## Горизонтальная линия

Национальный корпус русского языка



Лев Успенский (1900-1978)

Как один человек придумал словарь навыворот:

**медленное чтение с Национальным корпусом русского языка**

методический кейс №11

по развитию практик обучения русскому языку и языкам народов России в контексте межкультурной коммуникации

# Аннотация

В этом кейсе встретятся два уникальных культурных феномена: замечательная книга Льва Успенского «Слово о словах» (Лениздат, 1962) и Национальный корпус русского языка <http://www.ruscorpora.ru/>. Книжка Успенского оказала огромное влияние на воспитание нескольких поколений наших соотечественников, которые научились задавать себе вопросы о том, как живёт и развивается русский язык, осознали свою ответственность за его развитие. Поэтому Успенского будем читать медленно, внимательно. Будем отвечать на вопросы автора и проверять его версии. Может ли обычный современный школьник вступить в равноправный диалог с известным лингвистом? Может, ведь у современного школьника в распоряжении (кроме интереса к русскому языку) – уникальный ресурс. Национальный корпус русского языка (НКРЯ). В этом проекте учащиеся научатся пользоваться поисковой системой Корпуса для решения исследовательских задач.

Корпус — это электронная информационно-справочная система, основанная на собрании текстов на русском языке. Корпус содержит по возможности все типы письменных и устных текстов, представленные в данном языке (художественные разных жанров, публицистические, учебные, научные, деловые, разговорные, диалектные и т.п.). И, самое важное, — сделанная лингвистами разметка и система поиска, которые позволяют организовать поиск слова, его формы, сочетаний слов, контекста.

В этом кейсе мы обратимся к истории замысла обратного словаря русского языка. Вслед за Л.Успенским разберёмся, что это за штука, для решения каких задач она может пригодиться современному школьнику. Научимся работать с поисковой системой НКРЯ и получим в свои руки уникальный инструмент, позволяющий автоматизировать работу, которую лингвисты традиционно делали вручную.

Основная форма работы в этом проекте – приём «чтение с остановками» фрагментов книги Л.Успенского «Слово о словах» и микроисследования с использованием Национального корпуса русского языка.

Кейс рекомендован для внеурочной работы по русскому языку, дополнительного образования учащихся основной и старшей школы.

# Целевые установки методического кейса

1. Выращивание языковой картины мира.
2. Формирование готовности переносить имеющиеся знания о языке на межкультурные, исторические контексты и использовать для моделирования поведения в межкультурной ситуации.
3. Развитие умений с Национальным корпусом русского языка для решения исследовательских и практических задач.

# Планируемые результаты

## в области развития языковой культуры:

## участники проекта расширяют опыт ценностного отношения к языку как развивающемуся и формирующемуся с их личным участием явлению,

## знают и понимают, как устроен обратный словарь;

* умеют пользоваться поисковой системой НКРЯ;

## в области культуры чтения:

участники развивают культуру осмысленного (смыслового) чтения;

## в области развития мышления учащихся:

## участники проекта расширяют опыт проведения лингвистических микроисследований с использованием НКРЯ;

## в области практик обучения русскому языку и языкам народов России:

учителя русского языка получают методически оснащенный модуль программы дополнительного образования с использованием интерактивных методов, основанных на межличностной коммуникации.

# Целевая аудитория

Кейс рекомендован для внеурочной работы по русскому языку, дополнительного образования учащихся основной и старшей школы.

# Набор методических материалов и рекомендации по их применению

* Сценарий практикума по работе с НКРЯ;
* Материалы для изучения;
* Инструкция по работе с НКРЯ;
* Инструкции по выполнению микроисследований с использованием НКРЯ в pptx

**Количество и возраст участников:** дискуссия разработана для детей школьного возраста, от 12 лет и старше.

**Время для практикума:** зависит от возраста участников, рассчитывайте на 45-60 минут.

**Оборудование:** компьютеры с выходом в интернет на сайт НКРЯ <http://www.ruscorpora.ru/> для индивидуальной или парной работы

**Сценарий** **практикума**

* + - 1. **Постановка проблемы, её представление классу**

**Учитель.** Существует популярное анекдотическое утверждение, будто в русском языке есть лишь три слова, оканчивающиеся на «-со»: «мясо», «просо» и «колесо». Как можно доказать, так это или не так?

Методический комментарий.

В ходе фронтальной работы определяем, что для доказательства или опровержения этого утверждения нам необходимо: 1) обратиться большому количеству устных и письменных текстов или словарю русского языка. 2) В этом корпусе (словаре) надо найти все слова, оканчивающиеся на –со. 3) составить список слов и посмотреть, есть ли в нём слова, кроме трёх, названных в условии. Таким образом, чтобы доказать справедливость или ложность данного утверждения, нам нужен весь список слов русского языка и удобная система поиска.

**Учитель.** Замечательный учёный-лингвист Лев Успенский так же пытался в своё время решить эту задачку. И вот что он придумал. Читаем его книжку «Слово о словах».

Методический комментарий.

Ученики читают по очереди с остановками. Суть этого приёма заключается в том, что учитель заранее делит текст на смысловые части. В ходе каждой остановки в чтении читатели проверяют (сверяют) понимание и делают прогноз, что в тексте будет дальше. Остановки в чтении отмечены словом «Стоп».

**Текст для чтения**

До сих пор речь шла о словарях, которые так или иначе, в большом или малом числе, но уже существуют. Однако мне хочется коснуться одного совсем небывалого предложения (в других областях знания можно было бы употребить слово «изобретение»), с которым мне пришлось немало повозиться. Суть его можно передать, условно назвав его «словарем навыворот».

**СТОП.** ***Что это может значить?***

Когда слова того или другого языка попадают в словарь и располагаются в нем в каком-либо определенном порядке (обычно в алфавитном), языковед сразу же получает в руки возможность решать многие научные вопросы и задачи, абсолютно не разрешимые до этого.

