

# Автоматическая обработка текстов

---

## МОРФОЛОГИЧЕСКИЙ АНАЛИЗ

# Морфология (лингвистика)

---

Морфоло́гия (от др.-греч. *μορφή* — «форма» и *λόγος* — «слово, учение») — раздел грамматики, основными объектами которого являются слова естественных языков, их значимые части и морфологические признаки.

В задачи морфологии входит определение слова как особого языкового объекта и описание его внутренней структуры

# Морфология включает в себя:

---

Учение о словоизменении в языке, парадигмах, словоизменительных типах. Это обязательная составляющая морфологии, и именно с составления парадигм (таблиц склонения и спряжения) исторически началась лингвистика вообще (в древнем Вавилоне)

Учение о словообразовании, стоящее на границе морфологии и лексикологии

Учение о частях речи, при выделении которых задействованы не только морфологические (в узком смысле), но и синтаксические и семантические критерии

Грамматическую семантику, то есть учение о грамматических значениях

# Определение и задачи

---

**Морфологический анализ** – определение грамматических характеристик слова

Морфологический анализ включает задачи:

- Определение части речи
- Лемматизация
- Определение грамматических характеристик
- Генерация словоформ

---

**Лемма** – каноническая (начальная) словоформа лексемы: *стол, деревянный, стоять*

- для существительных: именительный падеж единственного числа
- для прилагательных: именительный падеж единственного числа мужского рода
- для глаголов, причастий и деепричастий: глагол в инфинитиве

**Лемматизация** – определение начальной формы слова

# Грамматические характеристики

---

- Часть речи
- Число
- Род
- Падеж
- Время
- Наклонение
- ...

# Части речи

---

Имя существительное (noun) *топор, философия*

Имя прилагательное (adjective) *красивый*

Местоимение (pronoun) *я, ты, этот, мой*

Глагол (verb) *петь, рисовать*

Причастие (participle) *бегущий, поющий*

Деепричастие (gerund) *пробегая, напевая*

Наречие (adverb) *красиво, быстро*

Имя числительное (numeral) *два, третий*

Предлог (preposition), союз (conjunction), частица (particle), междометие (interjection) *в, чтобы, не, ах*

# Пример

---

Дом [дом: S, муж, неод=(вин,ед | им,ед)]

был [быть: V, нп=прош,ед,изъяв,муж,несов]

построен [построить: V, сов, пе=прош,ед,прич,кр,муж,страд]

рабочими [рабочий: S, муж, од=твор,мн]



# Пример

Input Your text:

☐ English ☒ Russian ☐ German

☐ With paradigms

Submit Request

Found	Dict ID	Lemma	Grammems
+	од,	СЛОН	С мр,дт,ед,

Input Your text:

☐ English ☒ Russian ☐ German

☐ With paradigms

Submit Request

Found	Dict ID	Lemma	Grammems	
+	кач,	БОЛЬШОЙ	П но,од,ср,мр,пр,ед,	

<https://aot.ru/>

# Парадигма словоформы «решала»

ИНФИНИТИВ дст,
РЕША'ТЬ

ЛИЧНАЯ ФОРМА дст,	
нст	
РЕША'ЕМ	1л,мн
РЕША'Ю	1л,ед
РЕША'ЕТЕ	2л,мн
РЕША'ЕШЬ	2л,ед
РЕША'ЮТ	3л,мн
РЕША'ЕТ	3л,ед
прш	
РЕША'ЛИ	мн
РЕША'Л	мр,ед
РЕША'ЛА	жр,ед
РЕША'ЛО	ср,ед
пвл,2л	
РЕША'ЙТЕ	мн
РЕША'Й	ед

ДЕЕПРИЧАСТИЕ дст,	
РЕША'Я	нст
РЕША'В	прш
РЕША'ВШИ	прш

КРАТКОЕ ПРИЧАСТИЕ стр,но,од,нст,	
РЕША'ЕМЫ	мн
РЕША'ЕМО	ср,ед
РЕША'ЕМА	жр,ед
РЕША'ЕМ	мр,ед

