



Что такое Национальный корпус русского языка?

Корпус – это собрание текстов в электронной форме, в котором можно осуществлять поиск слов, словосочетаний, грамматических форм, значений слов с помощью определенной поисковой системы. Существует много разновидностей корпусов – например, корпус одного автора, корпус одной книги (в частности, первые корпуса делались для Библии). **Национальный корпус** некоторого языка представляет данный язык на определенном этапе его существования и во всём многообразии жанров, стилей, территориальных и социальных вариантов и т. п. Национальный корпус русского языка функционирует в Интернете с 2003 года и в данный момент включает самые разные русские тексты общим объемом 140 миллионов словоупотреблений. В перспективе Национальный корпус русского языка должен содержать тексты общим объемом порядка 200 миллионов словоупотреблений.

Национальный корпус русского языка, как и национальные корпуса других языков, имеет две важные особенности. Во-первых, он **представителен и сбалансирован**, т. е. содержит все типы письменных и устных текстов, представленных в данном языке (художественные разных жанров, публицистические, учебные, научные, деловые, разговорные, диалектные и т.п.), и все эти тексты входят в корпус по возможности пропорционально их доле в языке соответствующего периода. Во-вторых, корпус содержит особую дополнительную информацию о свойствах входящих в него текстов (так называемую **разметку**, или **аннотацию**). Разметка — главная характеристика корпуса; она отличает корпус от простых коллекций (или «библиотек») текстов, в изобилии представленных в современном Интернете. Чем богаче и разнообразнее разметка, тем выше научная и учебная ценность корпуса. В Национальном корпусе русского языка в настоящее время широко используются три типа разметки: метатекстовая (она характеризует текст в целом, с точки зрения его автора, жанровых и других особенностей), морфологическая и семантическая (морфологическая и семантическая разметки характеризуют не целый текст, а отдельное слово).

Зачем нужен национальный корпус?

Национальный корпус предназначен в первую очередь для научных исследований лексики и грамматики языка. Другая задача корпуса — предоставление всевозможных справок, относящихся к указанным областям (лексика, грамматика, акцентология, история языка). Раньше специалист мог лишь просматривать тексты и вручную выписывать из них нужные примеры; эта деятельность была очень трудоемкой и не позволяла обрабатывать большие массивы материала. Теперь ограничений на объем анализируемого материала и скорость поиска информации нет, а это означает, что в распоряжении ученого и преподавателя оказываются колоссальные массивы текстов самого разного типа.

Основными пользователями национальных корпусов являются, конечно, исследователи-лингвисты самого разного профиля. Однако ими круг пользователей корпуса вовсе не ограничивается! Надежные статистические данные о языке определенной эпохи или определенного автора могут интересовать литературоведов, историков и представителей многих других областей гуманитарного знания. Важнейшее значение национальные корпуса имеют и для преподавания языка

в качестве родного или иностранного; всё больше учебников и учебных программ в наше время оказываются ориентированы на корпус. Быстро и эффективно проверить с помощью корпуса особенности употребления незнакомого слова или грамматической формы у авторитетных авторов сможет и иностранец, и школьник, и учитель, и журналист, и писатель.

Как будет развиваться Национальный корпус?

Национальный корпус русского языка будет охватывать прежде всего период от начала XIX до начала XXI века: этот период представляет язык в разных социолингвистических вариантах — литературном, разговорном, просторечном, отчасти диалектном. В корпус включаются оригинальные (непереводные) произведения художественной литературы (проза и драматургия, поэзия), имеющие культурную значимость, а также представляющие интерес с точки зрения языка. Но Национальный корпус ни в коей мере не является только корпусом языка художественной литературы. Помимо художественных текстов, в корпус в большом количестве включаются и другие образцы письменного (а для современного этапа — и устного) языка: мемуары, эссеистика, публицистика, научно-популярная и научная литература, публичные выступления, частная переписка, дневники, документы и т. п.

Оглавление

Общие вопросы

1. Как найти сайт Национального корпуса русского языка в Интернете?.....	5
2. Как осуществлять поиск в Национальном корпусе русского языка?.....	6
3. Как попасть в Основной корпус?.....	7
4. Как найти Поэтический подкорпус Национального корпуса русского языка?.....	8
5. Как найти Параллельный корпус?.....	9
6. Как найти Диалектный корпус?.....	10
7. Как найти Корпус живой русской речи?.....	11

Как задать свой подкорпус

8. Как задать подкорпус текстов со снятой морфологической омонимией?.....	13
9. Как выбрать текст по названию?.....	14
10. Как отобразить тексты определенного автора?.....	15
11. Как отобразить тексты по полу автора?.....	16
12. Как отобразить тексты по году рождения автора?.....	17
13. Как отобразить тексты по году их создания?.....	18
14. Как отобразить все художественные прозаические тексты?.....	19
15. Как выбрать отдельный жанр художественной прозы?.....	20
16. Как выбрать отдельный тип художественного текста?.....	21
17. Как выбрать тексты в соответствии с местом и временем описываемых событий?.....	22
18. Как выбрать нехудожественные тексты определенной сферы функционирования?.....	23
19. Как выбрать отдельный тип нехудожественного прозаического текста?.....	24
20. Как отобразить нехудожественные тексты по тематике?.....	25
21. Как задать подкорпус Устной публичной речи?.....	26
22. Как задать подкорпус Устной непубличной речи?.....	27
23. Как задать подкорпус «Речь кино»?.....	28
24. Как выбрать тот или иной тип текста в Устной публичной речи?.....	29
25. Как выбрать тот или иной тип текста в Устной непубличной речи?.....	30
26. Как выбрать тот или иной тип текста в подкорпусе «Речь кино»?.....	31
27. Как отобразить художественные тексты по совокупности признаков?.....	32
28. Как отобразить нехудожественные тексты по совокупности признаков?.....	33
29. Как выбрать тексты из подкорпуса живой русской речи по совокупности признаков?.....	34
30. Как сохранить Ваш подкорпус для последующей работы?.....	35
31. Как вернуться на страницу поиска?.....	36

Поиск в Корпусе

32. Как найти словоформу в Корпусе?.....	37
Три технические замечания.....	38
Как действовать, если у Вас нет русской клавиатуры?.....	38
Как действовать, если информация не помещается на экране?.....	39
Как действовать, если Вы получили слишком короткий пример?.....	40
33. Как найти в Корпусе отдельную лексему?.....	41
34. Как найти в Корпусе все лексемы, имеющие одинаковую конечную часть?.....	42
35. Как выбрать из Корпуса все словоформы, имеющие одну и ту же грамматическую характеристику?.....	43

36. Как найти в Корпусе слова, имеющие одинаковую семантическую характеристику?	45
37. Как найти все словоформы данной лексики, имеющие одну и ту же грамматическую характеристику?	47
38. Как найти в Корпусе слово, которое характеризуется определенными грамматическими и семантическими признаками?	48
39. Как найти в Корпусе словосочетание?	49
40. Как найти словосочетание, в котором задано одно слово, а второе характеризуется какими-либо грамматическими признаками?	54
41. Как найти словосочетания, в которых одно слово точно определено и заданы семантические признаки второго слова?	56
42. Как найти в корпусе словосочетание, в котором заданы грамматические характеристики первого и второго слова?	58
43. Как найти словосочетание, в котором заданы грамматические признаки одного слова и семантические признаки второго слова?	59
44. Как найти в Корпусе словосочетания, первое и второе слово в которых характеризуются какими-либо семантическими признаками?	61
Особенности Поэтического корпуса	
45. Стандартные типы поиска в Поэтическом корпусе.	62
46. Нестандартные типы поиска в Поэтическом корпусе.	65
1. Отбор текстов по жанру	65
2. Отбор текстов по метру	66
3. Отбор текстов по стопности	67
4. Отбор текстов по клаузуле	68
5. Отбор текстов по типу строфы	69
6. Отбор текстов по типу рифмы	71
7. Поиск по всем параметрам	72
Особенности параллельного подкорпуса	
47. Как найти слово в русско-английском параллельном корпусе?	73
48. Как найти слово в англо-русском параллельном корпусе?	75
49. Как сортировать примеры в Параллельном корпусе?	77
50. Как найти словосочетание в русско-английском и в англо-русском параллельном корпусе?	80
Некоторые необходимые термины.....	82

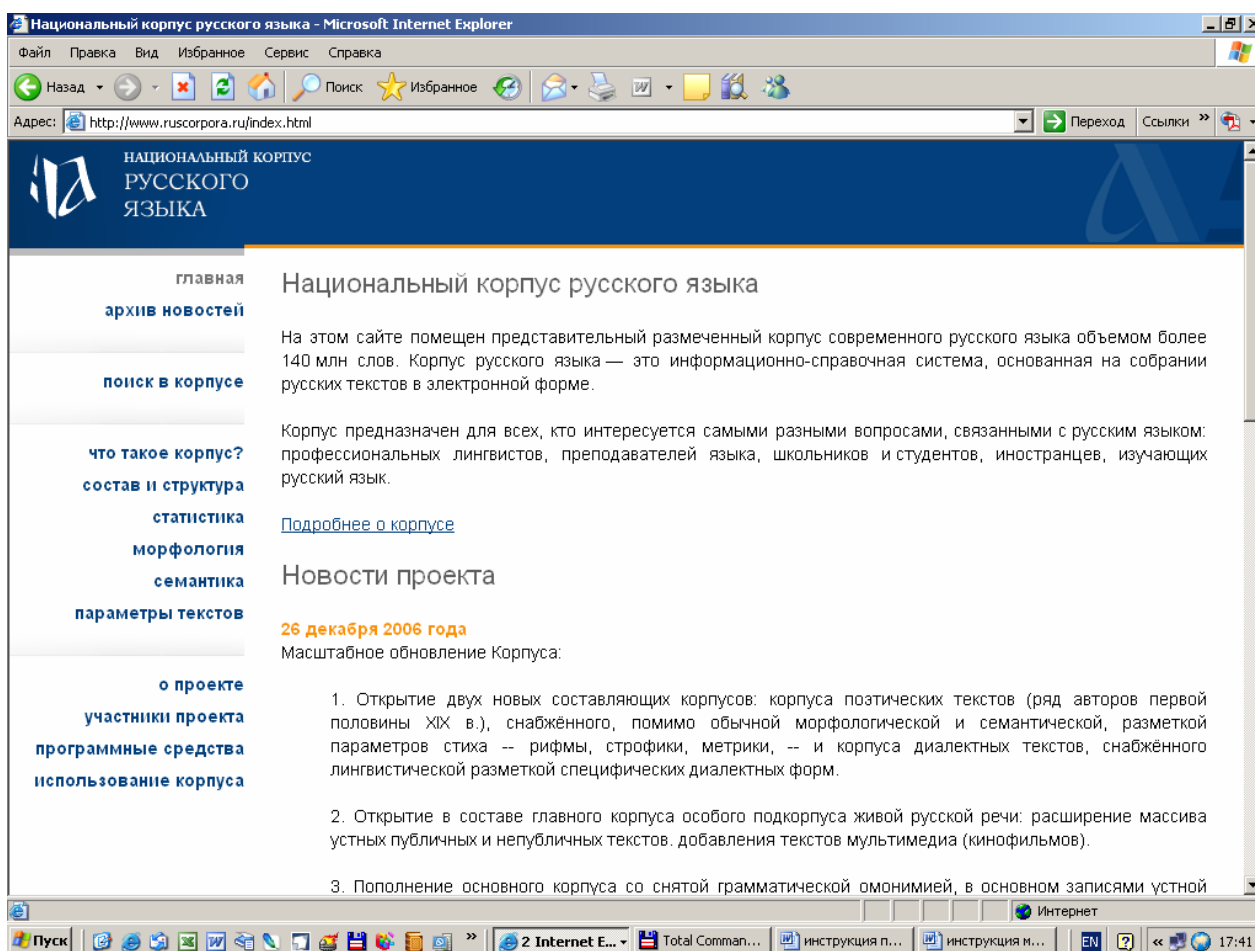
I шаг

Первый шаг. Как найти сайт Национального корпуса русского языка в Интернете?

Наберите адрес:

<http://ruscorpora.ru> или <http://www.ruscorpora.ru>

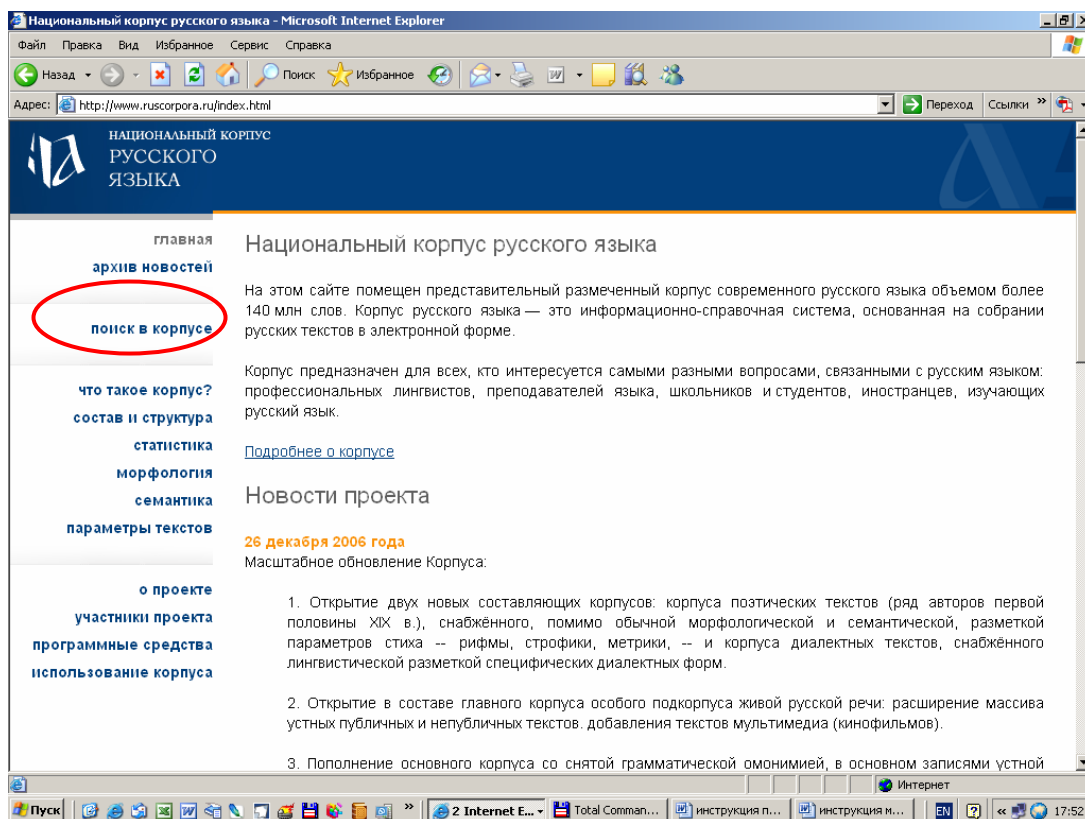
и Вы попадете на наш сайт



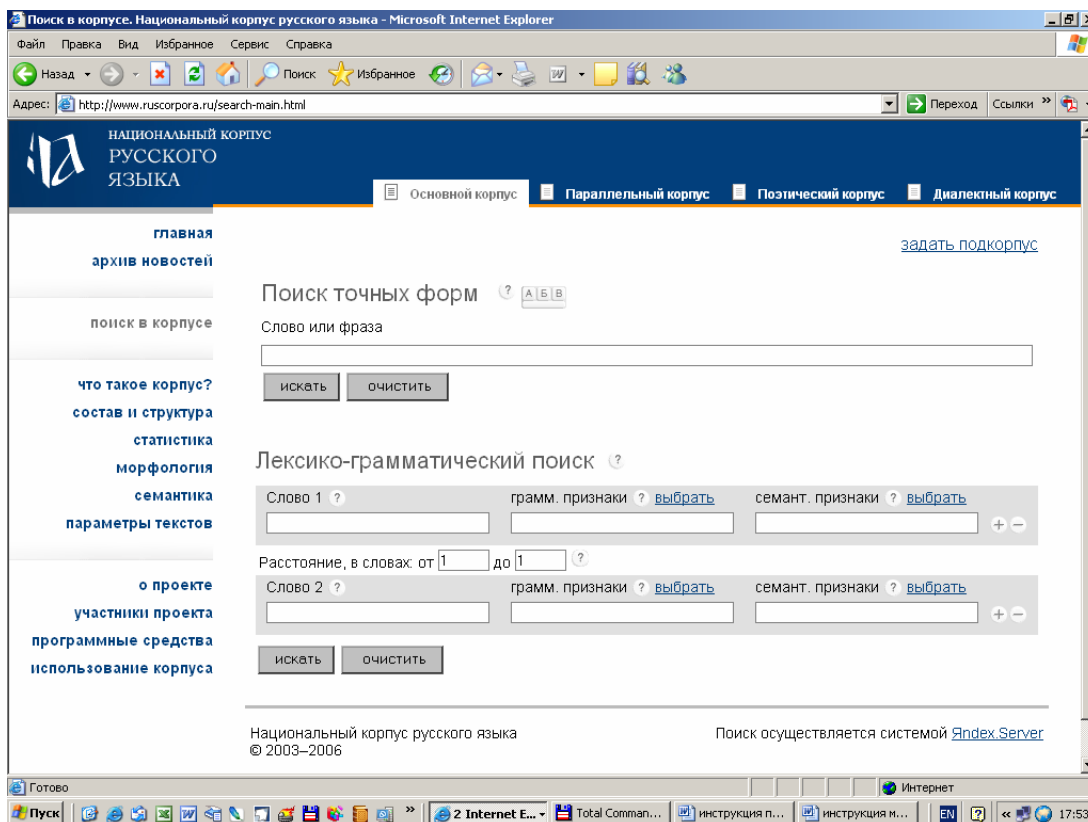
Здесь можно узнать, что такое Национальный корпус русского языка, ознакомиться с его составом и структурой, а также прочитать о том, кто его создает.

2 шаг

Второй шаг. Как осуществлять поиск в Национальном корпусе русского языка?
На главной странице сайта найдите ссылку «Поиск в корпусе» и щелкните на ней мышкой:



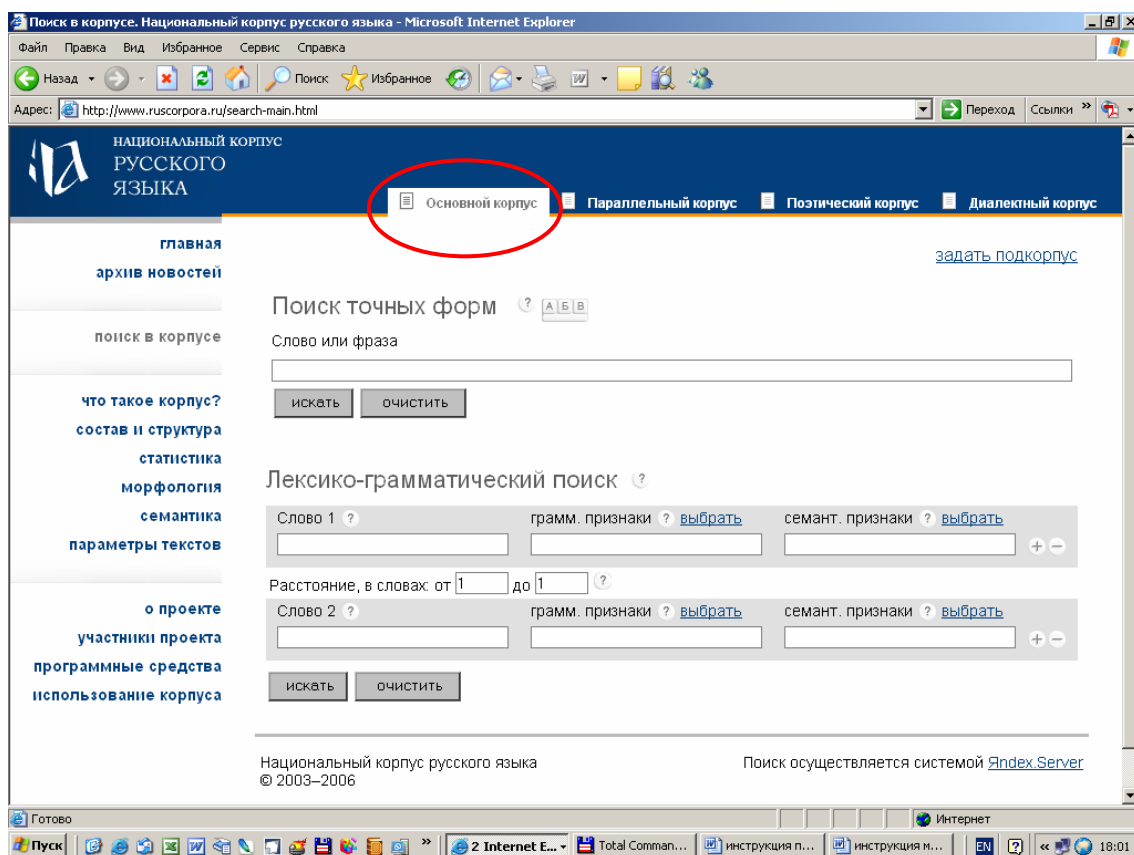
Перед Вами откроется страница «Поиск в корпусе»:



3 шаг

Третий шаг. Как попасть в Основной корпус?

Обратим внимание на то, что как только Вы осуществляете **второй шаг**, Вы автоматически попадаете в Основной корпус Национального корпуса русского языка, т.е. корпус, который содержит литературную прозу (художественную и нехудожественную, письменную и устную (живую)) 19-20 веков. Именно этот корпус является сердцевинной и основной частью Национального корпуса русского языка.



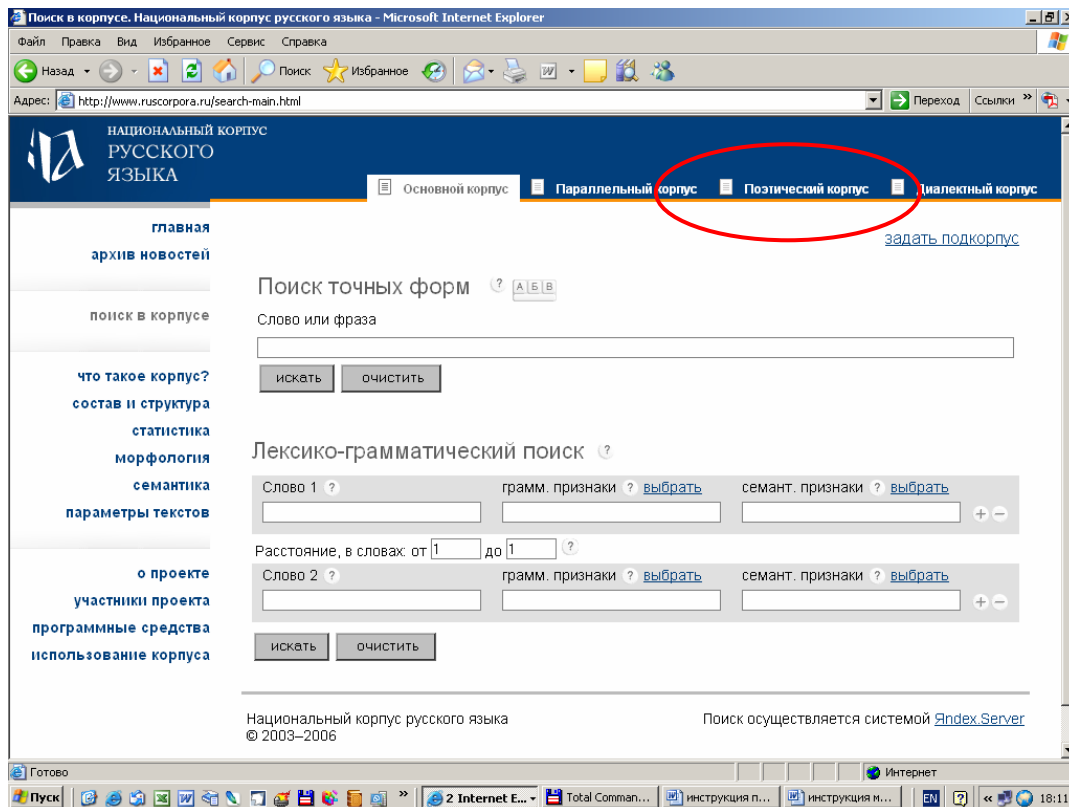
Если Вам необходимо работать с поэтической (а не прозаической), диалектной (а не литературной) речью, или, например, с Параллельным корпусом, или если Вы заинтересованы в тех данных, которые можно найти в подкорпусе устной (живой) русской речи, Вам нужно осуществить дополнительные шаги (**четвертый**, или **шестой**, или **пятый**, или **седьмой**).

Если, напротив, Вы хотите работать с Основным корпусом, то сразу переходите к **шагу восьмому**.

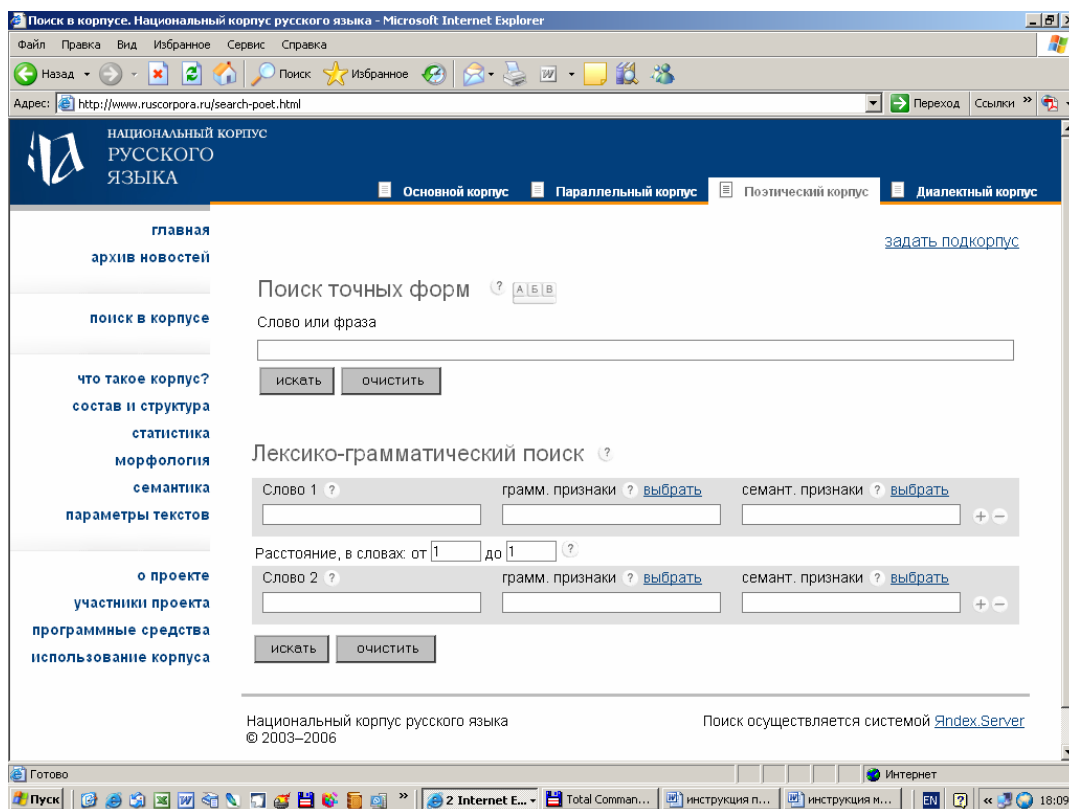
4 шаг

Четвертый шаг. Как найти Поэтический подкорпус Национального корпуса русского языка?