Приведу простейший пример. Ученого может заинтересовать, скажем, сравнительная употребительность в русском языке приставок «пре-» и «пере-». Действительно, какие слова более свойственны нашей речи, – такие, как «переплет», «переправить», или же такие, как «преткновение», «преступник», «предложить»? Вопрос вполне осмысленный: первая из этих приставок чисто русского происхождения, вторая заимствована из древнеславянского языка. А я, допустим, хочу оценить степень влияния этого языка на современный русский.

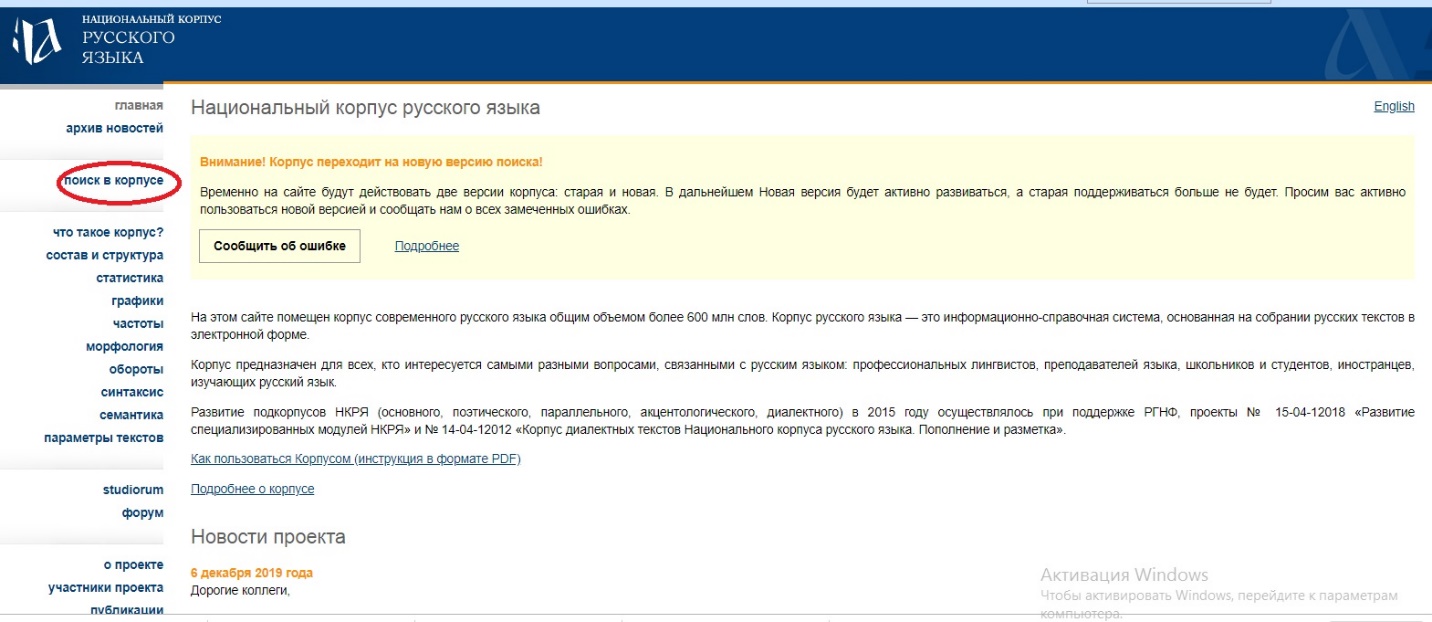
Сделать это легко, если есть словарь, в котором слова расположены по алфавиту; незачем объяснять, как должна идти работа.

**СТОП*. Давайте всё-таки объясним, как должна идти работа. Как можно оценить степень влияния древнерусского языка на современный русский язык? Сделайте предположение, каких слов в словаре будет больше: с приставкой при- или пре-? Что вам понадобится для проведения этого исследования?***

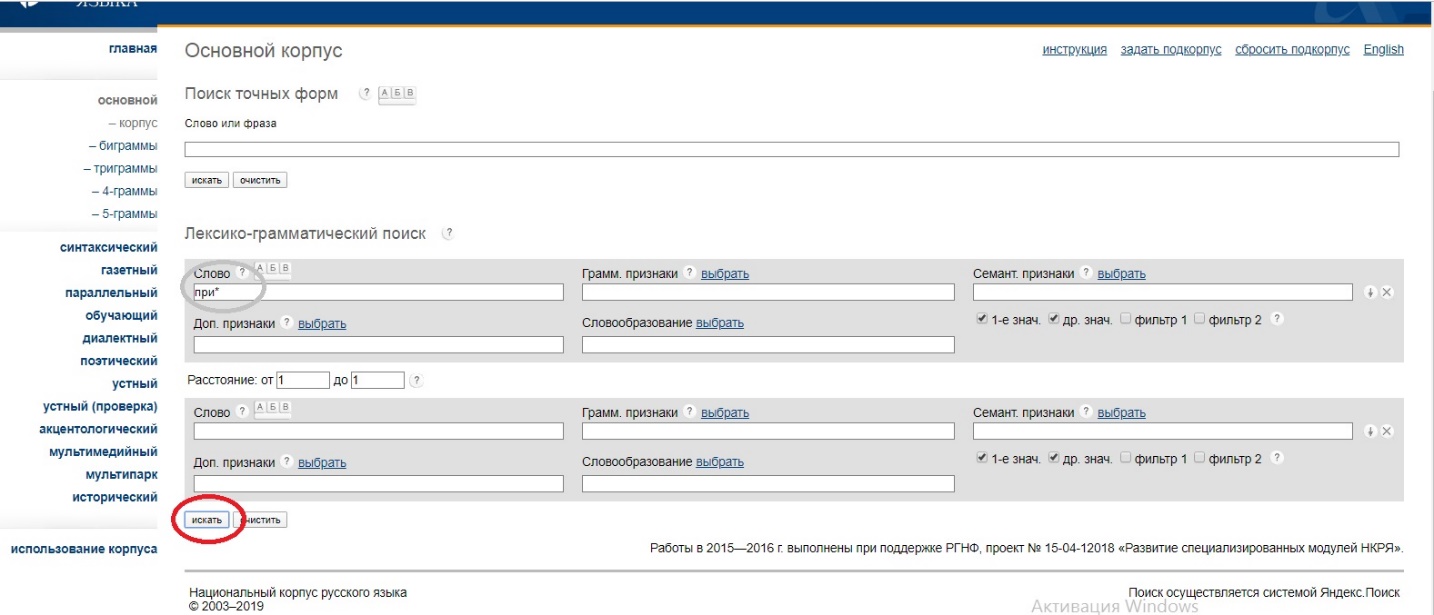
**Учитель.** Сегодня в нашей работе нам поможет уникальный электронный ресурс Национальный корпус русского языка (НКРЯ). Посмотрим, сколько слов с приставкой при- и с приставкой пре- выдаст нам поисковая система.

