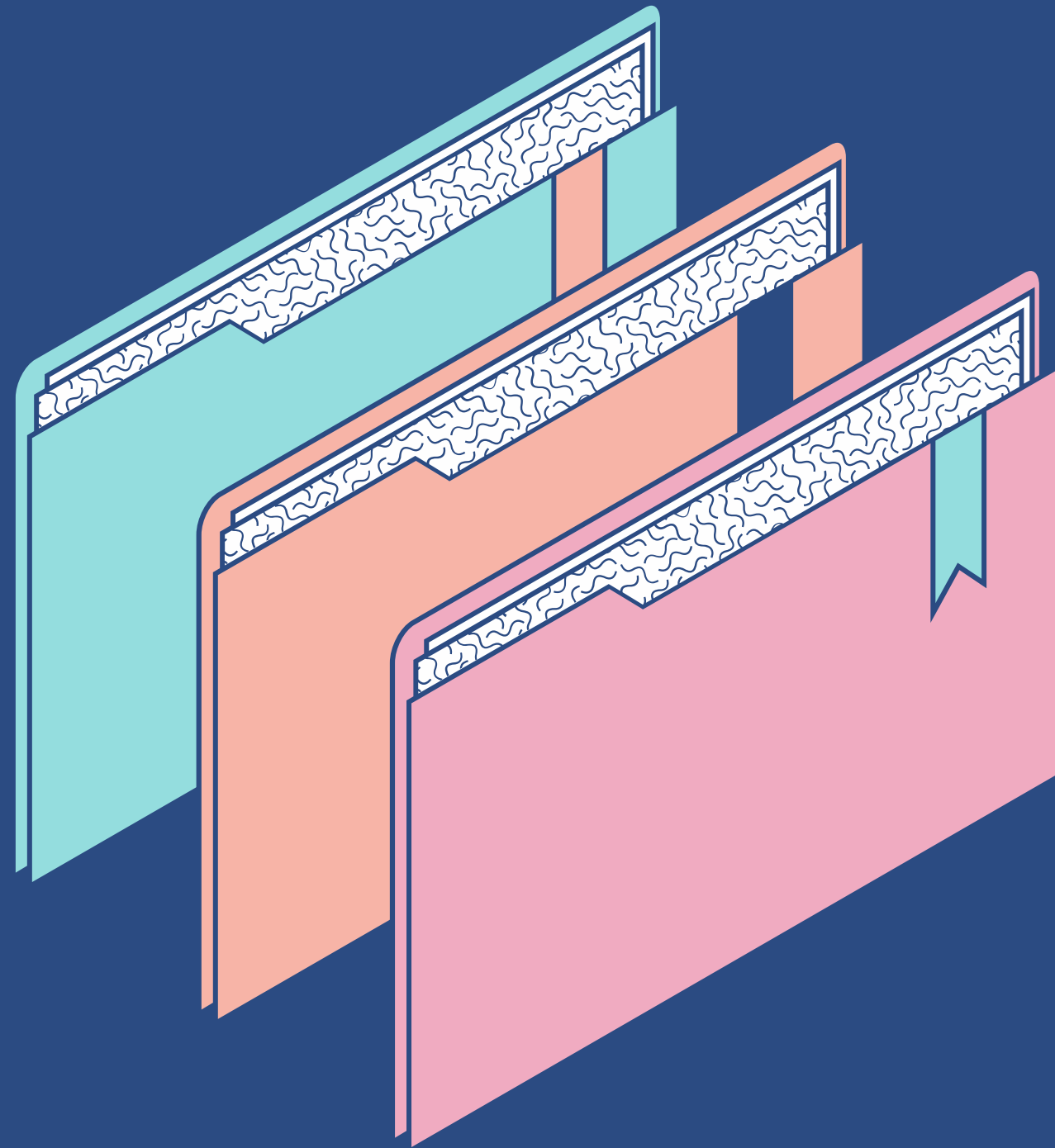




Инструменты Национального корпуса русского языка и их использование в исследованиях лексической сочетаемости прилагательных: диахронический аспект