После **шага второго** щелкните мышкой по закладке Поэтический корпус:



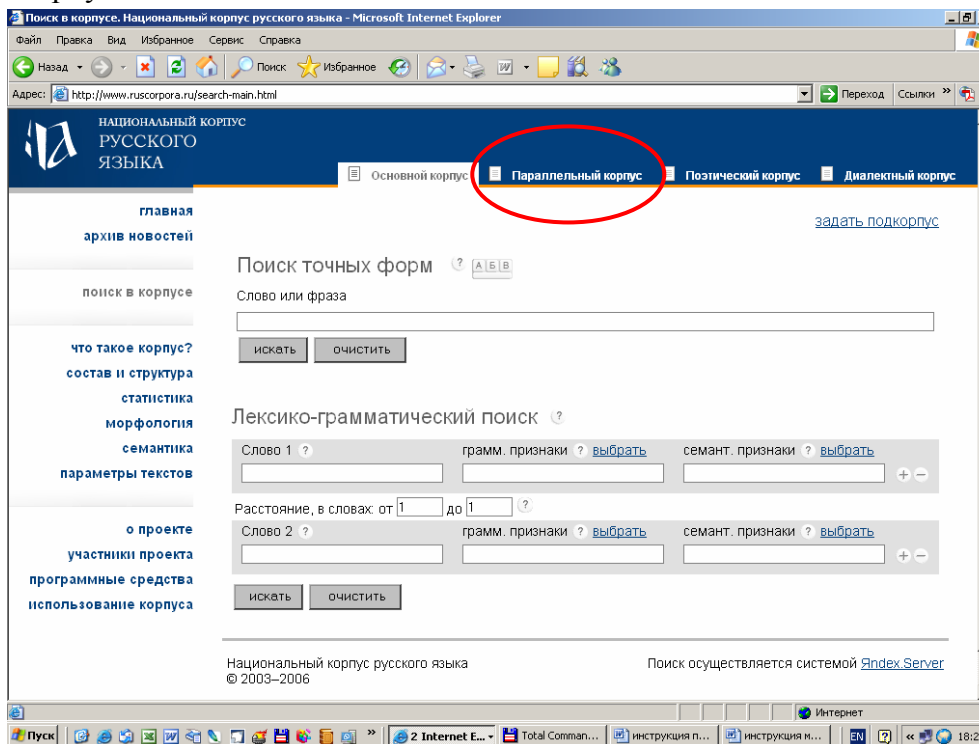
После этого перед Вами откроется поисковая страница Поэтического корпуса:



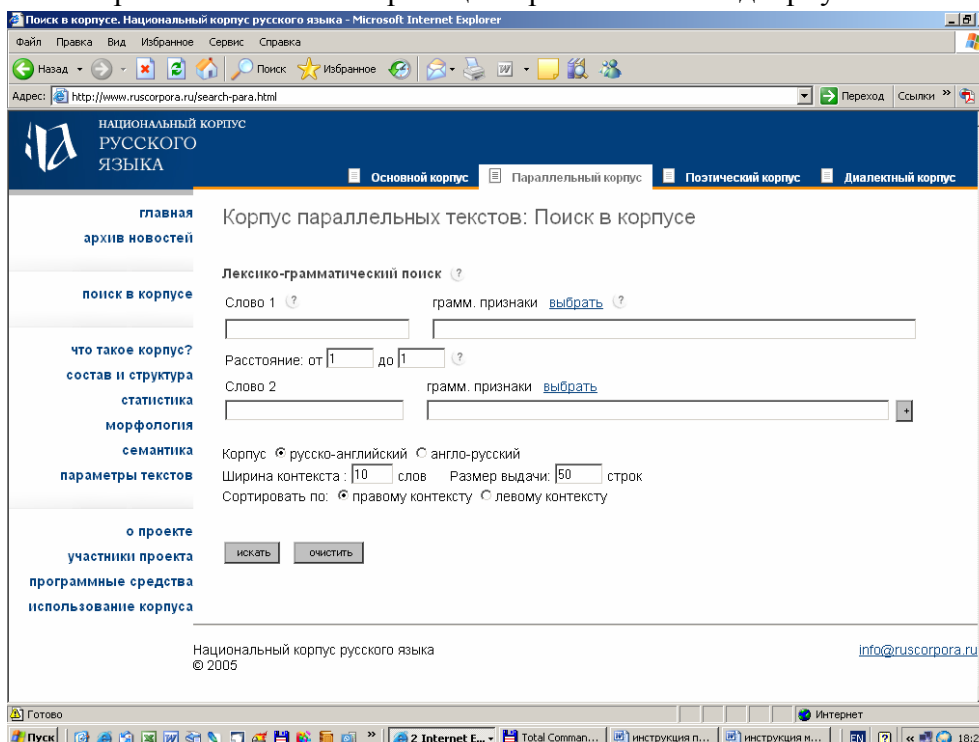
5 шаг

Пятый шаг. Как найти Параллельный корпус?

Параллельный корпус – это подкорпус Национального корпуса русского языка, в котором Вы можете найти русские тексты, переведенные на английский язык, и английские тексты, переведенные на русский язык. Параллельный подкорпус устроен так, что русские и английские тексты выровнены, т.е. определенному участку русского текста поставлен в соответствие перевод этого участка на английский язык, и наоборот, определенному участку английского текста поставлен в соответствие перевод этого участка на русский язык. После **шага второго** щелкните мышкой по закладке Параллельный корпус:



Перед Вами откроется поисковая страница Параллельного подкорпуса:

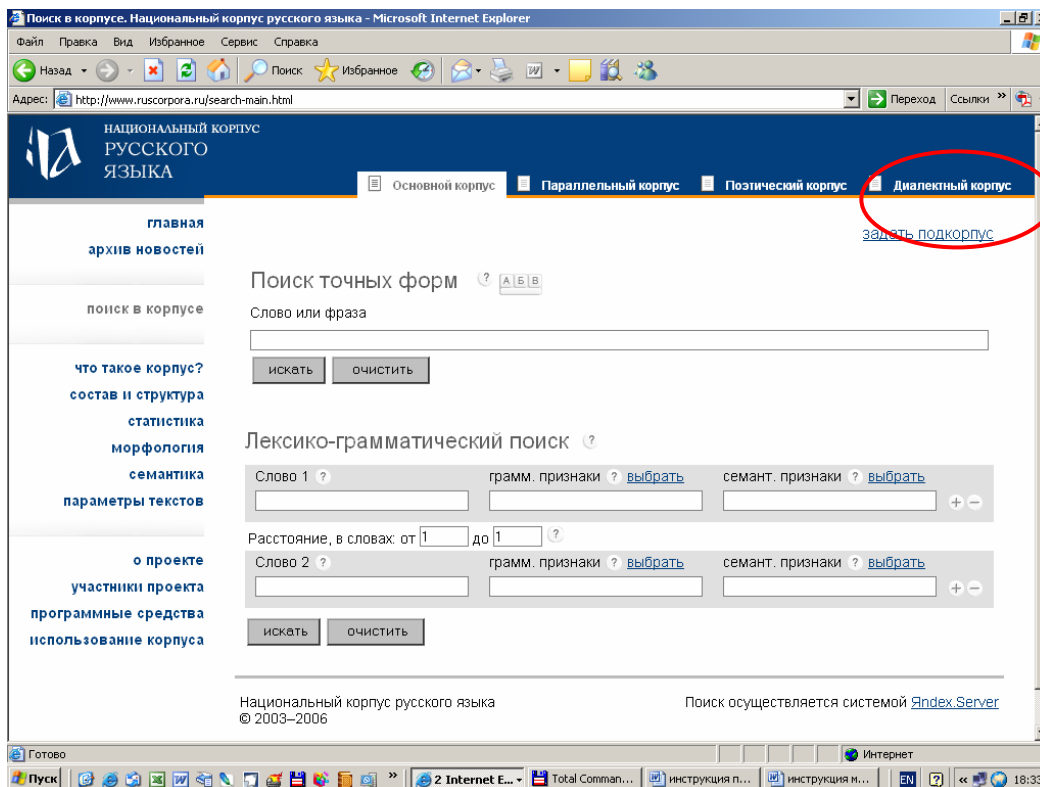


6 шаг

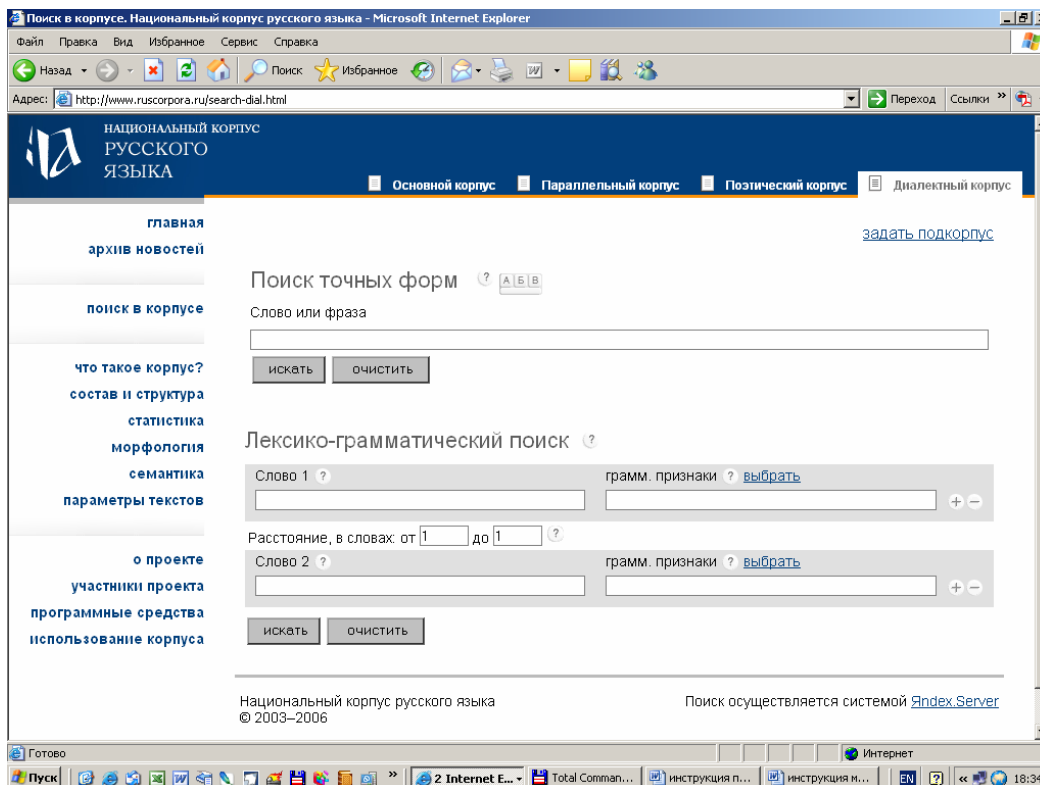
Шестой шаг. Как найти Диалектный корпус?

Диалектный корпус – это подкорпус Национального корпуса русского языка, который содержит тексты, написанные не литературным русским языком, а тем или иным из многочисленных русских диалектов.

После **шага второго** щелкните мышкой на закладке Диалектный корпус:



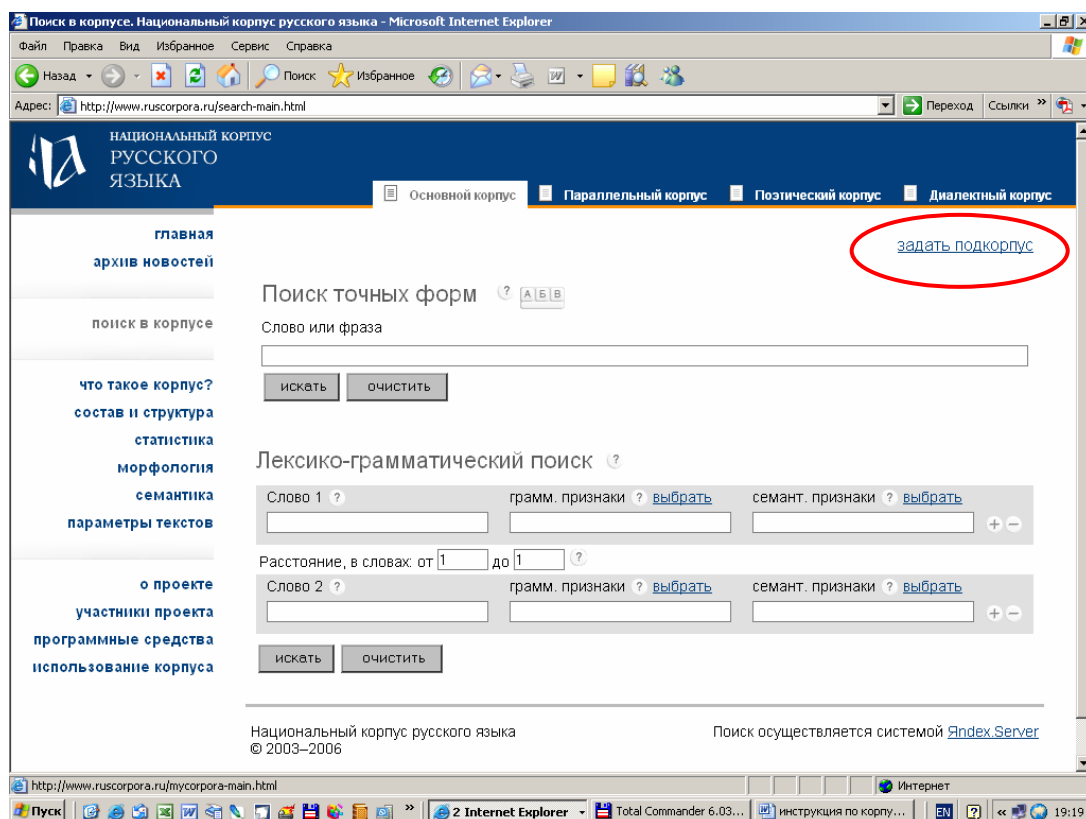
Перед Вами откроется поисковая страница Диалектного корпуса:



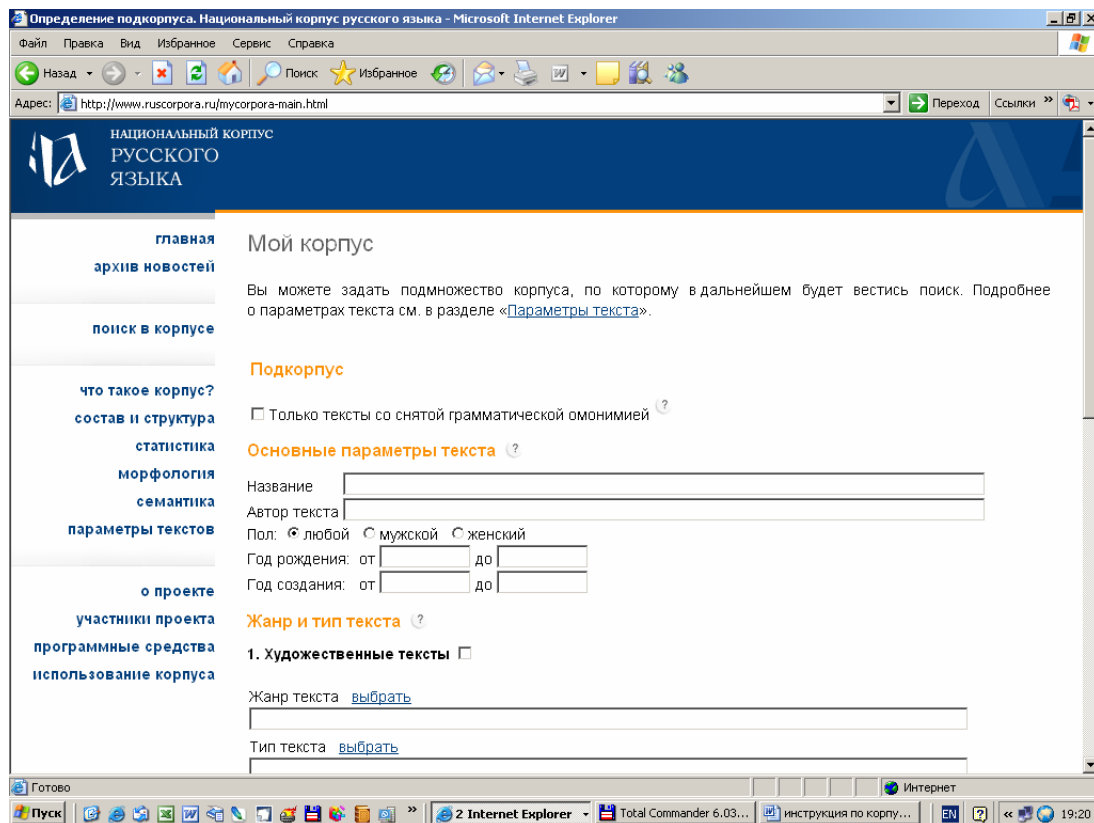
7 шаг

Седьмой шаг. Как найти Корпус живой русской речи?

1. После того, как Вы сделали **второй шаг**, на странице «Поиск в корпусе» найдите ссылку «Задать подкорпус» и щелкните на ней мышкой:

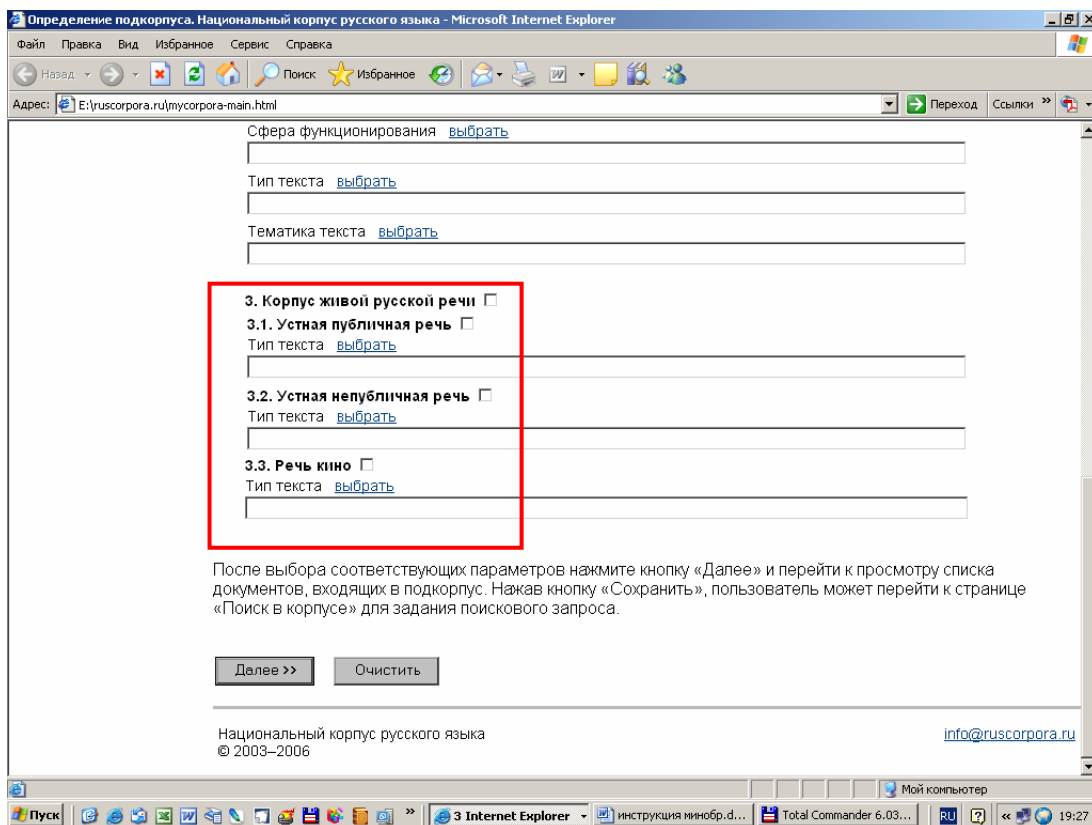


Перед Вами откроется окно «Мой корпус»

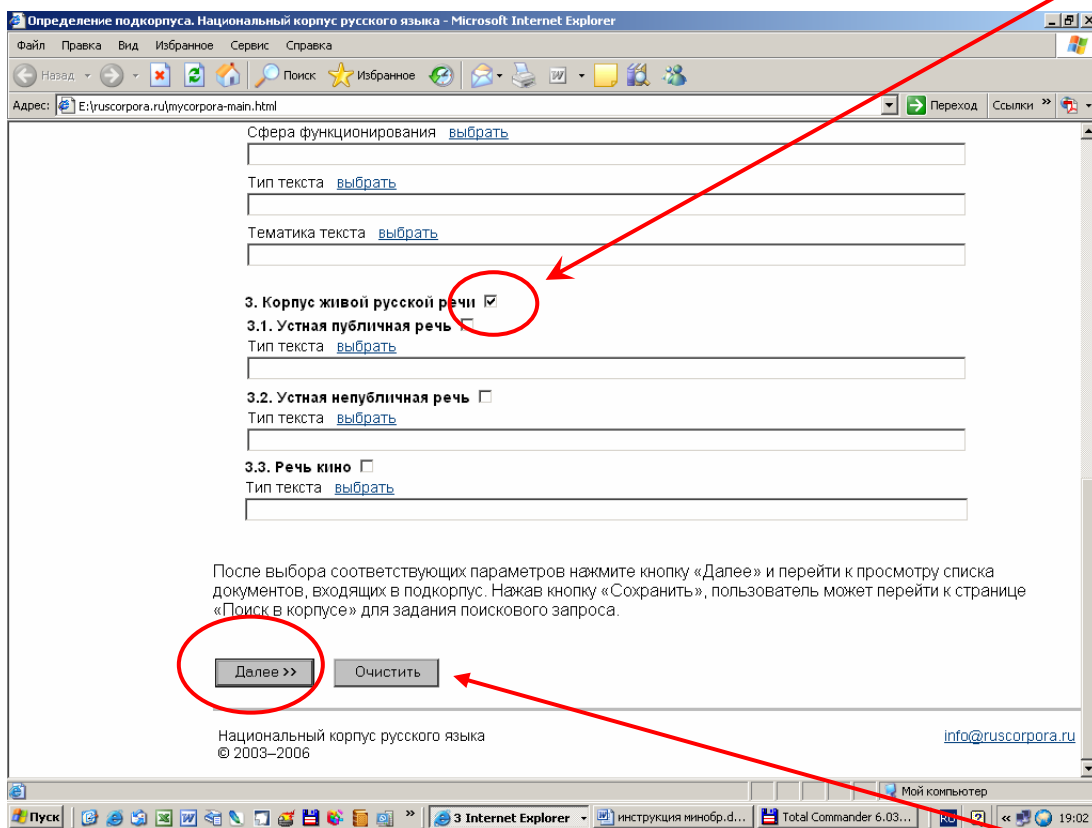


7 шаг

2. Пролитайте страницу «Мой корпус» и Вы увидите пункт 3 «Корпус живой русской речи»:



Если Вы хотите работать со всем Корпусом живой русской речи, то щелкните мышкой вот здесь, а потом щелкните на кнопке «Далее»:



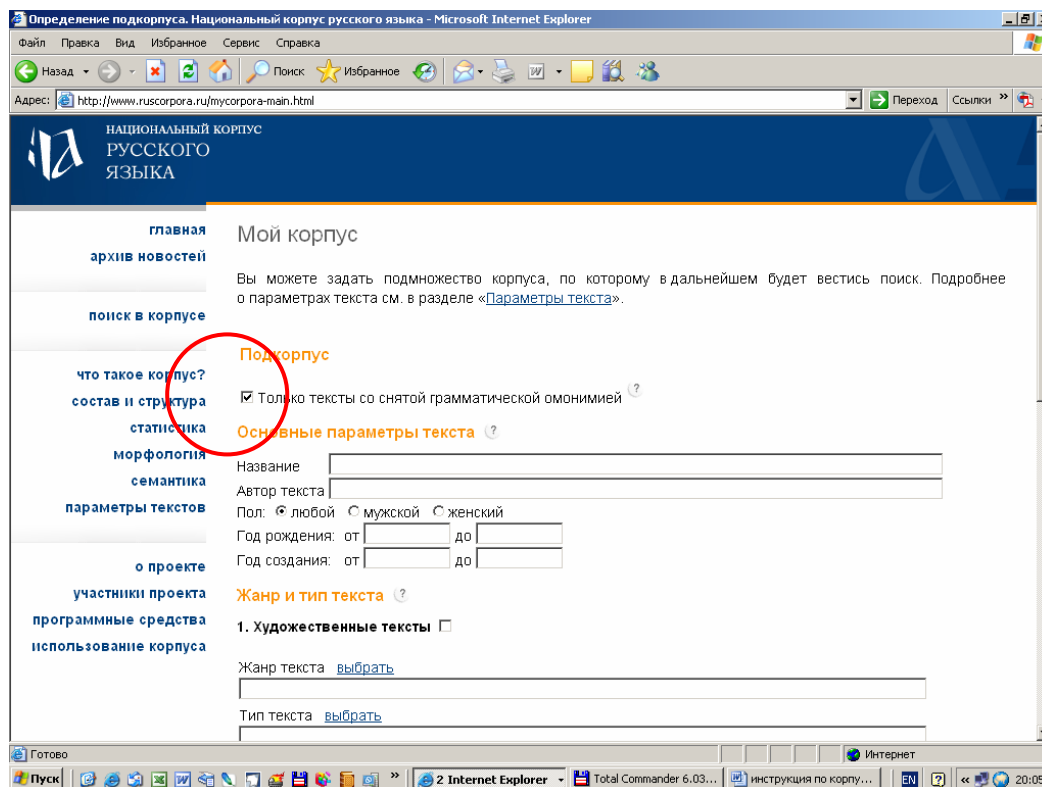
Если Вы передумали работать с Корпусом живой речи, щелкните кнопкой «Очистить» и переходите к тому подкорпусу, который нужен Вам в данный момент; если Ваше намерение не изменилось, переходите к **шагу тридцатому**.

8 шаг

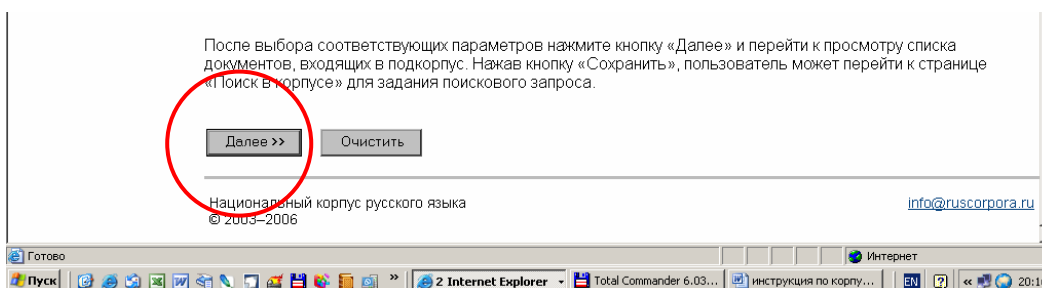
Восьмой шаг. Как задать подкорпус текстов со снятой морфологической омонимией?

Как известно, в русском языке одна и та же словоформа может выражать разные грамматические значения. Например, словоформа «сапог» может обозначать именительный падеж единственного числа и родительный падеж множественного числа, словоформа «порой» может обозначать творительный падеж единственного числа от слов «пОра» и «пОрА», наречие «порой», а также повелительное наклонение, единственное число от глагола «порыть». Это явление и называется **морфологической**, или **грамматической омонимией**. В основном Корпус содержит тексты, в которых омонимия не снята, т.е. словоформа «порой» имеет при себе все четыре перечисленных выше морфологических разбора, и, соответственно, будет выпадать при четырех разных запросах пользователя. Однако в небольшой части Корпуса морфологическая омонимия была снята вручную, так что при словоформе оставался только один, правильный разбор. Вот такие тексты и формируют подкорпус текстов со снятой омонимией. В этих текстах проставлено ударение.

Если Вам нужны для Вашей работы только такие тексты, осуществите **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, отметьте галочкой, щелкнув мышкой, позицию «Только тексты со снятой грамматической омонимией»:



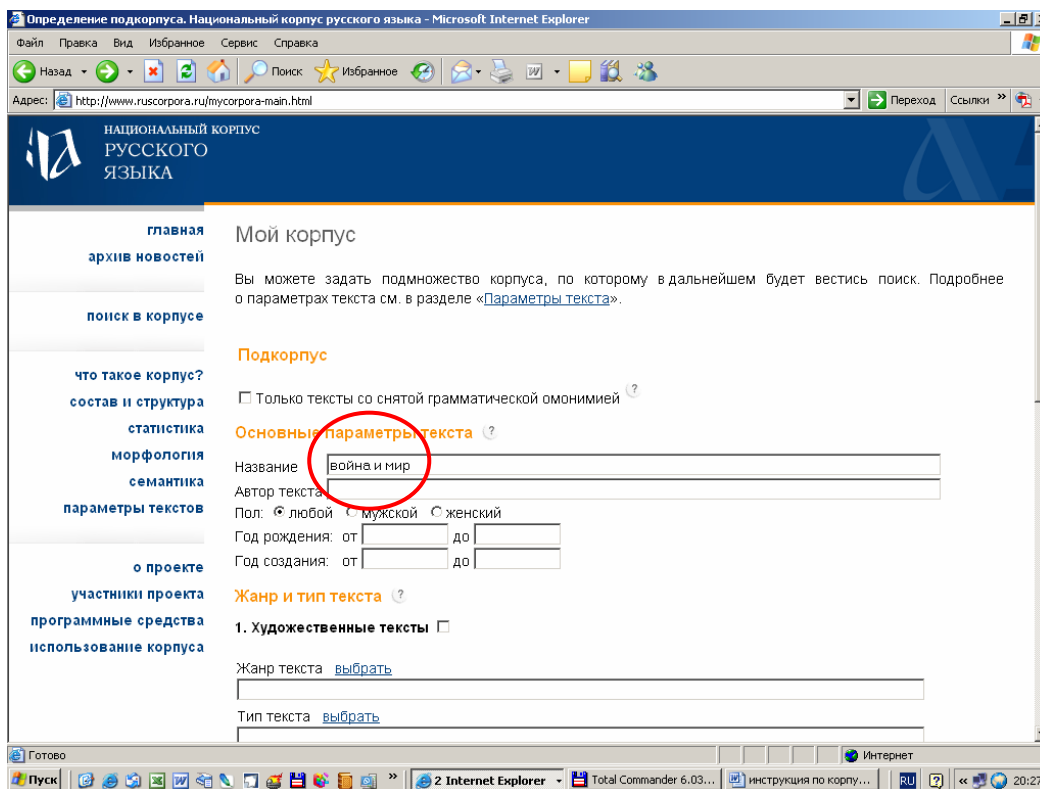
После этого пролистайте эту страницу до конца, найдите кнопку «Далее», щелкните по ней мышкой, а дальше переходите к **шагу тридцатому**.