**Инструкция №1 по работе с НКРЯ.**

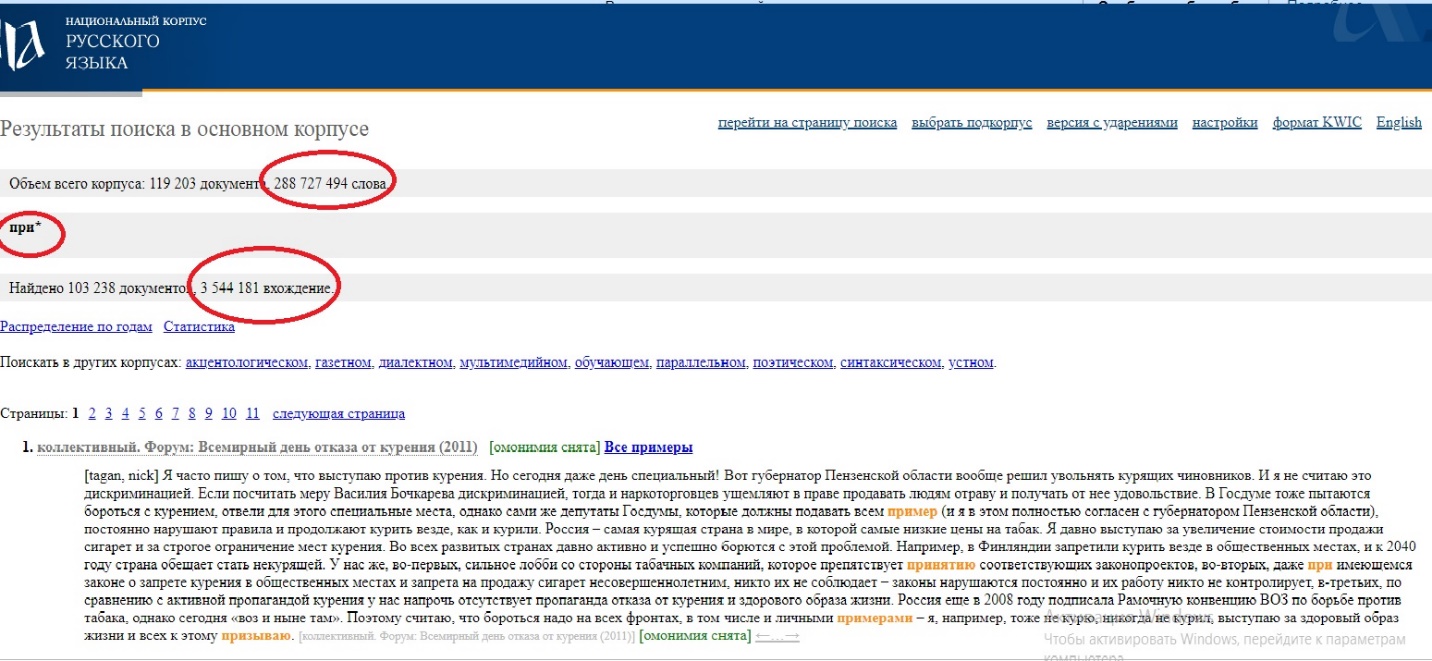
Зайдем на главную страницу ресурса. В меню слева найдем «Поиск в корпусе».

****

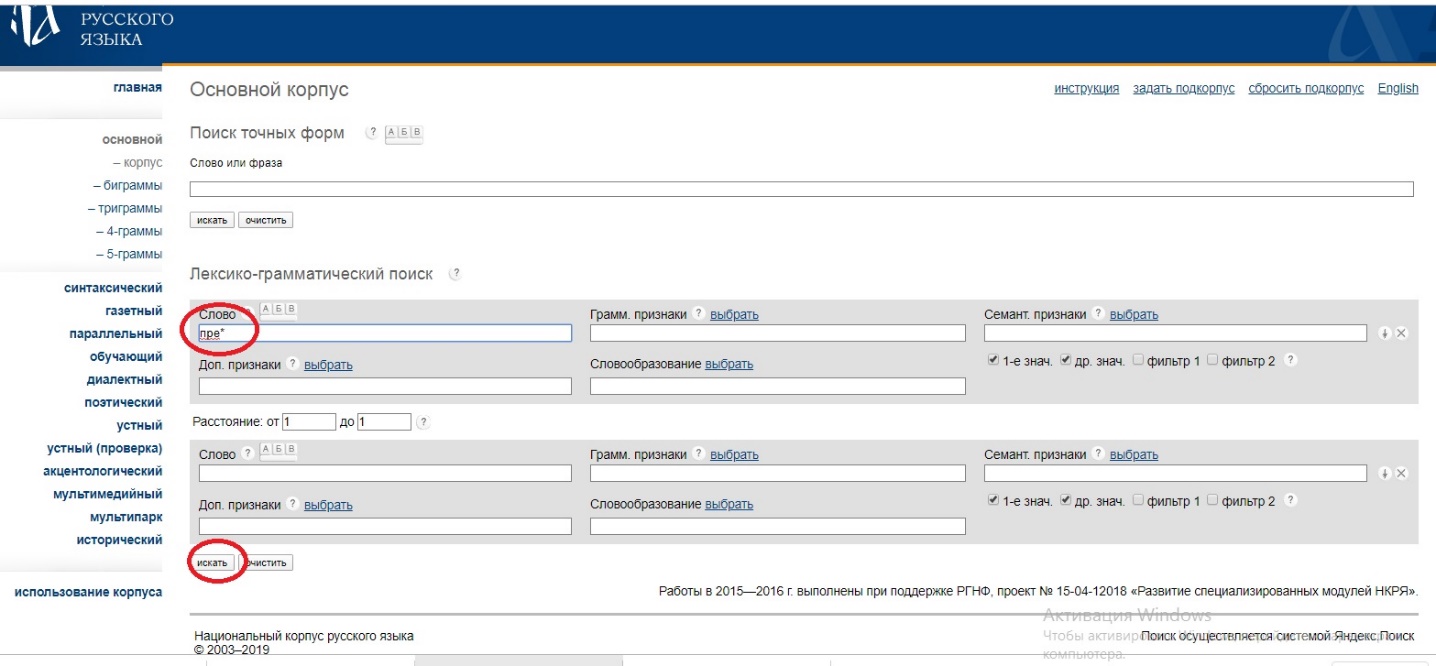
Выберем лексико-грамматический поиск. Введём при со звёздочкой\*. \* означает, что что часть слова после при- может быть любая. И жмём на кнопку искать».



