

Морфологический словарь русского языка (проект)

Russian morphological dictionary (project). This dictionary presents a new approach to Russian morphology, the material being pervadingly ordered in a non-trivial way. In addition, the word-forms are organized in a hierarchical structure. This makes it possible to define a large number of sets and subsets of word-forms, identified mainly in terms of grammatical characteristics. Each set is determined by its connections to other, associated sets.

Предлагаемая статья в сжатом виде описывает работу, скорее проект работы, которая не претендует на научную оригинальность. Новых решений фонологических или морфологических проблем в данном словаре не выдвигается. Тем не менее нельзя признать содержание планируемой книги полностью традиционным, повтором общеизвестных фактов. Новизна заключается в том, что материал представлен, прежде всего упорядочен, по принципам, резко отличающимся от ранее применяемых. Перечислю основные различия:

1. В основу упорядочения словоформ и окончаний положено не графическое, а (некое) фонологическое представление.
2. Используется не традиционный, условный алфавитный порядок сортировки, а специальный, научно обоснованный порядок символов фонем.
3. Сортировка словоформ проводится не с начала слова, как в традиционных словарях, и не с конца слова, как в обратных словарях, а начиная с того места внутри слова, который представляет наибольший интерес с морфологической точки зрения: со стыка между основой и окончанием.
4. Словоформы каждой лексемы образуют иерархическую структуру: одна форма выводится из другой, в первую очередь, на основе традиционного, но более или менее научно обоснованного упорядочения грамматических значений.

Остановимся более подробно на каждом из этих моментов.

1. Эксплицитной фонологической транскрипции подвергаются только основные единицы данного словаря, а именно окончания (используются латинские буквы). Применяются относительно общепризнанные принципы московской фонологической школы¹. Так, гласные

¹ См., например, А. А. Реформацкий. Из истории отечественной фонологии. М., 1970; М. В. Панов. Современный русский язык: Фонетика. М., 1979; Л. Л. Касаткин. О некоторых

передаются десятью символами с учетом того, что каждая из пяти фонем может быть под ударением или безударной. Окончание словоформ *глупый*, *летний* представлено как *-oj*, окончание словоформы *дядей* — как *-oj* или *-ej*, в зависимости от актуального грамматического значения: творительный падеж единственного числа или родительный множественного. Наличие фонологически неразрешимых позиций, конечно, не допускается; для каждой такой позиции должно быть найдено решение, в крайнем случае, совпадающее с орфографическим представлением. Беглый гласный не передается как таковой, а только через свою фонетическую реализацию. Это значит, например, что отрезок *-ок* в *знаток* и *кусок* передается одинаково.

2. Сортировка символов фонем опирается на традиционные фонетические таблицы, отдельно для гласных и согласных, причем гласные предшествуют согласным. Разумеется, однозначного линейного пути через двухмерное поле, какое представляет собой таблица, нет, поэтому принято условное (и традиционное) решение, заключающееся в том, что упорядочение идет сначала слева направо, а потом сверху вниз. Однако допускаются и отклонения от этого принципа. Так, в случае гласных первое место занимает самый отдаленный от согласных гласный *a*. Поэтому здесь чисто случайно осуществляется совпадение с алфавитным порядком (*a, e, и, o, y*). Что касается согласных, то и здесь допускается отклонения от упомянутого пути: например, парные по твердости/мягкости предшествуют непарным.

Символы упорядочены следующим образом:

Гласные: (нуль), *á, a, é, e, í, i, ó, o, ú, u*.

Согласные: (нуль), *b, b', p, p', v, v', f, f', d, d', t, t', z, z', s, s', m, m', n, n', l, l', r, r', k, g, x, c, ž, š, č, j*.

Как видно, последний критерий для гласных — ударение; для согласных предпоследний критерий состоит в том, что звонкие предшествуют глухим, а последний — в том, что твердые предшествуют мягким.

3. Как было сказано выше, главной единицей словаря является окончание. Поэтому сортировка, исходящая из флексийного стыка, идет сначала направо, в сторону окончания, а потом налево, в сторону основы. Сортировка окончаний опирается исключительно на их фонемный облик, не на значение. Таким образом, все употребления одного окончания, например, окончание *-a*, собраны и описаны в одном месте. Сортировка же основ опирается в первую очередь на грамматическое значение словоформы, а только в качестве второго критерия применяется фонемный состав. Это значит, что сначала находим, например, словоформы существительных мужского рода единственного числа именительного падежа на *-a*, например, *nána*, затем — с изменением падежной характеристики — классы

словоформ типа *ка́рна* (винительный и родительный падежи), *хле́ба* (только родительный падеж) и т.д.