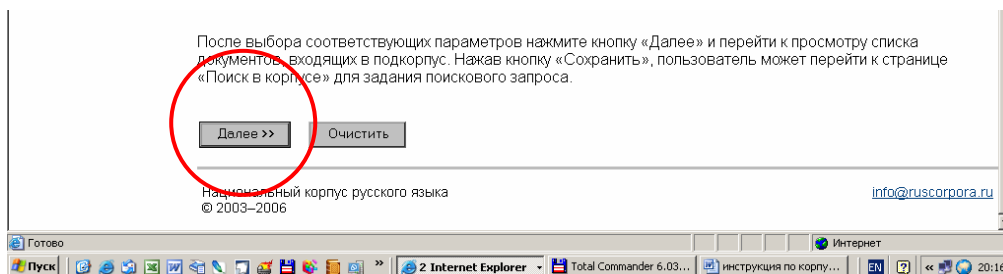
9 шаг

Девятый шаг. Как выбрать текст по названию?

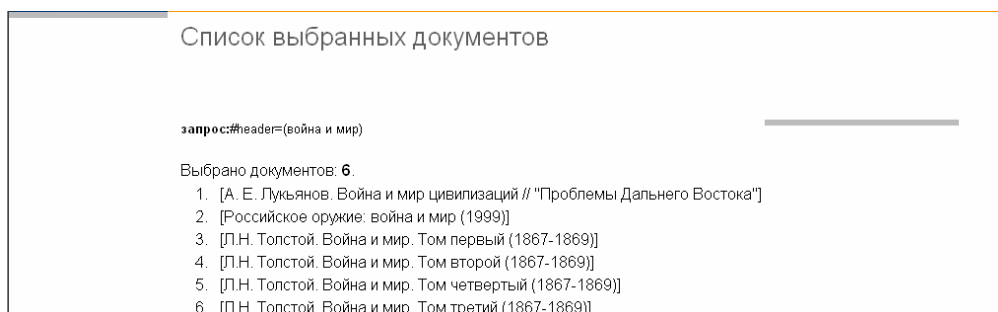
Осуществите **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «Основные параметры текста» в строке «Название» наберите название текста:



После этого пролистайте эту страницу до конца, найдите кнопку «Далее», щелкните по ней мышкой:



После этого перед Вами появится окошко, в котором Вы увидите список всех тексты, которые содержат в названии слова «война», «и», «мир»:

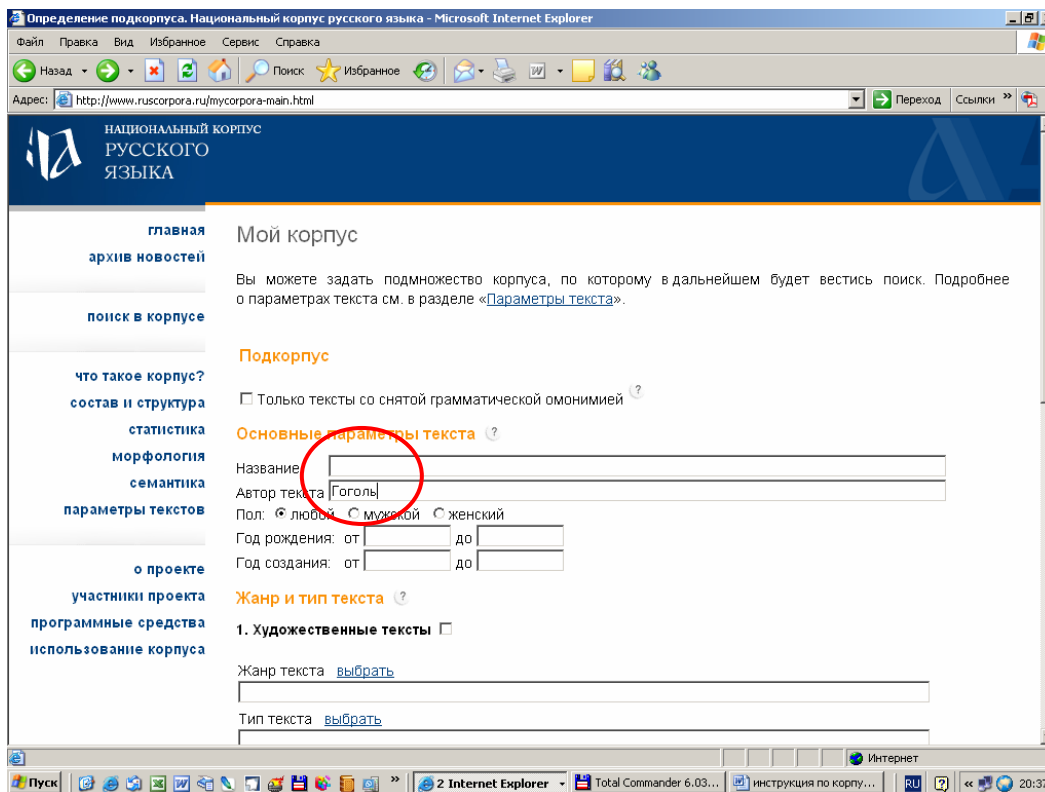


Теперь Вы можете переходить к **шагу тридцатому**.

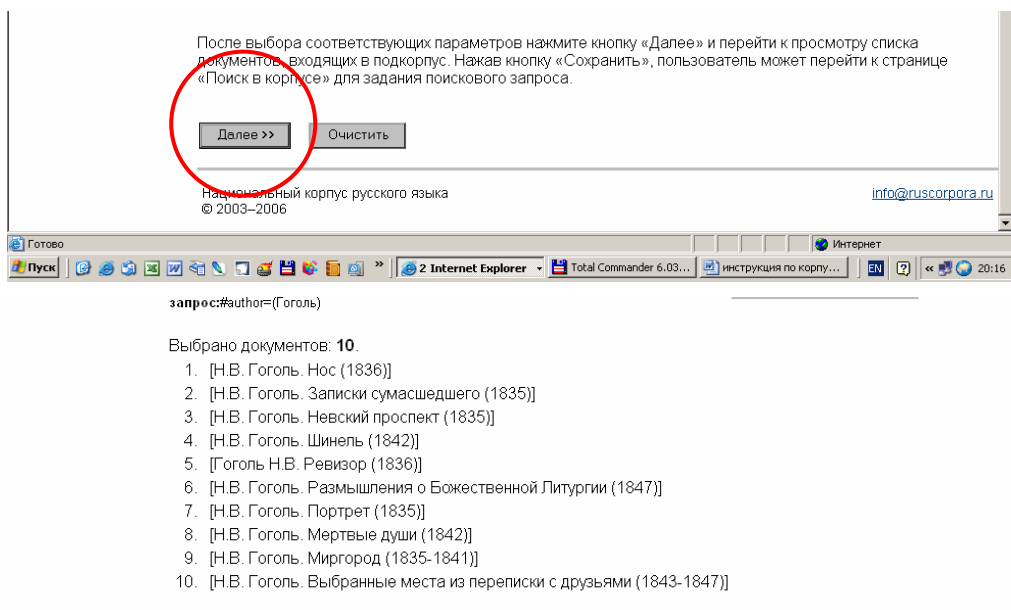
10 шаг

Десятый шаг. Как отобразить тексты определенного автора?

Осуществите **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «Основные параметры текста» в строке «Автор текста» наберите фамилию автора:



После этого пролистайте эту страницу до конца, найдите кнопку «Далее», щелкните по ней мышкой и Вы увидите список найденных текстов:

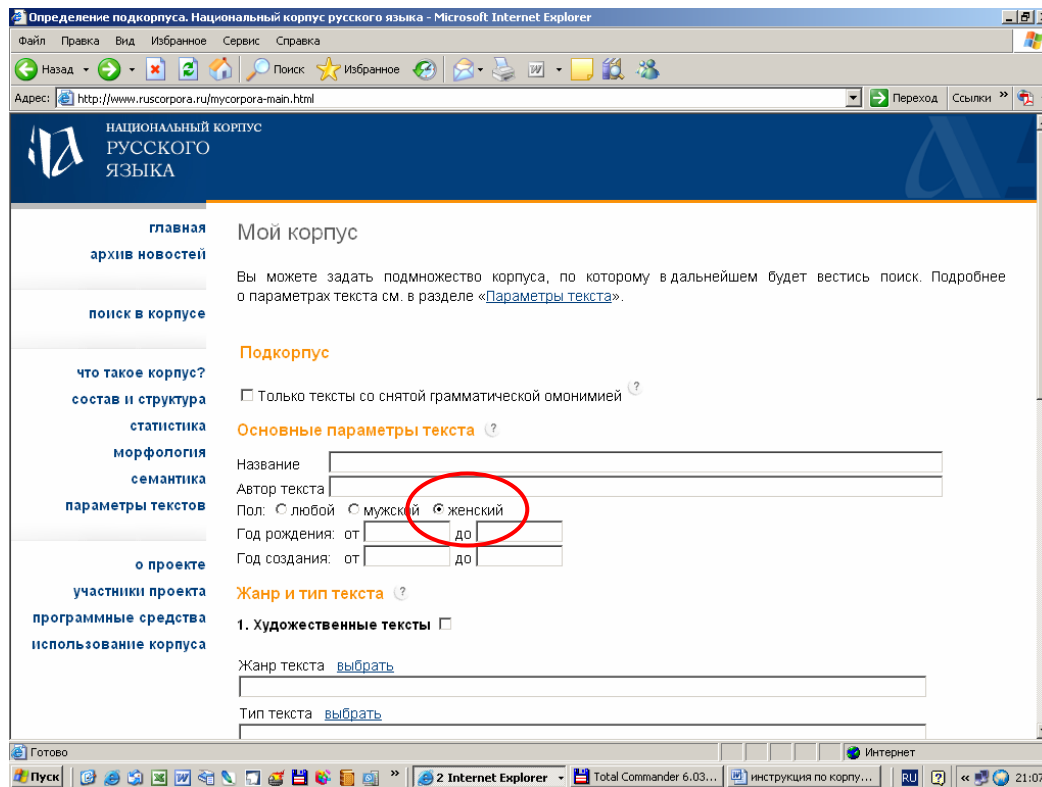


После этого можно переходить к **шагу тридцатому**.

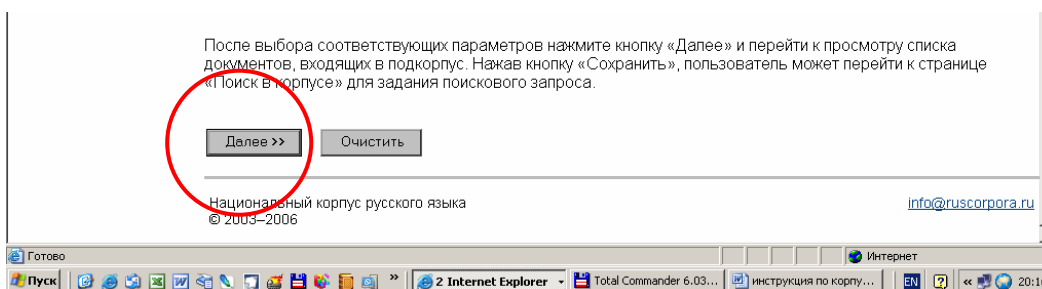
II шаг

Одиннадцатый шаг. Как отобрать тексты по полу автора?

Если Вы хотите работать с текстами, написанными только мужчинами, или только женщинами, Вам следует осуществить **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «Основные параметры текста» в строке «Пол» выберите «мужской» или «женский» (обращаем Ваше внимание, что по умолчанию выбирается «любой» пол автора):



После этого пролистайте эту страницу до конца, найдите кнопку «Далее», щелкните по ней мышкой:

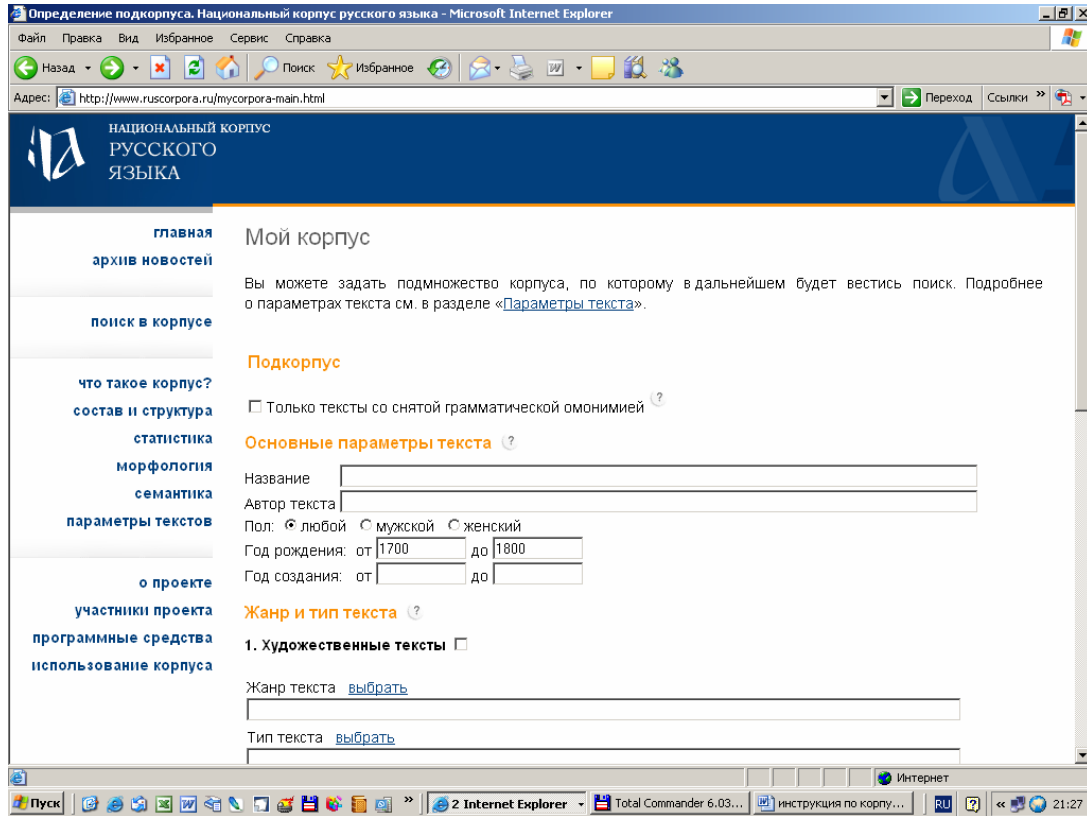


После этого можно переходить к **шагу тридцатому**.

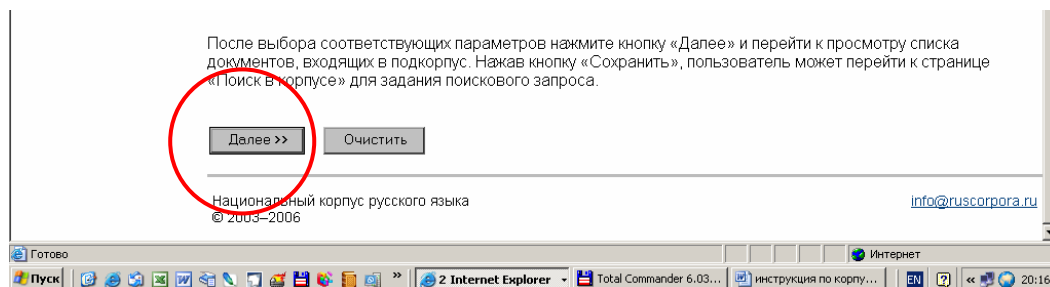
12 шаг

Шаг двенадцатый. Как отобразить тексты по году рождения автора?

Если Вы хотите работать с текстами, написанными авторами, которые родились в определенном промежутке времени или в конкретный год, то Вам следует осуществить **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «Основные параметры текста» в строке «Год рождения» выберите интересующий Вас интервал (предположим, Вас интересуют тексты авторов, родившихся в 18 веке):



После этого пролистайте эту страницу до конца, найдите кнопку «Далее», щелкните по ней мышкой:

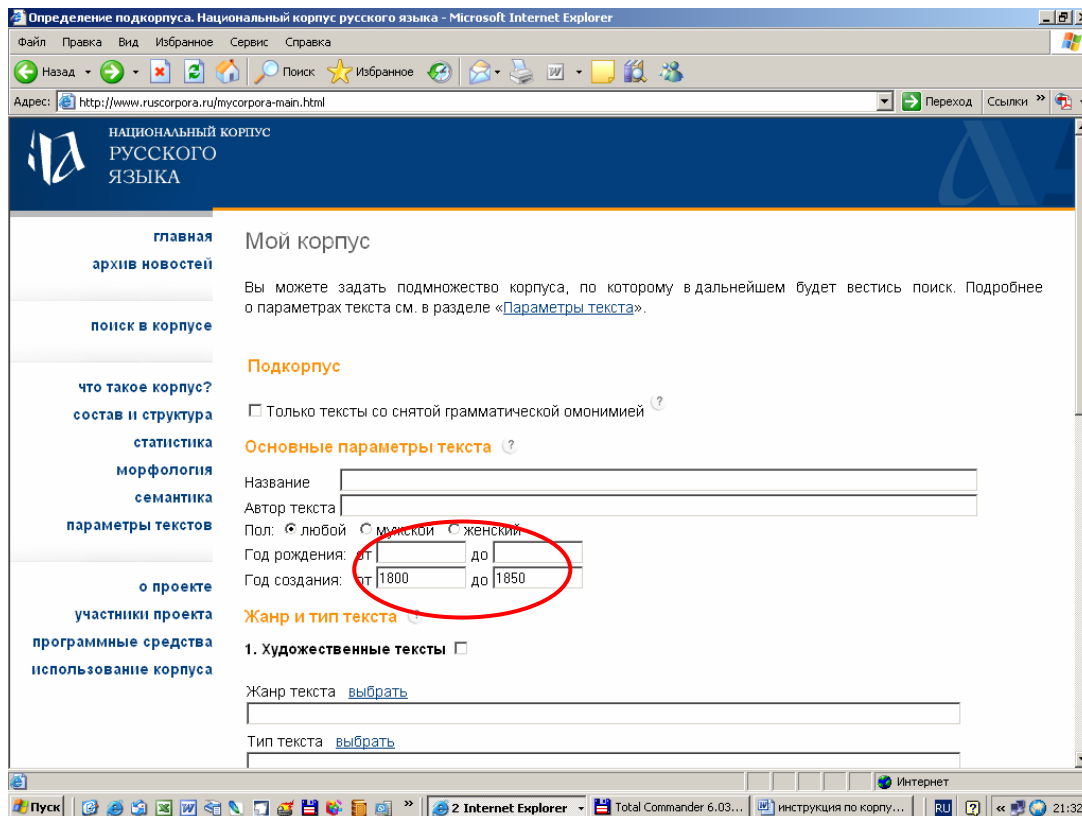


После этого можно переходить к **шагу тридцатому**.

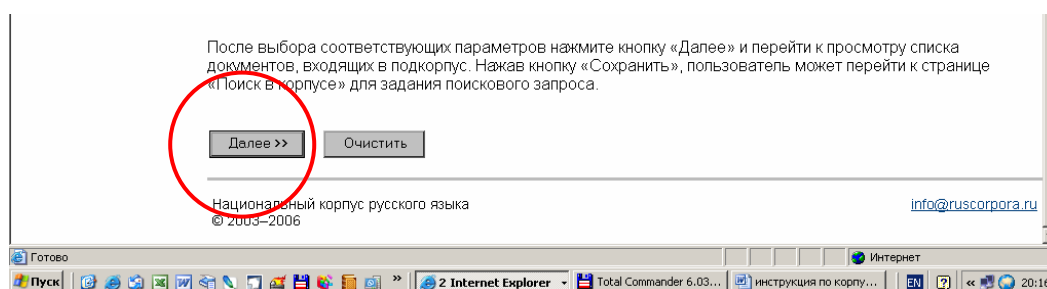
13 шаг

Тринадцатый шаг. Как отобразить тексты по году их создания?

Если Вы хотите работать с текстами, написанными в определенный год или в определенный промежуток времени, то Вам следует осуществить **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «Основные параметры текста» в строке «Год создания» выберите интересующий Вас интервал (предположим, Вас интересуют тексты, написанные в первой половине 19 века):



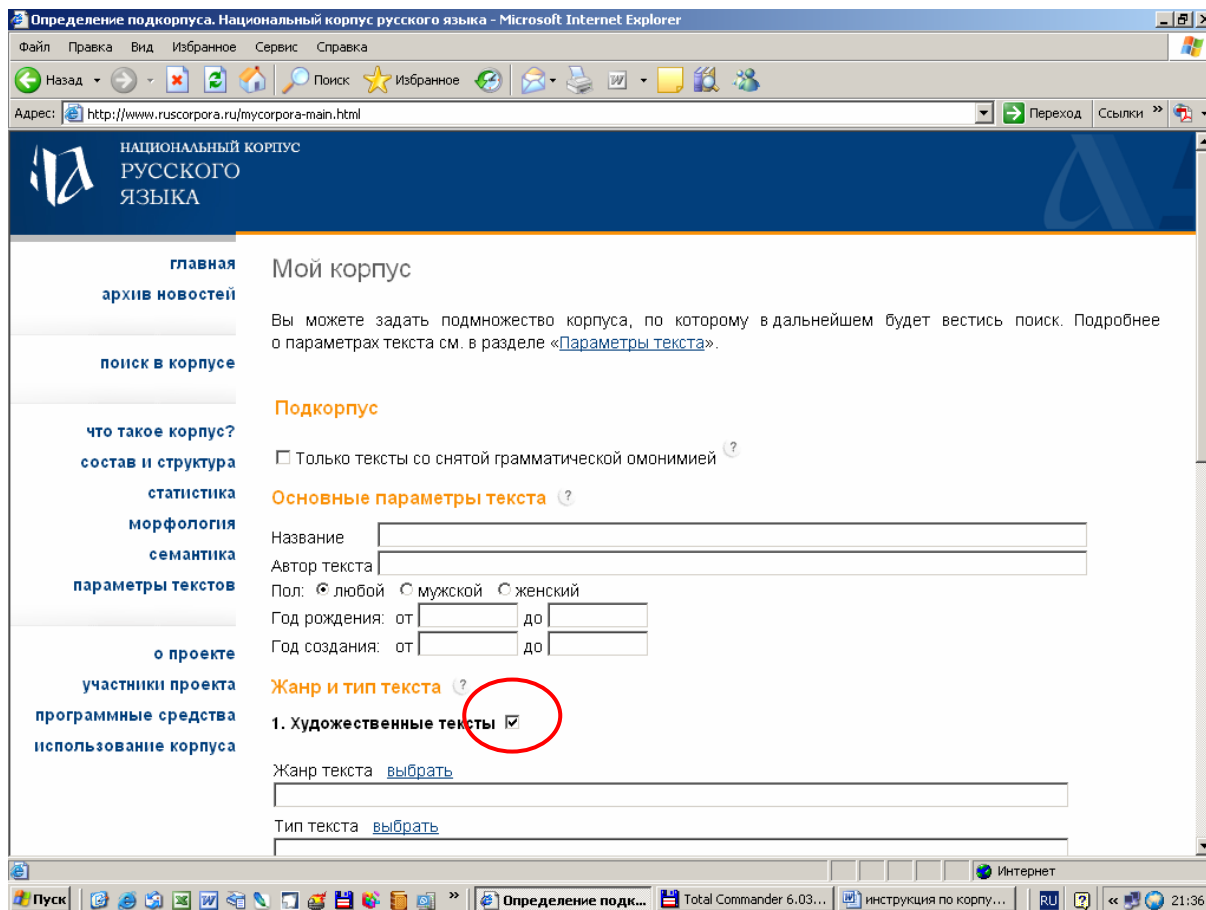
После этого пролистайте эту страницу до конца, найдите кнопку «Далее», щелкните по ней мышкой:



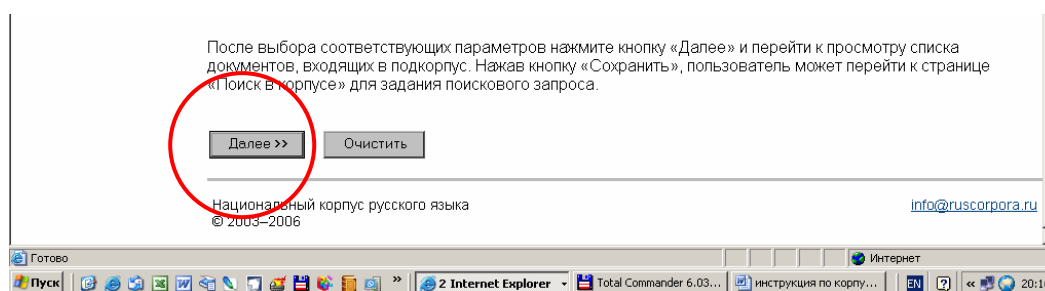
После этого можно переходить к **шагу тридцатому**.

14 шаг

Четырнадцатый шаг. Как отобразить все художественные прозаические тексты? Если Вы хотите работать только с художественной прозой, то Вам следует осуществить **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «Жанр и тип текста» щелкните мышкой напротив пункта 1 «Художественные тексты»:



После этого пролистайте эту страницу до конца, найдите кнопку «Далее», щелкните по ней мышкой:

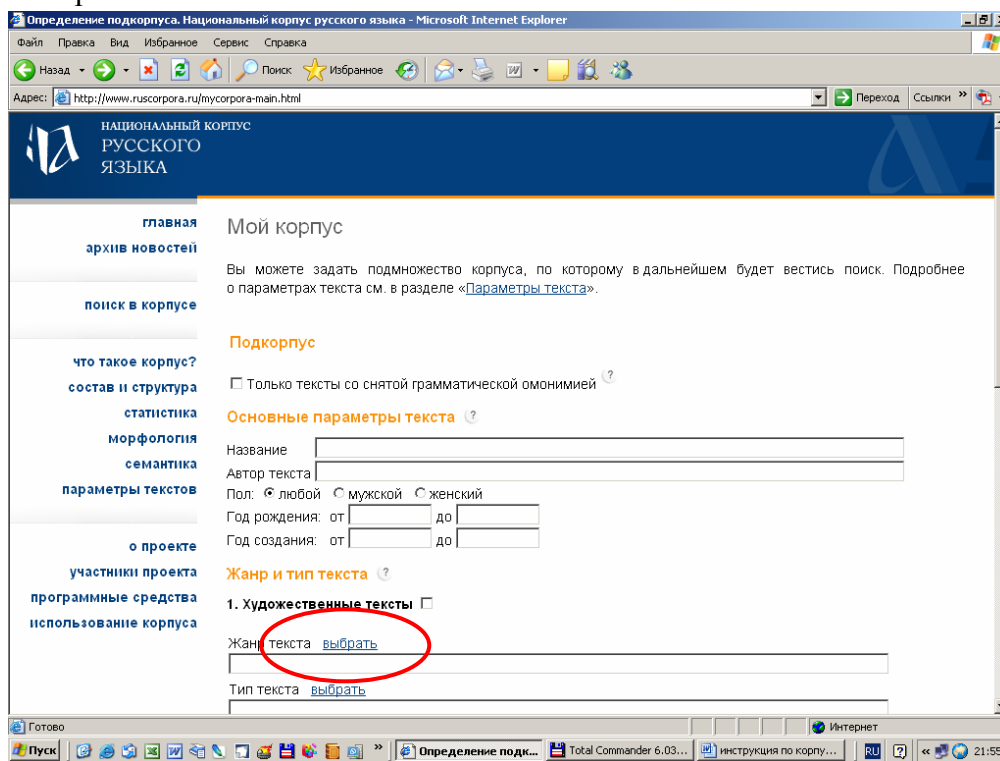


После этого можно переходить к **шагу тридцатому**.

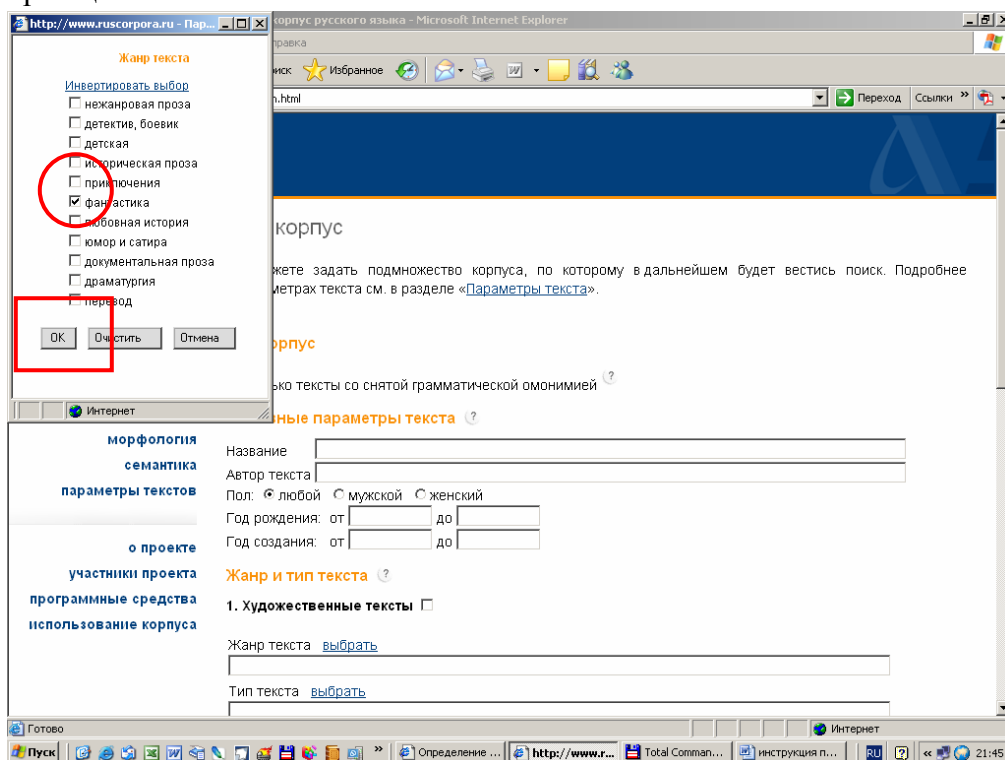
15 шаг

Пятнадцатый шаг. Как выбрать отдельный жанр художественной прозы?