Смотрим результаты. 3544181 вхождение.



Поищем слова с приставкой пре-



1. Посмотрим результаты поиска слов с приставкой пре-. 1812056 вхождений.



Методический комментарий

Сравним количество вхождений в корпус слов с приставками при- и пре-Мы видим, что степень влияния древнерусского языка велика. И всё же приставки при- используются в русском языке в два раза чаще.

**Чтение с остановками.**

Но представьте себе, что я захочу узнать что-либо в этом же духе и роде, только связанное не с началами, а с окончаниями слов. Ну, положим, какое значение имеет в русском языке суффикс «-л-» в словах среднего рода, вроде «зерка-л-о»? Или каких суффиксов «-чик» в нем больше: тех ли, которые образуют слова, означающие профессию, род занятий (вроде «лет-чик», «рез-чик», «пулемет-чик»), или образующих уменьшительные имена («маль-чик», «паль-чик» и пр.). Мне может понадобиться и сведение, какой суффикс более употребителен: «-чик» или «-ник» (а может быть, «-тель») в тех же словах, означающих род занятий («гранат-о-мет-чик» или «подрыв-ник»?).

Очень легко понять, что разрешить эти вопросы куда труднее, чем в случае с приставками: слова, оканчивающиеся на «-чик», «-ник» и «-ло», разбросаны по самым разным буквам алфавита: поди-ка собери их все! Занятие долгое и неточное: как поручишься за то, что выбрал их до последнего?

Короче, скажу так. Существует популярное анекдотическое утверждение, будто в русском языке есть лишь три слова, оканчивающиеся на «-со»: «мясо», «просо» и «колесо». Оставив в стороне несерьезность этого примера, попробуйте доказать, справедлив ли он или, наоборот, нелеп. Так это или не так?

Чтобы выяснить это, придется произвести работу, во много раз большую, чем при попытке найти в справочной книжке фамилию телефонного абонента по известному номеру его телефона. А все потому, что словари составляются в алфавите **начал** слов, а не их **концов** .

**СТОП*. Как вы представляете себе этот словарь не начал слов, а их концов?***

**Чтение с остановками.**

В самом деле, если бы существовал словарь, где на букву «о» шли бы слова в таком порядке:

*покрывало*

*зубило*

*рыло*

*зеркало*

*било*

*крыло*

*сало*

*мыло*

*одеяло,*

задача определения числа и состава слов среднего рода с суффиксом «-л» была бы уже наполовину разрешена: мы смогли бы окинуть их все одним взглядом, легко рассортировать на группы, сравнить между собою их значение и объяснить, почему в каждом данном случае в образовании этих слов принял участие суффикс «-л». В слове «покрывало» его значение вполне понятно: это «то, чем покрывают». Понятно оно и у «мыла»: «мыло» – «то, чем моют», и у «рыла» – того, чем «роют». А что вы скажете насчет «крыла» или «масла»?[[1]](#footnote-1) Представьте себе: и тут то же! «Шил-о» – то, чем сшивают, соединяют швом; «мас-л-о» – то, чем «мастят», «умащают» какую-нибудь поверхность…

Но в наш алфавит попадут и слова совсем другого рода: слово «стило» – знатный иностранец; «кресло» и т. п. Надо будет (и окажется очень удобным) проверить их все и разобраться во всех их свойствах. Так же легко можно будет разрешать и многие другие задачи, **когда будет составлен словарь навыворот**.

**СТОП*. Но разве так уж трудно его составить?***

**Чтение с остановками.**

Я сказал бы, что дело это не столько трудное, сколько долгое и кропотливое. Надо «расписать» на отдельные карточки любой достаточно полный словарь современного русского языка и разместить эти карточки в новом порядке, в порядке алфавита не начал, а окончаний слов. Тогда слово «лампочка» найдет себе место не между словами «ламентация» и «лампада», где оно помещается, скажем, в словаре Ушакова, а где-то возле «бабочка», «мордочка» и «тапочка», то есть в совершенно новом окружении. Так сказать: «акчобаб» «акчодром», «акчопмал», «акчопат»… Забавный словарь? Забавный, но, по моему глубокому убеждению, и весьма полезный. Если бы кто-либо взял на себя огромный груд по его составлению (не забудьте, что в словаре Д. Н. Ушакова 87 000 слов), он заслужил бы по окончании своей работы благодарность всех языковедов, уважение и даже восторг. А до окончания? А вот тут не ручаюсь… Вероятно, не обошлось бы без недоверия, пожимания плечами и даже иронических усмешек.

