



119270, Москва, Лужнецкая наб., д. 6,  
стр.1, офис 214, ООО «ЭР СИ О»  
Тел./факс: (495) 287-98-87  
E-mail: [info@rco.ru](mailto:info@rco.ru)  
<http://www.rco.ru>

**Руководство администратора  
RCO Morph – пакет настройки  
морфологического словаря**

Версия 2.1

(Microsoft Windows)

Москва, 2007

В содержание данного документа могут быть внесены изменения без предварительного уведомления. Названия организаций, имена и даты, используемые в качестве примеров, являются вымышленными, если не оговорено обратное.

© ООО «ЭР СИ О», 2000-2007. Все права защищены.

ЭР СИ О, Russian Context Optimizer, RCO являются охраняемыми товарными знаками.

Oracle 9i, Oracle 10g, Oracle 11g, Oracle, Oracle Text, SQL\*Plus, PL/SQL являются зарегистрированными торговыми марками корпорации Oracle.

ООО «ЭР СИ О» может являться правообладателем патентов и заявок, поданных на получение патента, товарных знаков и объектов авторского права, которые имеют отношение к содержанию данного документа.

Предоставление вам данного документа не означает передачи какой-либо лицензии на использование данных патентов, товарных знаков и объектов авторского права, за исключением использования, явно оговоренного в лицензионном соглашении ООО «ЭР СИ О».

Все другие названия юридических лиц и изделий являются охраняемыми товарными знаками или товарными знаками, принадлежащими их владельцам.

**Содержание**

Обзор .....	4
Введение .....	5
Установка пакета RCO Morph.....	6
Работа с пакетом RCO Morph.....	7
Морфологические словари .....	7
Базовые словари .....	7
Дополнительные словари .....	8
Сборка словаря: MorphDicCompiler .....	9
Пополнение словаря: MorphDicEx .....	10
Тестирование словаря: MorphDemo .....	10
Приложение 1. Описание парадигмы словоизменения .....	12
Приложение 2. Классификация частей речи.....	17

---

# Обзор

---

В этом руководстве содержатся основные сведения по установке и работе с пакетом **RCO Morph**, который позволяет настраивать словарь, используемый в продуктах линейки **RCO (Russian Context Optimizer)** для морфологического анализа слов русского языка и расширения запросов при контекстном поиске.

Целевой бинарный словарь **MorphDct.dat**, используемый в программах **RCO** для высокоэффективной работы, собирается из исходных текстовых файлов словаря при помощи программ, входящих в этот пакет.

Для установки и работы с пакетом **RCO Morph** необходимо располагать:

- компьютером с установленной операционной системой **Microsoft Windows NT** версии 4.0 и выше.

# Введение

---

Пакет **RCO Morph** предназначен для настройки словаря морфологического анализа, используемого при приведении слов русского языка к нормальной форме, построении словоформ и т.п.

Средства морфологического анализа **RCO** включают в себя:

- точный анализ известных слов по словарю объемом более 115 тысяч слов, что покрывает более 3-х миллионов словоформ;
- высоко достоверный анализ неизвестного слова на основе комплекса правил словообразования и словоизменения;
- вероятностный анализ посредством соотнесения с моделями словоизменения часто встречающихся слов на основе оценки флективной и суффиксальной частей слова.

Подход к описанию морфологической системы языка использует обучение на примерах словоизменения, в ходе которого автоматически выделяются окончания, основы, суффиксы и строятся модели, описывающие изменение различных классов слов.

Неизвестные слова анализируются по аналогии с известными, что обычно оказывается достаточным для корректной обработки большинства специальных терминов и постоянно появляющихся в языке неологизмов.

Однако, ввиду присутствия в языке явления морфологической омонимии, достоверный анализ слова в ряде случаев оказывается невозможен. Например, слово «бантука» может оказаться как грамматической формой родительного падежа существительного мужского рода «бантук», так и формой именительного падежа существительного женского рода «бантука». В подобном случае может потребоваться добавить слово в словарь со всеми грамматическими формами.

По нашим оценкам, из 110 тысяч слов русского языка около 2-х тысяч имеет совпадающие грамматические формы. Ввиду этого при работе с поисковой системой возможны ситуации, когда некоторые из словоформ, встречающиеся редко в предметной области, целесообразно удалить из словаря. Так, если объектом поиска в юридической базе данных часто является существительное «право», во избежание путаницы, краткую форму среднего рода у прилагательного «правый» можно удалить.

Для решения проблем, связанных с недостаточной полнотой словаря или омонимией словоформ, предназначены средства, поставляемые в настоящем пакете.