ПРИЧАСТИЕ	
стр,нст,мр,ед	
РЕША'ЕМОГО	од,вн
РЕША'ЕМЫЙ	но,вн
РЕША'ЕМЫЙ	но,од,им
РЕША'ЕМОГО	но,од,рд
РЕША'ЕМОМУ	но,од,дт
РЕША'ЕНЫМ	но,од,тв
РЕША'ЕНОМ	но,од,пр
стр,но,од,нст,жр,ед	
РЕША'ЕМАЯ	им
РЕША'ЕМОЙ	рд
РЕША'ЕМОЙ	дт
РЕША'ЕМУЮ	вн
РЕША'ЕМОЙ	тв
РЕША'ЕМОЮ	тв
РЕША'ЕМОЙ	пр
стр,но,од,нст,ср,ед	
РЕША'ЕМОЕ	им

# Разметка текста

## mystem

---

<b>A</b>	прилагательное
<b>ADV</b>	наречие
<b>ADVPRO</b>	местоименное наречие
<b>ANUM</b>	числительное-прилагательное
<b>APRO</b>	местоимение-прилагательное
<b>COM</b>	часть композита - сложного слова
<b>CONJ</b>	союз
<b>INTJ</b>	междометие
<b>NUM</b>	числительное
<b>PART</b>	частица
<b>PR</b>	предлог
<b>S</b>	существительное
<b>SPRO</b>	местоимение-существительное
<b>V</b>	глагол

# Разметка текста

## НКРЯ

---

<b>A</b>	прилагательное (коричневый, таинственный, морской)
<b>ADV</b>	наречие (сгоряча, очень)
<b>ADVPRO</b>	местоименное наречие (где, вот)
<b>ANUM</b>	числительное-прилагательное (один, седьмой, восьмидесятый)
<b>APRO</b>	местоимение-прилагательное (который, твой)
<b>CONJ</b>	союз (и, чтобы)
<b>INTJ</b>	междометие (увы, батюшки)
<b>NUM</b>	числительное (четыре, десять, много)
<b>PARENTH</b>	вводное слово (кстати, по-моему)
<b>PART</b>	частица (бы, же, пусть)
<b>PR</b>	предлог (под, напротив)
<b>PRAEDIC</b>	предикатив (жаль, хорошо, пора)
<b>PRAEDICPRO</b>	местоимение-предикатив (некого, нечего)
<b>S</b>	существительное (яблоня, лошадь, корпус, вечность)
<b>SPRO</b>	местоимение-существительное (она, что)
<b>V</b>	глагол (пользоваться, обрабатывать)

# Разметка текста Opencorpora

<b>NOUN</b>	СУЩ	имя существительное
<b>ADJF</b>	ПРИЛ	имя прилагательное (полное)
<b>ADJS</b>	КР_ПРИЛ	имя прилагательное (краткое)
<b>COMP</b>	КОМП	компаратив
<b>VERB</b>	ГЛ	глагол (личная форма)
<b>INFN</b>	ИНФ	глагол (инфинитив)
<b>PRTF</b>	ПРИЧ	причастие (полное)
<b>PRTS</b>	КР_ПРИЧ	причастие (краткое)
<b>GRND</b>	ДЕЕПР	деепричастие
<b>NUMR</b>	ЧИСЛ	числительное
<b>ADVB</b>	Н	наречие
<b>NPRO</b>	МС	местоимение-существительное
<b>PRED</b>	ПРЕДК	предикатив
<b>PREP</b>	ПР	предлог
<b>CONJ</b>	СОЮЗ	союз
<b>PRCL</b>	ЧАСТ	частица
<b>INTJ</b>	МЕЖД	междометие

# Универсальная разметка проекта Universal Dependencies CoNLL-U Format

---

<b>ADJ</b>	adjective	прилагательное
<b>ADP</b>	adposition	предлог
<b>ADV</b>	adverb	наречие
<b>AUX</b>	auxiliary	вспомогательный
<b>CCONJ</b>	coordinating conjunction	соединительный союз
<b>DET</b>	determiner	
<b>INTJ</b>	interjection	междометие
<b>NOUN</b>	noun	существительное
<b>NUM</b>	numeral	числительное
<b>PART</b>	particle	частица
<b>PRON</b>	pronoun	местоимение
<b>PROPN</b>	proper noun	имя собственное
<b>PUNCT</b>	punctuation	пунктуация
<b>SCONJ</b>	subordinating conjunction	подчинительный союз
<b>SYM</b>	symbol	символ
<b>VERB</b>	verb	глагол
<b>X</b>	other	другое

# Определение части речи

---

методы:

- правила
- методы классификации (например, дерево решений)
- вероятностные методы
- нейронные сети

# Stanford Log-linear Part-Of-Speech Tagger

It is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife.