Но настоящие энтузиасты ничего этого не боятся: посмеивались и над Далем. А как к нему относятся теперь?

Только что прочитанная вами главка была впервые опубликована в 1956 году. И вот что случилось с того недавнего времени.

Во-первых, я получил великое множество предложений: десятки энтузиастов и юного, и среднего, и совсем пожилого возраста выразили желание посвятить свои досуги составлению и словников (то есть полных коллекций слов) для словарей различных писателей, и самого «Зеркального словаря» русского языка…

В 1958 году уже вышел в свет первый такой словарь русского языка. Он появился через два года после того, как мое предложение было впервые опубликовано в «Слове о словах». К сожалению, издан он оказался в ГДР, под редакцией профессора берлинского университета Г. Бильфельдта, и, естественно, при всех достоинствах содержит некоторые недочеты, которых легко избежал бы составитель русский. Вскоре вслед за первым таким словарем появился второй, тоже зарубежный, несколько большего объема (в словарь Бильфельдта вошло около 80 000 русских слов); приступила к подготовке еще более объемистого и солидного «инверсионного» словаря русского языка и Академия наук СССР.

**СТОП*. Совсем по-другому работают в век информационных технологий. Время карточек ушло в прошлое. Основную техническую работу делает за человека компьютерная программа. Уже в 1974 году был издан «Обратный словарь русского языка», который был подготовлен лингвистами совместно со специалистами вычислительного центра Академии наук СССР.***

**Чтение с остановками.**

Таким образом, надобность в помощи сотрудников-добровольцев внезапно отпала; в этом смысле я оказался плохим пророком, – идея словаря нашла нежданно быстрое признание. Зато хорошим пророком я могу счесть себя в другом отношении: первый же вышедший в свет словарь привлек всеобщий интерес и заслужил уважение. И, как всегда бывает, выяснилось, что нужен он вовсе не одним только специалистам-лингвистам: он понадобился во множестве других профессий.

**СТОП*. Подумайте, в каких случаях обратный словарь окажется незаменимым помощником.***

**Чтение с остановками.**

Приведу единственный пример: вы корпите над расшифровкой старинной, попорченной временем рукописи, важного документа. Перед вами там и здесь проступающие концы слов, лишенные начал. Вы видите сочетание букв «ерок». Какое слово могло стоять тут? Почти немыслимо решить эту задачу наобум: кто знает, сколько слов в русском языке оканчивается на «ерок» и каковы они? Но у вас в руках словарь «зеркального типа», пусть хоть несовершенный бильфельдтовский. Вы находите в нем колонку слов, оканчивающихся на это самое «ерок». Вы видите: их всего 8, от «зверок» до «вечерок». А видя их сразу все, вам по смыслу всего контекста, всего документа уже вовсе не трудно подобрать нужное и возможное слово: конечно, не «ветерок» и не «зверок», а – «недомерок». Хотел бы я видеть, сколько сил и времени потратите вы, если придется вам изыскивать это единственное слово не из восьми, а из двухсот тысяч возможных! Разве вы подозревали, что таких слов действительно всего лишь восемь? Я и сам об этом представления не имел.

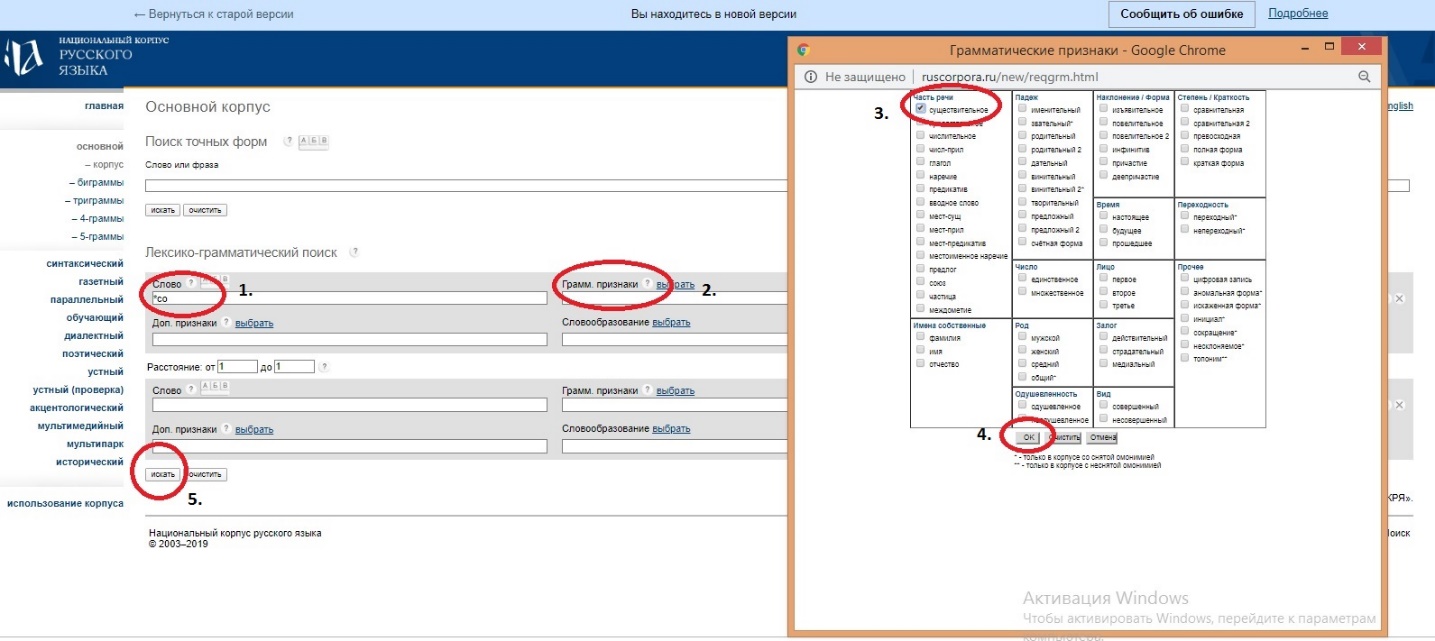
С этой позиции и «нелепый» вопрос о «просе, мясе, колесе» (есть и другая такая же старая задача: найти третье слово, оканчивающееся на «-зо», к «пузо» и «железо»; когда-то ходила легенда, будто Академия наук учредила даже крупную премию тому, кто выполнит это задание) становится не таким уже нелепым. Теперь оба эти вопроса решаются очень просто: вот список русских слов, оканчивающихся на «-со»: со (в смысле «со мною»), кюрасо, колесо, плёсо, накосо, просо, серсо, инкассо, лассо, мясо.