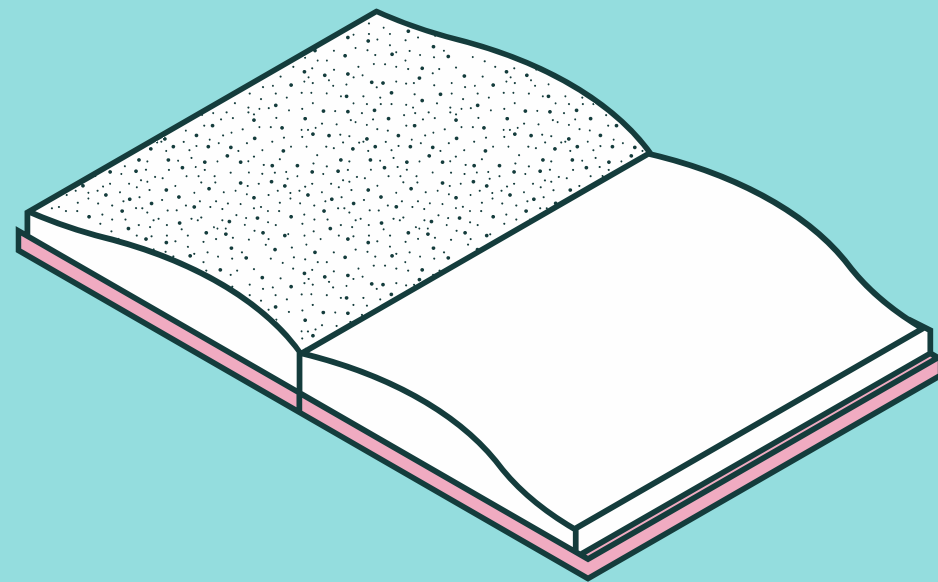
Загребина Ульяна Сергеевна,
преподаватель кафедры русского языка,
теоретической, прикладной лингвистики и
русского языка как иностранного
Удмуртский государственный
университет, Ижевск

Текстовый корпус -



«большой, представленный в электронном виде, структурированный и размеченный, филологически представительный массив языковых данных, предназначенных для решения определенных лингвистических задач» (Захаров В.П.).

**Задача: рассмотрение
первого употребления
языковой единицы**



**Панхронический
корпус**

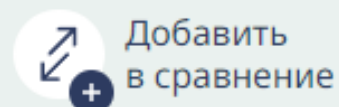
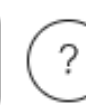

Подкорпус

[Задать](#)

Запрос



[Вернуться к поиску](#)

• 9316 текстов • 33 498 примеров

[злой](#)[Конкорданс](#) [KWIC](#) [График](#)[Скачать](#) 

1. Изборник 1076 г. (перевод X в. [Болгария])

Все примеры — 44

Роуцѣ съгъвен<ѣ> имѣи на събъраникѣ *зълааго* имѣниа свѣта сего· простъртѣ же на пригатию оубогъныхъ: Не стыднѣа въсакомоу съзъданомоу въ образъ божни· главы своа покланати· старѣишааго днѣми почъстити не лѣниса· и покони старость юго потъчиса: Съвѣретьник<ы> своа миромъ  

таковъимъ главы своа покланати: Аште тако можеша творити· не сътворить ти пакости богатъство твоа: нъ такоже и прѣди рекоу вождь ти боудеть црѣствѣ нѣсьноуоумоу· акы ходатан и другъ добръ: А зълѣ прикълломо и





Подкорпус

Задать

Запрос

Вернуться к поиску

• 2766 текстов • 5020 примеров

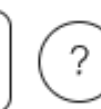
злбный

Добавить
в сравнение

Конкорданс KWIC График



Скачать



1. История Иудейской войны Иосифа Флавия (XI—XIII вв.)



оугодно собѣ. и събравъ дрѣжиноу свою близкы и не встави ниединоу мысли **злбныа** на александра. тѣмъ замысломъ и чарованіемъ. и тако же съ застраши црь, тако мнѣшетъ видѣти александра на(д) собою съ мечемъ. и емъ сваза и. нача мочити дрѣжиноу его. и малчѣщен же оумирахоутъ мнози. не оумѣюще что изреци лжи. дрѣсни же не трѣпѣще мочкъ. ноудими быша болѣстию сългати глахѣ. тако доумашета вѣа брата на него. и лавѣшета оубити его ловы дѣюща. тѣмъ рѣчемъ аще и неправымъ соушимъ. но ноужею строеномъ. вѣроу гатъ црь. и оутѣшашетса имъ да не мнитса безъ правды свазавъ сна. александръ же вида тако не мѣжетъ ѿць его възвратити гнѣва ѿ него. и посреди зла идѣше. и написавъ четверы книгы на врагы. и вѣа ѿць исповѣдагса, тако мыслаше на нь. причѣстници же дѣмѣ быша ѿ воларѣ. и ферора. и салѣми. тако во въ единоу ноць прѣиде к нему на лѣже, и поигра с