However little known the feelings or views of such a man may be on his first entering a neighbourhood, this truth is so well fixed in the minds of the surrounding families, that he is considered the rightful property of some one or other of their daughters.

1	PRP	VBZ	DT	NN	RB	VBD	,	IN	DT	JJ	NN	IN	NN	IN	DT	JJ	NN	,	MD	VB	IN	VB	IN	DT	NN	.
	It	is	a	truth	universally	acknowledged	,	that	a	single	man	in	possession	of	a	good	fortune	,	must	be	in	want	of	a	wife	.
2	RB	RB	JJ	DT	NNS	CC	NNS	IN	POT	DT	NN	MD	VB	IN	PRPS	JJ	VBG	DT	NN	,	DT	NN	VBZ	RB		
	However	little	known	the	feelings	or	views	of	such	a	man	may	be	on	his	first	entering	a	neighbourhood	,	this	truth	is	so		
	RB	VBN	IN	DT	NNS	IN	DT	VBG	NNS	.																
	well	fixed	in	the	minds	of	the	surrounding	families	,																
	IN	PRP	VBZ	VBN	DT	JJ	NN	IN	DT	CD	CC	JJ	IN	PRPS	NNS	.										
	that	he	is	considered	the	rightful	property	of	some	one	or	other	of	their	daughters	.										



# Омонимия

---

Омонимия – совпадение по форме двух разных по значению слов

**Лексическая** омонимия: **ключ** – источник воды и объект для открытия замка

**Морфологическая** омонимия: словоформа **круг** соответствует именительному и винительному падежам

**Лексико-морфологическая** омонимия: **стих** – глагол в единственном числе мужского рода (начальная форма – стихать) и существительное в единственном числе, именительном падеже (начальная форма – стих)

# Задание

---

Выполнить морфологический разбор слов:

- Черепах
- Печь
- Стоит

# Морфологический анализ

---

Словарь

Стемминг

Лемматизация на основе правил

Статистические модели

# <http://odict.ru/> - Открытый грамматический словарь русского языка

---

бадминтон,м,бадминтона,бадминтону,бадминтон,бадминтоном,бадминтоне,бадминтоне,бадминтон		
бадминтонист,мо,бадминтониста,бадминтонисту,бадминтониста,бадминтонистом,бадминтонисте,ба		
бадьа,ж,бадьи,бадье,бадью,бадьёй,бадье,бадье,бадьи,бадей,бадьям,бадьи,бадьями,бадьях,бадьях,		
база,ж,базы,базе,базу,базой,базе,базе,базы,баз,базам,базы,базами,базах,базах,базою		
базальт,м,базальта,базальту,базальт,базальтом,базальте,базальте,базальты,базальтов,базальтам,баз		
базальтовый,п,базальтового,базальтовому,базальтовый,базальтовым,базальтовом,базальтовом,баз		
базар,м,базара,базару,базар,базаром,базаре,базаре,базары,базаров,базарам,базары,базарами,базар		
базарить,нсв,базарил,базарила,базарило,базарили,базарю,базаришь,базарит,базарим,базарите,баз		
базарный,п,базарного,базарному,базарный,базарным,базарном,базарном,базарен,базарная,базарно		
базаровщина,ж,базаровщины,базаровщине,базаровщину,базаровщиной,базаровщине,базаровщине,		
базарщина,ж,базарщины,базарщине,базарщину,базарщиной,базарщине,базарщине,базарщины,баз		

