поскольку передают четкие семантико-синтаксические отношения между знаменательными словами (*устройство с автономным питанием* vs *устройство без автономного питания*), между характеристиками предиката (*говорить с акцентом* vs *говорить без акцента; переводить с английского* vs *переводить на английский*), между пространственно-временными спецификациями пропозиций (*с 1 января* vs *до 1 января*; *на верхней площадке* vs *под верхней площадкой*). И приведенные выше примеры, и устойчивая корпусная встречаемость предложных конструкций являются четким показателем их важности для построения текстов на русском языке. Вторым важным фактором недостаточного использования предлогов является неясность и неструктурированность их значений, например, для предлога «в» в Викисловаре приводится 21 значение, при этом синонимия предлогов указана частично, неясно, в чем различия близких, но несинонимичных предлогов.

В рамках проекта по квантитативному описанию грамматики предлогов мы начинаем систематическое описание предлогов со значениями из определенных семантических рубрик. В данной статье это конструкции транзитива, темпоратива и медиатива.

Семантические рубрики (термин, близкий к понятию «синтаксема»), представленные в статье, помогают обобщить называемые часто по-разному предложные значения.

## 2. Инвентарь предложных конструкций

Предлоги в русских текстах неоднородны и разнообразны: есть небольшая группа первообразных предлогов и обширная, с неясными границами, группа производных предлогов, которые мотивированы знаменательными частями речи (существительными, наречиями и глагольными формами деепричастий), возможно, в сочетании с первообразными предлогами. Отсюда корпусная частотность первообразных предлогов обычно завышена за счет того, что часть их используется в структуре производных предлогов. При исчислении последних используются зачастую противоречивые параметры. Во-первых, важно отвлеченное значение существительных (у наречий и деепричастий оно безусловно такое), во-вторых, обязательным признаком производного предлога является непроницаемость линейной последовательности такого грамматикализованного знаменательного слова и первообразного предлога, например *в течение, во исполнение* и др. Косвенным и непоследовательным показателем производного предлога является отличное от формы существительного орфографическое написание: *в течение года* vs *в течении реки; несмотря на непогоду* vs *не смотря по сторонам; во исполнение приказа* vs *в исполнении детского хора.*

Самым важным показателем «предложного» характера сочетаний со знаменательными компонентами будет их полная или частичная функциональная синонимичность первообразным предлогам или «подтвержденным» производным предлогам (предлоги с модифицированной орфографией или просто высокочастотные производные предлоги).

Первообразные предлоги определяются в грамматических источниках весьма нечетко: «Предлог — это служебная часть речи, оформляющая подчинение одного знаменательного слова другому в словосочетании или в предложении и тем самым выражающая отношение друг к другу тех предметов и действий, состояний, признаков, которые этими словами называются» [1, с. 706]. В этом определении опущен важный элемент «синтаксического поведения» предлогов — то, что они оформляют предложно-падежную форму, т. е. могут «оформлять подчинение» только падежных знаменательных слов: в первую очередь, имен существительных, во вторую очередь, местоимений-существительных, числительных и, возможно, прилагательных при их окказиональной субстантивации. Скорректированное определение: «Предлог — служебная часть речи, обозначающая отношение между объектом и субъектом, выражающая синтаксическую зависимость имен существительных, местоимений, числительных от других слов в словосочетаниях и предложениях», представленное в Википедии, может нас еще более запутать указанием на «отношения между объектом и субъектом». Разве только объектные и субъектные позиции могут быть оформлены предложно-падежными конструкциями? Безусловно, обстоятельственные значения превалируют в текстах (и в словосочетаниях, и в предложениях), но ими не исчерпывается весь спектр предложных значений.