Грамматические значения упорядочены в синтагматическом и парадигматическом плане. На первом месте находим части речи: Сущ., Прил., Мест., Числ., Глаг. (в указанном порядке). На втором месте следует субкатегоризация для прилагательных (пол.степ., ср. степ., пр.степ.) и глаголов (личн., пов., инф., прич., дееспр.). После этого отмечаются чисто флексийные категории рода (м.р., ж.р., с.р.), числа (ед.ч., мн.ч.), падежа (И, В, Р, Д, Т, П, З), времени (наст.вр., прош.вр.) и лица (1 л., 2 л., 3 л.). Таким образом, грамматическое значение формы *на́на* представлено как «Сущ. м.р. ед.ч. И» и словоформы *сидя́т* как «Глаг. личн. наст.вр. 3 л.».

Звательный падеж выступает только в двух словоформах, *Бо́же* и *Го́споди*. Формы типа *мам*, *Надь* отношу, вслед за Реформацким², к явлениям фонетической реализации (ср. звонкое произношение конечного звука в формах *Надь*, *Серёж*).

Принимается условие, что словоформа *гриб* представляет собой одну, а не две единицы, несмотря на то, что в ней дизъюнктивно представлены две граммемы одной категории (И и В). Другими словами, эта словоформа занимает две позиции в парадигме. Это весьма важное решение означает, например, что парадигма слова *вещь* содержит не двенадцать, а семь единиц. Дизъюнкция грамматических значений обозначается косой чертой; например, словоформа *вёщи* представлена как «Сущ. ж.р. ед.ч. Р/Д/П / мн.ч. И/В».

Окончания делятся на собственно окончания и т.наз. окончания-суффиксы. К последним (они перечисляются отдельно) относятся окончания «островов неизменяемости» в общей системе словоизменения, а именно неизменяемая форма сравнительной степени прилагательных (*новее*) и форма дееричастия глаголов (*крича́*).

4. Упорядочение словоформ в пределах одной лексемы опирается, в первую очередь, на систему грамматических значений, т.е., на парадигму, причем здесь, в отличие от фонетических таблиц, путь идет сначала сверху вниз, а потом слева направо (по крайней мере это верно для именных парадигм). Например, для существительного *гриб* прослеживается такой путь:

гриб - гриба́; гриба́ - грибу́; грибу́ - грибо́м; грибу́ - грибе́; гриба́ - грибы́; грибы́ - грибо́в; грибо́в - гриба́м; гриба́м - гриба́ми; гриба́м - гриба́х.

Как видно, дается предпочтение цепочечной последовательности, а не радиальной. Но это не объясняет, почему словоформа *грибы́* выводится из *гриба́*, а не из *гриб*. Здесь вступает в силу другой принцип, заключающийся в стремлении к минимизации деривационных шагов с изменением основы. Вместо того, чтобы два раза перейти от ударения на основе к ударению на окончании, можно обойтись только одной такой операцией. Однако это осуществляется ценой нарушения направления сверху вниз (от В/Р ед.ч. к И/В мн.ч.).

² А. А. Реформацкий. Очерки по фонологии, морфонологии и морфологии. М., 1979, с. 50-51.

Очень часто приходится идти на такие компромиссы. Поэтому важно установить степень категоричности каждого запрета.

Следует еще отметить, что установление ударения всегда опирается на уровень фонетической реализации. Не принимается, например, абстрактное решение, что ударение в словоформе *гриб* падает на нулевое окончание³.

Структура настоящего словаря сложнее традиционного. Весь материал делится на 63 «главы», в каждой из которых представлено ровно одно окончание; например, пятая глава описывает окончание *-а*. В этой главе находим как вариант под ударением, так и вариант не под ударением: *-á, -а* (*гриба́, хле́ба*). Впрочем, есть и случаи с двумя вариантами под ударением: *-óго / -огó, -óти / -отí*. К окончанию *-а* относятся также сложные варианты, например, *-t-á* (*взята́*), но не другие окончания, начинающиеся с гласного *-а*, например, *-am, -am'i*.