# Словарь: основа + окончания

---

бадминтон: бадминтон -а; -у; -; -ом; -е; -е; -ы; -ов; -ам; -ы; -ами; -ах; -ах

# Словарь OpenCorpora

## <http://opencorpora.org/dict.php>

---

```
<?xml version="1.0" encoding="utf8" standalone="yes"?>
<dictionary version="0.8" revision="403605">
  <grammmemes>
    <grammeme parent="">POST</grammeme>
    <grammeme parent="POST">NOUN</grammeme>
    ...
  </grammmemes>
  <lemmata>
    <lemma id="1" rev="402007">
      <l t="абажыр"><g v="NOUN"/><g v="inan"/><g v="masc"/></l>
      <f t="абажыр"><g v="sing"/><g v="nomn"/></f>
      <f t="абажыра"><g v="sing"/><g v="gent"/></f>
      ...
    </lemma>
    ...
  </lemmata>
</dictionary>
```

# Викисловарь — многофункциональный многоязычный словарь и тезаурус <https://ru.wiktionary.org/>

---

## Морфологические и синтаксические свойства

боль-шóй

Прилагательное, качественное, тип склонения по классификации А. Зализняка — 4b.  
Сравнительная степень — [бóльше](#), [побóльше](#), [бóльший](#) (бóльшее, бóльшая).

Корень: **-больш-**; окончание: **-ой**

## Значение

1. [имеющий](#) крупный, значительный [размер](#), превосходящий среднее [значение](#) ♦ Озёр здесь более двухсот двадцати — [больших](#), поменьше и совсем крохотных.

## Синонимы

1. [крупный](#), [значительный](#)

## Антонимы

1. [маленький](#), [мелкий](#), [малый](#)

# Проблемы словаря

---



# Проблемы словаря

---

- Известные слова
- Новые слова
- Объем

# Стемминг

---

В больничном дворе стоит небольшой флигель, окруженный целым лесом репейника, крапивы и дикой конопли.

в больничн двор сто небольш флигел, окружен цел лес репейник,  
крапив и дик конопл

# Стемминг

---

Восстановите предложение:

он устав на нос лошада

# Стемминг

---

Восстановите предложение:

он устав на нос лошад

они уставились на носившуюся лошадь

# Стеммep

---

It is a truth univers~~ally~~ acknowledged~~ed~~, that a singl~~e~~ man in possessi~~on~~ of a good fortun~~e~~, must be in want of a wife.

Howev~~er~~ littl~~e~~ known the feeli~~ngs~~ or view~~s~~ of such a man may be on his first enteri~~ng~~ a neighbourhood, this truth is so well fix~~ed~~ in the mind~~s~~ of the surroundi~~ng~~ famili~~es~~, that he is consider~~ed~~ the right~~ful~~ properti~~y~~ of some one or other of their daughters~~s~~.

# Стеммер

---

Все знают, что молодой человек, располагающий средствами, должен подыскивать себе жену.

Как бы мало ни были известны намерения и взгляды такого человека после того, как он поселился на новом месте, эта истина настолько прочно овладевает умами неподалеку живущих семейств, что на него тут же начинают смотреть как на законную добычу той или другой соседской дочери.

# Стеммер Портера

## Проект Snowball

---

**RV** – часть слова после первой гласной или пустая строка для слова без гласных

**R1** – часть слова после первой согласной, стоящей после первой гласной, или пустая строка, если таких согласных нет

**R2** – часть слова после первой согласной, следующей после первой гласной в части R1

Слово	RV	R1	R2
важность	жность	ность	ть
вагон	гон	он	
вес	с		
да			
в			

# Стеммер Портера

---

## Существительное

- а, ев, ов, ие, ье, е, иями, ями, ами, еи, ии, и, ией, ей, ой, ий, й, иям, ям, ием, ем, ам, ом, о, у, ах, иях, ях, ы, ь, ию, ью, ю, ия, ья, я

## Прилагательное

- ее, ие, ые, ое, ими, ыми, ей, ий, ый, ой, ем, им, ым, ом, его, ого, ему, ому, их, ых, ую, юю, ая, яя, ою, ею

## Причастие

- ем, нн, вш, ющ, щ,
- ивш, ывш, ующ

## Глагол

- ла, на, ете, йте, ли, й, л, ем, н, ло, но, ет, ют, ны, ть, ешь, нно
- ила, ыла, ена, ейте, уйте, ите, или, ыли, ей, уй, ил, ыл, им, ым, ен, ило, ыло, ено, ят, ует, уют, ит, ыт, ены, ить, ыть, ишь, ую, ю