В нашем проекте будем считать предлогами «стереотипные способы уточнения падежных значений имен (чаще всего, имен существительных) при выражении валентных позиций знаменательных слов (чаще всего, глаголов и отглагольных дериватов) и/или различных обстоятельственных квалификаторов в предложении». «Стереотипность» значения будет доказываться высокой частотностью выражения определенного значения в корпусе, кроме того синонимия и частичная синонимия между первообразными и производными предлогами будет описана в терминах семантических классов зависимых существительных (оформленных грамматически в виде предложных конструкций) и главных слов, предсказывающих появление предложно-падежных структур.

Первоначально синонимия между первообразными и производными предлогами анализировалась нами по разным словарям, в результате чего были составлены базовые таблицы синонимии (в ряде случаев, естественно, частичной) для предлогов, являющихся объектами анализа в данной статье. При этом еще раз напомним, что значения предлогов во многом являются контекстно-зависимыми.

## 3. Семантическая структура предложных конструкций в рамках определенной семантической рубрики

Поскольку задачей нашего проекта является создание квантитативной грамматики предложных конструкций, мы используем понятие «семантической рубрики» в качестве обобщенного названия группы значений предлогов. Часто такими рубриками являются названия семантических актантов (или падежей, ролей), однако ввиду разнообразной трактовки этих понятий в литературе и отсутствия четких границ между элементами списка, мы начинаем описание «снизу», от некоторого наиболее частотного базового предлога, вариативность значений которого отражается в группе сопряженных с ним предлогов (в основном, производных), образуя ряды синонимичных или квазисинонимичных конструкций. Поэтому названия семантических рубрик отчасти носят «условный» характер.

**Таблица 1.** Частотные характеристики предлога «через» и некоторых синонимов

| **предлог** | **значение** | **НКРЯ со снятой омонимией, 6 млн. слов (200 случайных контекстов)** | **НКРЯ газетный, 228,5 млн. слов (200 случайных контекстов)** | **Araneum Russicum веб-корпус 120 млн. (200 случайных контекстов)** |
|---|---|---|---|---|
| **Через** | **медиатив** | **ipm 173,53 20,5% (41 употр.) от всех употребл.** | **ipm 221,30 34,5% (69 употр.) от всех употребл.** | **ipm 185,21 32,5% (65 употр.) от всех употребл.** |
| С помощью | медиатив | ipm 76,63 460 употр. | ipm 113,21 25870 употр. | ipm 228,64 27437 употр. |
| При помощи | медиатив | ipm 22,32 134 употр. | ipm 27,66 6322 употр. | ipm 82,60 9916 употр. |
| Посредством | медиатив | ipm 16,49 111употр. | ipm 12,31 2813 употр. | ipm 42,90 5154 употр. |
| Через посредство | медиатив | ipm 0,99 6 употр. | ipm 0,09 22 употр. | ipm 0,42 50 употр. |
| **Через** | **транзитив** | **ipm 245,39 29,5% (57 употр.) от всех употребл.** | **ipm 118,65 18,5% (37 употр.) от всех употребл.** | **ipm 157,71 27,5% (55 употр.) от всех употребл.** |
| сквозь | транзитив | ipm 119,83 719 употр. | ipm 19,25 4398 употр. | . ipm 23,90 2869 употр |
| Поперек | транзитив | ipm 12,00 72 употр. | ipm3,15 720 употр. | ipm3,40 408 употр. |
| **Через** | **темпоратив** | **ipm 390,96 47% (94 употр.) от всех употребл.** | **ipm 319,95 50% (100 употр.) от всех употребл.** | **ipm 222,25 27,5% (55 употр.) от всех употребл.** |
| Спустя | темпоратив | ipm 54,67 328 употр. | ipm 82,28 18800 употр. | ipm 52,50 6300 употр. |
| По истечении | темпоратив | ipm 2,17 13 употр. | ipm 5,81 1328 употр. | ipm 6,88 826употр. |
| Чрез | темпоратив | ipm 14,00 84 употр. | ipm 0,16 37 употр. | ipm 3,29 395 употр. |