**СТОП*. Сегодня поиск слов, оканчивающихся на –со займёт у нас буквально несколько секунд. Правда, просмотр результатов потребует времени.***

**Инструкция №2 по работе с НКРЯ.**

Выберем лексико-грамматический поиск;

* введём \****со***. \* означает, что что часть слова до ***- со*** может быть любая;
* определим грамматические признаки: ограничим поиск именами существительными;
* в раскрывшемся окне отметим, что будем искать только существительные;
* не забудем сохранить фильтр;
* и жмём на кнопку «искать».

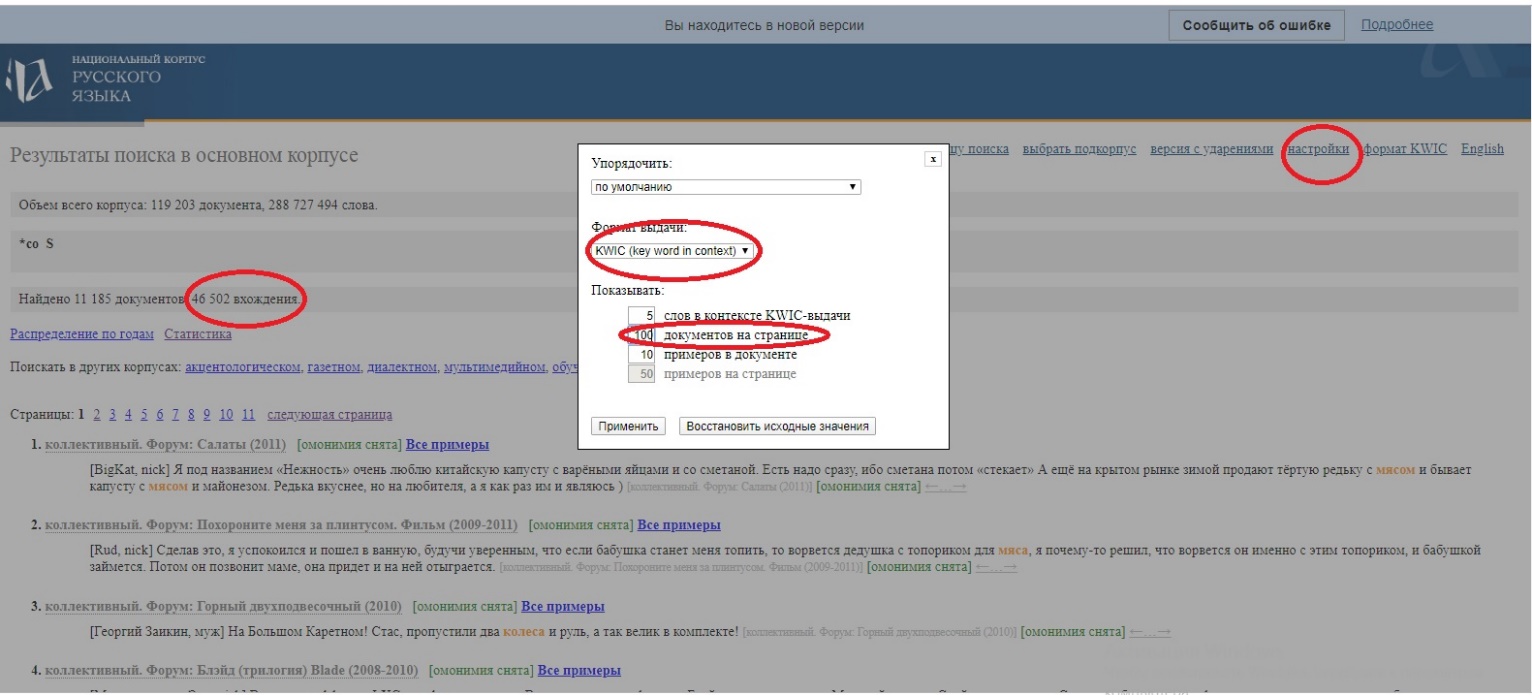


Изучаем результаты поиска. Мы получили 45502вхождения. Определим настройки просмотра.