## Герундий

- в, вши, вшись
- ив, ивши, ившись, ыв, ывши, ывшись



# Стеммер Портера (Алгоритм)

---

**Все действия выполняются только для RV.**

1. Поиск окончаний для герундия.  
Если найдено – удалить и перейти на шаг 2, иначе удалить возвратное окончание ("ся", "сь"),  
выполнить поиск окончаний для прилагательного+причастия, глагола, существительного, как только будет найдено одно из окончаний, удалить его и перейти к шагу 2
2. Если слово заканчивается на "и", удалить "и"
3. Поиск и удаление окончаний "ост" и "ость" в R2
4. Удалить повторное "н";  
если слово заканчивается на "ейш" или "ейше" – удалить их,  
а затем удалить удвоение "н";  
если слово заканчивается на "ь" – удалить его

# Задание

---

Найти основы для следующих слов:

- замечать, замечая, замечающий, замечание
- подушка, подушками, подушек
- стала, сталью
- остров, острый, острие, острить
- благодать, благодатью
- память, памяти
- день, дню, дне
- бежит, бегу
- косившуюся, косил, коса

# Задание

Словоформа	Основа	Словоформа	Основа
замечать	замеча	память	памя
замечая	замеч	памяти	памят
замечающий	замеча	день	ден
замечание	замечан	дню	дню
подушка	подушк	дне	дне
подушками	подушк	бежит	беж
подушек	подушек	бегу	бег
стала	стал	косившуюся	кос
сталью	стал	косил	кос
остров	остр	коса	кос
острый	остр		
острие	остр		
острить	остр		
благодать	благода		
благодатью	благодат		

# Словарь Зализняка

---



Андрей Анатольевич Зализняк

(1935 – 2017) – российский лингвист, действительный член (академик) Российской академии наук по Отделению литературы и языка, доктор филологических наук.

# Словарь Зализняка

---

- Словарь слов с индексом словоизменения
- Таблицы словоизменений
- Таблицы ударений

# Словарь слов с индексом словоизменения

---

ста́дия	ж	7а
трагед́ия	ж	7а
комед́ия	ж	7а
трагикомед́ия	ж	7а
кйнокомед́ия	ж	7а
интермед́ия	ж	7а
логопед́ия	ж	7а
энциклопед́ия	ж	7а
гипнопед́ия	ж	7а
ортопед́ия	ж	7а
мид́ия	ж	7а

ре́зус-фа́ктор	м	1а
сельфа́ктор	м	1а
ве́ктор	м	1а
конве́ктор	м	1а
ра́диус-ве́ктор	м	1а
инже́ктор	м	1а
проже́ктор	м,	1а // 1с①
эже́ктор	м	1а
прозе́ктор	мо	1а
ле́ктор	мо	1а
селе́ктор	м	1а
киле́ктор	м	1а

# Обозначения в индексе

---

Часть речи (склонение):

- м, мо, ж, жо, с, со, мо-жо – существительное
- п – прилагательное
- мс, мс-п – местоимение
- прич. – причастие

Латинская буква – схема ударения

- а – постоянное ударение на основе
- b – ударение на окончаниях
- c-f – разные виды подвижного ударения

Цифра – тип основы:

- 1 – с основой на твердый согласный (топор)
- 2 – с основой на мягкий согласный (тюлень)
- 3 – с основой на г, к или х (петух)
- 4 – с основой на ж, ш, ч, щ (галоша, жилище)
- 5 – с основой на ц (конец, девица)
- 6 – с основой на гласный (кроме и) или й (бой, шея)
- 7 – с основой на и (полоний, сложение)
- 8 – с традиционным "3-м склонением" (боль, тетрадь, зыбь)

# Таблицы словоизменений

Стандартные окончания субстантивного склонения (существительные)

		Ед.			Мн.		
		муж.	ср.	жен.	муж.	ср.	жен.
И.		нуль/-ь	-о/безуд. -е	-а/-я	-ы/-и	-а/-я	-ы/-и
			-о/уд. -ё				
Р.		-а/-я		-ы/-и	-ов/-ей	нуль/безуд. -ь	
						нуль/уд. -ей	
Д.		-у/-ю		-е	-ам/-ям		
В.	неод.	=И.	=И.	-у/-ю	=И.		
	одуш.	=Р.			=Р.		
Т.		-ом/безуд. -ем		-ой/безуд. -ей	-ами/-ями		
		-ом/уд. -ём		-ой/уд. -ёй			
П.		-е			-ах/-ях		