Для выявления семантической структуры необходимо выяснить базовое ядерное значение, которое представлено частотным употреблением первообразного падежа, в отдельных случаях производного предлога или характерной падежной формы. Подбор синонимических конструкций производится частично за счет информации, представленной в толковых словарях. Синонимия дополнительно проверяются через выявление характерных конструкций реализации значения в случайных выборочных

совокупностях корпусных контекстов. В случае несовпадения сочетаемости предлогов и семантических классов имен или появления иных ограничений на реализацию значений семантической рубрики они рассматриваются как частичные синонимы, различия в употреблении и значении будут оценены в соответствии с корпусной статистикой (см. табл. 1 и табл. 2).

**Таблица 2.** Частотные характеристики предлога «по» и некоторых синонимов

| **предлог** | **значение** | **НКРЯ со снятой омонимией, 6 млн. слов (200 случайных контекстов)** | **НКРЯ газетный, 228,5 млн. слов (200 случайных контекстов)** | **Araneum Russicum веб-корпус 120 млн. (200 случайных контекстов)** |
|---|---|---|---|---|
| **По** | **медиатив** | **ipm 332,95 7% от всех употребл.** | **ipm 82,57 1% от всех употребл.** | **ipm 237,03 3% от всех употребл.** |
| Посредством | медиатив | ipm 16,49 111 употр. | ipm 12,31 2813 употр. | ipm 42,90 5154 употр. |
| Через посредство | медиатив | ipm 0,99 6употр. | ipm 0,09 22употр. | ipm 0,42 50 употр. |
| **По** | **транзитив** | **ipm 245,39 16% от всех употребл.** | **ipm 908,35 11% от всех употребл.** | **ipm 869,11 11% от всех употребл.** |
| **По** | **темпоратив** | **ipm 390,96 4% от всех употребл.** | **ipm 82,57 1% от всех употребл.** | **ipm474,06 6% от всех употребл.** |
| после | темпоратив | ipm 810,83 4865 употр. | ipm 1446,41 330504 употр. | ipm 1245,90 149504 употр. |
| вслед за | темпоратив | ipm 39,00 234употр. | ipm 34,36 7851употр. | ipm 14,53 1744употр. |
| по истечении | темпоратив | ipm 2,17 13 употр. | ipm 5,81 1328 употр. | ipm 6,88 826 употр. |
| по завершении | темпоратив | ipm 0,67 4 употр. | ipm 3,68 842 употр. | ipm 2,96 355 употр. |
| по окончании | темпоратив | ipm 9,33 56 употр. | ipm18,86 4310 употр. | ipm 16,80 2916 употр. |

### 3.1. Структура предложных конструкций медиатива

Медиатив как семантическая рубрика имеет узкую и широкую интерпретацию. В большинстве случаев она рассматривается как определенная семантическая роль в предикативной структуре глагола. В узком смысле медиатив понимается как средство, т. е. вещество или предмет, которые расходуются при выполнении действия или процесса [2]. В более широком понимании медиатив является инструментом и включает их вещественные и отвлеченные реализации [3]. В русском языке инструмент и медиатив, как правило, выражаются формой творительного падежа (*красить стены валиком, рисовать картину красками*), однако в виде предложной конструкции мы видим более сложные варианты их синкретизма.

Предлог «через» используется в рубрике медиатива с высокой корпусной частотой 163 ipm в корпусе «Бокренок» кафедры математической лингвистики СПбГУ, при этом нюансы его значений весьма разнообразны: *гладить брюки через влажную ткань* (инструмент *утюг*); *настроить протокол TCP/IP через вкладку «Конфигурация»; ультразвук воздействует на организм через воздух*; *ставить горчичники через газету*; *отмывать деньги через другие фирмы; протереть творог через дуршлаг; получить кредит через знакомых* и др. Во всех таких случаях четкая обстоятельственная