---

# Установка пакета RCO Morph

---

Для установки пакета просто скопируйте все файлы дистрибутива в папку **RCO Morph** или любую иную папку на вашем компьютере.

В состав пакета входят следующие файлы:

<b>morph.ini</b>	файл с перечнем исходных текстовых словарей, используемых для сборки бинарного словаря <b>MorphDct.dat</b>
<b>*.txt</b>	набор исходных текстовых файлов словаря, перечисленных в файле <b>morph.ini</b>
<b>MorphDct.dat</b>	бинарный файл морфологического словаря, создаваемый из текстовых файлов при помощи приложения-компилятора <b>MorphDicCompiler.exe</b>
<b>MorphDicCompiler.exe</b>	компилятор бинарного словаря – консольное приложение
<b>MorphDemo.exe</b>	графическое приложение для тестирования бинарного морфологического словаря <b>MorphDct.dat</b>
<b>MorphDicEx.exe</b>	утилита для автоматизированного расширения словаря – консольное приложение
<b>RCO Morph.pdf</b>	файл настоящего руководства

---

# Работа с пакетом RCO Morph

---

## Морфологические словари

Набор стандартных словарей словоизменения общей лексики русского языка, входящий в состав лингвистического обеспечения **RCO**, хранится в текстовом виде в файлах с расширением **\*.txt** в базовой папке **RCO Morph**. Список файлов-словарей перечислен в файле **morph.ini**.

Утилита **MorphDicCompiler.exe** собирает из текстовых файлов словарей двоичный файл **MorphDct.dat**. Словарь **MorphDct.dat** используется программными продуктами из серии **RCO** для морфологического анализа.

Существует возможность внесения изменений в словарь посредством редактирования файлов в текстовом редакторе, а также посредством включения при компиляции дополнительных словарей, в которых указываются все словоформы новых слов. Дополнительно, для удобства добавления новых слов, в состав пакета входит приложение **MorphDicEx.exe**, которое позволяет автоматизировать процесс построения словоформ.

---

**Замечание.** В зависимости от условий поставки файлы словарей могут быть либо доступны для просмотра и редактирования, либо зашифрованы. В последнем случае изменение базовых словарей невозможно.

---

## Базовые словари

Текстовые файлы базовых морфологических словарей содержат слова в нормальных формах, снабженные парой маркеров – кодом парадигмы словоизменения и кодом части речи, например:

---

«АБАЖУР s2 NounInanimateM»

---

Первый символ кода парадигмы означает класс словоизменения:

- s – существительное,
- d – прилагательное,
- v – глагол,
- n – неизменяемое.

Номер, следующий за ним, указывает конкретную парадигму изменения, описанную в словаре **MorphDct.dat**, т.е. способ построения всех грамматических форм.

Подробная информация о парадигмах словоизменения приведена в разделе [Приложение 1](#). Коды частей речи описаны в [Приложении 2](#).

Базовые словари допускают удаление слов, а также пополнение новыми словами, которые изменяются в соответствии с имеющимися парадигмами. Например, можно добавить «БУЖУР s2 GeographyName», которое будет иметь такие же окончания всех грамматических форм, как слово «abajур» и прочие слова с кодом парадигмы s2.

Все подключаемые базовые морфологические словари должны быть перечислены в файле **morph.ini** с указанием ключевого слова *brief\_dictionary*, которое определяет описанный формат записи словаря, например: “*brief\_dictionary: GeoName.txt*”.

## Дополнительные словари

Файлы дополнительных словарей позволяют вводить новые слова с неизвестными парадигмами изменения. Формат описания парадигмы следующий:

---

```
* s GeoNameF
АМУДАРЬЯ // И. ед.
АМУДАРЬИ // Р. ед.",
АМУДАРЬЕ // Д. ед.",
АМУДАРЬЮ // В. ед.",
АМУДАРЬЕЙ/АМУДАРЬЕЮ // "Т. ед.",
АМУДАРЬЕ //П. ед.",
-
-
```

---

и т.д.

Здесь после звездочки, означающей начало нового слова, указывается класс словоизменения (**s**, **d**, **v**, **n**) и код части речи, после чего следует список грамматических форм в порядке, указанном в [Приложении 1](#). После словоформы в каждой строке допускаются произвольные комментарии, игнорируемые компилятором.

Случаю отсутствия формы соответствует символ «←». Возможные варианты одной грамматической формы (допустимо не более двух) разделяются символом “/”.