Если Вы хотите работать только с конкретным жанром художественной прозы, то Вам следует осуществить **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «Жанр и тип текста» щелкните мышкой на ссылке «Выбрать» в строке «Жанр текста»:



Перед вами откроется маленькое окошко со списком жанров. Выберите в нем нужный Вам жанр и щелкните мышкой по кнопке «ОК»:

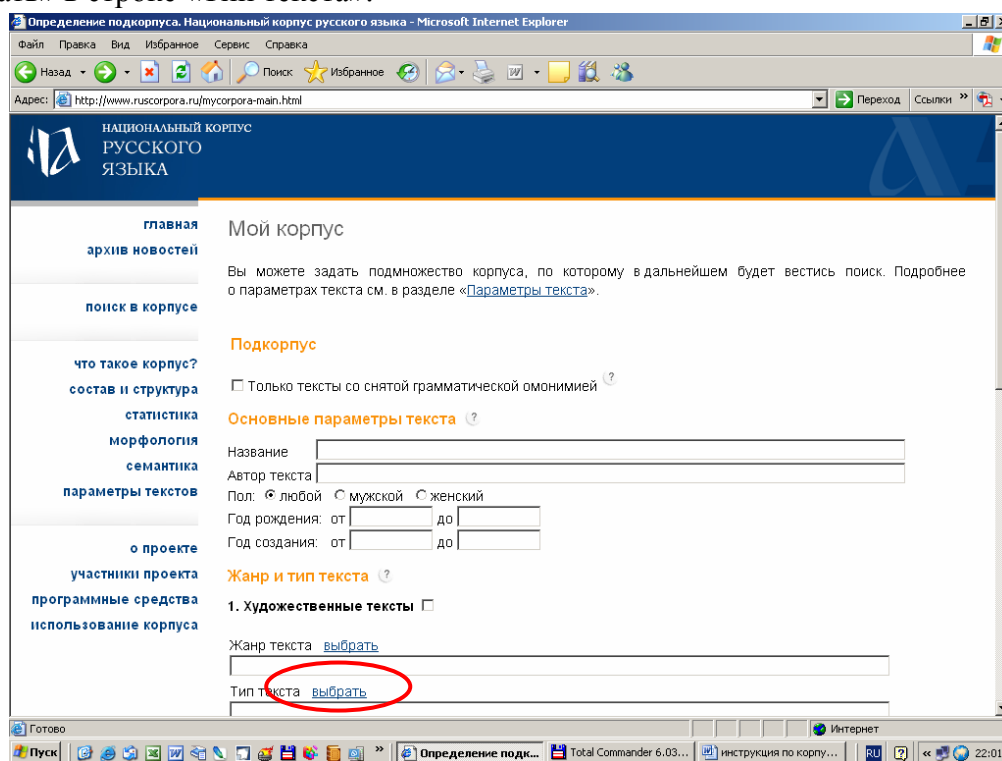


После этого пролистайте страницу вниз, найдите кнопку «Далее» (см. выше, **шаги с седьмого по четырнадцатый**) и щелкните по ней мышкой. После этого переходите к **шагу тридцатому**.

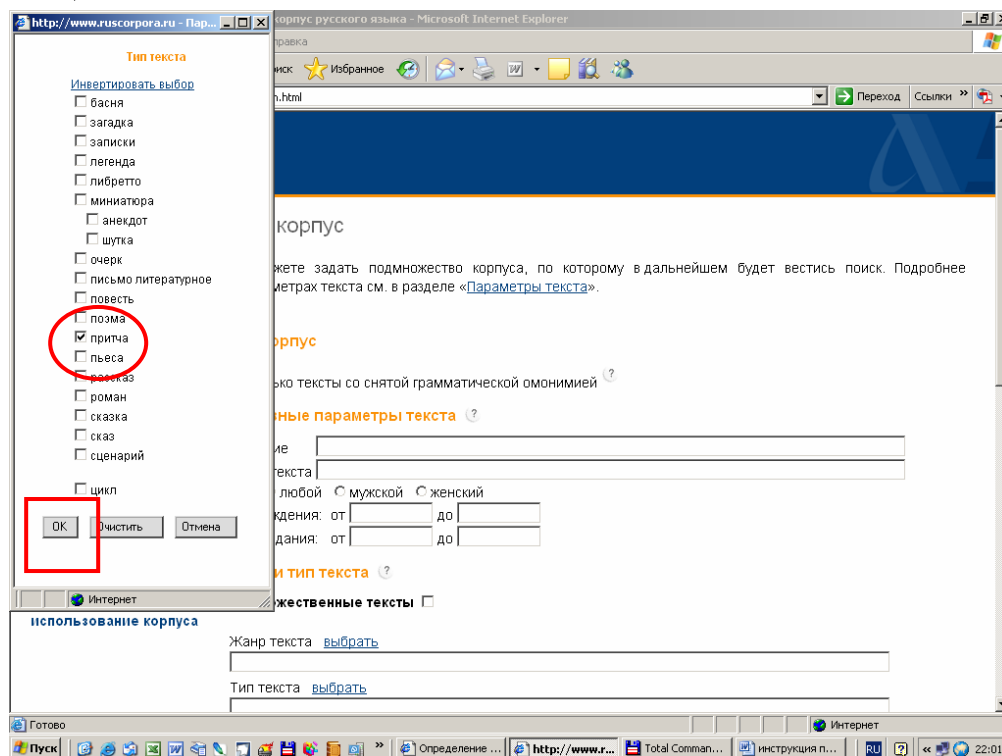
16 шаг

Шестнадцатый шаг. Как выбрать отдельный тип художественного текста?

Если Вы хотите работать только с определенным типом художественного прозаического текста, то Вам следует осуществить **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «Жанр и тип текста» щелкните мышкой на ссылке «Выбрать» в строке «Тип текста»:



Перед вами откроется маленькое окошко со списком типов. Выберите в нем нужный Вам тип и щелкните мышкой по кнопке «ОК»:

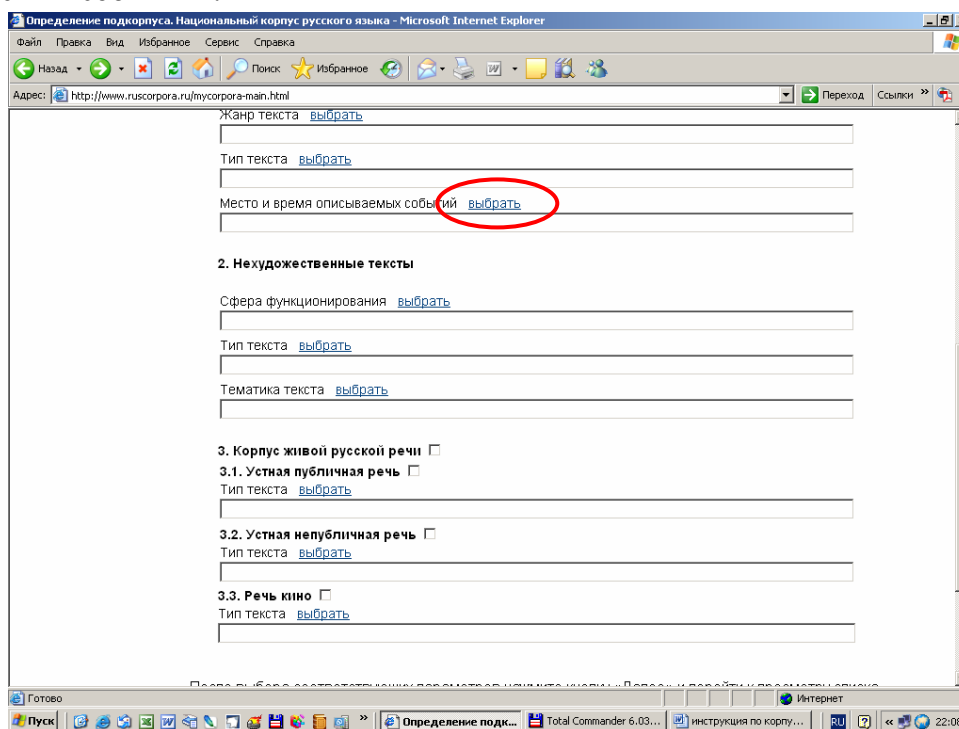


После этого пролистайте страницу вниз, найдите кнопку «Далее» (см. выше, **шаги с седьмого по четырнадцатый**) и щелкните по ней мышкой. После этого переходите к **шагу тридцатому**.

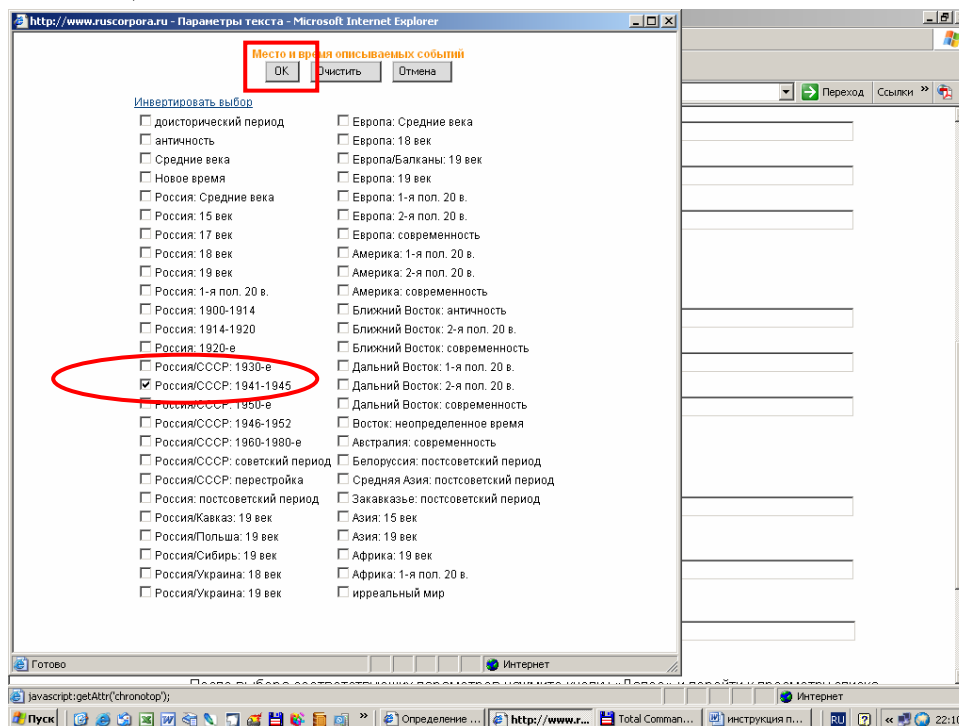
17 шаг

Семнадцатый шаг. Как выбрать тексты в соответствии с местом и временем описываемых событий?

Если Вы хотите работать только с художественными текстами, сюжет которых относится к определенному **хронотопу** (месту и времени), то Вам следует осуществить **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «Жанр и тип текста» щелкните мышкой на ссылке «Выбрать» в строке «Место и время описываемых событий»:



Перед вами откроется маленькое окошко с видами хронотопа. Выберите в нем нужный Вам хронотоп и щелкните мышкой по кнопке «ОК»:

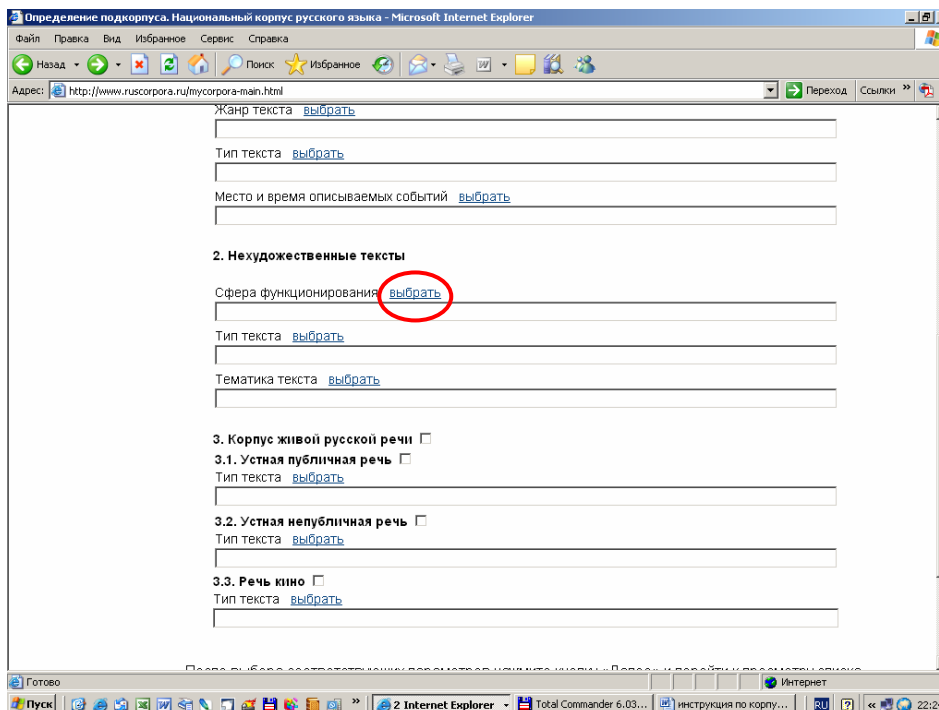


После этого пролистайте страницу вниз, найдите кнопку «Далее» (см. выше, **шаги с седьмого по четырнадцатый**) и щелкните по ней мышкой. После этого переходите к **шагу тридцатому**.

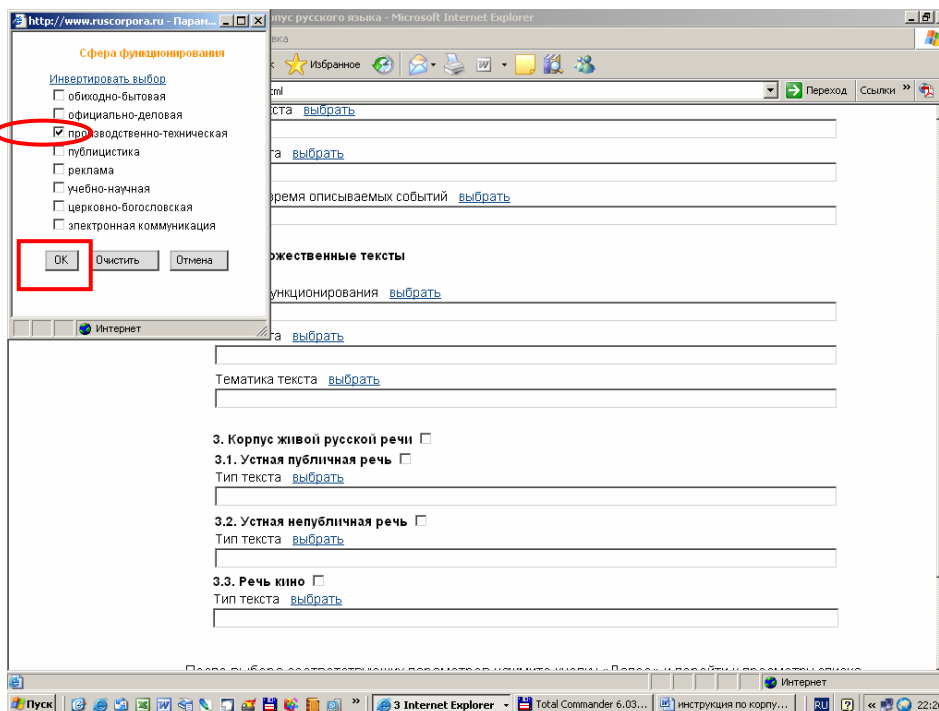
18 шаг

Восемнадцатый шаг. Как выбрать нехудожественные тексты определенной сферы функционирования?

Если Вы хотите работать только с нехудожественными текстами, относящимися к определенной сфере функционирования, то Вам следует осуществить **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «2. Нехудожественные тексты» щелкните мышкой на ссылке «Выбрать» в строке «Сфера функционирования»:



Перед Вами откроется окошко, где будут перечислены основные сферы функционирования нехудожественных текстов. Выберите сферу, которая Вас интересует и щелкните мышкой на кнопке ОК:

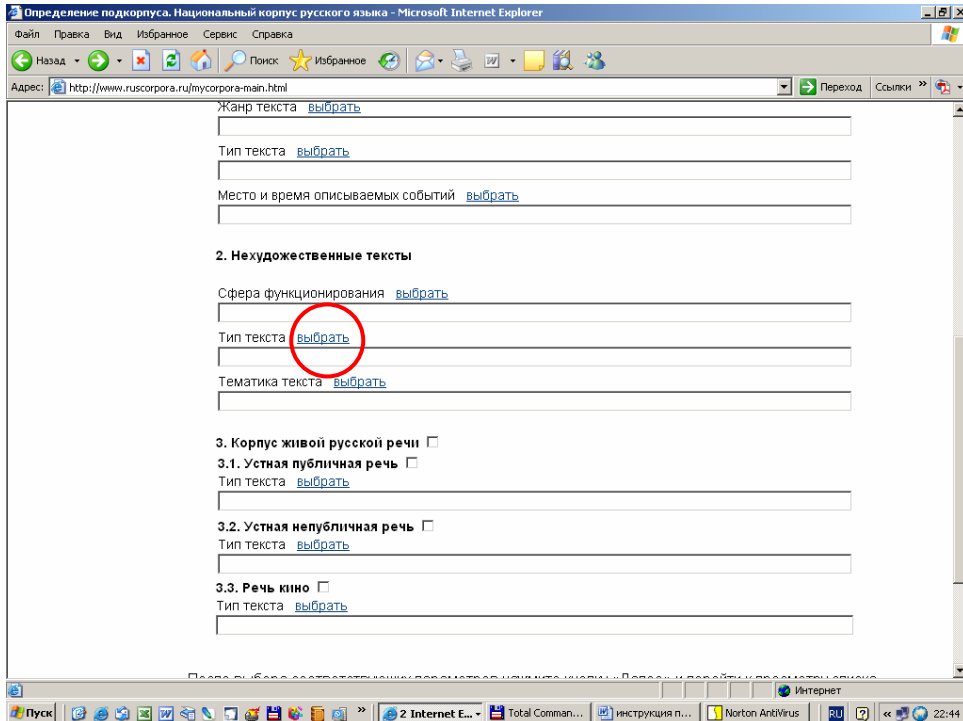


После этого пролистайте страницу вниз, найдите кнопку «Далее» (см. выше, **шаги с седьмого по четырнадцатый**) и щелкните по ней мышкой. После этого переходите к **шагу тридцатому**.

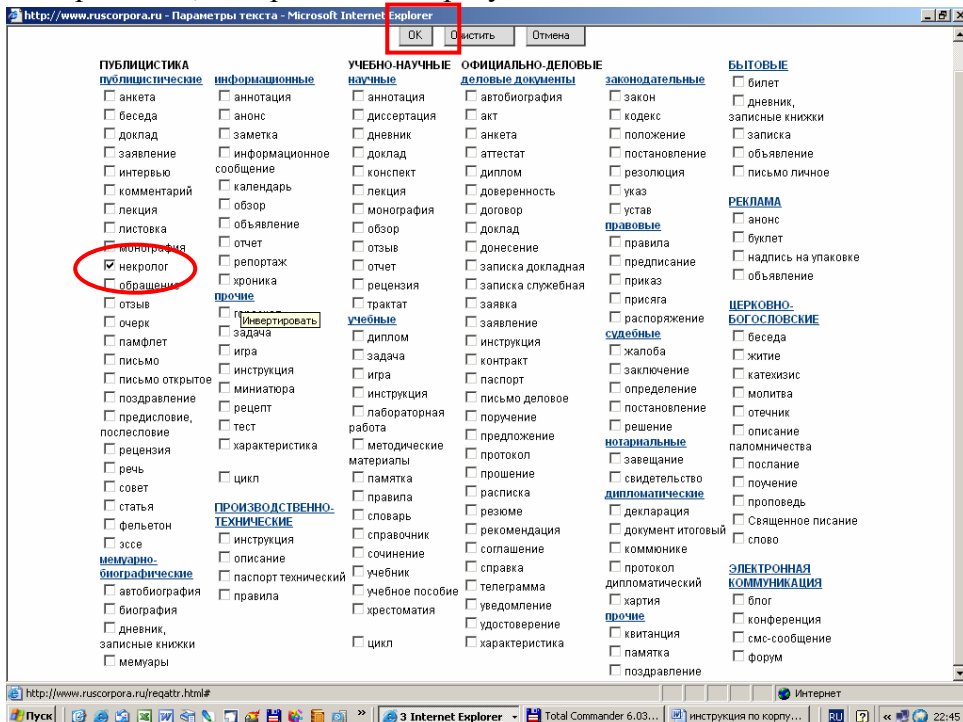
19 шаг

Девятнадцатый шаг. Как выбрать отдельный тип нехудожественного прозаического текста?

Если Вы хотите работать только с определенным типом нехудожественных текстов, то Вам следует осуществить **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «2. Нехудожественные тексты» щелкните мышкой на ссылке «Выбрать» в строке «Тип текста»:



Перед Вами откроется окно, где будут перечислены основные типы нехудожественных текстов. Выберите тип, которая Вас интересует и щелкните мышкой на кнопке ОК:

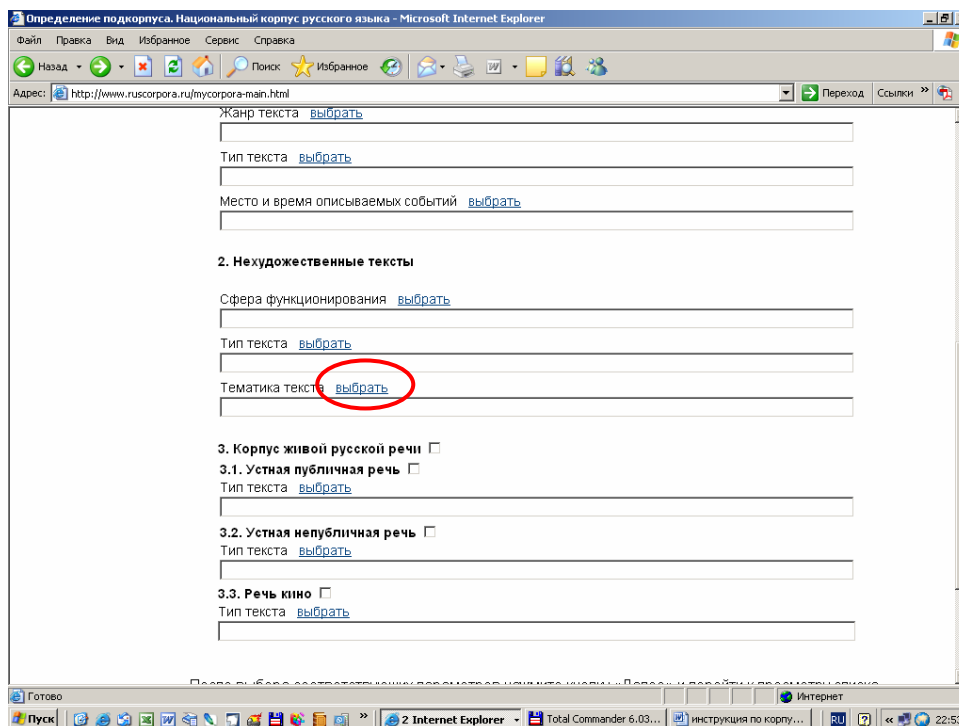


После этого пролистайте страницу вниз, найдите кнопку «Далее» (см. выше, **шаги с седьмого по четырнадцатый**) и щелкните по ней мышкой. После этого переходите к **шагу тридцатому**.

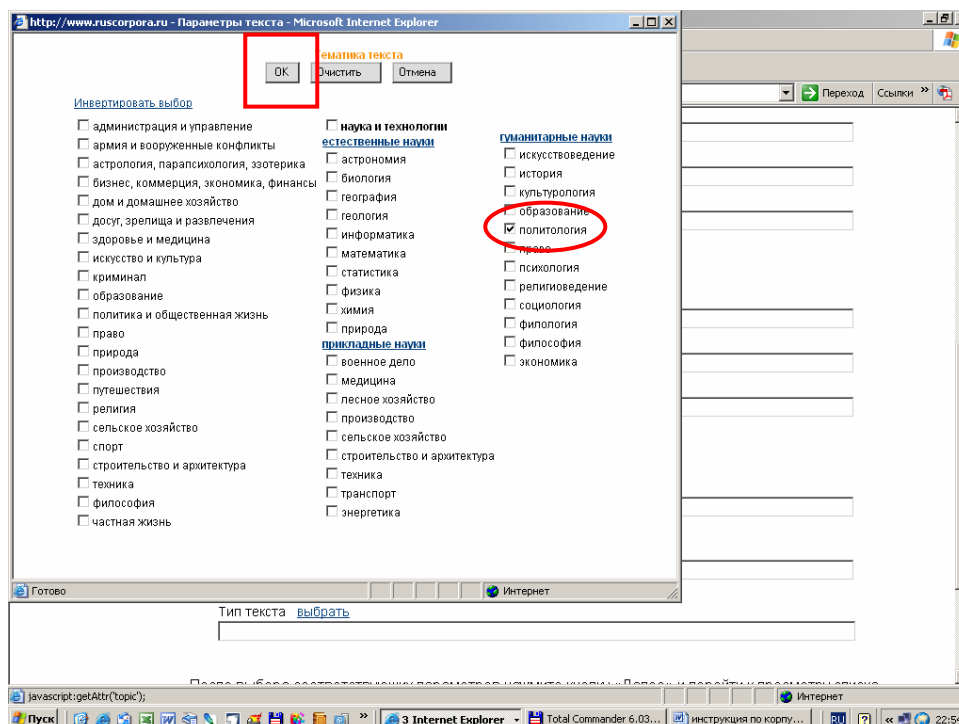
20 шаг

Двадцатый шаг. Как отобразить нехудожественные тексты по тематике?

Если Вы хотите работать с нехудожественными текстами определенной тематики, то Вам следует осуществить **шаг седьмой, часть первую**. В окошке, которое перед Вами откроется, в разделе «2. Нехудожественные тексты» щелкните мышкой на ссылке «Выбрать» в строке «Тематика текста»:



Перед Вами откроется окно, где будут перечислены основные тематические типы нехудожественных текстов. Выберите тип, которая Вас интересует и щелкните мышкой на кнопке ОК:



После этого пролистайте страницу вниз, найдите кнопку «Далее» (см. выше, **шаги с седьмого по четырнадцатый**) и щелкните по ней мышкой. После этого переходите к **шагу тридцатому**.

21 шаг

Двадцать первый шаг. Как задать подкорпус Устной публичной речи?

Вернитесь к шагу **седьмому, второй части** и поставьте галочку напротив пункта 3.1. Устная публичная речь. После этого щелкните на кнопке «Далее»:

Определение подкорпуса. Национальный корпус русского языка - Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

Назад Поиск Избранное

Адрес: <http://www.ruscorpora.ru/mycorpora-main.html> Переход Ссылки

Сфера функционирования [выбрать](#)

Тип текста [выбрать](#)

Тематика текста [выбрать](#)

3. Корпус живой русской речи

3.1. Устная публичная речь

Тип текста [выбрать](#)

3.2. Устная непубличная речь

Тип текста [выбрать](#)

3.3. Речь кино

Тип текста [выбрать](#)

После выбора соответствующих параметров нажмите кнопку «Далее» и перейти к просмотру списка документов, входящих в подкорпус. Нажав кнопку «Сохранить», пользователь может перейти к странице «Поиск в корпусе» для задания поискового запроса.

[Далее >>](#) [Очистить](#)

Национальный корпус русского языка info@ruscorpora.ru
© 2003–2006

Интернет

Пуск 2 Internet Explorer Total Commander 6.03... 2 Microsoft Office Word RU 23:08

После этого сделайте **шаг тридцатый**.

22 шаг

Двадцать второй шаг. Как задать подкорпус Устной непубличной речи?

Вернитесь к шагу **седьмому, второй части** и поставьте галочку напротив пункта 3.2. Устная непубличная речь. После этого щелкните на кнопке «Далее»:

Сфера функционирования [выбрать](#)

Тип текста [выбрать](#)

Тематика текста [выбрать](#)

3. Корпус живой русской речи

3.1. Устная публичная речь

Тип текста [выбрать](#)

3.2. Устная непубличная речь

Тип текста [выбрать](#)

3.3. Речь кино

Тип текста [выбрать](#)

После выбора соответствующих параметров нажмите кнопку «Далее» и перейти к просмотру списка документов, входящих в подкорпус. Нажав кнопку «Сохранить», пользователь может перейти к странице «Поиск в корпусе» для задания поискового запроса.

[Далее >>](#) [Очистить](#)

Национальный корпус русского языка
© 2003–2006 info@ruscorpora.ru

После этого сделайте **шаг тридцатый**.

23 шаг

Двадцать третий шаг. Как задать подкорпус «Речь кино»?