Подкорпус

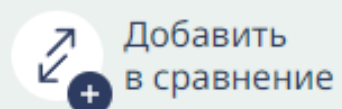
Задать

Запрос

Вернуться к поиску

• 897 текстов • 1194 примера

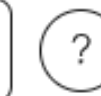
ЗЛОСТНЫЙ

Добавить
в сравнение

Конкорданс KWIC График



Скачать

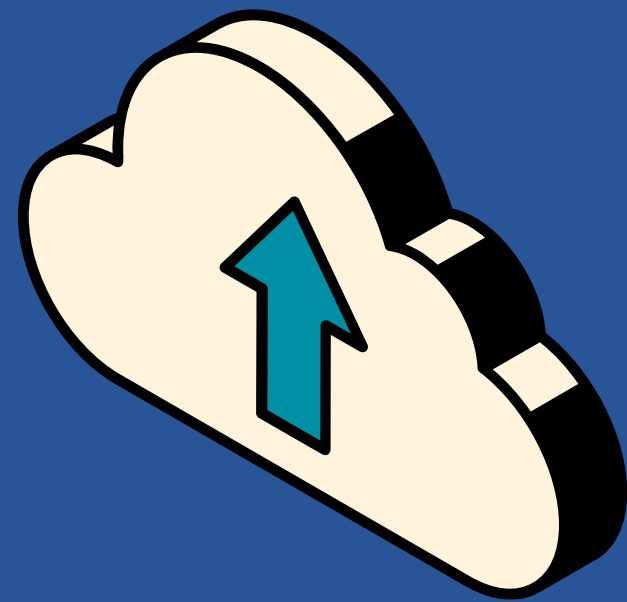
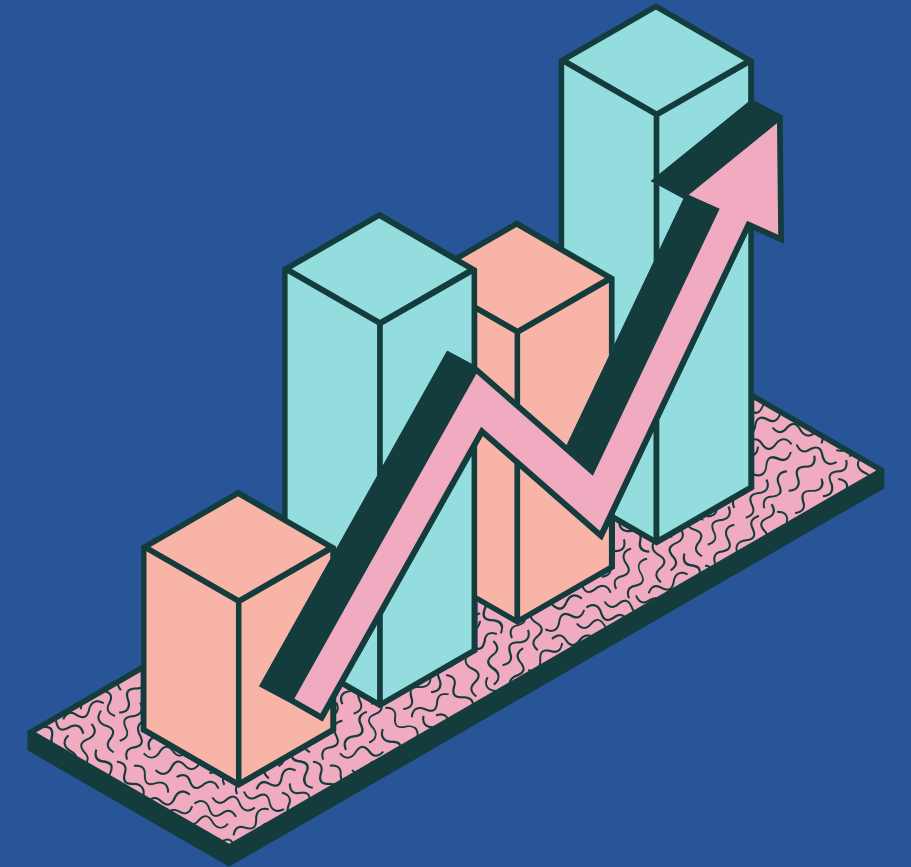


1. В. Н. Татищев. Произвольное и согласное разсуждение и мнение собравшегося шляхетства русского о правлении государственном (1730-1735)

Тайная канцелярия хотя весьма давно и сусче если не при Августе, то при Тиверии, наследнике его, для безопасности монарха вымышлено, и оное, если токмо человеку благочестному и справедливому поручится, нимало не вредно, а *злостные* и нечестивые, недолго тем наслаждаясь, сами исчезают, как всех, так таких по истории преждних и при нас бывших видим.