словар-ь

рам-а



# Таблицы словоизменений

Стандартные окончания адъективного склонения (прилагательные)

		Ед.			Мн.
		муж.	ср.	жен.	(всех родов)
И.		безуд. -ый/-ий	-ое/-ее	-ая/-яя	-ые/-ие
		уд. -ой/-ий			
Р.		-ого/-его		-ой/-ей	-ых/-их
Д.		-ому/-ему		-ой/-ей	-ым/-им
В.	неод.	=И.	=И.	-ую/-юю	=И.
	одуш.	=Р.			=Р.
Т.		-ым/-им		-ой/-ей	-ыми/-ими
П.		-ом/-ем		-ой/-ей	-ых/-их

словарн -ный

# Таблицы ударений

## Типы 1—7

### Схема ударения а

Индекс		1а	2а	3а	4а	5а	6а	7а
Образцы	$\frac{\text{м}}{\text{мо}}$	заво́д а́ртист	портфе́ль жи́тель	ча́йник бульдóг	марш това́рищ	ме́сяц принц	случа́й геро́й	сценáрий вика́рий
Ед.	И. Р. Д. В. у неод. у одуш. Т. П.	заво́д заво́да заво́ду заво́д (а́ртиста) заво́дом о заво́де	жи́тель жи́теля жи́телю (портфе́ль) жи́теля жи́телем о жи́теле	ча́йник ча́йника ча́йнику ча́йник (бульдóга) ча́йником о ча́йнике	марш ма́рша ма́ршу марш (това́рища) ма́ршем о ма́рше	ме́сяц ме́сяца ме́сяцу ме́сяц (принца) ме́сяцем о ме́сяце	геро́й геро́я геро́ю (случа́й) геро́я геро́ем о геро́е	сценáрий сценáрия сценáрию сценáрий (вика́рия) сценáрием о сценáрии
Мн.	И. Р. Д. В. у неод. у одуш. Т. П.	заво́ды заво́дов заво́дам заво́ды (а́ртистов) заво́дами о заво́дах	жи́тели жи́телей жи́телям (портфе́ли) жи́телей жи́телями о жи́телях	ча́йники ча́йников ча́йникам ча́йники (бульдóгов) ча́йниками о ча́йниках	ма́рши ма́ршей ма́ршам ма́рши (това́рищей) ма́ршами о ма́ршах	ме́сяцы ме́сяцев ме́сяцам ме́сяцы (принцев) ме́сяцами о ме́сяцах	геро́и геро́ев геро́ям (случа́и) геро́ев геро́ями о геро́ях	сценáрии сценáриев сценáриям сценáрии (вика́риев) сценáриями о сценáриях

# Таблицы ударений

Схема ударения b

Индекс		1b	2b	3b	4b	5b	6b	7с//7b
Образцы	м	топóр	слова́рь	пиро́г	нож	костре́ц	лиша́й	кий <sup>1</sup>
	мо	бегу́н	кара́сь	рыба́к	бога́ч	кузне́ц	холу́й	—
Ед.	И. Р. Д. В. у неод. у одуш. Т. П.	топóр топора́ топору́ топóр (бегуна́) топорóм о топорé	слова́рь слова́ря слова́рю слова́рь (карася́) слова́рем о слова́ре	рыба́к рыбака́ рыбаку́ (пиро́г) рыбака́ рыбакóм о рыбаке́	бога́ч богача́ богачу́ (нож) богача́ богачóм о богачé	кузне́ц кузнеца́ кузнецу́ (костре́ц) кузнеца́ кузнецо́м о кузнеце́	лиша́й лишайа́ лишайо́ лиша́й (холуя́) лишаём о лишаé	кий кия́//кня́ кью́//кию́ кий — кйём//киём о кйи́//о кйé
Мн.	И. Р. Д. В. у неод. у одуш. Т. П.	топоры́ топоров́ топорáм топоры́ (бегунóв) топорáми о топорáх	слова́ри слова́рей слова́рям слова́ри (карасей́) слова́рями о слова́рях	рыбаки́ рыбако́в рыбакáм (пиро́ги) рыбако́в рыбакáми о рыбака́х	богачи́ богачей́ богачáм (ножи́) богачей́ богачáми о богачáх	кузнецы́ кузнецо́в кузнеца́м (костре́цы) кузнецо́в кузнеца́ми о кузнеца́х	лишай лишаёв лишайáм лишай (холуёв) лишайáми о лишайáх	кий киёв ки́ям кий — кня́ми о кия́х