Поскольку большинство глагольных форм может быть однозначно построено по правилам (возвратные формы, склонения причастий), отсутствует необходимость явно задавать все формы (около 230) в словаре. Вместо этого достаточно указать лишь 26 основных (первых) форм глагола. Наличие форм прямого/возвратного инфинитива, форм именительного падежа действительных и страдательных причастий настоящего и прошедшего времени указывает компилятору на то, что остальные варианты этих форм следует достроить по правилам.

Все подключаемые дополнительные словари должны быть перечислены в файле **morph.ini** с указанием ключевого слова *full\_dictionary*, которое определяет описанный формат записи словаря, например: “*full\_dictionary: MyWords.txt*”.

## Сборка словаря: MorphDicCompiler

Сборка новой версии двоичного словаря **MorphDct.dat**, используемого **RCO** для морфологического анализа, производится запуском приложения **MorphDicCompiler**. Для этого перейдите в папку **RCO Morph** и запустите файл **MorphDicCompiler.exe**.

Компилятор словаря обрабатывает все текстовые файлы, перечисленные в файле **morph.ini**, и собирает новый бинарный словарь **MorphDctNew.dat**, а также новые текстовые словари, одноименные словарям, включенным в процесс сборки, с расширением **.new**.

---

**Внимание!** После компиляции список парадигм в результирующем бинарном словаре **MorphDctNew.dat** модифицируется – могут возникнуть новые парадигмы и исчезнуть старые.

В результате новые словари содержат все слова из файлов, участвовавших в процессе сборки, с указанием нового номера парадигмы словоизменения, отсылающего к парадигме из нового бинарного словаря **MorphDctNew.dat**. Таким образом, номера парадигм в текстовых файлах строго привязаны к данным в соответствующем **MorphDct.dat**.

---

Сообщения об ошибках в ходе компиляции помещаются в файл **\_errors.txt**. Ошибки могут быть вызваны:

- указанием несуществующего кода парадигмы в файле словаря типа *brief\_dictionary* или несоответствием указанной парадигмы слову (конец слова не совпадает с окончанием нормальной формы в заданной парадигме);
- отсутствием требуемого количества форм для слова в файле словаря типа *full\_dictionary*;
- указанием несуществующего кода части речи.

В случае отсутствия сообщений об ошибках следует заменить файлы **MorphDct.dat** на **MorphDctNew.dat**, а также все файлы, включенные в процесс компиляции, на их новые версии, получившие расширение **.new**. Файлы дополнительных словарей типа *full\_dictionary* можно оставлять прежними, сохраняя новые слова в формате с развернутой парадигмой словоизменения.

## Пополнение словаря: MorphDicEx

Если требуется пополнить словарь большим количеством новых слов, целесообразно воспользоваться утилитой автоматизированного пакетного пополнения словаря **MorphDicEx**.

Для этого добавляемые слова следует ввести в текстовый файл, после чего из базовой папки **RCO Morph** запустить консольное приложение **MorphDicEx.exe**, указав ему в качестве параметра имя созданного файла.

---

**Внимание!** Все добавляемые слова должны стоять в нормальной форме, т.е. в той форме, в которой они приводятся в словарях. Например, для глаголов это – инфинитив, для существительных – это форма именительного падежа единственного числа, для прилагательных – форма именительного падежа единственного числа мужского рода. Имена собственные должны быть написаны с большой буквы.

---

В результате обработки в базовой папке **RCO Morph** будут созданы два файла: **AddDct.txt** и **known.txt**.

Первый файл содержит список слов, не входящих в текущий бинарный словарь **MorphDct.dat**, с развернутыми парадигмами словоизменения (списками словоформ), в формате, готовом для сборки (см. раздел «Сборка словаря: **MorphDicCompiler**»).

В файл **known.txt** заносятся уже присутствующие в словаре слова, которые, скорее всего, добавлять не следует.

---

**Замечание.** Словоформы неизвестных слов (файл **AddDct.txt**) строятся алгоритмом вероятностного морфологического анализа, по аналогии с уже присутствующими в словаре словами. Поэтому перед включением словаря в компиляцию рекомендуется просмотреть содержимое файла в текстовом редакторе и, если потребуется, подправить (или добавить) некоторые словоформы, которые оказались неверно построены.