**Задача: рассмотрение
общих сочетательных
возможностей**



Поиск коллокаций

Запрос

Вернуться к поиску

• 2557 текстов • 252 коллокации

предпочтительные разборы, сов...

зловещий

S, на расстоянии от -2 до 2 от Кл...

Коллокации



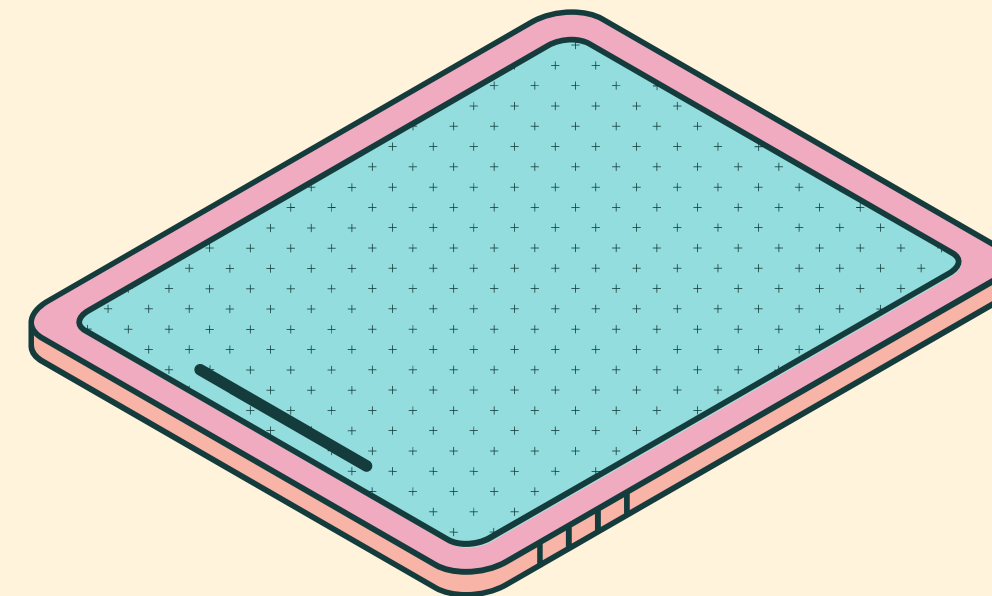
Скачать



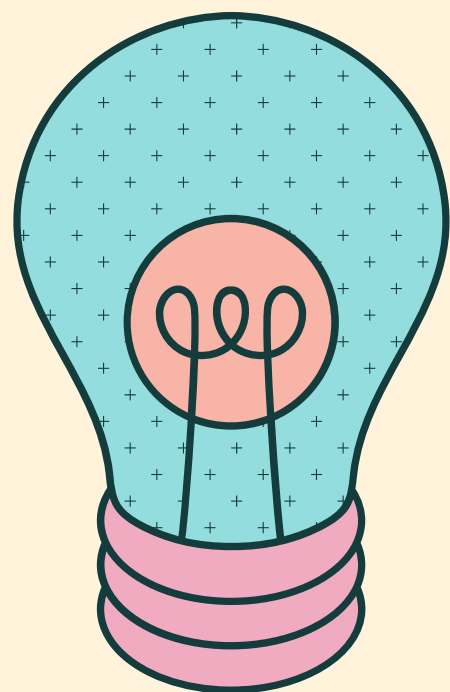
Ключ	Коллокат	Совместная частота	Частота ключа	Частота коллоката	LogDice ▼	Loglikelihood ▼	MI ³ ▼	t-score ▼	Агр. мера ▼	Конкорданс
зловещий	зарев	40	5282	2346	9.44	488.84	14.48	6.32	9.56	Примеры
зловещий	предзнаменование	31	5282	778	9.42	432.12	14.81	5.57	9.08	Примеры
зловещий	тишина	156	5282	31268	9.24	1525.23	15.97	12.45	14.10	Примеры
зловещий	шепот	81	5282	16811	9.08	785.14	14.62	8.97		
зловещий	предсказание	26	5282	3546	8.87	273.60	12.77	5.09		
зловещий	ворон	21	5282	2713	8.75	223.24	12.40	4.57		
зловещий	предчувствие	31	5282	6557	8.75	299.02	12.68	5.55		
зловещий	туча	53	5282	15404	8.73	477.81	13.44	7.25		
зловещий	карканье	13	5282	300	8.63	183.39	13.16	3.60		



**Задача: рассмотрение
конкретных коллокаций**



**Примеры в
конкордансе**



Запрос

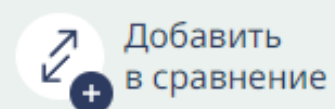
Вернуться к поиску

• 41 текст • 41 пример

предпочтительные разборы, сов...