# Таблицы ударений

Типы 1—7  
Схема ударения а

Индекс		1а	2а	3а	4а	5а	6а	7а
Образцы	ж	ка́рта	неде́ля	кни́га	ту́ча	улица	ста́туя	ли́ния
	жо или мо-жо	коро́ва	боги́ня	соба́ка	касси́рша	волчи́ца	фе́я	фу́рия
Ед.	И. Р. Д., П. В. Т.	ка́рта ка́рты ка́рте ка́рту ка́ртой	неде́ля неде́ли неде́ле неде́лю неде́лей	соба́ка соба́ки соба́ке соба́ку соба́кой	ту́ча ту́чи ту́че ту́чу ту́чей	у́лица у́лицы у́лице у́лицу у́лицей	ста́туя ста́туи ста́туе ста́тую ста́туей	ли́ния ли́нии ли́нии ли́нию ли́нией
Мн.	И. Р. Д. В. у неод. у одуш. Т. П.	ка́рты карт картам ка́рты (коро́в) картами о картах	неде́ли неде́ль неде́лям неде́ли (боги́нь) неде́лями о неде́лях	соба́ки соба́к соба́кам (кни́ги) соба́к соба́ками о соба́ках	ту́чи туч ту́чам ту́чи (касси́рш) ту́чами о ту́чах	у́лицы у́лиц у́лицам у́лицы (волчи́ц) у́лицами об у́лицах	ста́туи ста́туй ста́туям ста́туи (фей) ста́туями о ста́туях	ли́нии ли́ний ли́ниям ли́нии (фу́рнй) ли́ниями о ли́ниях

# mystem

---

<https://yandex.ru/dev/mystem/>

MyStem умеет строить разборы для множества слов, включая гипотетические разборы слов, которых нет в словаре

Первую версию программы написали Илья Сегалович и Виталий Титов

Запуск: `mystem [опции] [входной файл] [выходной файл]`

- Применить контекстное снятие омонимии
- Использовать файл с пользовательским словарём
- Генерировать все возможные гипотезы для несловарных слов

# mystem

---

Эти типы стали есть в цехе

# mystem

---

Эти типы стали есть в цехе

## **Без снятия омонимии**

```
./mystem -e UTF-8 -gni
```

эти{этот}  
типы{тип}  
стали{становиться|сталь}  
есть{быть|есть}  
в{в}  
цехе{цех}

## **Со снятием омонимии**

```
./mystem -e UTF-8 -gnd
```

эти{этот}  
типы{тип}  
стали{становиться}  
есть{есть}  
в{в}  
цехе{цех}

# Неизвестные слова

---

Бармаглот

Варкалось. Хливкие шорьки

Пырялись по наве,

И хрюкотали зелюки,

Как мюмзики в мове.



# mystem

---

варкалось{варкаться?=V,несов,нп=прош,ед,изъяв,сред}

хливкие{хливкий?=A=вин,мн,полн,неод|?=A=им,мн,полн}

шорьки{шорька?=S,жен,неод=вин,мн|?=S,жен,неод=род,ед|?=S,жен,неод=им,мн}

пырялись{пыряться=V,несов,нп=прош,мн,изъяв}

по{по=PR=}

наве{нав=S,имя,муж,од=пр,ед}

и{и=CONJ=}

хрюкотали{хрюкотать?=V,несов,нп=прош,мн,изъяв}

зелюки{зелюк?=S,муж,од=им,мн}

как{как=CONJ=}

мюмзики{мюмзик?=S,имя,муж,од=им,мн}

в{в=PR=}

мове{мов?=S,гео,муж,неод=пр,ед}

# Статистический стеммер – Stemka