Вернитесь к шагу **седьмому, второй части** и поставьте галочку напротив пункта 3.3. Речь кино. После этого щелкните на кнопке «Далее»:

Определение подкорпуса. Национальный корпус русского языка - Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

Назад Поиск Избранное

Адрес: E:\ruscorpora.ru\mycorpora-main.html

Сфера функционирования [выбрать](#)

Тип текста [выбрать](#)

Тематика текста [выбрать](#)

3. Корпус живой русской речи

3.1. Устная публичная речь

Тип текста [выбрать](#)

3.2. Устная непубличная речь

Тип текста [выбрать](#)

3.3. Речь кино

Тип текста [выбрать](#)

После выбора соответствующих параметров нажмите кнопку «Далее» и перейти к просмотру списка документов, входящих в подкорпус. Нажав кнопку «Сохранить», пользователь может перейти к странице «Поиск в корпусе» для задания поискового запроса.

[Далее >>](#) [Очистить](#)

Национальный корпус русского языка © 2003–2006 info@ruscorpora.ru

Мой компьютер

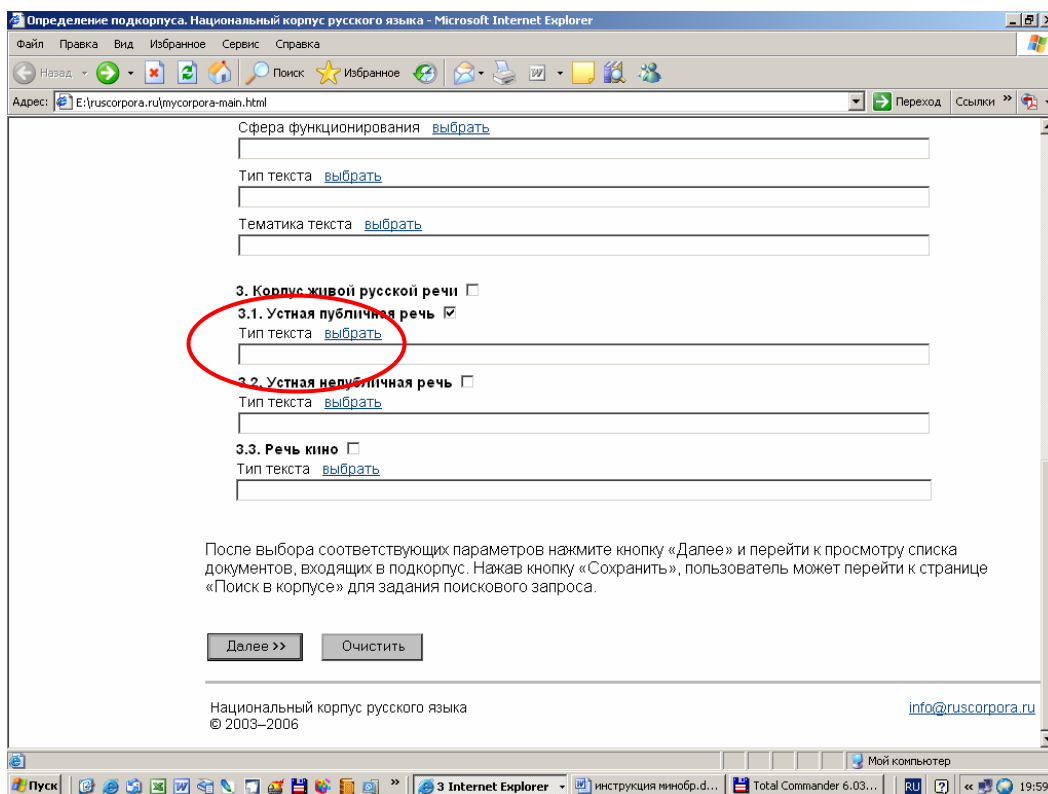
3 Internet Explorer инструкция минобр.d... Total Commander 6.03... RU 19:55

После этого сделайте **шаг тридцатый**.

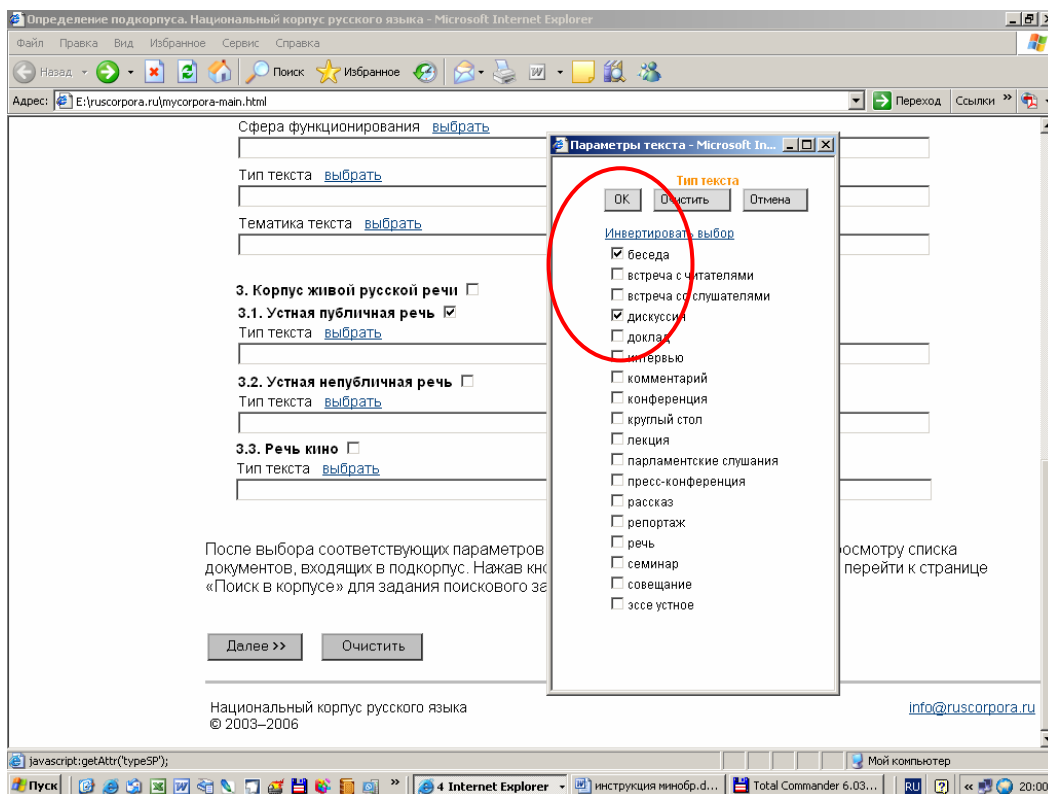
24 шаг

Двадцать четвертый шаг. Как выбрать тот или иной тип текста в Устной публичной речи?

Вернитесь к **шагу седьмому** и щелкните мышкой на ссылке «Выбрать тип текста» в пункте 3.1. Устная публичная речь:



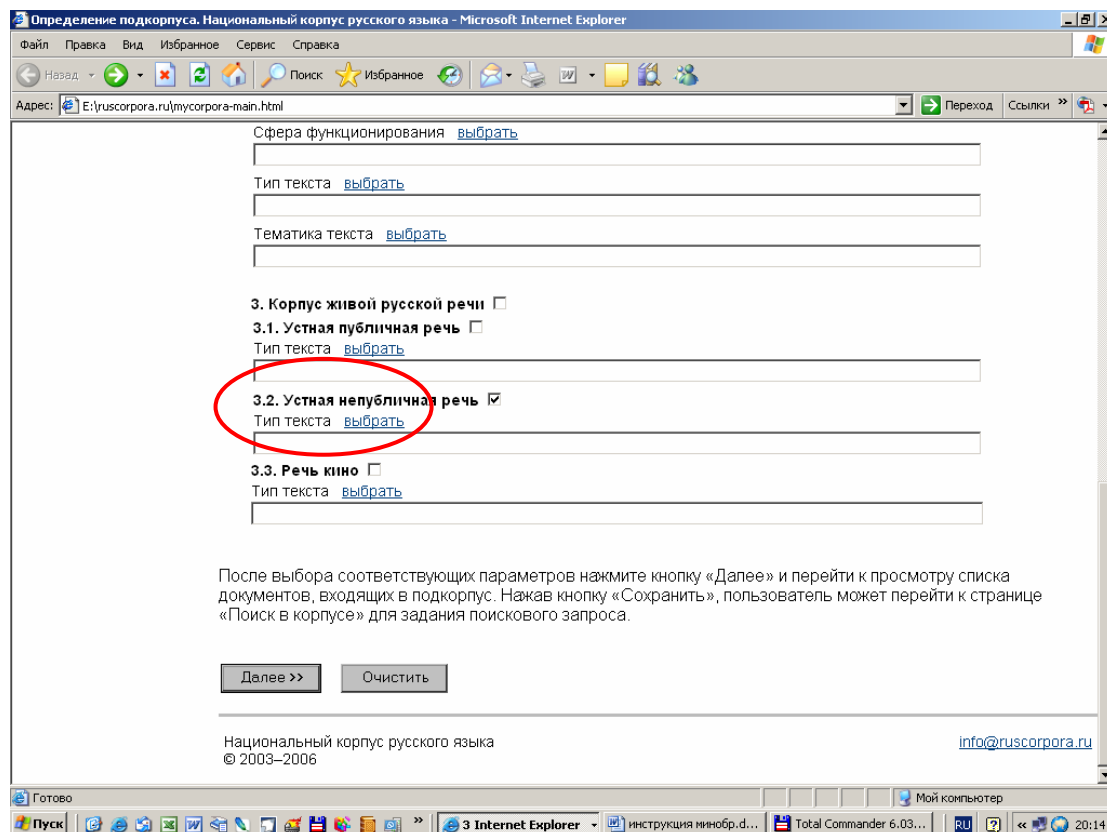
Перед Вами откроется дополнительное окно, где Вы сможете выбрать нужный Вам тип текста:



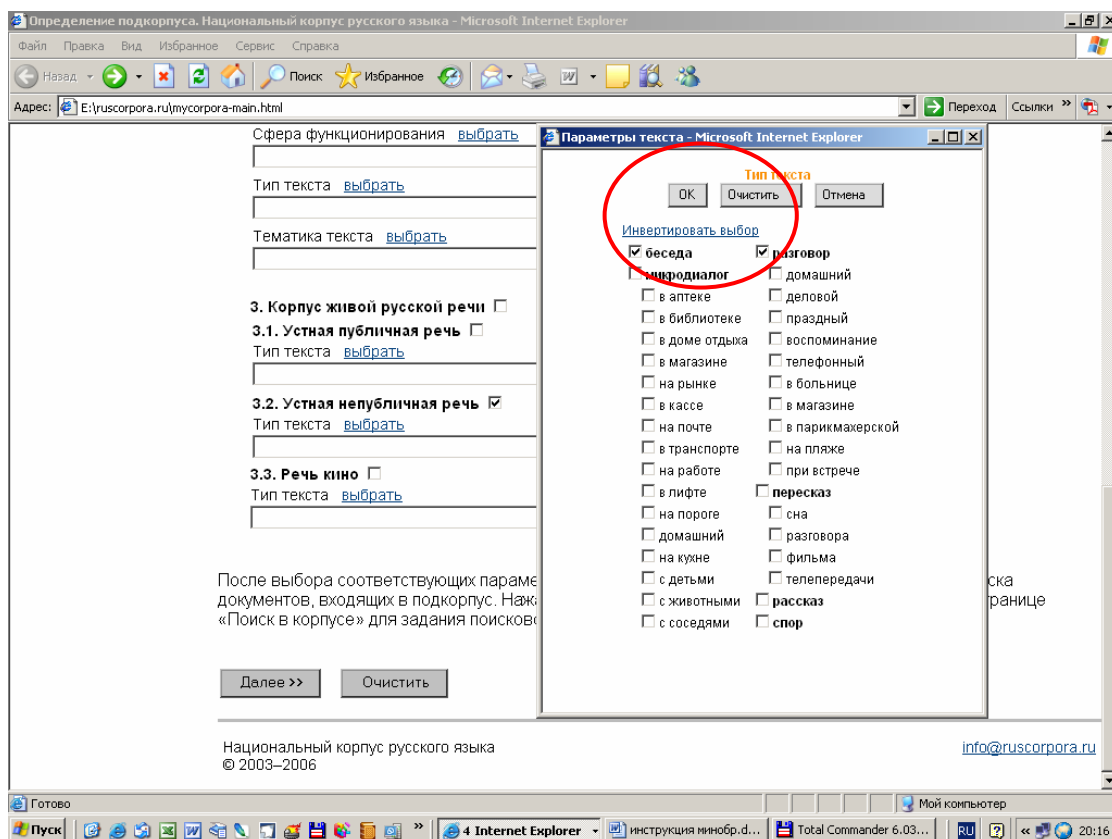
После этого щелкните на кнопке ОК и сделайте **шаг тридцатый**.

25 шаг

Двадцать пятый шаг. Как выбрать тот или иной тип текста в Устной непубличной речи? Вернитесь к **шагу седьмому** и щелкните мышкой на ссылке «Выбрать тип текста» в пункте 3.2. Устная непубличная речь:



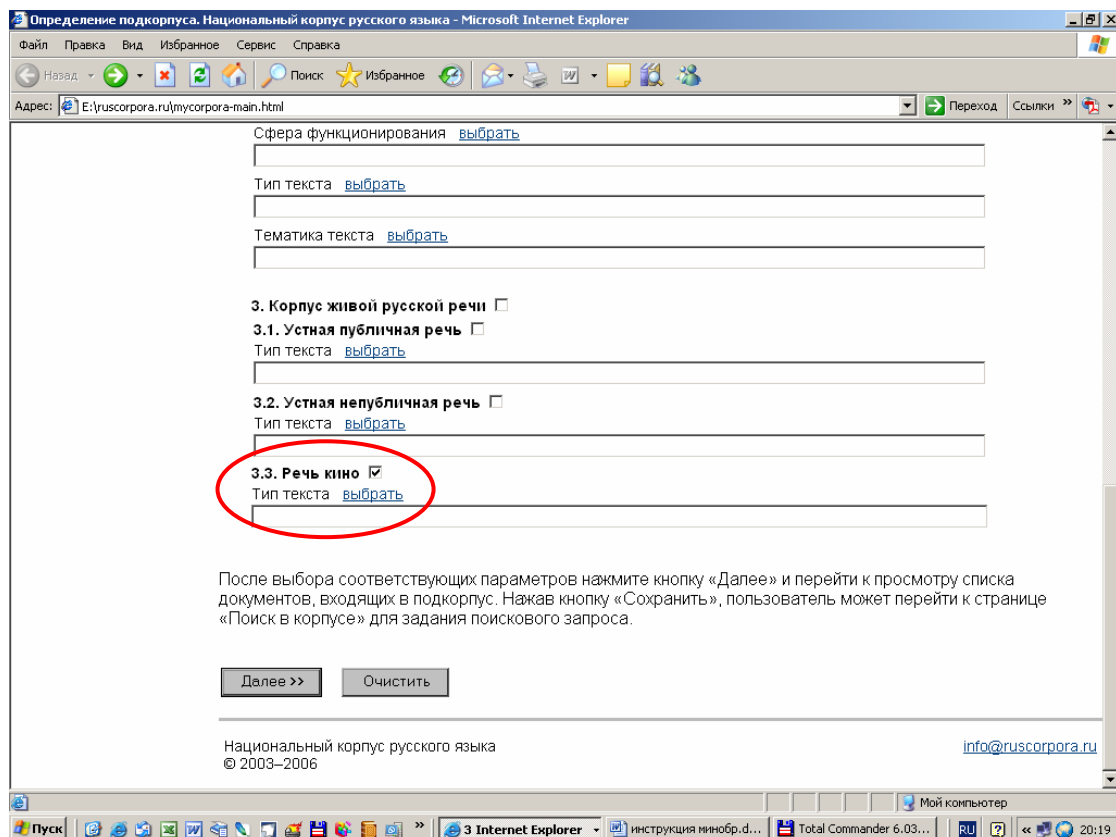
Перед Вами откроется дополнительное окно, где Вы сможете выбрать нужный Вам тип текста:



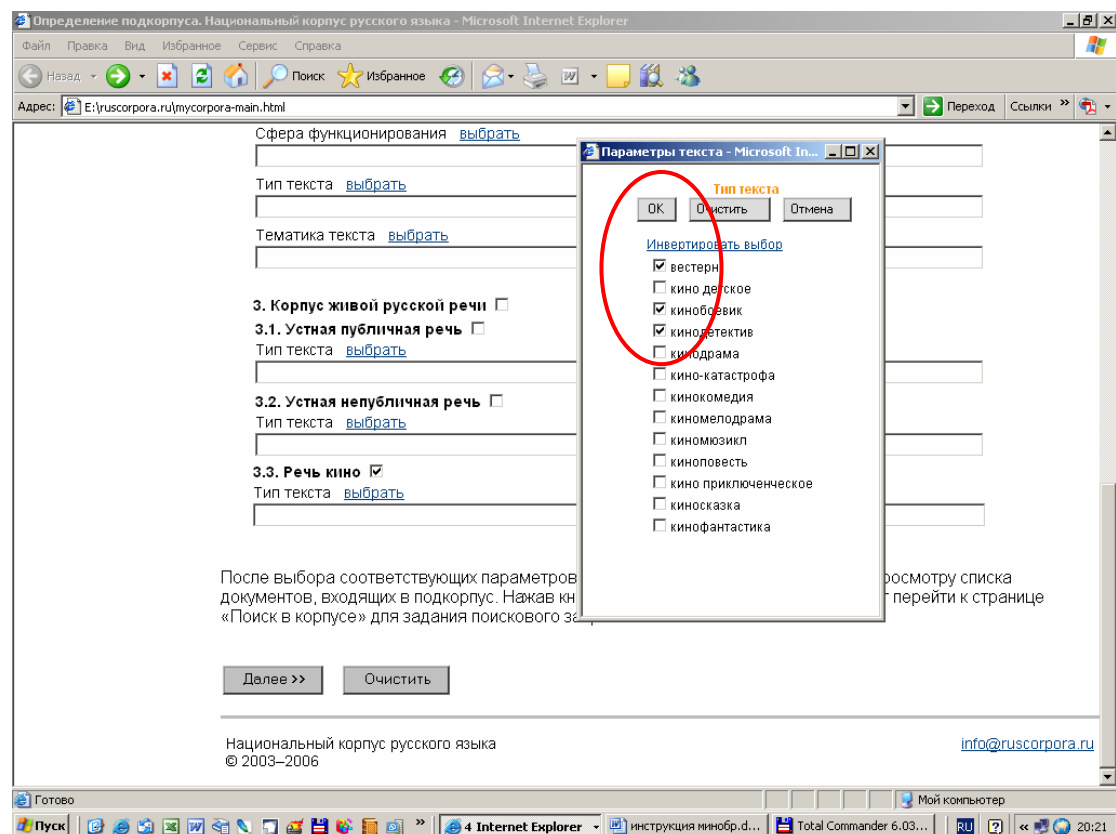
После этого щелкните на кнопке ОК и сделайте **шаг тридцатый**.

26 шаг

Двадцать шестой шаг. Как выбрать тот или иной тип текста в подкорпусе «Речь кино»? Вернитесь к **шагу седьмому** и щелкните мышкой на ссылке «Выбрать тип текста» в пункте 3.3. Речь кино:



Перед Вами откроется дополнительное окно, где Вы сможете выбрать нужный Вам тип текста:



После этого щелкните на кнопке ОК и сделайте **шаг тридцатый**.

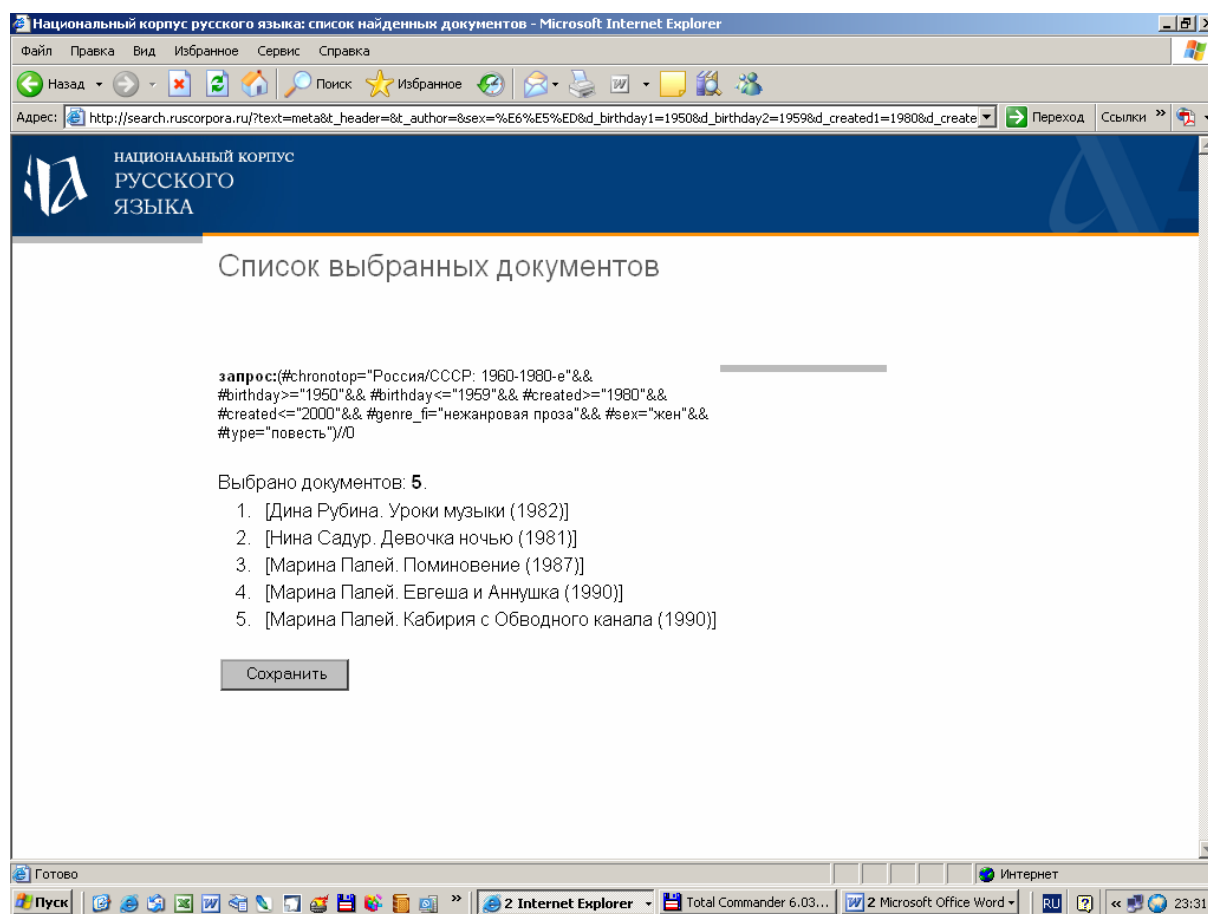
27 шаг

Двадцать седьмой шаг. Как отобрать художественные тексты по совокупности признаков?

Пройдите последовательно все **шаги с одиннадцатого по тринадцатый** и с **пятнадцатого по семнадцатый**.

Например, отберите все 1) художественные прозаические тексты, 2) написанные женщинами, 3) родившимися в 1950-е годы; 4) при этом сами тексты должны быть написаны в 1980-1990-е годы, 5) по жанру являться **нежанровой прозой** (т.е. прозой, которую трудно подвести под какие-то жанровые определения), 6) по типу текста – повестями, а 7) хронотоп их должен быть «Россия/СССР: 1960-1980-е».

Список отобранных в результате этой операции текстов Вы сможете увидеть после того, как, выбрав соответствующие параметры так, как описано выше, щелкнете мышкой на кнопке «Далее»:



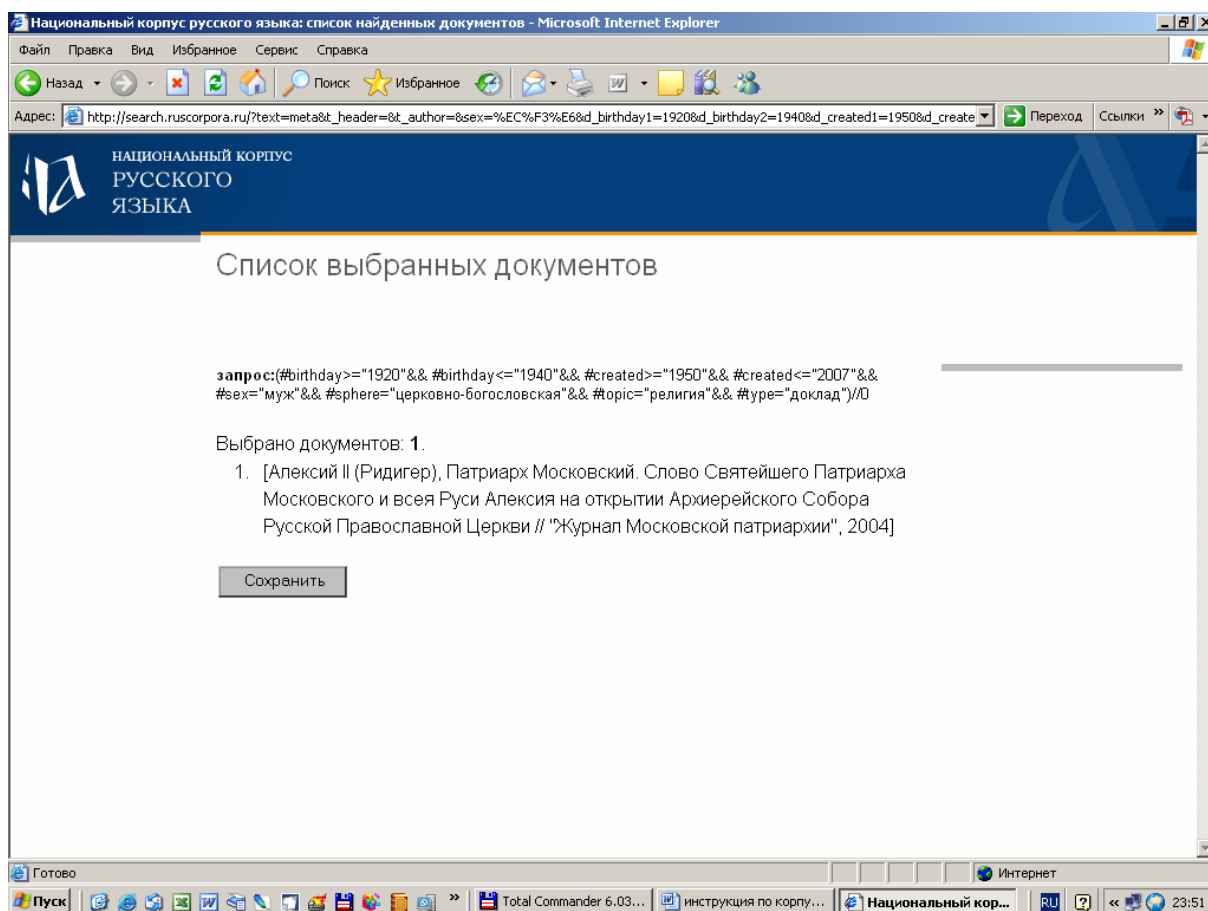
Двадцать восьмой шаг. Как отобразить нехудожественные тексты по совокупности признаков?

Пройдите последовательно все **шаги с восемнадцатого по двадцатый**.

Используйте также **шаги с восьмого по тринадцатый**.

Например, отберите 1) все нехудожественные прозаические тексты, 2) написанные мужчинами, 3) родившимися в 1920-1930-е годы; 4) при этом сами тексты должны относиться к церковно-богословской сфере функционирования, 5) по типу являться докладом, 6) а по тематике относиться к религиоведению.

Список отобранных в результате этой операции текстов Вы сможете увидеть после того, как, выбрав соответствующие параметры так, как описано выше, щелкнете мышкой на кнопке «Далее»:



29 шаг

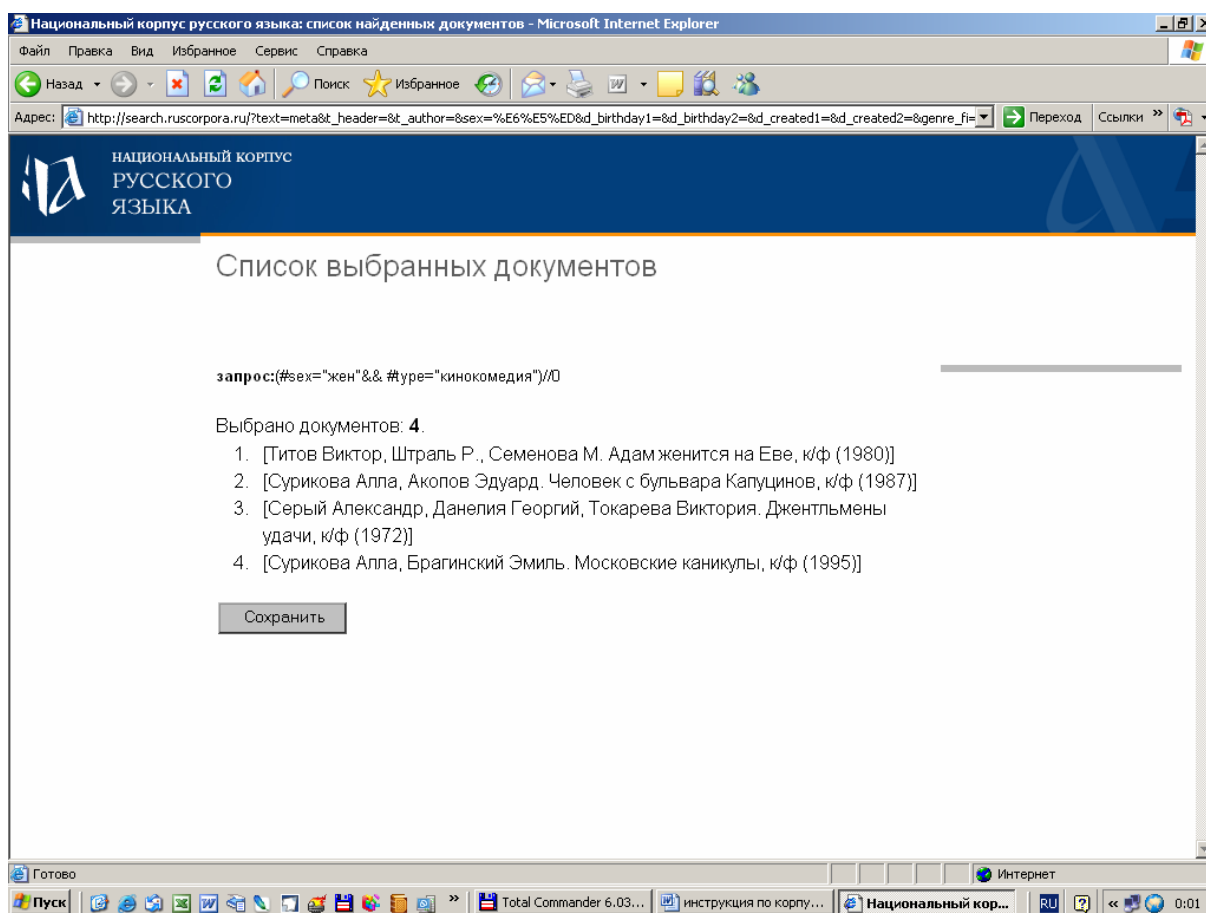
Двадцать девятый шаг. Как выбрать тексты из подкорпуса живой русской речи по совокупности признаков?