зловещий

крик, S, на расстоянии от -2 до 2 ...



Добавить
в сравнение

Конкорданс KWIC График Статистика Частотность 2-граммы 3-граммы 4-граммы 5-граммы



Скачать ▾



1. В. В. Мезенцев, К. С. Абиляханов. Чудеса: Популярная энциклопедия. Том 2. Книга 4 (1991) ↩

Вчера ночная птица в самый полдень На площади уселась, оглашая Окрестности своим *зловещим криком*. 📄 ↔

2. Митьки. Папуас из Гондураса (1987) ↩

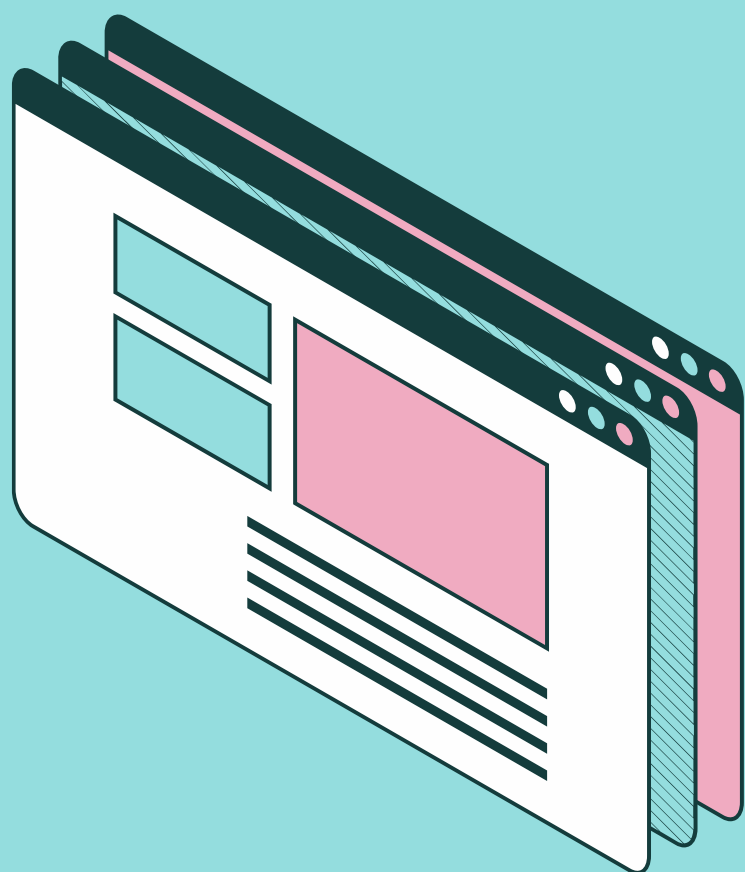
Выждав, бандиты со *зловещими криками* бросились на веселящихся чернокожих. 📄 ↔

3. Е. Климов. М. В. Добужинский // «Зарубежье», 1972 ↩

Зловещий крик птицы, померкнувшее солнце и стремительный поход дружины сплелись с терзаемым бурей деревом в выразительный графический узор. 📄 ↔



**Задача: рассмотрение
конкретных коллокаций
за определенный период**



**Создание
пользовательского
подкорпуса**

Подкорпус

[Изменить](#)[Очистить](#)

• 5697 текстов • 41 276 662 слова

1801-1900

Нехудожественные тексты



Запрос

[Вернуться к поиску](#)

• 192 текста • 14 коллокаций

предпочтительные разборы, сов...

зловещий

S, на расстоянии от -2 до 2 от Кл...