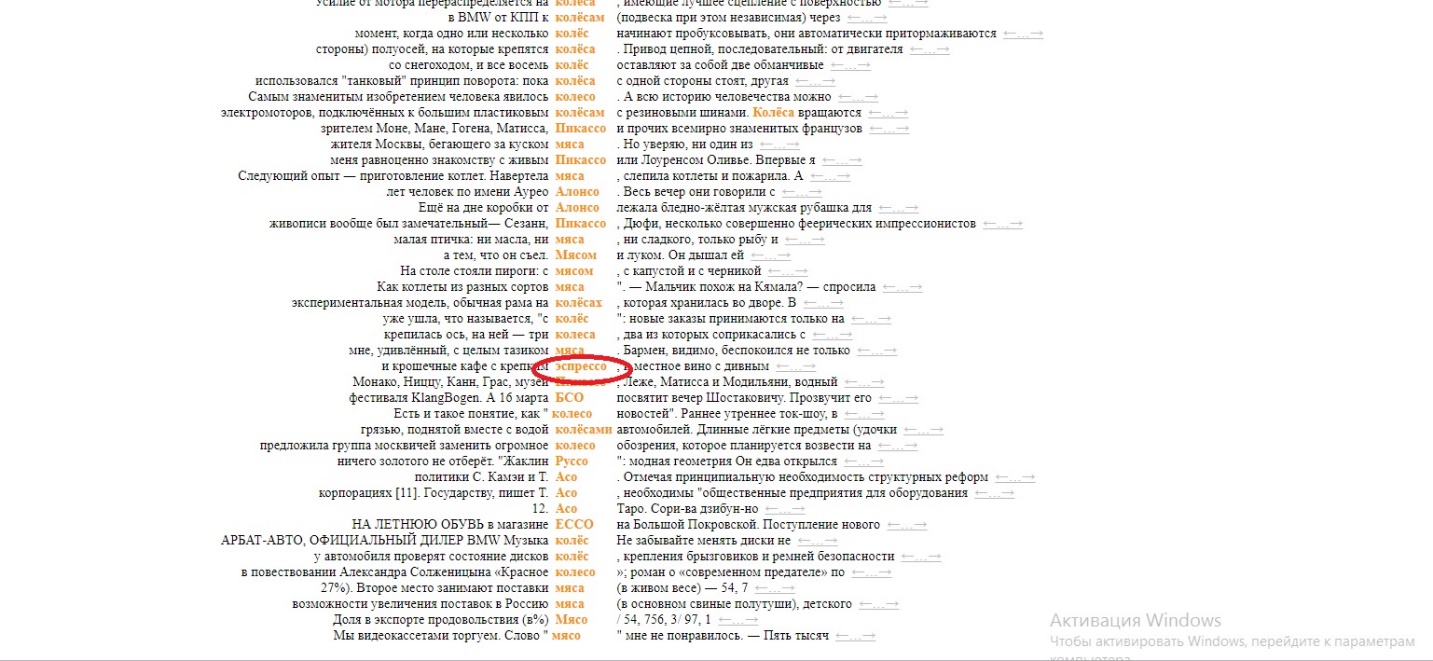
* Перейдем в верхнее меню, найдем ссылку «настройки»;

В раскрывшемся окне выберем форму выдачи KWIC, когда от каждого найденного примера показывается лишь одна строчка, и первое искомое слово помещается в ее середину;

* Определим показ 100 документов на странице.



1. Получаем результаты. Имена собственные в расчет не берём. Дальше терпение и внимательность. Смотрим, нет ли в результатах слов, которые не назвал Успенский?



Методический комментарий.

Приходится признать, что и исследование вручную оказалось довольно полным. Наш улов – иностранные слова, которые вошли в язык по мере усиления культурного обмена: *экспрессо (напиток), мисо(блюдо), песо(денежная единица).* И авторские новообразования *догадалсо, собралсо…*

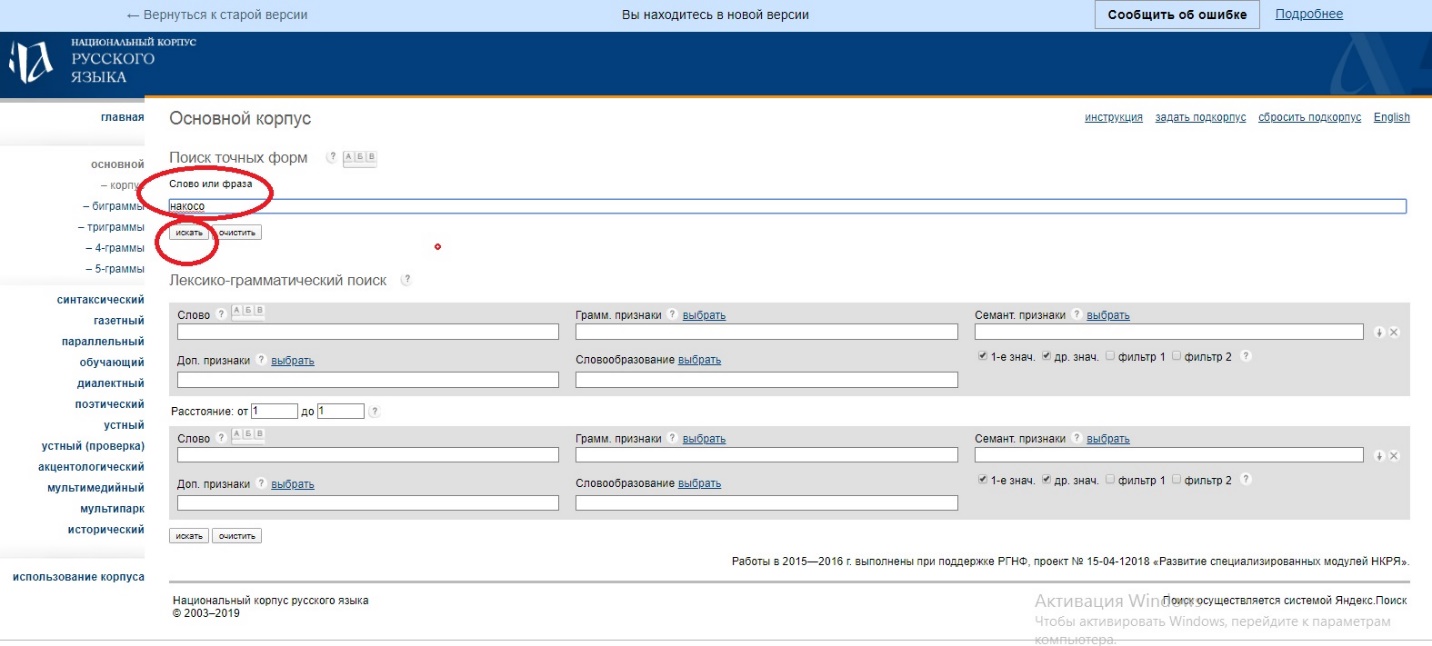
**Чтение с остановками.**

Как видите, их далеко не три, хотя многим из них можно, пожалуй, дать нечто вроде отвода: серсо, инкассо и лассо – как явным иностранцам, наречию «накосо» – как крайне неупотребительному в литературной речи. Все же вместо трех остаются пять несомненных.