характеристика затруднительна, а зачастую — амбивалентна. Например медиативная конструкция *пропустить мясо через мясорубку* одновременно указывает также на реальное перемещение объекта (см. транзитив ниже). Первое значение предлога «через» в Викисловаре трактуется как конструкция с транзитивным значением при помощи синонимов «сквозь, поперек», приводится пример *перевести женщину через дорогу,* но ни один из указанных предлогов мы не можем подставить в этот пример, а в конструкции с амбивалентным значением «через», наоборот, можем заменить его на «сквозь»: *протереть творог через/сквозь сито/дуршлаг*, аналогично в конструкциях наблюдения *видеть через/сквозь стекло.*

Предлог «сквозь» рассматривается как производный (отнаречный), хотя употреблений «сквозь» в качестве наречия в корпусе просто нет. Частотность предлога «сквозь» в медиативно-транзитивном употреблении существенно ниже (20 ipm), чем «через». В большинстве случаев «сквозь» вряд ли может быть заменен на «через»: *видеть сквозь дымку / туман / крону деревьев*. В указанных контекстах объекты, вводимые предлогом «сквозь», обозначают не столько то, что способствует осуществлению действия, сколько то, что не смогло помешать ему. Обозначим такое употребление как медиативно-неактивное. Это и другие наблюдения говорят о том, что, возможно, номенклатура предложных значений требует расширения. Менее частотным (7 ipm) является употребление «сквозь» для (обычно) слухового восприятия на фоне каких-то помех: *слышать смех /крик сквозь шум / телефонный звонок / сон.*

Частотные производные предлоги в данной рубрике мотивированы существительным «помощь» в комбинации с первообразными предлогами: *с помощью* (98 ipm), *при помощи* (27 ipm). В качестве наиболее частотного семантического класса выступают обозначения абстрактных понятий (*игры, техники, приема, публикации* и т.д.) и реальных предметов-инструментов или устройств (*зеркала, микроскопа, проволоки* и т.д.), обозначения людей-помощников медиативно-активных (*друзей, секундантов* и т.д.).

Аналогичный низкочастотный предлог на базе лексемы «помощь» — *без помощи*, своеобразный антоним для предлогов *с помощью* и *при помощи.*

Производный предлог «путем» обладает высокой частотностью (51 ipm), присоединяет абстрактные понятия: *путем угроз / обмана / обещаний* и т.д., тяготеет стилистически к официально-деловым текстам. Вообще говоря, особо следует изучить поведение одних и тех же предлогов (предложных значений) в текстах разных функциональных стилей.

Еще одна знаменательная лексема «посредство» мотивирует целый ряд производных предлогов, среди которых есть один частотный *посредством* (19 ipm) и целая группа низкочастотных (*при посредстве, через посредство, благодаря посредству*). Как и в случае с наречием «сквозь», само существительное «посредство» в корпусах или не представлено, или имеет крайне низкую частоту, не сопоставимую с частотностью предлога. Предлог «посредством» присоединяет в равной степени как абстрактные понятия (*посредством воображения / скобочной записи* и т.д.), так и обозначения предметов-инструментов и веществ (*посредством специальных приборов / электромузыкальных инструментов / анилиновой краски / нажатия правой кнопки мыши* и т. д.)

Частотный производный предлог «благодаря» (71 ipm) обладает сходной сочетаемостью с приведенными выше предлогами (*благодаря деятельности / сыну / судье / компьютеру*), его традиционно относят к другой рубрике — каузативу. Амбивалентность инструментатива (и, следовательно, медиатива) неоднократно подчеркивалась [3], однако в нашем проекте мы будем рассматривать каузативные предлоги в виде самостоятельной рубрики. При этом предлог «благодаря» тяготеет к данной рубрике, что подтверждается производным образованием « *благодаря посредству*».