Коллокации

[Скачать](#)

Ключ	Коллокат	Совместная частота	Частота ключа	Частота коллоката	LogDice ▼	Loglikelihood ▼	MI ³ ▼	t-score ▼	Агр. мера ▼	Конкорданс
зловещий	предзнаменование	6	347	179	10.22	87.80	11.87	2.45		
зловещий	предчувствие	6	347	596	9.64	73.23	10.67	2.45		
зловещий	блеск	8	347	1947	9.03	83.32	10.35	2.82		
зловещий	тишина	9	347	2374	8.98	92.31	10.51	2.99		
зловещий	крик	9	347	3537	8.63	85.14	10.11	2.99		
зловещий	признак	12	347	4940	8.61	112.52	10.64	3.45		
зловещий	слух	9	347	5413	8.23	77.51	9.68	2.98		



Подкорпус

[Изменить](#)[Очистить](#)

• 72 356 текстов • 116 888 805 слов

1901-2000



Нехудожественные тексты

Запрос

[Вернуться к поиску](#)

• 782 текста • 79 коллокаций

предпочтительные разборы, сов...

зловещий

S, на расстоянии от -2 до 2 от Кл...

Коллокации ?



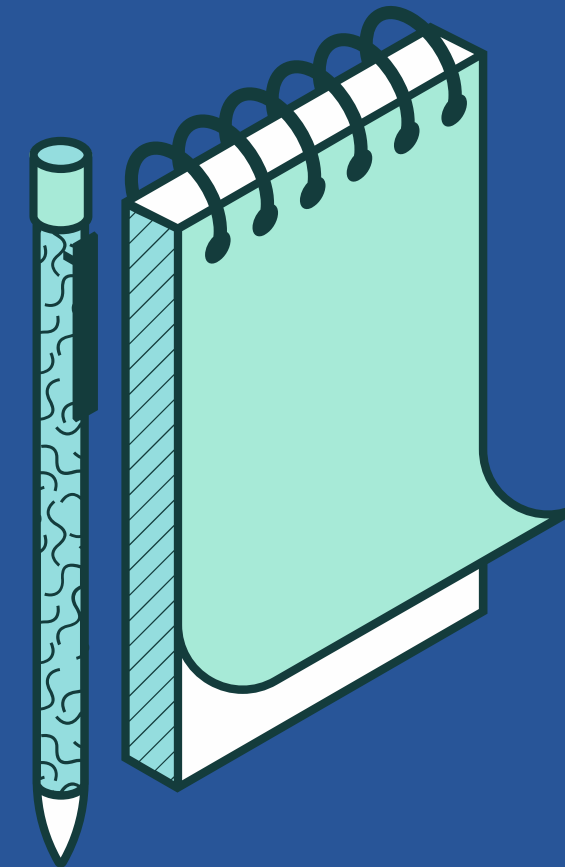
Скачать ▼



Ключ	Коллокат	Совместная частота	Частота ключа	Частота коллоката	LogDice ▼	Loglikelihood ▼	MI ³ ▼	t-score ▼	Агр. мера ▼	Конкорданс
зловещий	шепот	26	1508	1828	9.84	313.14	13.52	5.09		
зловещий	зарев	15	1508	468	9.81	205.18	13.23	3.87		
зловещий	тишина	45	1508	6790	9.48	473.57	13.85	6.70		
зловещий	туча	26	1508	3826	9.37	274.59	12.78	5.09		
зловещий	затишье	11	1508	850	9.33	130.28	11.71	3.31		
зловещий	призрак	14	1508	2237	9.10	145.44	11.46	3.73		



**Задача: выявление
семантико-
грамматических
особенностей**



**Лексико-
грамматический
ПОИСК**

Поиск точных форм ? ▼

Поиск коллокаций ? ▼

Лексико-грамматический поиск ? ^

По умолчанию поиск ведется по предпочтительным разборам, слова могут совпадать.

Поиск учитывает:

- все разборы
- предпочтительные разборы

При поиске:

- совпадения слов исключаются
- слова могут совпадать

Слово 1 × +

Слово 2 × +

Лемма

зловещий

Лемма

Словоформа

Словоформа

Грамм. признаки

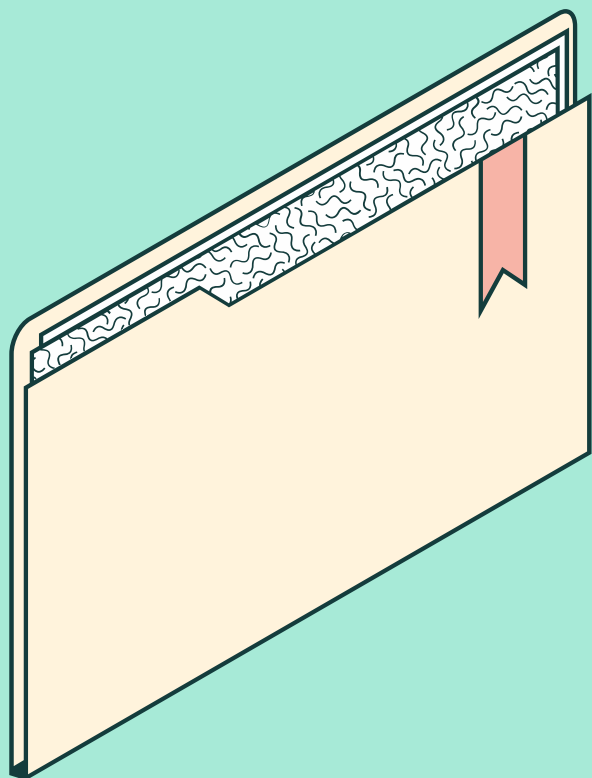
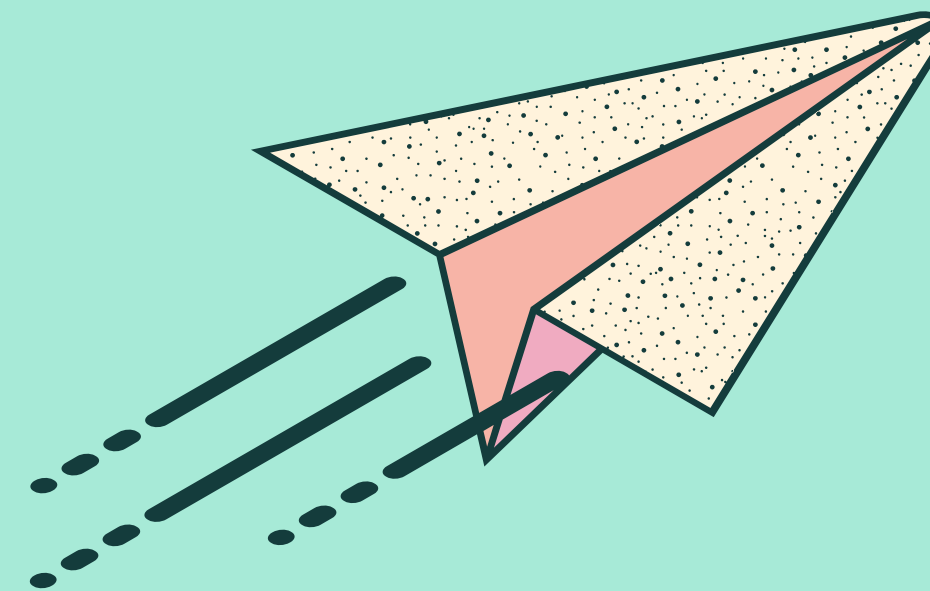
[выбрать](#)

Грамм. признаки

[выбрать](#)

(S) & (anim)

**Задача: комплексный
анализ языковой
единицы**



Портрет слова

Скетчи ?

злой Прилагательное

Определяемые слова

1. умысел	9,98
2. дух	9,87
3. шутка	8,93
4. воля	8,79
5. гений	8,26
6. рок	8,22
7. язык	7,81
8. насмешка	7,78
9. собака	7,73
10. ирония	7,67

Словосочетания с наречием

1. страшно	6,26
2. ужасно	5,79
3. отвратительно	5,28
4. подчас	5,21
5. донельзя	5,2
6. больно	5,14
7. невероятно	5,05
8. чертовски	5,04
9. необыкновенно	4,85
10. очень	4,61

Сочиненные прилагательные

1. добрый	8,85
2. злой	8,35
3. жестокий	8,28
4. коварный	8,08
5. <u>мстительный</u>	7,92
6. завистливый	7,92
7. глупый	7,88
8. насмешливый	7,4
9. презлый	7,37
10. несправедливый	7,34

[Показать все коллокации](#)



О слове ?

злой

Лемма	злой (см. словарь)
Грамматика	прилагательное
Семантика	качественные, свойство человека

Частотность слова ?



Морфемный разбор β ? [Оценить](#)

злой

Однокоренные слова β ? [Оценить](#)

сгенерировано НейроКРЯ

злой
зло
злоба
злодей
злость
злбный
злиться
назло
разозлиться
озлбить
озлобление

Похожие слова ?