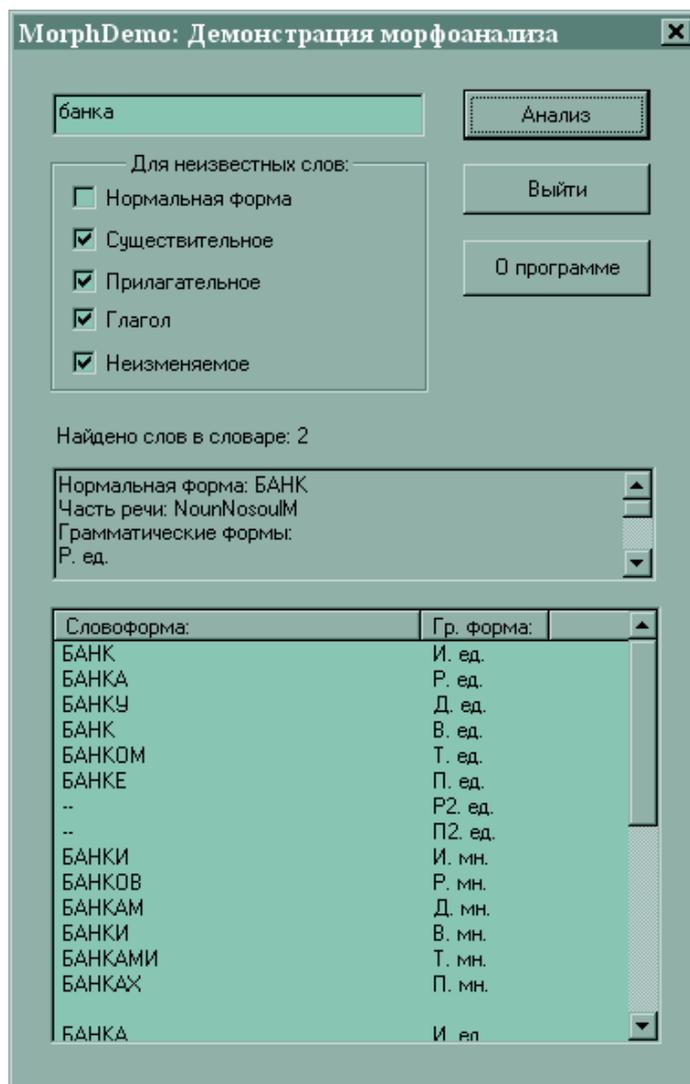
---

## Тестирование словаря: MorphDemo

Для тестирования морфологического анализатора на основе текущего бинарного словаря **MorphDct.dat** найдите в меню **Program** раздел **RCO Morph** и выберите пункт **Morphology Demo**. Окно запущившегося приложения будет иметь вид, представленный на рисунке ниже.

Чтобы проверить, как обрабатывается интересующее слово, наберите его в верхнем текстовом поле, после чего нажмите кнопку «**Анализ**».

В нижнем окне будет представлен список всех построенных словоформ, а в среднем окне – информация о нормальной форме или основе (для неизвестных слов), части речи и грамматических формах, в которых может стоять введенная словоформа. Для известных слов в словаре может быть найдено несколько вариантов изменения и, соответственно, построено несколько возможных нормальных форм.



Для повышения точности анализа неизвестных слов **RCO** позволяет использовать информацию о допустимой части речи (если она известна). Например, крайне редко встречаются отсутствующие в словаре глаголы, а если поиск ведется по базе медицинских терминов, то подавляющее большинство неизвестных слов окажется существительными – названиями препаратов. Так, слово «ортопилл», скорее всего, не следует обрабатывать как глагол, несмотря на характерное глагольное окончание.

Установка флажков в разделе «Для неизвестных слов» напротив названий классов частей речи позволяет включить в анализ только выбранные классы моделей словоизменения.

Установка флажка «Нормальная форма» позволяет построить все словоформы неизвестного слова с более высокой точностью, если слово задано в нормальной форме.

# Приложение 1.

## Описание парадигмы словоизменения

Парадигма словоизменения представляет множество словоформ, которые может принимать слово в определенных грамматических формах. В зависимости от части речи группы слов могут потенциально иметь ограниченный набор грамматических форм. Вследствие этого все парадигмы разделяются на классы таким образом, что каждый класс парадигм описывает изменение слов, относящихся к одному грамматическому разряду.

В принятой здесь системе описания словоизменения существуют три полноценных класса парадигм: существительного, прилагательного и глагола. К четвертому, вырожденному классу, относятся неизменяемые слова. Словоизменение всех прочих частей речи (см. [Приложение 2](#)) строится по парадигмам одного из этих классов. Причастия не рассматриваются, как самостоятельная часть речи, и все их формы включены в парадигму глагола.

Описание кодов парадигм.