Вернитесь к **шагу седьмому** и, в зависимости от поставленной Вами перед собой задачи, пройдите **шаги с двадцать первого по двадцать шестой**.

Используйте также **шаги с восьмого по тринадцатый**.

Например, отберите все кинофильмы, **транскрипты** (транскрипт фильма – запись реплик персонажей фильма вместе с небольшим количеством необходимых ремарок) которых входят в Корпус, а 1) их режиссерами или авторами сценариев являются женщины; 2) при этом сами кинофильмы по жанру должны быть кинокомедиями.

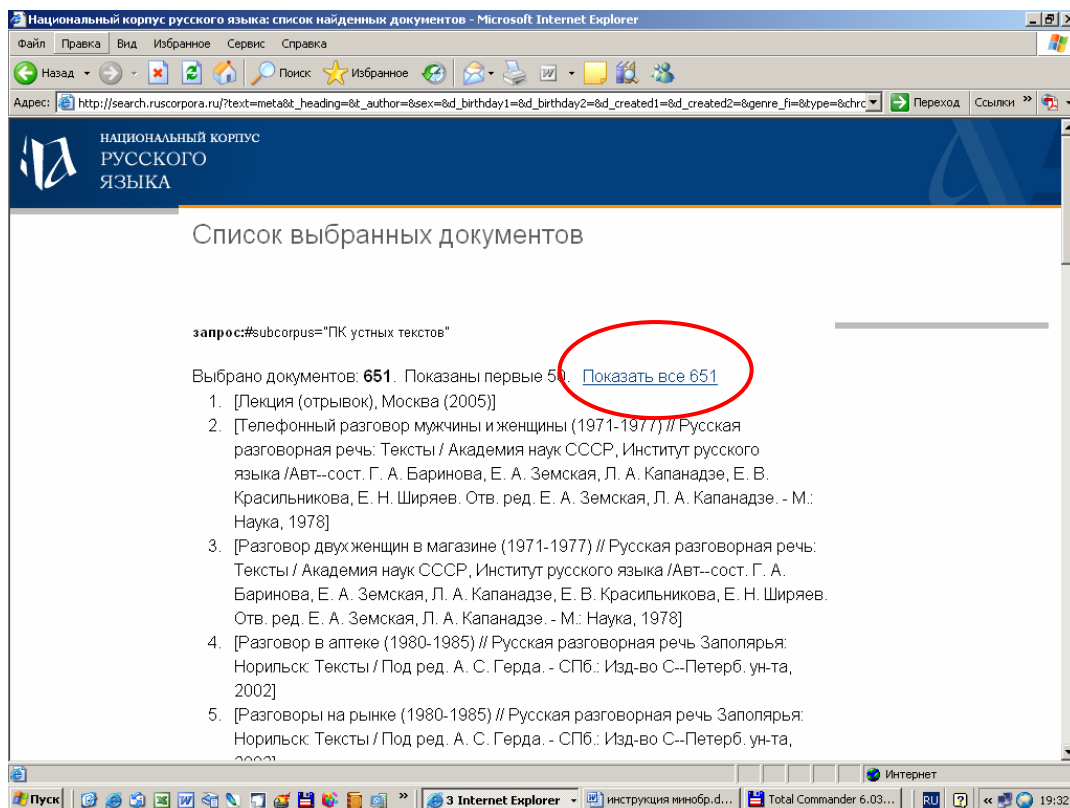
Список отобранных в результате этой операции текстов Вы сможете увидеть после того, как, выбрав соответствующие параметры так, как описано выше, щелкнете мышкой на кнопке «Далее»:



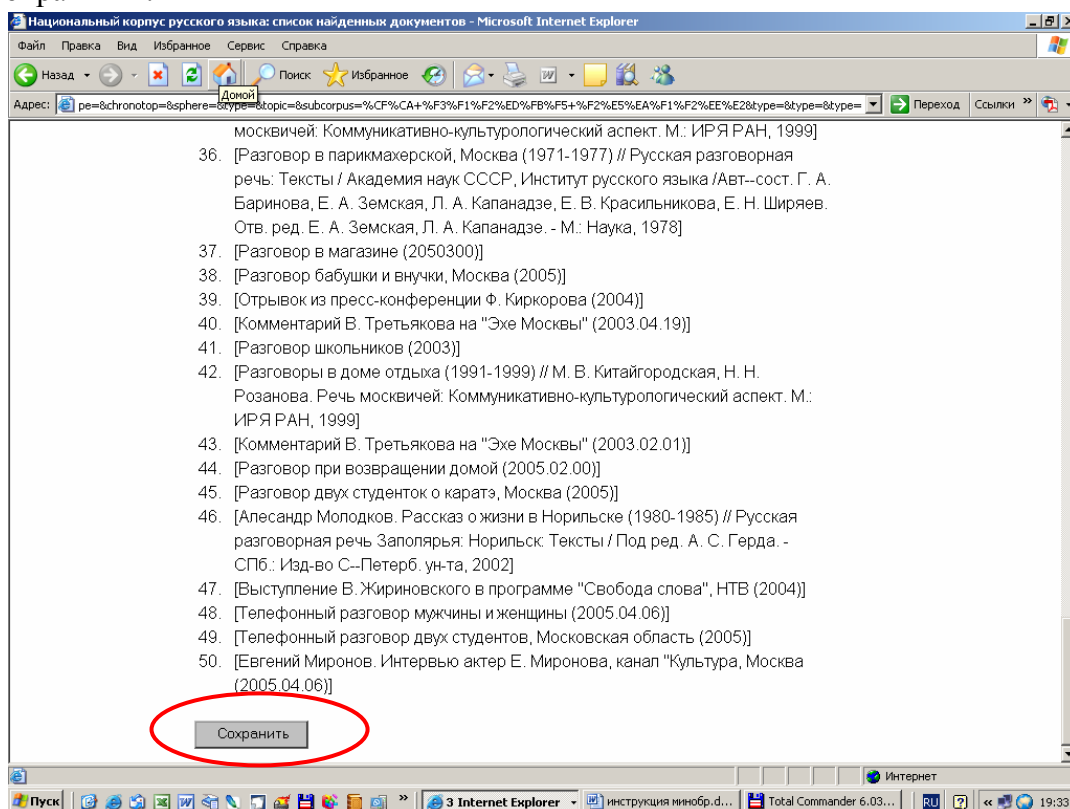
30 шаг

Тридцатый шаг. Как сохранить Ваш подкорпус для последующей работы?

После того, как Вы осуществили **шаги**, начиная с **седьмого** и заканчивая **двадцать девятым**, перед Вами открывается список документов, который соответствует Вашему запросу. Если Вы хотите просмотреть список полностью, то щелкните мышкой на кнопке «Показать все...»



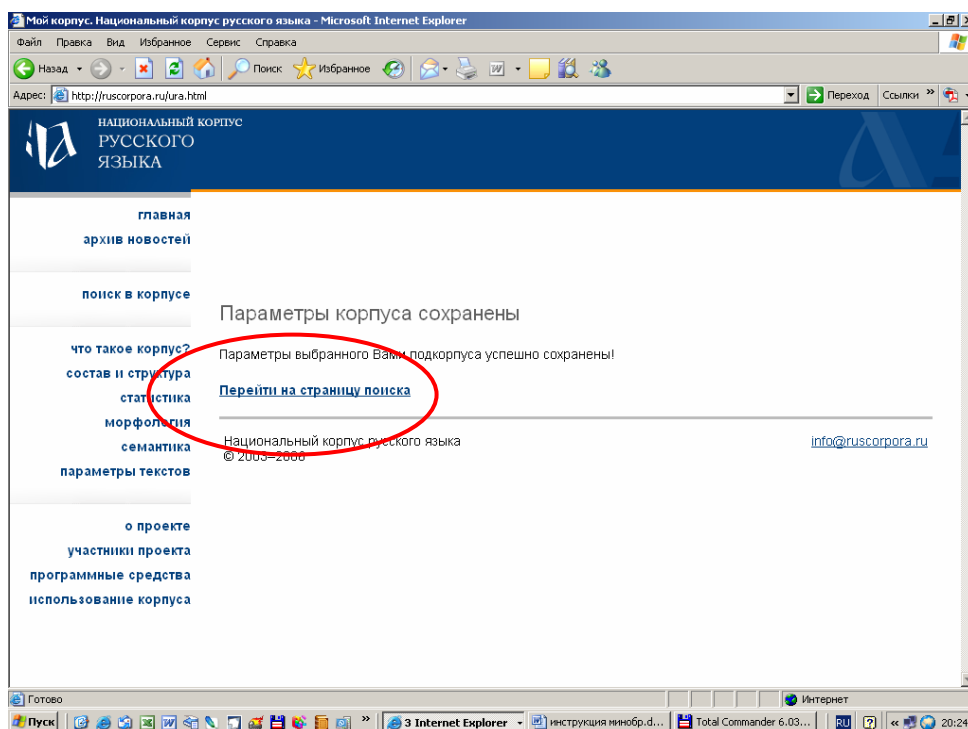
Чтобы сохранить отобранный Вами подкорпус, пролистайте страницу вниз и щелкните мышкой на кнопке «Сохранить»:



31 шаг

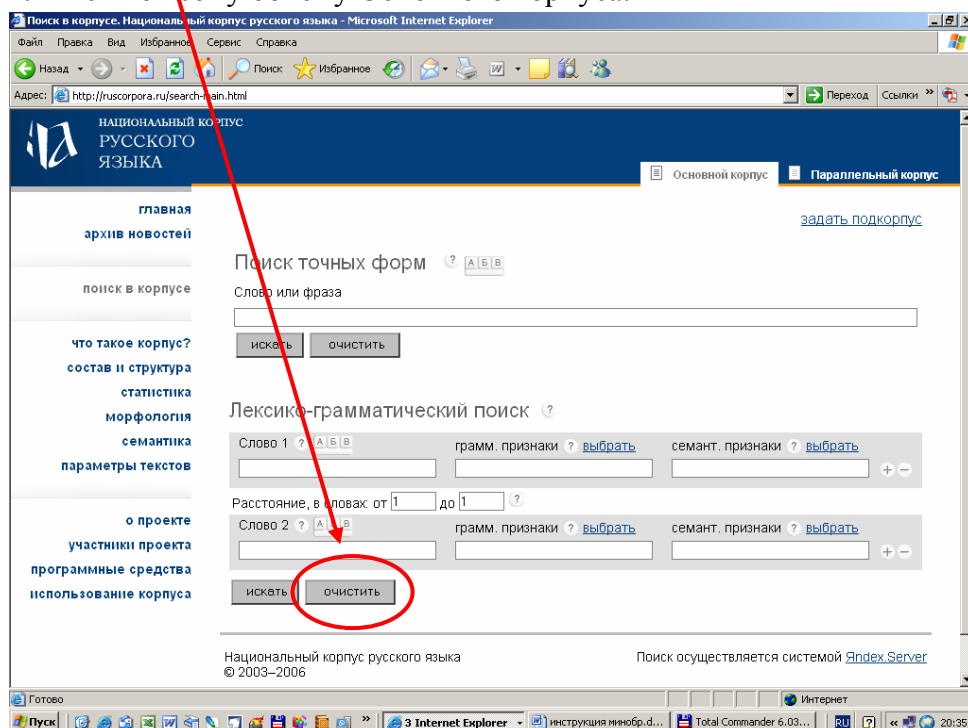
Тридцать первый шаг. Как вернуться на страницу поиска?

После того, как Вы сделали **тридцатый шаг**, т.е. щелкнули на кнопке «Сохранить», перед Вами открывается следующее окно:



Чтобы вернуться на страницу поиска, щелкните на ссылке «Перейти на страницу поиска», и перед Вами снова откроется окно поиска (см. **шаг второй**). Теперь все Ваши вопросы и требования будут адресованы к тому подкорпусу, который Вы задали в ходе **шагов с седьмого по двадцать девятый**.

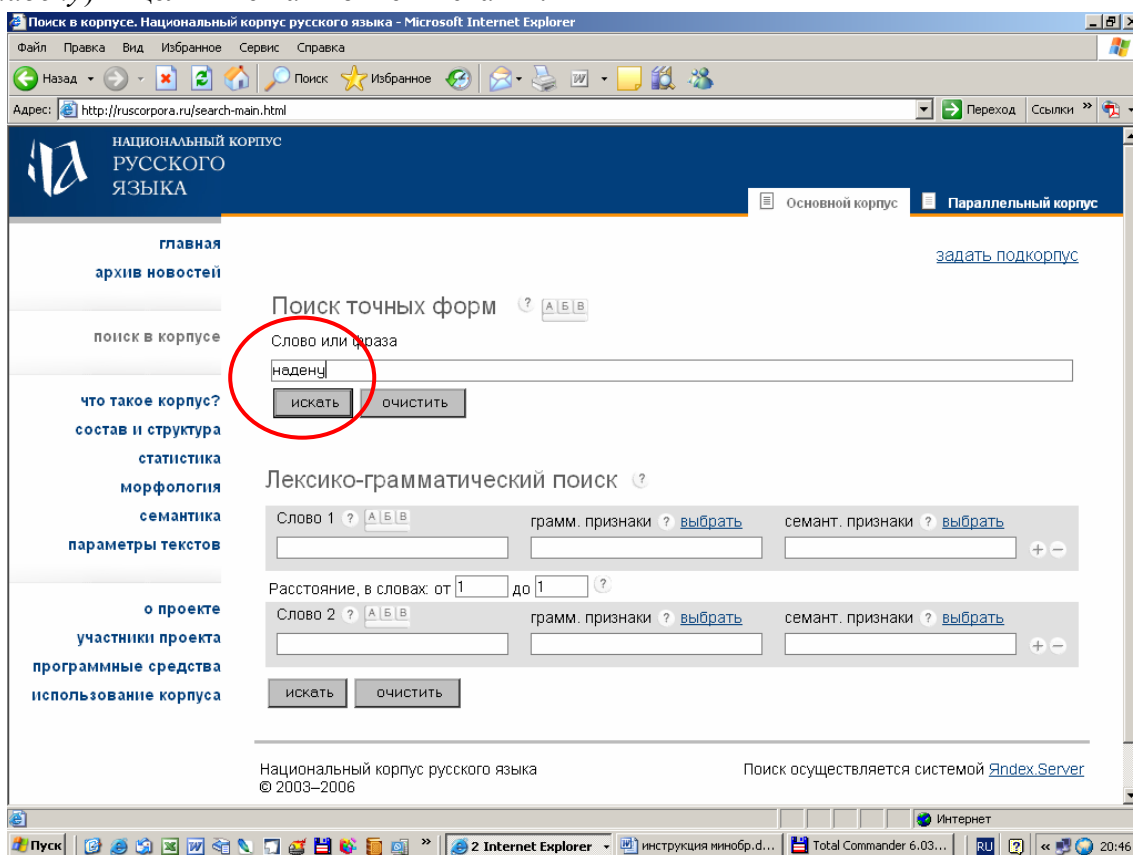
Если Вас перестал устраивать тот подкорпус, который Вы сформировали, Вы должны щелкнуть мышкой по кнопке «Очистить» на странице поиска. Если Вы хотите сформировать новый подкорпус, то Вам снова придется осуществить некоторые **шаги с седьмого по двадцать девятый**. Если Вы этого не сделаете, то весь дальнейший поиск после использования кнопки «Очистить» будет осуществляться по всему объему Основного корпуса.



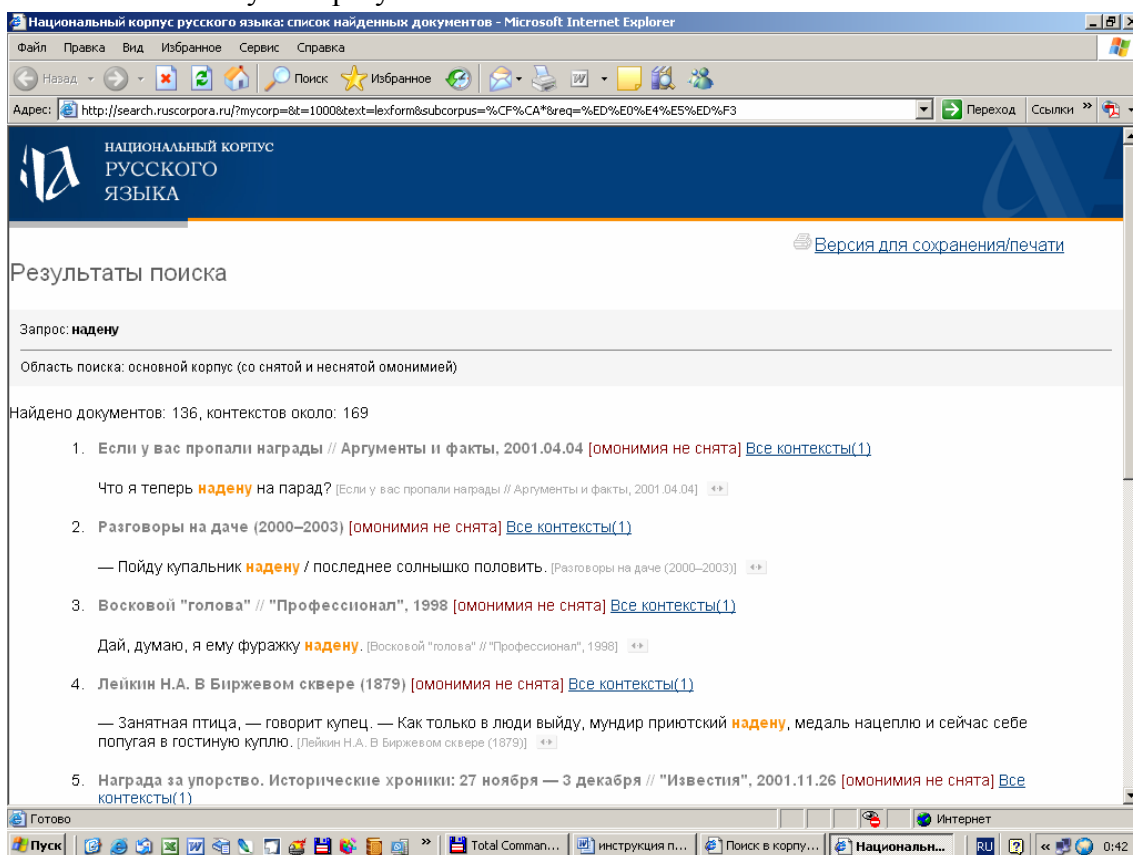
32 шаг

Тридцать второй шаг. Как найти словоформу в Корпусе?

В окне поиска в строке «Поиск точных форм» наберите нужную словоформу (например, словоформу *надену*) и щелкните на кнопке «Искать»:



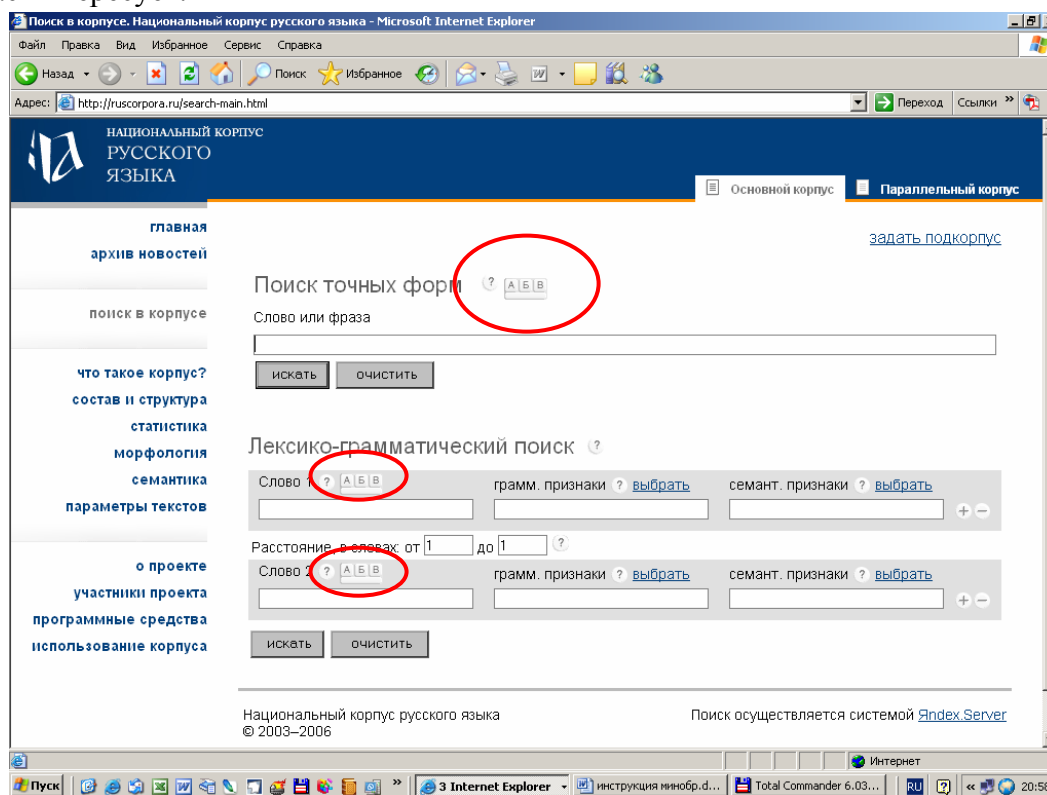
В отдельном окне Вы получите результаты поиска:



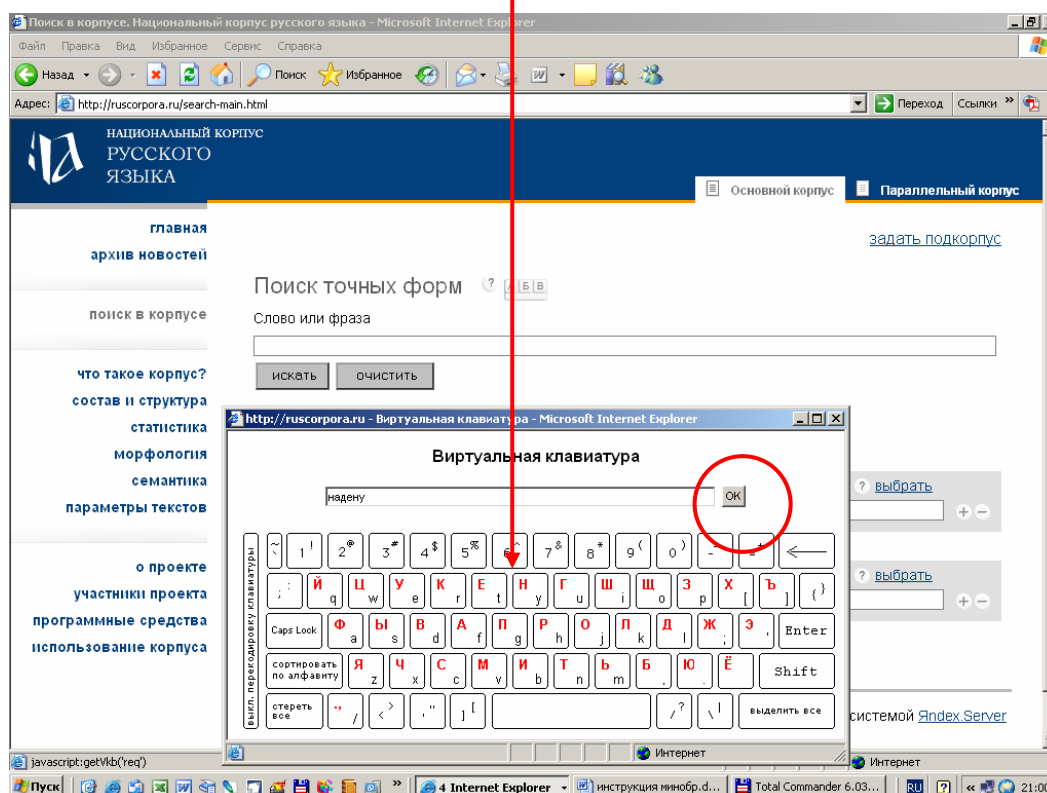
Технические замечания

Три технических замечания. Чтобы Вы могли с большей легкостью осуществлять дальнейшие шаги, мы должны снабдить Вас дополнительной информацией технического характера.