В пределах каждой главы перечисляются словоформы, образованные при помощи данного окончания. Эти словоформы делятся на классы, определяемые соответствующим набором грамматических значений. Каждый такой класс образует некую словарную статью. Например, класс «Сущ. ж.р. ед.ч. И» содержит словоформы типа *ко́мната*. Одновременно любая словоформа входит в иерархическую цепочку, где она занимает начальную, промежуточную или конечную позицию. Словоформа *ко́мната* вместе со всем классом таких словоформ занимает, конечно, начальную позицию, поэтому соответствующий класс снабжается пометой «Первичный класс». Этот класс находим в статье, которая в очень сокращенном виде выглядит так:

5.1.9.2. (-а) ры́ба. Первичный класс, >200: ба́ба, осо́ба, па́губа, сва́дьба, бо́мба, ве́рба, ре́па, [...] кало́рия, о́ргия, хво́я, ста́туя, го́стья, плясу́нья, лгу́нья, келья.

→ 12.6.1. водé, 16.5. ничьéй, 18.1.11.1. губы́, [...]

Это надо понимать так: данный класс образуется при помощи пятого (5.) окончания, т.е. *-а*, причем в его простом (1.), а не в сложном виде. В пределах главы данный класс является девятым (9.) по очереди. Девятый класс охватывает как словоформы с окончанием под ударением, так и словоформы с безударным окончанием, как в нашем примере (2.). Первичные классы словоформ перечисляются *in extenso*, если они относительно маленькие, или же приводится в качестве иллюстрации какое-то число словоформ. В данном случае приводится максимальное число словоформ, а именно 200, с указанием, что словоформ на самом деле больше (>). Последняя строка указывает на статьи, содержащие словоформы, которые в иерархической системе следуют непосредственно после (всех или некоторых) словоформ данного класса.

Сейчас рассмотрим, но уже в несокращенном виде, более типичную статью с тем же окончанием, а именно статью, содержащую класс «Сущ. м.р. ед.ч. Р» (вариант под ударением):

³ Ср. А. А. Зализняк. Грамматический словарь русского языка. М., 1977, с. 31.

5.1.4.1. (-á)

а. & 1.1.3. удар. гриба́ ← гриб, >10: снопа́, следá, куста́, холма́, рубля́, ноября́, пирога́, языка́, ножа́, карандаша́; удар. + бегл. рта́ ← рот, >10: рва́, льда́, хребта́, овса́, дня́, угла́, бугра́, куска́, конца́, ручья́. *Искл.*: пол (по́ла).

б. & 1.1.4. удар. + бегл. чулка́ ← чуло́к (1). *Искл.*: боти́нок (боти́нка).

→ 18.1.7.1. грибы́, 39.2.1. табаку́, 39.3.1. плоту́, 39.4. углу́.

Мы видим, что этот класс словоформ в иерархической структуре занимает промежуточную позицию. Когда один класс выводится из другого, он сопоставляется с ним (обозначается знаком &), в результате чего возникают четыре возможных логических соотношения: совпадение, включение (два случая, один для каждого направления), пересечение. С точки зрения своего происхождения класс 5.1.4.1. не совпадает ни с одним другим классом; он составлен из двух других классов. Первый из них, 1.1.3., имеет грамматическую характеристику «Сущ. м.р. ед.ч. И/В» (*гриб*), а второй, 1.1.4. — грамматическую характеристику «Сущ. м.р. ед.ч. И/В / мн.ч. Р» (*чуло́к*). Более того, ни один из этих классов полностью, без остатка, не входит в класс 5.1.4.1. В результате имеет место пересечение как по отношению к классу 1.1.3., так и по отношению к классу 1.1.4. Те части обоих этих классов, которые не входят в класс 5.1.4.1., представлены в статье примерами с пометой «Искл.». После такой пометы приводится всегда только по одному примеру. В данном случае получаем информацию, что есть словоформы, входящие в классы 1.1.3. и 1.1.4., соответственно, но не входящие в класс 5.1.4.1.