<i>MO_PARADIGM_SUBJECT</i>	класс существительных
<i>MO_PARADIGM_DEFINITION</i>	класс прилагательных
<i>MO_PARADIGM_VERB</i>	класс глаголов
<i>MO_PARADIGM_UNCHANGED</i>	неизменяемые слова

Ниже приведен порядок следования грамматических форм в парадигмах словоизменения.

Слова, изменяющиеся по типу существительного (код *MO\_PARADIGM\_SUBJECT*), могут иметь формы, указанные в таблице ниже.

Индекс	Грамматическая форма	Пример
0	имен. падеж, един.	река
1	род. падеж, един.	реки
2	дат. падеж, един.	реке
3	вин. падеж, един.	реку
4	твор. падеж, един.	рекой, рекою
5	пред. падеж, един.	реке
6	2-ой. род. падеж, един.	–
7	2-ой. пред. падеж, един.	–
8	им. падеж, множ.	реки
9	род. падеж, множ.	рек
10	дат. падеж, множ.	рекам
11	вин. падеж, множ.	реки
12	твор. падеж, множ.	реками
13	пред. падеж, множ.	реках

Слова, изменяющиеся по типу прилагательного (код *MO\_PARADIGM\_DEFINITION*), могут иметь формы, указанные в таблице ниже.

Индекс	Грамматическая форма	Пример
0	имен. падеж, един., муж.	быстрый
1	род. падеж, един., муж.	быстрого
2	дат. падеж, един., муж.	быстрому
3	вин. падеж, един., муж., одуш.	быстрого
4	вин. падеж, един., муж., неодуш.	быстрый
5	твор. падеж, един., муж.	быстрым
6	пред. падеж, един., муж.	быстром
7	краткая форма, муж.	быстр
8	имен. падеж, един., жен.	быстрая
9	род. падеж, един., жен.	быстрой
10	дат. падеж, един., жен.	быстрой
11	вин. падеж, един., жен.	быструю
12	твор. падеж, един., жен.	быстрой, быстрою
13	пред. падеж, един., жен.	быстрой
14	краткая форма, жен.	быстра
15	имен. падеж, един., сред.	быстрое
16	род. падеж, един., сред.	быстрого
17	дат. падеж, един., сред.	быстрому
18	вин. падеж, един., сред.	быстрое
19	твор. падеж, един., сред.	быстрым
20	пред. падеж, един., сред.	быстром
21	краткая форма, сред.	быстро
22	имен. падеж, множ.	быстрые
23	род. падеж, множ.	быстрых
24	дат. падеж, множ.	быстрым
25	вин. падеж, множ., одуш.	быстрых
26	вин. падеж, множ., неодуш.	быстрые
27	твор. падеж, множ.	быстрыми
28	пред. падеж, множ.	быстрых
29	краткая форма, множ.	быстры
30	сравнительная степень	быстрее, быстреей

Слова, изменяющиеся по типу глагола (код *MO\_PARADIGM\_VERB*), могут иметь формы, указанные в таблице ниже.

Индекс	Грамматическая форма	Пример
0	Инф.	перифотографировать
1	Возв. инф.	перифотографироваться
2	Повел. Накл.	перифотографируй
3	Множ. повел.	перифотографируйте
4	Буд. 1-е ед.	перифотографирую
5	Буд. 1-е мн.	перифотографируем
6	Буд. 2-е ед.	перифотографируешь
7	Буд. 2-е мн.	перифотографируете
8	Буд. 3-е ед.	перифотографирует
9	Буд. 3-е мн.	перифотографируют
10	Наст. 1-е ед.	перифотографирую
11	Наст. 1-е мн.	перифотографируем
12	Наст. 2-е ед.	перифотографируешь
13	Наст. 2-е мн.	перифотографируете
14	Наст. 3-е ед.	перифотографирует