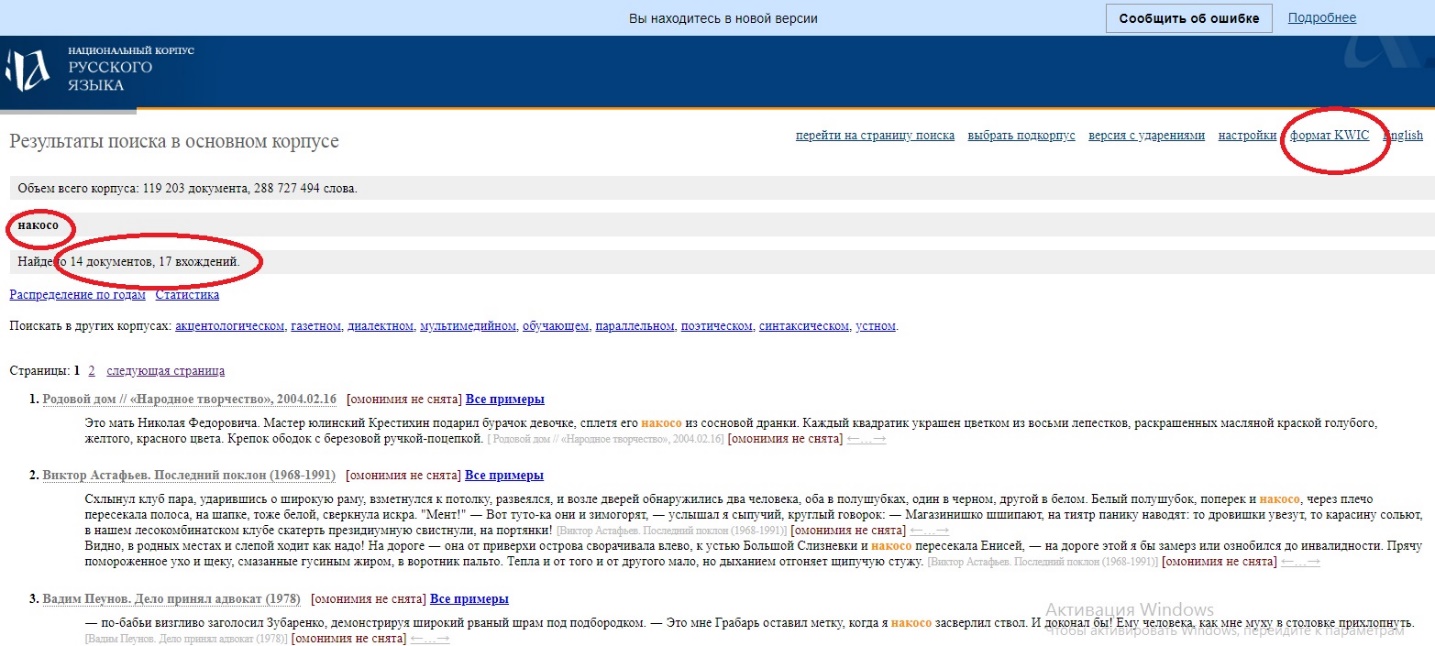
**СТОП*. Давайте проверим утверждение Льва Успенского, что слово «накосо» крайне редко употребляется в литературной речи. Ведь автор этой книги не имел возможностей электронного поиска для подтверждения своих гипотез.***

**Инструкция №3 по работе с НКРЯ.**

Вернемся на страницу поиска. Теперь нас интересует точная форма слова «накосо». Воспользуемся поиском точных форм.



Посмотрим результаты поиска: 17 вхождений из 288 млн слов.



Методический комментарий

Действительно, ничтожно малое количество вхождений. Рассмотрим внимательно случаи употребления этого слова. Большинство случаем вполне Успенский мог знать. Поисковая система нашла слово «накосо» в трудах М.В.Ломоносова, Д.Н.Мамина\_Сибиряка, С.А.Есенина, Д.С.Мережковского, М.А.Кузмина, Б.С.Житкова, П.П.Бажова. Из свеженького во 2-1 половине 20 века найдем только знаменитый роман В.Астафьева «Последний поклон». Успенский в который раз оказывается прав.

**Чтение с остановками.**

А слова, оканчивающиеся на «-зо»? Пожалуйста, оставя в стороне «безо», почти никогда не являющееся в качестве самостоятельного слова, а только в виде фонетического варианта к «без», мы находим в словаре:

*железо*

*ариозо*

*изо[[2]](#footnote-2)*

*пузо*

*авизо*

Ничего не скажешь: и бухгалтерский термин «авизо» и театральный «ариозо» вошли в русский язык, их из него не выкинешь. Академии наук пришлось бы раскошеливаться на премию, если бы пустая легенда была справедлива.

Вы, может быть, спросите: а стоило ли об этом говорить? Прошу прощения: поскольку у меня есть основания считать себя, так сказать, посажённым отцом словарей этого типа, понятно, что мне захотелось поделиться с вами неожиданным даже и для меня по своей быстроте успехом забавной идеи и о их создании.

(Успенский Л.В. Слово о словах.-Л.:Детская литература.-С.57-59.)

**Учитель.** И ещё, почему об этом стоило говорить. Внимательно читая эту старую книжку, мы освоили новый цифровой инструмент для исследования русского языка. Национальный корпус русского языка.

1. Если вы не забыли, о словах с этим суффиксом уже рассказывалось. [↑](#footnote-ref-1)
2. В смысле «изобразительное искусство». [↑](#footnote-ref-2)