сгенерировано НейроКРЯ

зл**о**бный мстительный завистливый
лукавый сварливый
озлобленный **злой** презлый
свирепый презлой
злопамятный

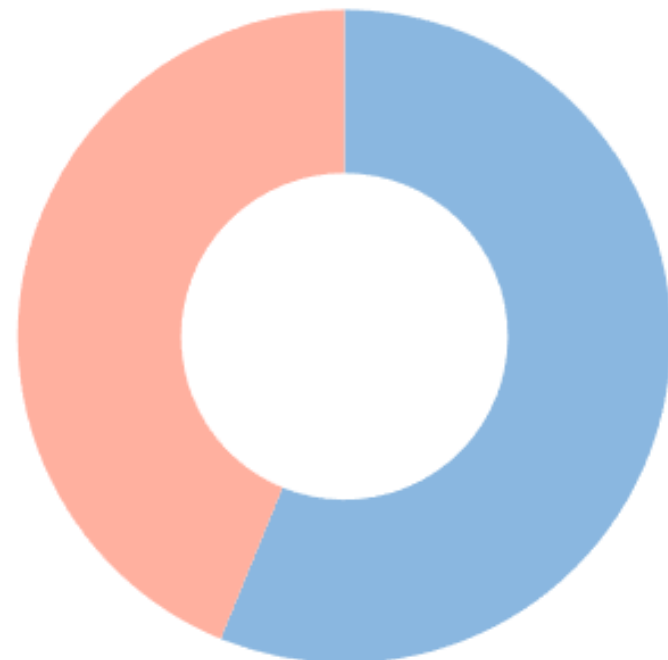
Статистика текстов ?

Метаатрибут

Вид текста ▾

Показывать

слова ▾



художественная (56.18%)

нехудожественная (43.82%)

Похожие слова ?

сгенерировано НейроКРЯ

озлобленный мстительный
раздраженный злобный злой ехидный
злорадный
гневный остервенелый завистливый свирепый

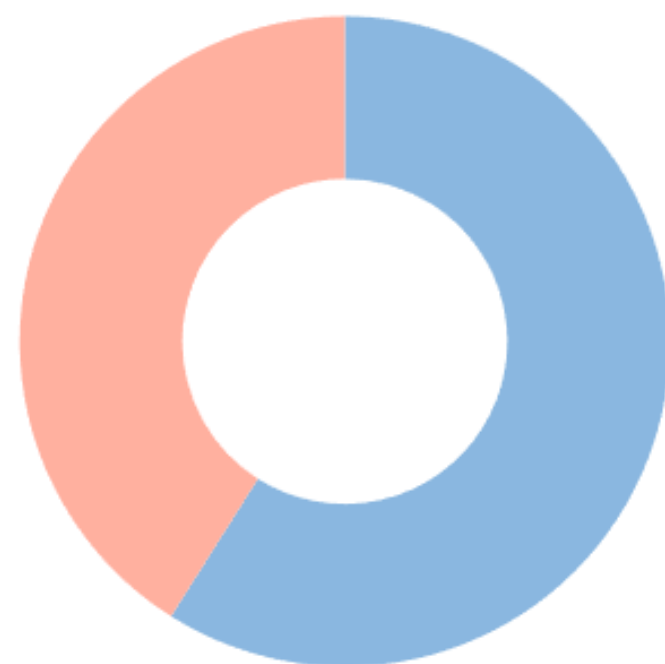
Статистика текстов ?

Метаатрибут

Вид текста ▾

Показывать

слова ▾



художественная (58.98%)

нехудожественная (41.02%)



Проблемы при работе с корпусом

1. Разница в объеме текстов как в разных корпусах, так и в пределах одного;
2. Формирование коллокаций не всегда достоверно, поскольку синтаксические связи могут быть не так учтены;
3. Несмотря на снятую омонимию, при поиске прилагательных нужно вручную отбирать примеры, т.к. краткая форма среднего рода часто всего совпадает с наречием.



Вывод

Таким образом, можно сделать вывод о том, что для проведения лингвистических исследований НКРЯ – очень ценный и удобный ресурс. Он вмещает в себя тексты разных жанров и периодов развития языка, что позволяет на его базе изучить историю определенного слова. Благодаря подробной метаразметке исследователь имеет возможность выбирать контексты с изучаемой лексической единицей, задавая разные параметры (жанр, дата создания, тематика текстов и др.). Несмотря на то что при использовании НКРЯ исследователь сталкивается с рядом проблем, а функционал большинства корпусов ограничен, этот ресурс является одним из наиболее значимых лингвистических проектов в России.