15	Наст. 3-е мн.	перифотографируют
16	Наст. прич. действ. И. ед.	перифотографирующий
17	Наст. прич. страд. И. ед.	перифотографируемый
18	Наст. дееприч.	перифотографируя
19	Прош. муж.	перифотографировал
20	Прош. жен.	перифотографировала
21	Прош. ср.	перифотографировало
22	Прош. множ.	перифотографировали
23	Прош. прич. действ. И. ед.	перифотографировавший
24	Прош. прич. страд. И. ед.	перифотографированный
25	Прош. дееприч.	перифотографировав, перифотографировавши
26	Наст. прич. действ. Р. муж.	перифотографирующего
27	Наст. прич. действ. Д. муж.	перифотографирующему
28	Наст. прич. действ. В. муж. одуш.	перифотографирующего
29	Наст. прич. действ. В. муж. неодуш.	перифотографирующий
30	Наст. прич. действ. Т. муж.	перифотографирующим
31	Наст. прич. действ. П. муж.	перифотографирующем
32	Наст. прич. действ. кф. муж.	–
33	Наст. прич. действ. И. жен.	перифотографирующая
34	Наст. прич. действ. Р. жен.	перифотографирующей
35	Наст. прич. действ. Д. жен.	перифотографирующей
36	Наст. прич. действ. В. жен.	перифотографирующую
37	Наст. прич. действ. Т.1 жен.	перифотографирующей
38	Наст. прич. действ. Т.2 жен.	перифотографирующей
39	Наст. прич. действ. П. жен.	перифотографирующей
40	Наст. прич. действ. кф. жен.	–
41	Наст. прич. действ. И. ср.	перифотографирующее
42	Наст. прич. действ. Р. ср.	перифотографирующего
43	Наст. прич. действ. Д. ср.	перифотографирующему
44	Наст. прич. действ. В. ср.	перифотографирующее
45	Наст. прич. действ. Т. ср.	перифотографирующим
46	Наст. прич. действ. П. ср.	перифотографирующем
47	Наст. прич. действ. кф. ср.	–
48	Наст. прич. действ. И. мн.	перифотографирующие
49	Наст. прич. действ. Р. мн.	перифотографирующих
50	Наст. прич. действ. Д. мн.	перифотографирующим
51	Наст. прич. действ. В. мн. одуш.	перифотографирующих
52	Наст. прич. действ. В. мн. неодуш.	перифотографирующие
53	Наст. прич. действ. Т. мн.	перифотографирующими
54	Наст. прич. действ. П. мн.	перифотографирующих
55	Наст. прич. действ. кф. мн.	–
56	Прош. прич. действ. Р. муж.	перифотографировавшего
57	Прош. прич. действ. Д. муж.	перифотографировавшему
58	Прош. прич. действ. В. муж. одуш.	перифотографировавшего
59	Прош. прич. действ. В. муж. неодуш.	перифотографировавший
60	Прош. прич. действ. Т. муж.	перифотографировавшим
61	Прош. прич. действ. П. муж.	перифотографировавшем
62	Прош. прич. действ. кф. муж.	–

63	Прош. прич. действ. И. жен.	перифотографировавшая
64	Прош. прич. действ. Р. жен.	перифотографировавшей
65	Прош. прич. действ. Д. жен.	перифотографировавшей
65	Прош. прич. действ. В. жен.	перифотографировавшую
67	Прош. прич. действ. Т.1 жен.	перифотографировавшей
68	Прош. прич. действ. Т.2 жен.	перифотографировавшею
69	Прош. прич. действ. П. жен.	перифотографировавшей
70	Прош. прич. действ. кф. жен.	–
71	Прош. прич. действ. И. ср.	перифотографировавшее
72	Прош. прич. действ. Р. ср.	перифотографировавшего
73	Прош. прич. действ. Д. ср.	перифотографировавшему
74	Прош. прич. действ. В. ср.	перифотографировавшее
75	Прош. прич. действ. Т. ср.	перифотографировавшим
76	Прош. прич. действ. П. ср.	перифотографировавшем
77	Прош. прич. действ. кф. ср.	–
78	Прош. прич. действ. И. мн.	перифотографировавшие
79	Прош. прич. действ. Р. мн.	перифотографировавших
80	Прош. прич. действ. Д. мн.	перифотографировавшим
81	Прош. прич. действ. В. мн. одуш.	перифотографировавших
82	Прош. прич. действ. В. мн. неодуш.	перифотографировавшие
83	Прош. прич. действ. Т. мн.	перифотографировавшими
84	Прош. прич. действ. П. мн.	перифотографировавших
85	Прош. прич. действ. кф. мн.	–
86	Наст. прич. страд. Р. муж.	перифотируемого
87	Наст. прич. страд. Д. муж.	перифотируемому
88	Наст. прич. страд. В. муж. одуш.	перифотируемого
89	Наст. прич. страд. В. муж. неодуш.	перифотируемый
90	Наст. прич. страд. Т. муж.	перифотируемым
91	Наст. прич. страд. П. муж.	перифотируемом
92	Наст. прич. страд. кф. муж.	перифотируем
93	Наст. прич. страд. И. жен.	перифотируемая
94	Наст. прич. страд. Р. жен.	перифотируемой
95	Наст. прич. страд. Д. жен.	перифотируемой
96	Наст. прич. страд. В. жен.	перифотируемую
97	Наст. прич. страд. Т.1 жен.	перифотируемой
98	Наст. прич. страд. Т.2 жен.	перифотируемой
99	Наст. прич. страд. П. жен.	перифотируемой
100	Наст. прич. страд. кф. жен.	перифотируема
101	Наст. прич. страд. И. ср.	перифотируемое
102	Наст. прич. страд. Р. ср.	перифотируемого
103	Наст. прич. страд. Д. ср.	перифотируемому
104	Наст. прич. страд. В. ср.	перифотируемое
105	Наст. прич. страд. Т. ср.	перифотируемым
106	Наст. прич. страд. П. ср.	перифотируемом
107	Наст. прич. страд. кф. ср.	перифотируемо
108	Наст. прич. страд. И. мн.	перифотируемые
109	Наст. прич. страд. Р. мн.	перифотируемых
110	Наст. прич. страд. Д. мн.	перифотируемым
111	Наст. прич. страд. В. мн. одуш.	перифотируемых
112	Наст. прич. страд. В. мн. неодуш.	перифотируемые