Отдельная задача, которая может решаться «попутно» — выявление нерегулярных (фразеологизированных, метафоризированных) употреблений предлогов. В первую очередь так используются первообразные предлоги и некоторые производные. В

частности для центра нашей рубрики «через» и «сквозь» имеется довольно большое число употреблений «со сдвигом» значений, которые частично упоминаются в толковых словарях:

*бить* ***через край*** — «слишком сильно»;

*обратиться* ***через голову X*** — «минуя X в социальной иерархии»;

*переступить* ***через себя*** **—** *«*преодолеть в себе какое-то чувство»;

*пропустить* ***через себя*** **—** *«*осмыслить»;

*сматывать* ***через локоть*** **—** *«*на локоть»;

*улыбнуться* ***через силу*** — *«*превозмогая себя»;

*готов/ хочет /...* ***провалиться сквозь землю*** **—** *«*скрыться, испытывая чувство стыда»;

***смотреть сквозь пальцы на X*** **—** *«*делать вид, что не замечаешь X»;

*цедить/ говорить* ***сквозь зубы*** **—** *«*говорить неотчетливо / нехотя, выражая недоброжелательное отношение к собеседнику».

### 3.2. Структура предложных конструкций транзитива

Транзитив — один из вариантов локализации пропозиции. В отличие от характеристики местоположения, которая применима практически к разнообразным действиям, состояниям и процессам, эта спецификация чаще всего ассоциируется с «рамочной» конструкций в паре с префиксом *пере-* для глаголов движения и их дериватов: *перейти через дорогу, перевозки нефти через Атлантику* и т. д. Аналогично предыдущей рубрике (медиативу) значение транзитива может быть выражено падежной формой, ср.: *перейти дорогу / площадь / реку через дорогу*, причем, когда возможны такие варианты выражения падежный вариант более частотный. Таким образом, для нас важно обозначить случаи, когда значение транзитива почему-то выражается только предложно-падежной формой. В Викисловаре приводятся 2 значения, которые можно связать с семантической рубрикой «транзитива»: (1) «сквозь, поперёк» (*Он помог слепой женщине перейти через дорогу на другую сторону*); (2) «поверх чего-либо» (*Я легко перепрыгнул через забор*). В примере для первого значения невозможно заменить «через» на приводимые синонимы, что указывает, что данная формулировка неточна. В МАС [4] первое значение «через» сформулированы точно: «1. Употребляется при обозначении какого-л. пространства, места и т.п., поперек которого располагается что-л., с одной стороны которого на другую совершается движение, действие» (*Перейти через улицу. Переправиться через реку. Мост через Неву*). То есть это расположение, движение или действие, в которое вовлекается объект, имеющий сторону меньшего размера. Лексический эквивалент значения при глаголах движения — «пересекая» для приглагольных и «пересекает» для присубстантивных употреблений: *воздух проходит через нос; пройти через вращающиеся двери*; для глаголов перемещения объекта эквивалент — «перемещая»: Предлог «сквозь» синонимичен отчасти: идея поперечника не существенна, важен пространственный характер того, что пересекается: *пройти / пролететь / прохождение) сквозь (атмосферу / толпу / пространство)*. В Википедии в качестве синонима для «сквозь» приводится «насквозь», однако это слово регулярно употребляется при тех же глаголах движения, но обозначает характеристику действия, что подчеркивается совместной встречаемостью с предлогом «через»: *винт проходит насквозь через гриф, мы прошли насквозь через весь холл.*

Употребление «сквозь» и «через» при глаголах, обозначающих перемещение с преодолением препятствия, имеет лексикализацию «проникая»: *росток пробился сквозь асфальт.*

### 3.3. Структура предложных конструкций темпоратива

Темпоратив — такая же сложная семантическая рубрика, как и локатив, поэтому все варианты ее реализации сразу представить довольно сложно. Предлог «через» обозначает

«по истечении некоторого отрезка времени». Это самое частотное употребление предлога (285 ipm в корпусе «Бокренок» кафедры математической лингвистики СПбГУ), причем оно мало зависит от выражаемой пропозиции, регулярно стоит в позиции детерминанта в абсолютном начале предложения и присоединяет существительные, обозначающие временные интервалы, в сочетании с числительными: *через день / 2 дня / 10 минут / 20 столетий*. Эквивалентный производный предлог — «спустя». Он употребляется с теми же существительными, но помимо предложной позиции (до группы существительного) может стоять после группы как послелог: *спустя несколько лет / несколько лет спустя; спустя три недели / три недели спустя* и т. д.