Далее: в том случае, когда переход от одной словоформы к другой осуществляется при помощи операции над основой, это оговаривается отдельно. Например, вывод формы *гриба́* из формы *гриб* сопровождается сдвигом ударения («удар.»). Кроме того, есть случаи, когда добавляется еще одна операция, а именно превращение беглого гласного в нуль или наоборот («удар. + бегл.»). Этот подкласс перечисляется отдельно. В результате, как мы видим, в пределах одной статьи определяется не только соответствующий ей общий класс словоформ, но и разного рода подклассы. Все они представлены иллюстративным материалом и количественной характеристикой. Производный класс перечисляется *in extenso*, если число членов не более, чем одиннадцать, в противном случае приводится, кроме примера-образца, десять примеров и количественная информация «>10». Если класс содержит ровно один член, цифра «1» дается в скобках. Классы словоформ, представленных единственным примером после пометы «Искл.», количественной информации не содержат.

Сейчас опять обратимся к приведенной статье. Последняя строка содержит информацию о том, какие классы непосредственно выводятся из словоформ класса 5.1.4.1. Мы видим, что сюда, в частности, относятся три класса с окончанием *-и*, представленным в главе 39. Эти три класса имеют следующие грамматические характеристики: 39.2.1. «Сущ. м.р. ед.ч. Р/Д» (*табаку́*), 39.3.1. «Сущ. м.р. ед.ч. Д» (*плоту́*), 39.4. «Сущ. м.р. ед.ч. Д/П» (*углу́*). Приведу статью последней из них:

39.4. Сущ. м.р. ед.ч. Д/П (-и́) & 5.1.4.1. углу́ ← угла́, 7: лбу, льду́, мосту́, рту, сукú, полку́. *Искл.*: столá (столу́ только Д).

Данный класс целиком включен в класс, из которого он выведен, т.е., класс 5.1.4.1., но он гораздо меньше: в нем всего семь членов. Вариант не под ударением отсутствует.

Относительно редко встречаются случаи совпадения классов. Приведу один пример, а именно окончание *-ám'i* в сочетании с грамматической характеристикой «Сущ. м.р. мн.ч. Т» (данный класс занимает конечную позицию в иерархии):

9.1.1. (*-ám'i*) & 8.1.1. *глазáми* ← *глазám*, >10: *голубя́ми, льва́ми, цвета́ми, парня́ми, стола́ми, рубля́ми, берега́ми, ключа́ми, воробья́ми, сыновья́ми.*

Помимо тех операций над осовой, которые были уже упомянуты, «удар.» и «бегл.» (*гриб - гриба́, па́лка - па́лок*), есть еще несколько: «усеч.» (*судно - суда́*); «удлин.» (*брат - братья́*); «смягч.» (*сосед - соседи́*); «отверд.» (*изли́шний - изли́шен*); «измен.» (*ухо - у́ши*); «замен.» (*котёнок - котя́та*); «суппл.» (*челове́к - лю́ди*). Эти операции играют двоякую роль в установлении деривационных отношений. Во-первых, как уже было сказано, стремимся к минимизации числа шагов, сопровождаемых какой-либо операцией. Во-вторых, стараемся учитывать направленность операций над основой. Некоторые из этих операций отличаются более или менее явной направленностью, асимметричностью. Например, смягчение более естественный процесс, чем отверждение, усечение встречается чаще, чем удлинение. Естественность направленности в некоторой степени связана с потерей информации. В типичном случае исходная форма более информативна, чем производная.

Особенно большую роль играют эти операции при установлении направленности у глаголов. В спряжении операций над основой больше, чем в склонении. Поэтому критерий традиционного пути через парадигму отодвигается на задний план. В качестве первого звена деривации выступает не только инфинитив (традиционная форма), но часто и личная форма настоящего времени в первом лице. Так, инфинитив *помо́чь* посредством операции «измен.» выводится из более информативной личной формы *помогу́*, а не наоборот; зато личная форма *куру́* выводится из инфинитива *куру́ть*, в силу естественной направленности операции «усеч.»⁴.

Несмотря на определенный эклектизм при выборе и применении критериев, выигрыш от иерархизации словоформ очень большой. Это важнейший компонент предлагаемого словаря, позволяющий определить большое число флексивных классов разного типа и соотношений между ними и, таким образом, под другим углом зрения рассмотреть сферы регулярности и нерегулярности в русской морфологии.

⁴ Ср. понятие «полной основы» в классической работе Р. Якобсона о русском спряжении: R. Jakobson. Russian Conjugation // Word 4:3, 1948.