113	Наст. прич. страд. Т. мн.	перифотографируемыми
114	Наст. прич. страд. П. мн.	перифотографируемых
115	Наст. прич. страд. кф. мн.	перифотографируемы
116	Прош. прич. страд. Р. муж.	перифотографированного
117	Прош. прич. страд. Д. муж.	перифотографированному
118	Прош. прич. страд. В. муж. одуш.	перифотографированного
119	Прош. прич. страд. В. муж. неодуш.	перифотографированный
120	Прош. прич. страд. Т. муж.	перифотографированным
121	Прош. прич. страд. П. муж.	перифотографированном
122	Прош. прич. страд. кф. муж.	перифотографирован
123	Прош. прич. страд. И. жен.	перифотографированная
124	Прош. прич. страд. Р. жен.	перифотографированной
125	Прош. прич. страд. Д. жен.	перифотографированной
126	Прош. прич. страд. В. жен.	перифотографированную
127	Прош. прич. страд. Т.1 жен.	перифотографированной
128	Прош. прич. страд. Т.2 жен.	перифотографированную
129	Прош. прич. страд. П. жен.	перифотографированной
130	Прош. прич. страд. кф. жен.	перифотографирована
131	Прош. прич. страд. И. ср.	перифотографированное
132	Прош. прич. страд. Р. ср.	перифотографированного
133	Прош. прич. страд. Д. ср.	перифотографированному
134	Прош. прич. страд. В. ср.	перифотографированное
135	Прош. прич. страд. Т. ср.	перифотографированным
136	Прош. прич. страд. П. ср.	перифотографированном
137	Прош. прич. страд. кф. ср.	перифотографировано
138	Прош. прич. страд. И. мн.	перифотографированные
139	Прош. прич. страд. Р. мн.	перифотографированных
140	Прош. прич. страд. Д. мн.	перифотографированным
141	Прош. прич. страд. В. мн. одуш.	перифотографированных
142	Прош. прич. страд. В. мн. неодуш.	перифотографированные
143	Прош. прич. страд. Т. мн.	перифотографированными
144	Прош. прич. страд. П. мн.	перифотографированных
145	Прош. прич. страд. кф. мн.	перифотографированы

## Приложение 2.

# Классификация частей речи

Ниже приведены коды частей речи (лексико-грамматических разрядов).

В последнем столбце указан класс парадигмы словоизменения, который определяет возможные грамматические формы слов (см. [Приложение 1](#)):

Н – неизменяемое, С – существительное, П – прилагательное, Г – глагол.