Производный крайне частотный (825 ipm) предлог «после» толкуется примерно таким же образом «по окончании чего-либо», но в подавляющем числе случаев указывает следование относительно некоторого события (*после обеда / прогулки / инъекции* и т. д.) и даже если присоединяют существительное со значением интервала времени, то, как правило, оно используется в событийном значении: *после своего рабочего дня / десятичасового замачивания*.

Производный предлог «по истечении» (3 ipm) сочетается с существительными по модели «через» (*по истечении года / месяца / двух лет*), «по окончании» (9 ipm) следует модели «после» и лишь небольшое число сочетаний аналогичны «через»: *по окончании месяца / года.* Приводимый в Викисловаре вариант перефразирования «вслед за» используется в основном для указания на несамостоятельное движение или аналогично «после» указывает на завершение события (2 ipm): *вслед за заключением мира*.

## Заключение

Семантические рубрики, представленные в статье, помогают организовать довольно расплывчатые предложные значения. Их близость и различие могут быть объяснены путем анализа семантических классов управляющих и подчиненных слов («хозяева» и «слуги»). Но это уже тема отдельной статьи. Вся структура предложных частот и распределения семантических (лексических) классов призваны стать базой для создания квантитативной грамматики предложных конструкций русского языка.

Следующие шаги в реализации проекта — анализ всего множества предлогов с разбиением их по семантическим рубрикам и составление реестра «хозяев» и «слуг» в терминах лексико-семантических классов и грамматических категорий.

Отдельная задача — описание синонимии первичных и производных предлогов как некоторого градуированного пространства в терминах семантических контекстов.

**Благодарности.** Работа поддержана грантом РФФИ № 17-29-09159 «Квантитативная грамматика русских предложных конструкций».

## Литература

[1] Грамматика русского языка. М., 1980.
[2] Всеволодова М.В., Потапова Г.В. Способы выражения временных отношений в современном русском языке. М., 1975.
[3] Мустайоки, А. Теория функционального синтаксиса: от семантических структур к языковым средствам. М., 2006.
[4] Словарь русского языка в 4 томах. Т. 4 / Под ред. А.П. Евгеньевой и Г.А. Разумниковой. Изд. 3-е, стереотип. М., 1988.

## Semantic Structure of Russian Prepositional Constructions

I.V. Azarova, V.P. Zakharov, A.D. Moskvina

Saint Petersburg State University

In this paper we present a technique for specification of Russian prepositional constructions. We bring together a number of primary and secondary prepositions expressing roughly similar relations between the main and the subordinate full words. The whole group of prepositional constructions demonstrates several implicit variants of the broad sense rubric, which we nominate the semantic domain. In the paper we describe 3 such domains: temporal, transitive and mediative. For each domain we range appropriate prepositions according corpus statistics, show contextual restrictions and subtle nuances of the dominant rubric meaning, which are expressed in collocations with particular semantic types of lexemes occupying head and daughter positions of prepositional relations. The most ambivalent domain described is mediative. Usually this meaning is connected with a notion of an instrument, however, it is complicated by notions of material, device, and even an assistant and space transition. The temporal domain is described partially as a reference to the time moment, the other temporal relations will be the matter of the future research. The transitive domain is a part of the heterogeneous locative relations, which are to be the developed later.

**Keywords:** corpus lexical statistics, semantic domain, Russian prepositional constructions, prepositional meanings, transitive prepositional constructions, temporal constructions, mediative constructions