Часть речи	Код	Пример	Класс
неизвестна	MO_POS_UNKNOWN		
Вводное слово	MO_POS_INTRODUCTORY	просто	Н
Междометье	MO_POS_INTERJECTION	ах	Н
Предикатив	MO_POS_PREDICATIVE	просто	Н
Предлог	MO_POS_PREPOSITION1	у	Н
Предлог	MO_POS_PREPOSITION2	к	Н
Предлог	MO_POS_PREPOSITION3	в	Н
Предлог	MO_POS_PREPOSITION4	с	Н
Предлог	MO_POS_PREPOSITION5	на	Н
Союз сочинительный	MO_POS_CONJUNCTIONCOORDINATIVE	и	Н
Союз	MO_POS_CONJUNCTION1	что	Н
Союз	MO_POS_CONJUNCTION2	что	Н
Союз	MO_POS_CONJUNCTION3	что	Н
Частица	MO_POS_PARTICLE	же	Н
Наречие	MO_POS_ADVERB	просто	Н
Сущ. сокращенное	MO_POS_ABBREVIATEDNOUN	выкл.	Н
Прил. сокращенное	MO_POS_ABBREVIATEDADJECTIVE	др.	Н
Сокращенное вводное слово	MO_POS_ABBREVIATEDINTRODUCTORY	напр.	Н
Обособленная сравнительная степень	MO_POS_COMPARATIVE	побольше	Н
Притяжательное прилагательное	MO_POS_POSSESSIVEADJECTIVE	птичий	П
Мест. муж.	MO_POS_PRONOUN	он	С
Мест. жен.	MO_POS_PRONOUNFEMININE	она	С
Мест. сред.	MO_POS_PRONOUNNEUTER	оно	С
Местоименное прилагательное	MO_POS_PRONOUNADJECTIVE	такой	П
Числ.	MO_POS_NUMERIC	три	С
Собирательное числ.	MO_POS_NUMERICCOLLECTIVE	трое	С
Числ. порядковое	MO_POS_NUMERICORDINAL	третий	П
Аббр. муж	MO_POS_ABBREVIATURE	ВАК	Н
Аббр. жен.	MO_POS_ABBREVIATUREFEMININE	РФ	Н
Аббр. ср.	MO_POS_ABBREVIATURENEUTER	ЦРУ	Н
Имя собств. муж.	MO_POS_NAMEMASCULINE	Саша	С
Имя собств. жен.	MO_POS_NAMEFEMININE	Мария	С
Отчество муж.	MO_POS_MIDDLENAMEMASCULINE	Васильевич	С

Отчество жен.	MO_POS_MIDDLENAMEFEMININE	Васильевна	С
Фамилия	MO_POS_LASTNAME	Иванов	П
Географ. название муж.	MO_POS_GEOGRAPHYNAME MASCULINE	Орел	С
Географ. название жен.	MO_POS_GEOGRAPHYNAMEFEMININE	Обь	С
Географ. название сред.	MO_POS_GEOGRAPHYNAME NEUTER	Внуково	С
Географ. название множ.	MO_POS_GEOGRAPHYNAME PLURAL	Карпаты	С
Глагол переходный несов. вида	MO_POS_VERBIMPERFECT	смотреть	Г
Глагол непереходный несов. вида	MO_POS_VERBINTRANSITIVEIMPERFECT	чихать	Г
Глагол переходный совершенного вида	MO_POS_VERBPERFECT	посмотреть	Г
Глагол непереходный совершенного вида	MO_POS_VERBINTRANSITIVE PERFECT	чихнуть	Г
Глагол переходный двувидовой	MO_POS_VERBTWOWAY	эмитировать	Г
Глагол непереходный двувидовой	MO_POS_VERBINTRANSITIVE TWOWAY	эмигрировать	Г
Прил.	MO_POS_ADJECTIVE1	фонетический	П
Прил.	MO_POS_ADJECTIVE2	фонетический	П
Сущ. неодуш. муж.	MO_POS_NOUNINANIMATE MASCULINE	трактор	С
Сущ. одуш. муж.	MO_POS_NOUNANIMATE MASCULINE	зяблик	С
Сущ. одуш.-неодуш. муж.	MO_POS_NOUNANIMATE INANIMATE MASCULINE	кодировщик	С
Сущ. неодуш. жен.	MO_POS_NOUNINANIMATE FEMININE	весна	С
Сущ. одуш. жен.	MO_POS_NOUNANIMATE FEMININE	птичка	С
Сущ. одуш.-неодуш. жен.	MO_POS_NOUNANIMATE INANIMATE FEMININE	матрешка	С
Сущ. неодуш. сред.	MO_POS_NOUNINANIMATE NEUTER	озеро	С
Сущ. одуш. сред.	MO_POS_NOUNANIMATE NEUTER	дитя	С
Сущ. одуш.-неодуш. сред.	MO_POS_NOUNANIMATE INANIMATE NEUTER	детище	С
Сущ. неодуш. множ.	MO_POS_NOUNINANIMATE PLURAL	ножницы	С
Наименование муж. рода	MO_POS_APPELLATION_MASCULINE	Гарант-Парк-интернет	С
Наименование жен. рода	MO_POS_APPELLATION_FEMININE	Формоза	С
Наименование сред. рода	MO_POS_APPELLATION_NEUTER	Алеко	С
Наименование множ. числа	MO_POS_APPELLATION_PLURAL	Инфо-Технологии	С