① Во-первых, как Вам поступать, если у Вас по какой-то причине нет русской клавиатуры (например, Вы живете или работаете не в России)? В этом случае Вы можете набирать слово прямо с экрана. Для этого щелкните мышкой на буквах АБВ рядом с тем полем, где Вы печатаете слово, которое Вас интересует:

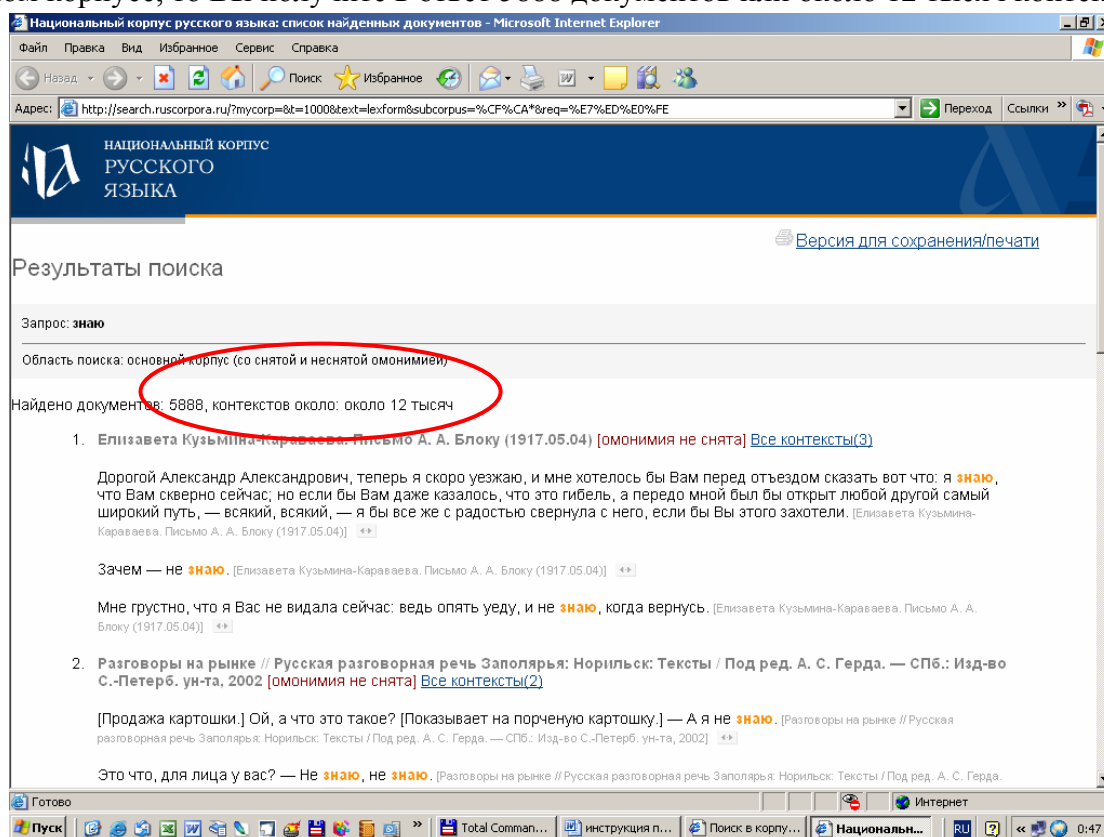


После этого перед Вами появится **виртуальная клавиатура**, на которой Вы можете, щелкая мышкой на соответствующих буквах, набрать нужное слово или словосочетание, щелкнуть на кнопке ОК, и слово появится в нужной строке (см. **шаг тридцать второй**):

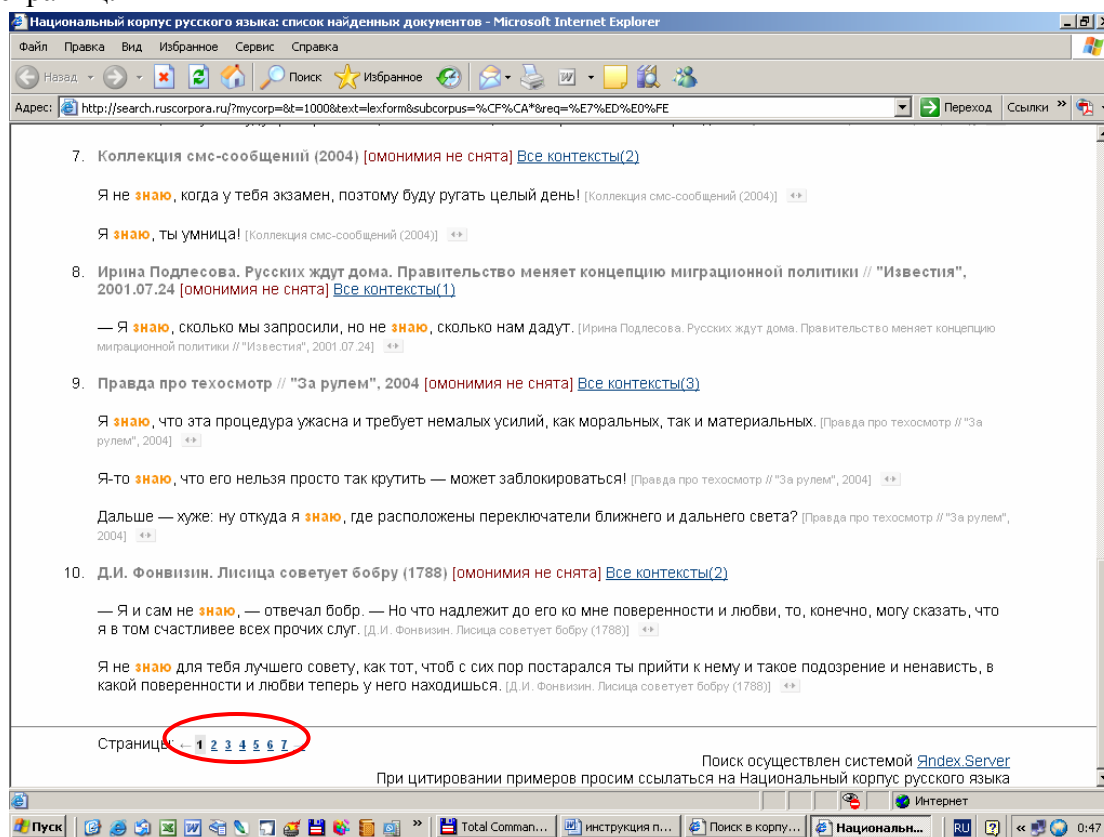


Технические замечания

② Во-вторых, как поступить, если на Ваш запрос Вы получили столько ответов, что они не умещаются на экране? Например, если Вы хотите найти все случаи употребления словоформы «знаю» в Основном корпусе, то Вы получите в ответ 5888 документов или около 12 тысяч контекстов:

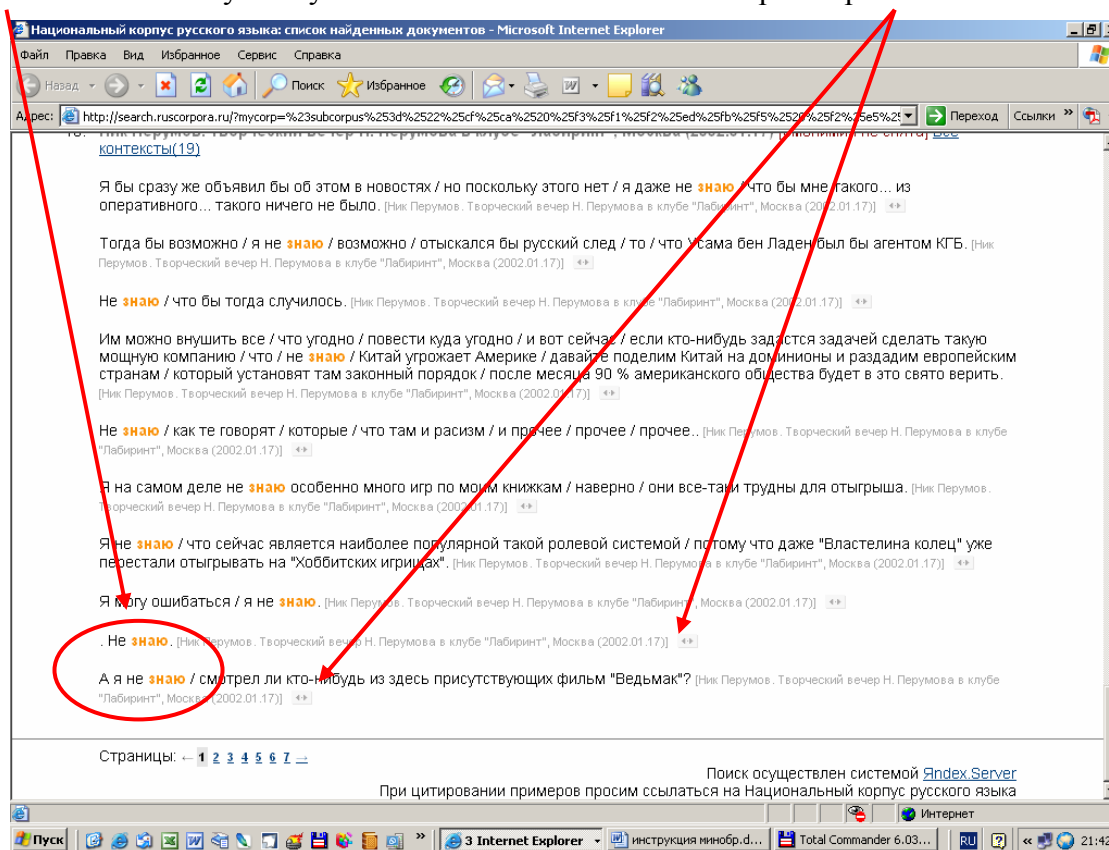


Чтобы просмотреть примеры, пролистайте страницу вниз и щелкайте мышкой на номерах следующих страниц:

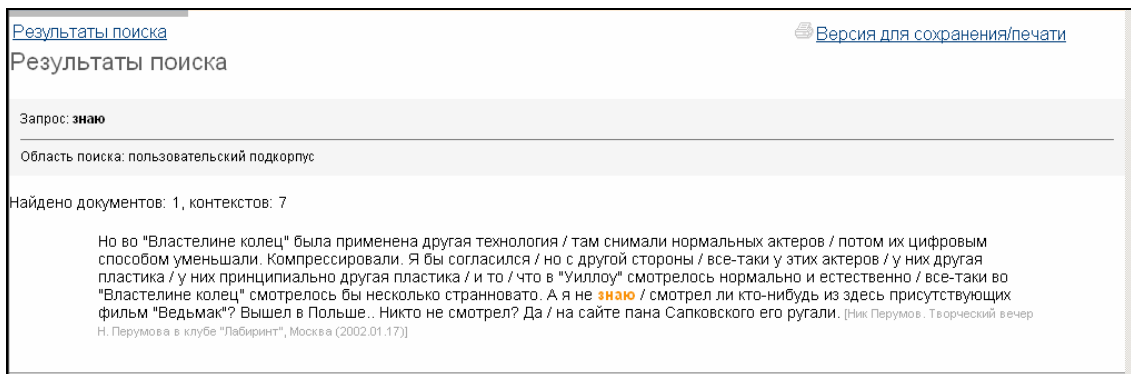
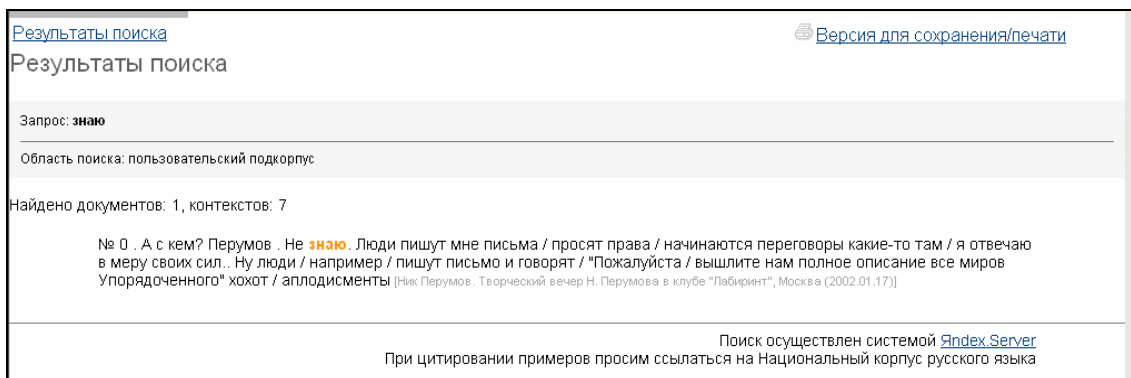


Технические замечания

③ В-третьих, что делать, если предложение, которое Вы получили в ответ на свой запрос, слишком короткое? В этом случае нужно воспользоваться кнопкой «расширение контекста»:



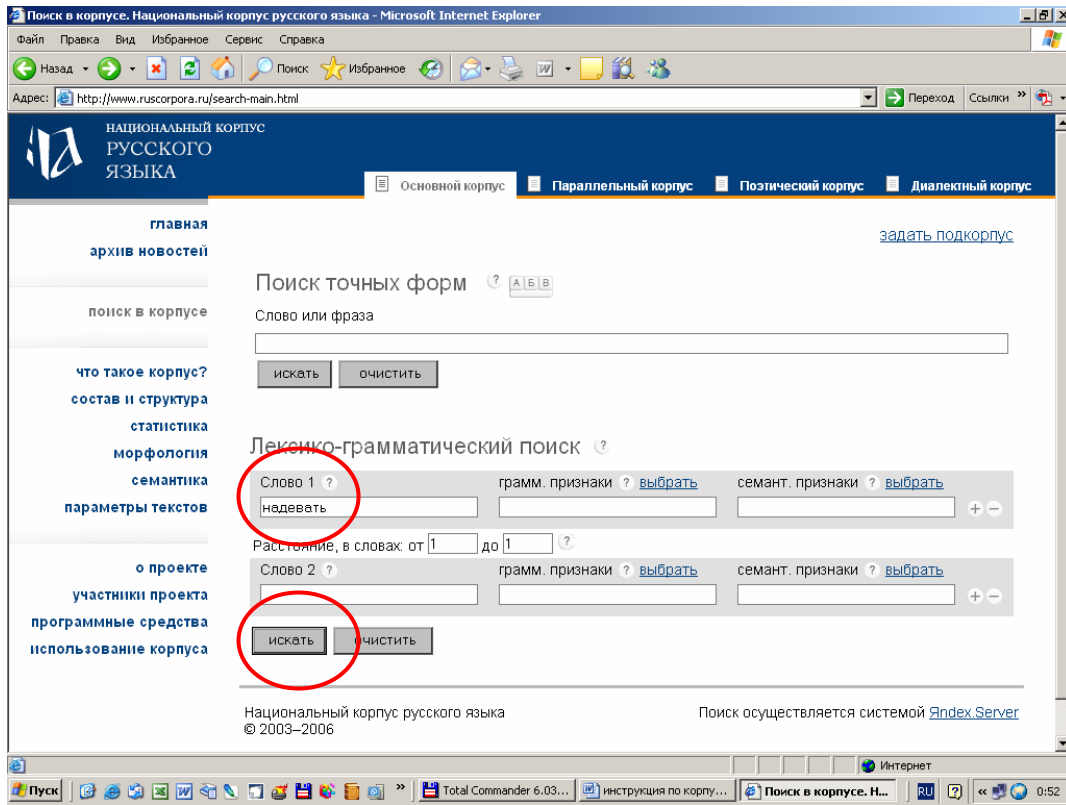
Тогда короткая фраза попадет в более широкий контекст и, соответственно, станет более понятной:



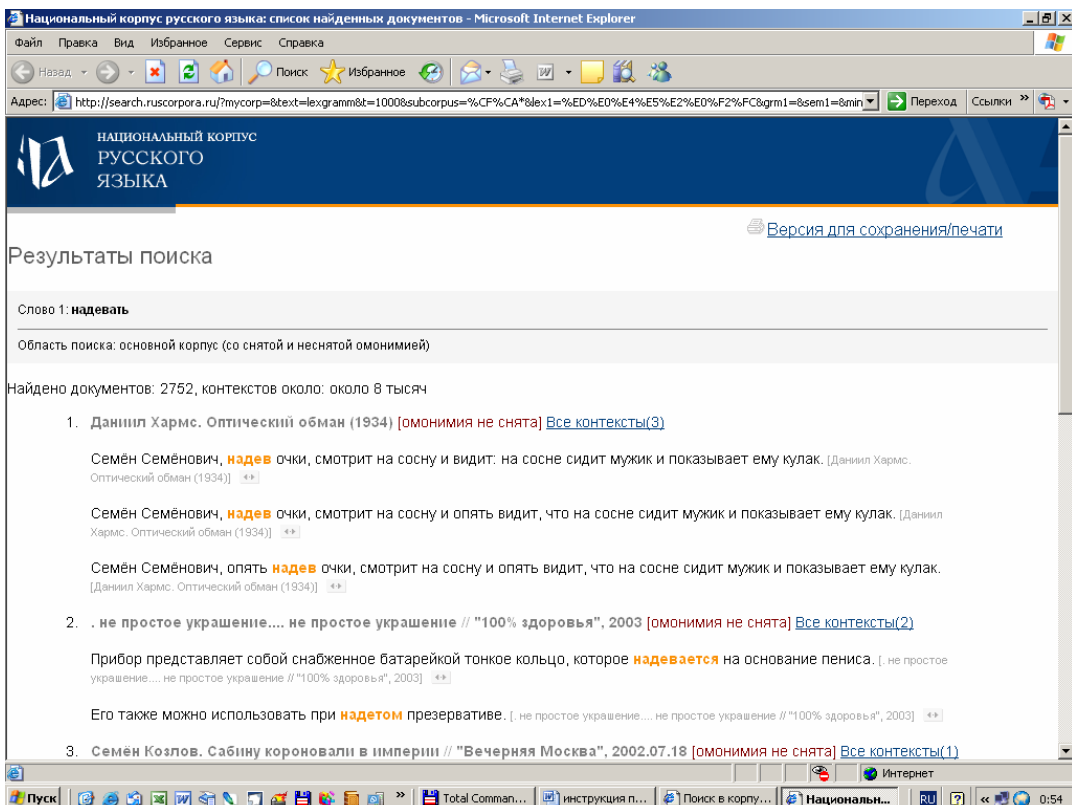
33 шаг

Тридцать третий шаг. Как найти в Корпусе отдельную лексему?

На странице поиска в зоне «Лексико-грамматический поиск» в поле «Слово 1» наберите нужное слово и щелкните на кнопке «Искать» (обратите внимание на то, что лексему нужно набирать в основной, словарной форме, т.е. *надевать*, а не, например, *надену*):

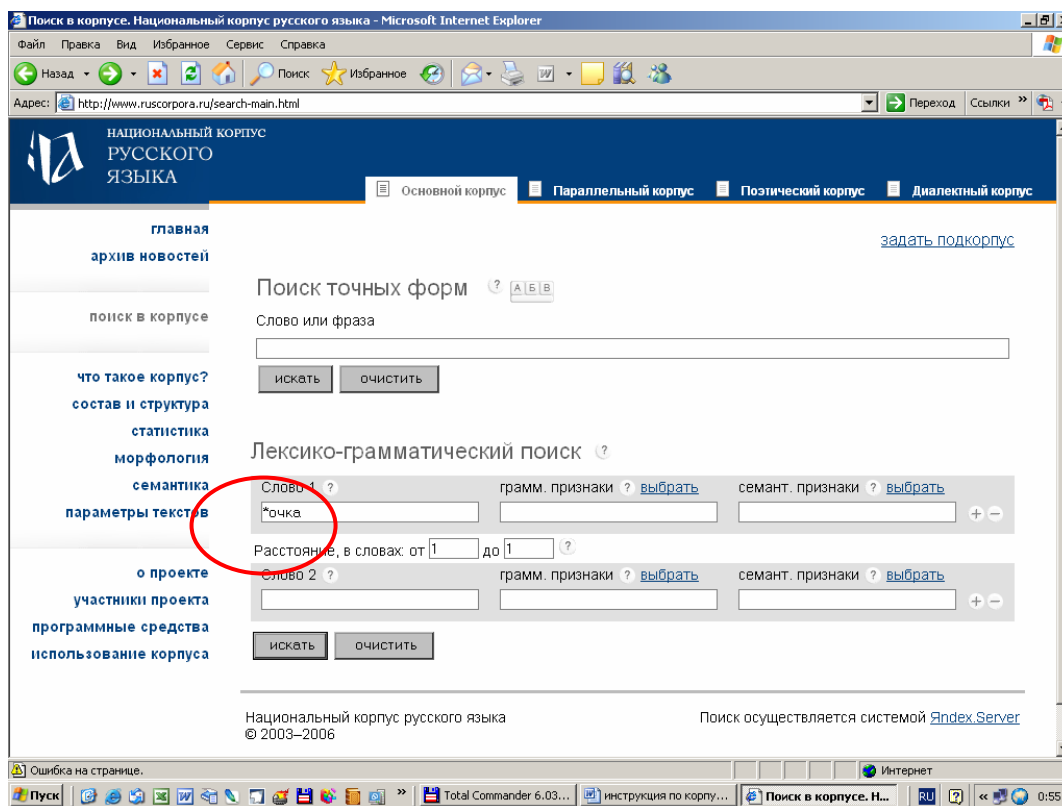


В ответ на этот запрос Вы получите все словоформы, относящиеся к лексеме *надевать* и входящие в Корпус (Основной или Ваш личный, пользовательский):

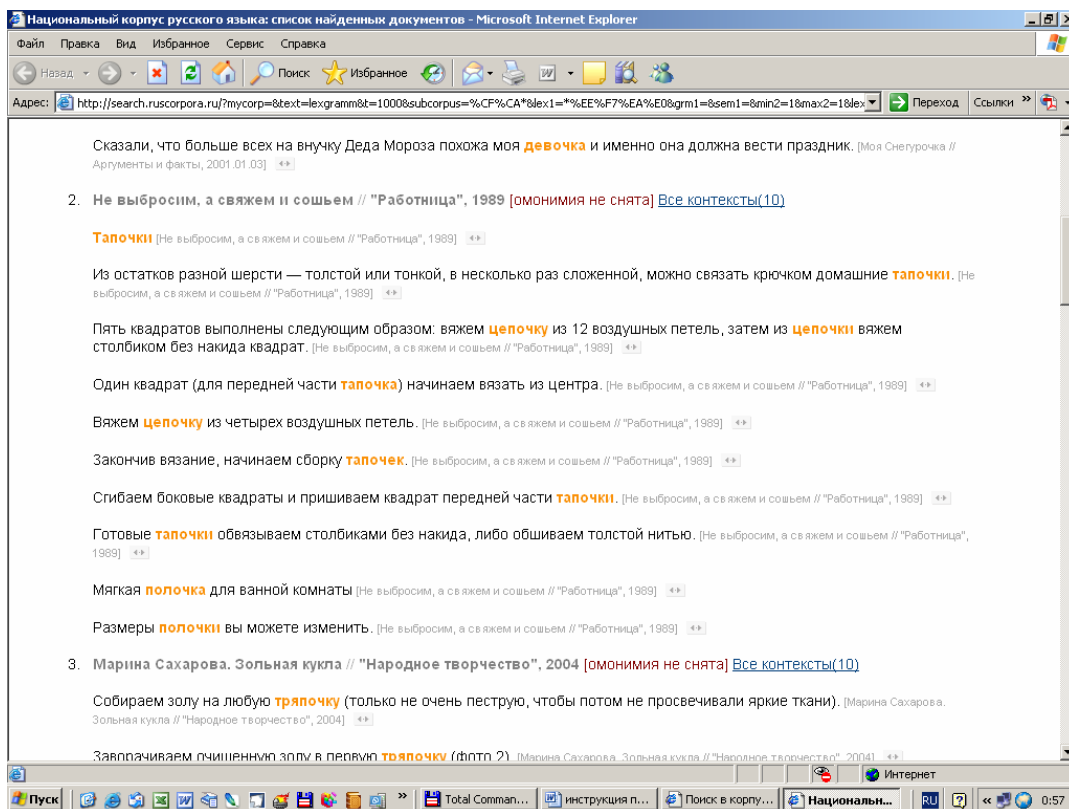


34 шаг

Тридцать четвертый шаг. Как найти в Корпусе все лексемы, имеющие одинаковую конечную часть? В зоне «Лексико-грамматический поиск» нужно набрать эту конечную часть после знака «звездочка»:

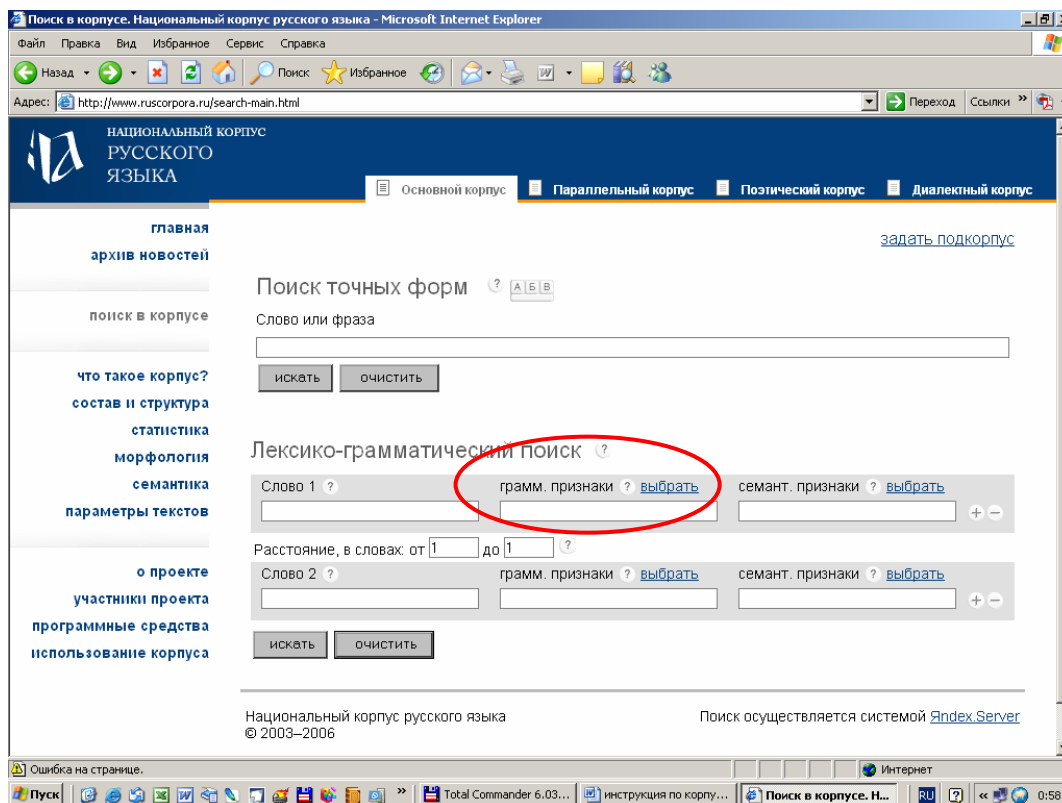


Вот пример того, что можно получить в результате такого запроса:

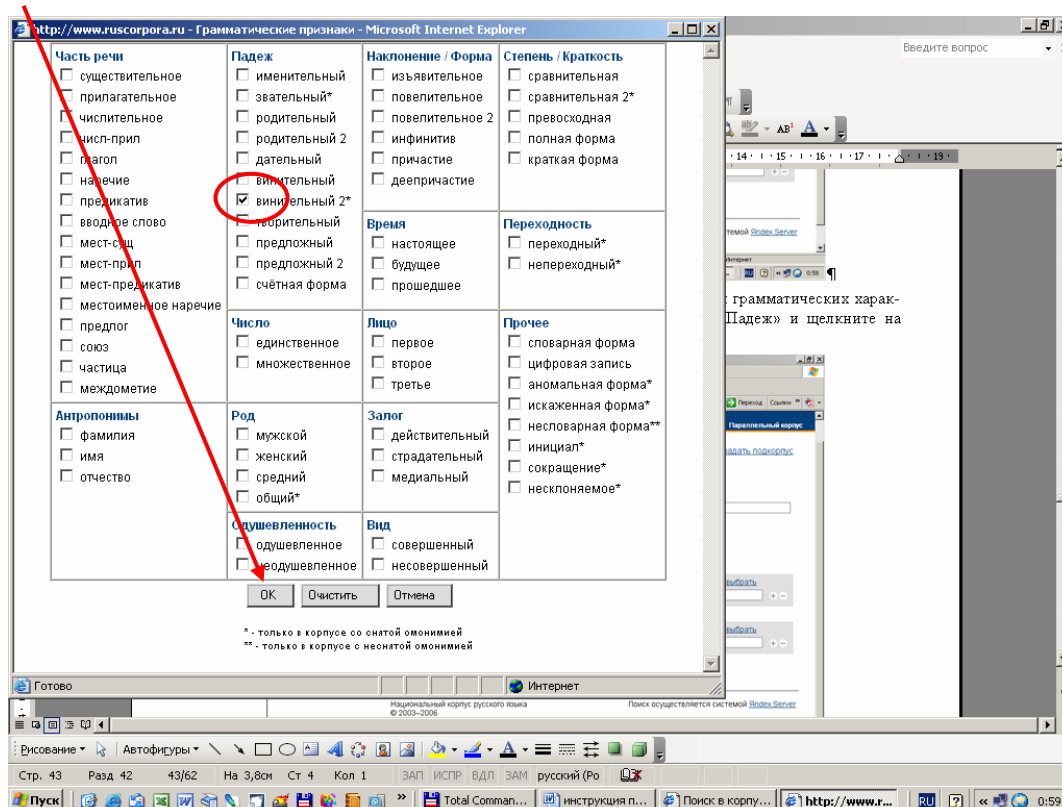


35 шаг

Тридцать пятый шаг. Как выбрать из Корпуса все словоформы, имеющие одну и ту же грамматическую характеристику? Для этого в поле «грамм. признаки» нужно щелкнуть по ссылке «Выбрать»:

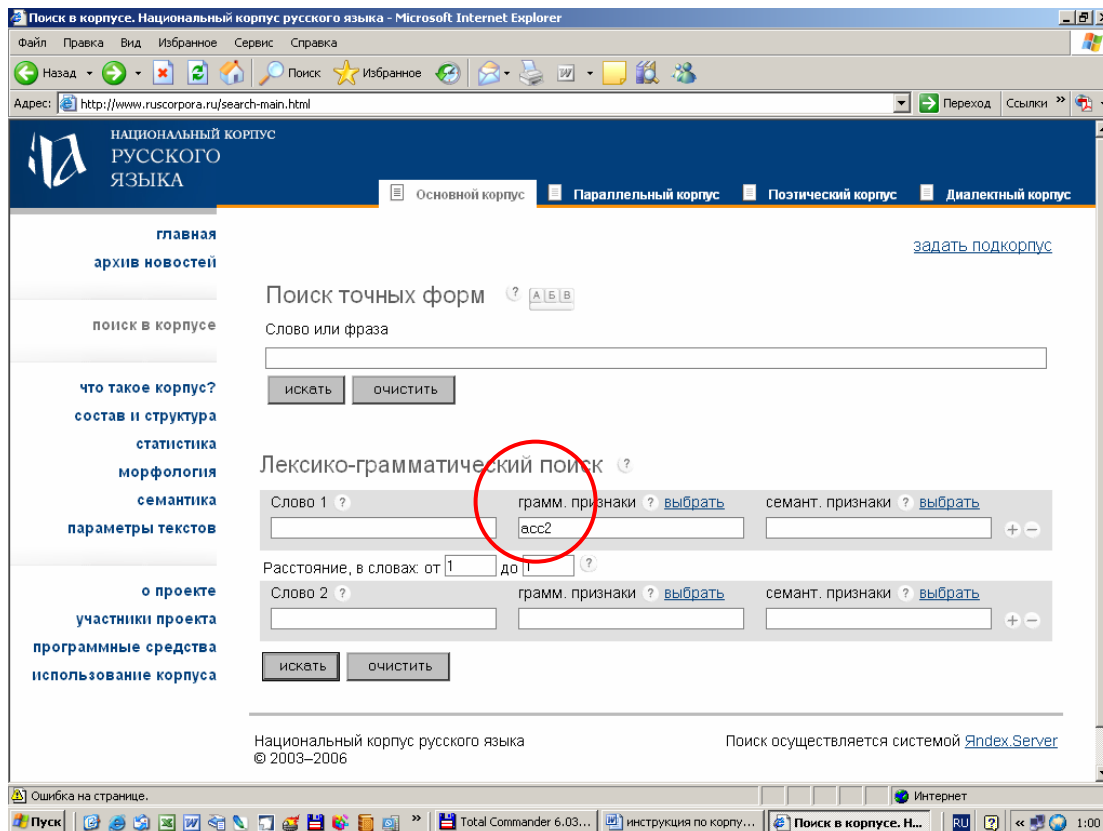


После этого откроется дополнительное окно со списком всех возможных грамматических характеристик. Выберите, например, позицию «винительный 2» в колонке «Падеж» и щелкните на кнопке ОК:

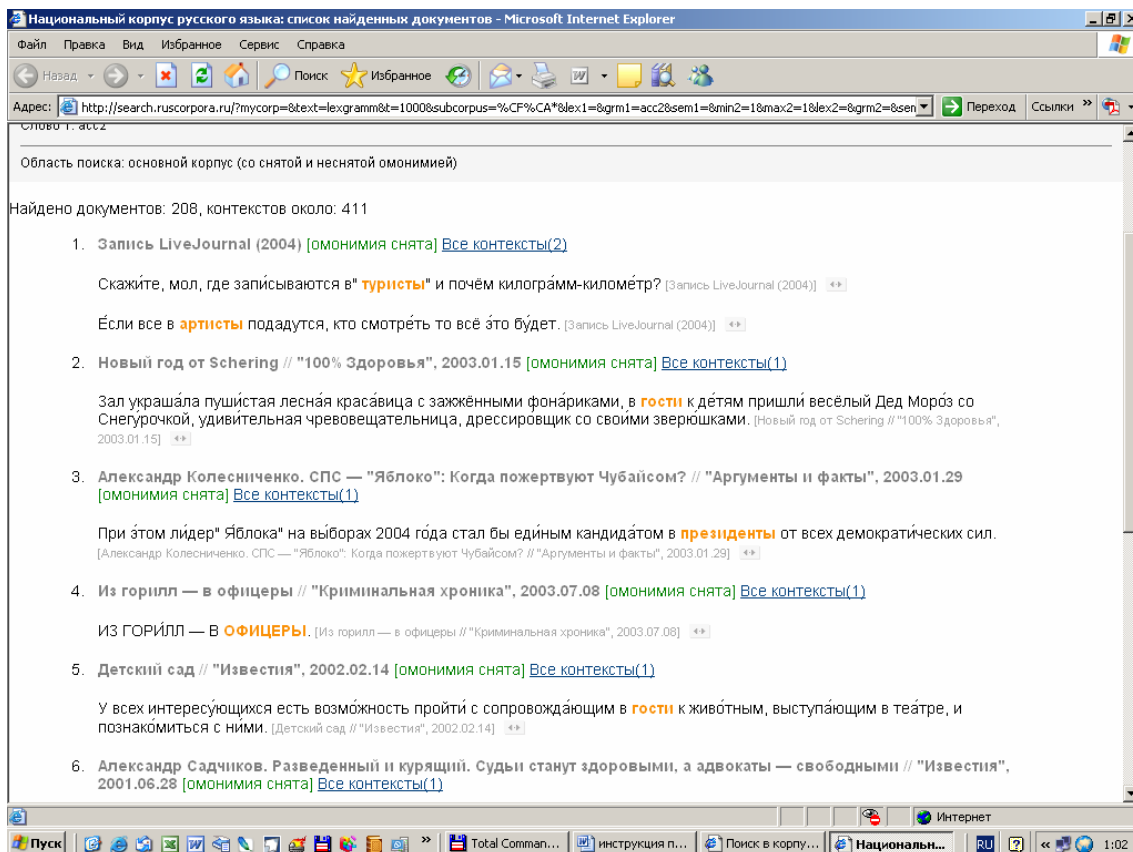


35 шаг

После этого окно выбора закрывается, и остается только окно поиска с зашифрованным обозначением соответствующей грамматической категории:

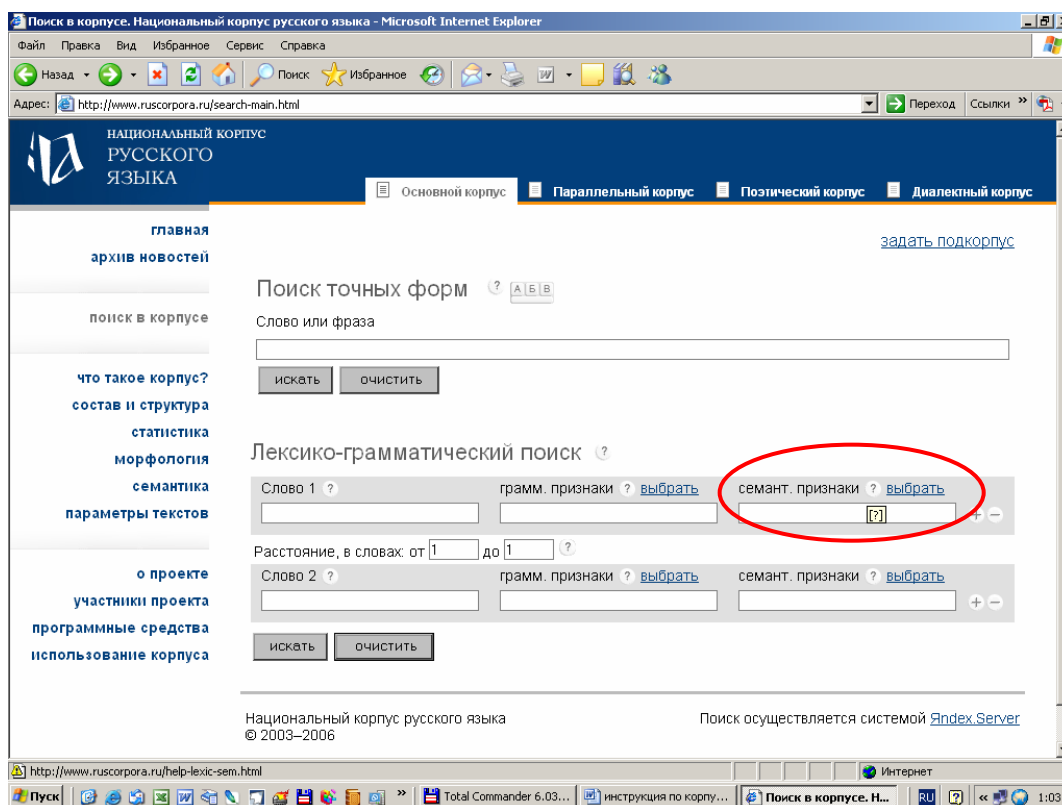


Если после этого щелкнуть на кнопке «Искать», то корпус выдаст все случаи вхождения второго винительного:

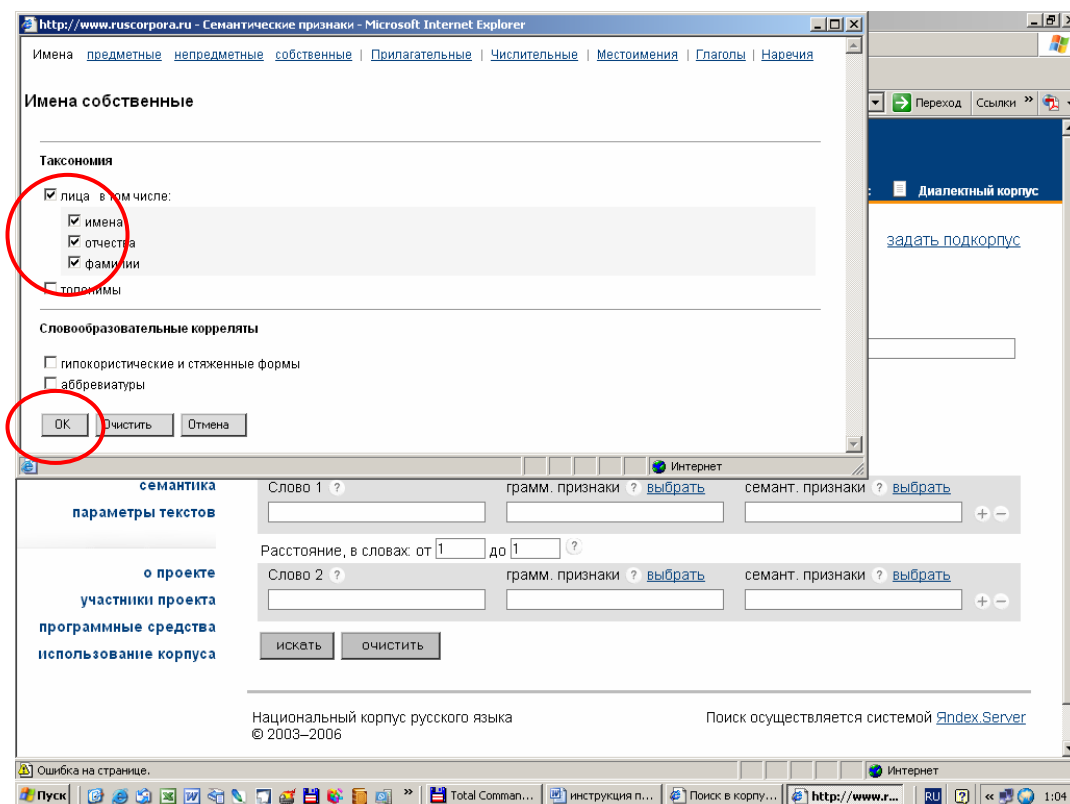


36 шаг

Тридцать шестой шаг. Как найти в Корпусе слова, имеющие одинаковую семантическую характеристику? На странице поиска в поле «семант. признаки» щелкните на ссылке «Выбрать»:

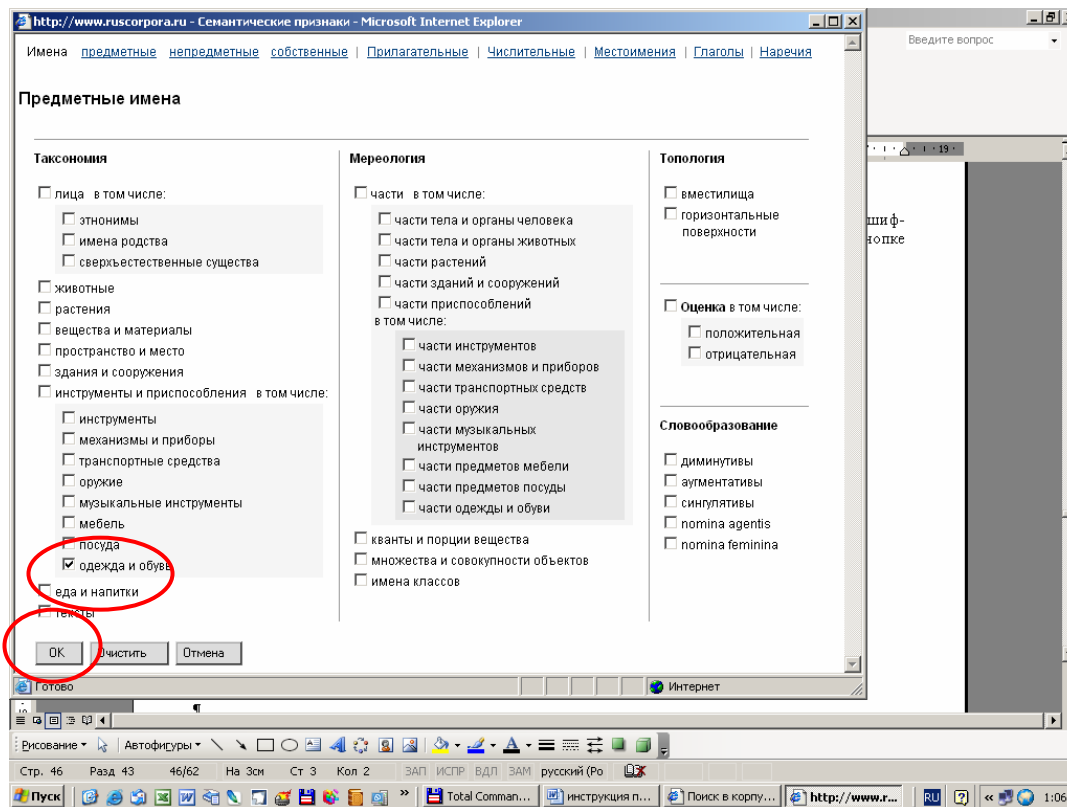


После этого откроется дополнительное окно со списком всех возможных семантических характеристик. Выберите, например, позицию «лица» в разделе «Имена собственные» и щелкните на кнопке ОК:

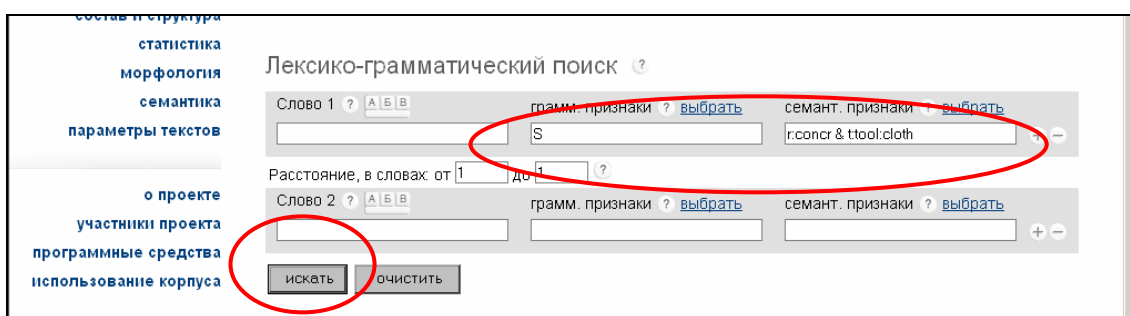
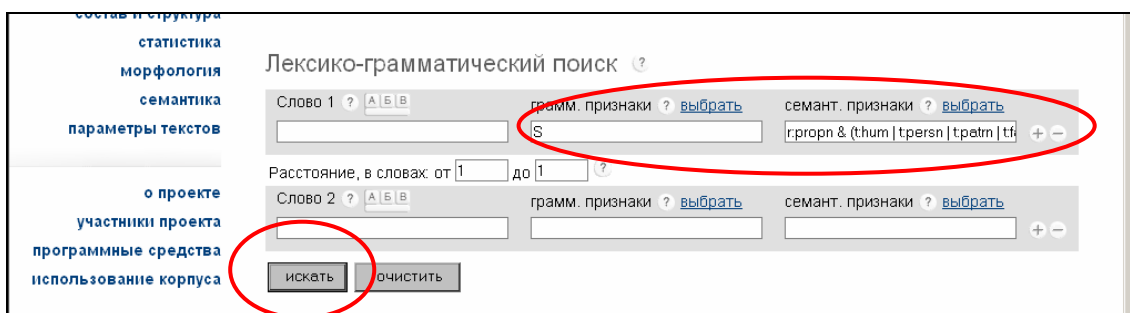


36 шаг

Или, например, Вы можете выбрать позицию «Одежда и обувь» в разделе «Предметные имена» и снова щелкнуть на кнопке ОК:

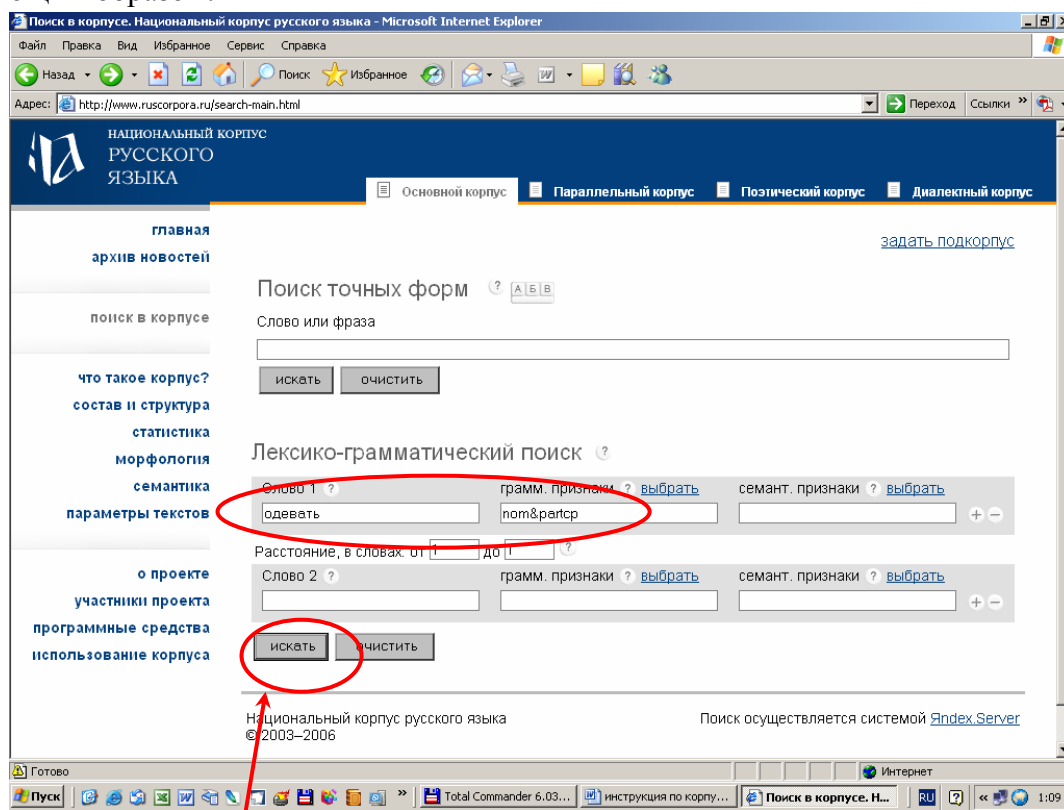


После этого дополнительное окно закроется, и перед Вами останется только окно поиска с зашифрованной записью Вашего запроса, после чего Вам останется только щелкнуть мышкой на кнопке «Искать»:

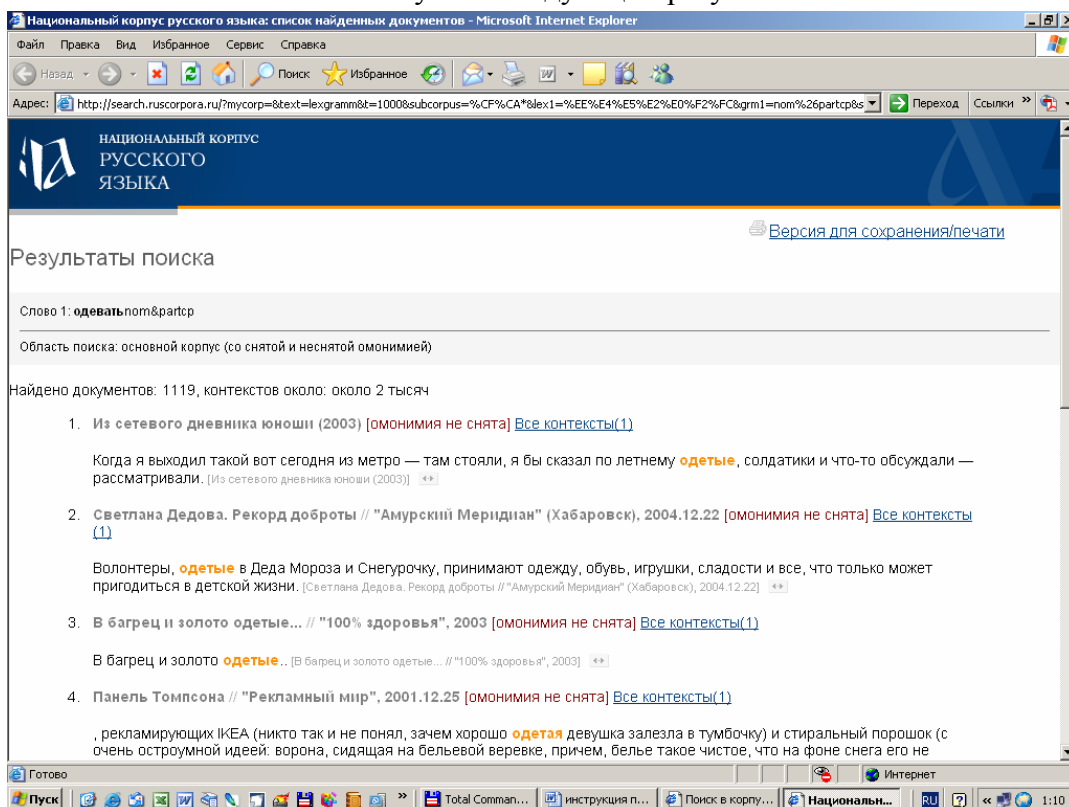


37 шаг

Тридцать седьмой шаг. Как найти все словоформы данной лексемы, имеющие одну и ту же грамматическую характеристику? В поле «Лексико-грамматический поиск» наберите искомое слово (**тридцать третий шаг**), а в поле «грамм. признаки» выберите нужную грамматическую характеристику (**тридцать пятый шаг**). Например, если Вы хотите найти все причастия в именительном падеже от лексемы *одевать*, то после шагов 33 и 35 Ваша страница поиска будет выглядеть следующим образом:

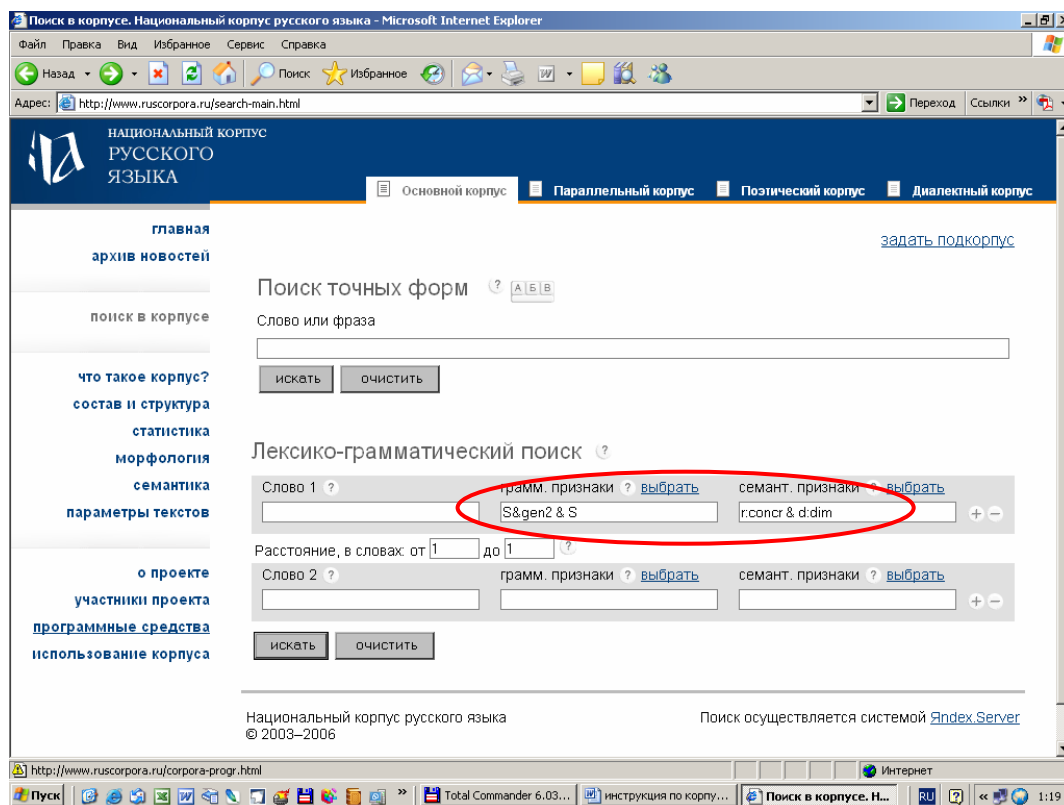


Щелкните на кнопке «Искать» и Вы получите следующий результат:

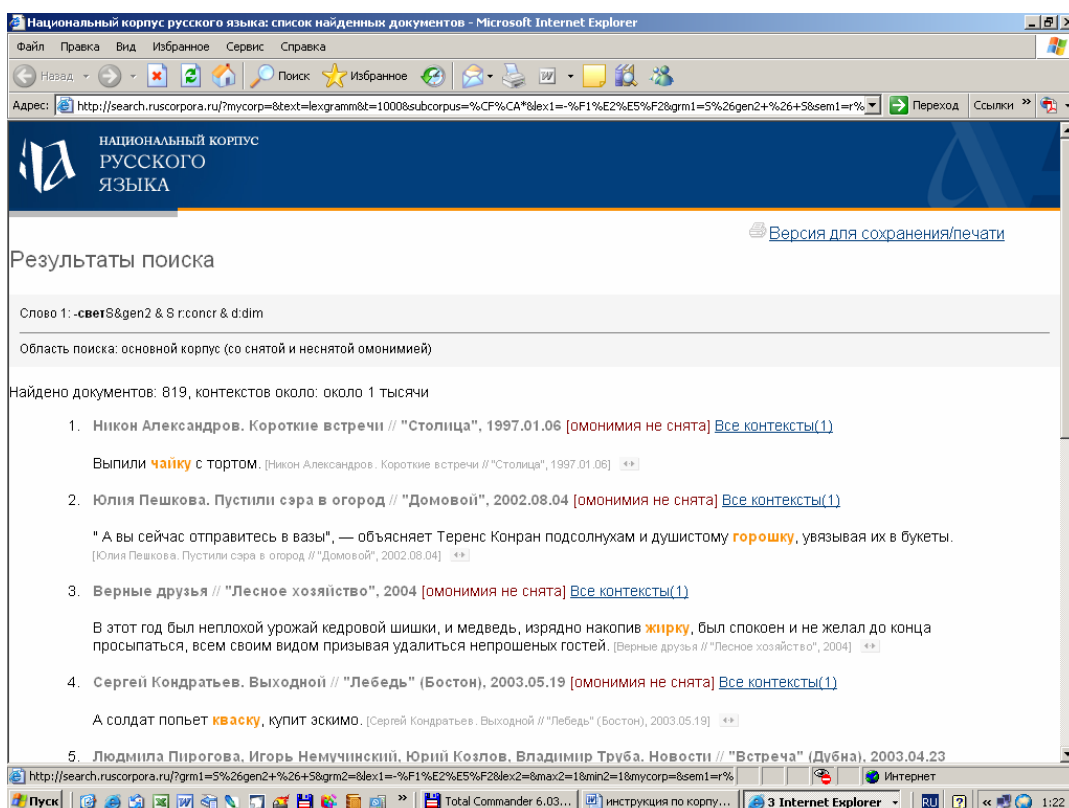


38 шаг

Тридцать восьмой шаг. Как найти в Корпусе слово, которое характеризуется определенными грамматическими и семантическими признаками? Например, как найти **дими́нтивы** (уменьшительные формы) существительных во втором родительном падеже (т. е. род. п. с окончанием -у)? Для этого в зоне «грамм. признаки» в разделе «Падеж» выбираем «родительный 2» (см. **шаг тридцать пятый**), а в поле «семант. признаки» в разделе «Предметные имена» выбираем параметр «Дими́нтивы» (см. **шаг тридцать шестой**). Страница поиска имеет следующий вид:

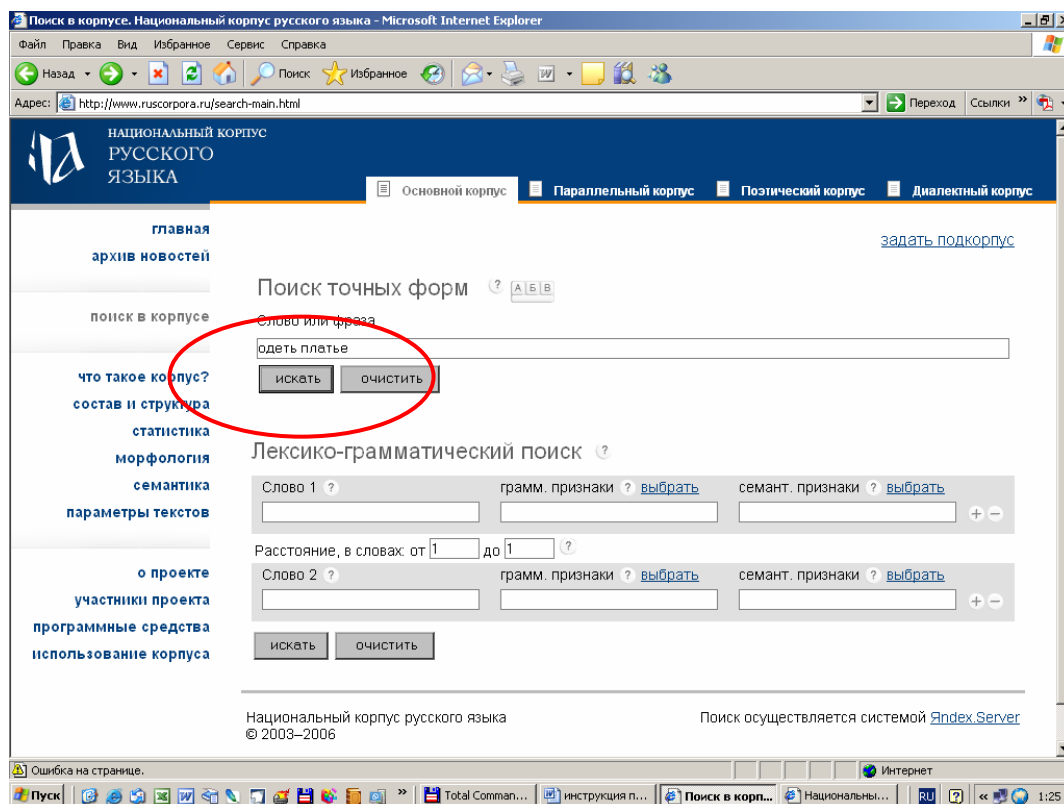


Результаты поиска выглядят следующим образом:

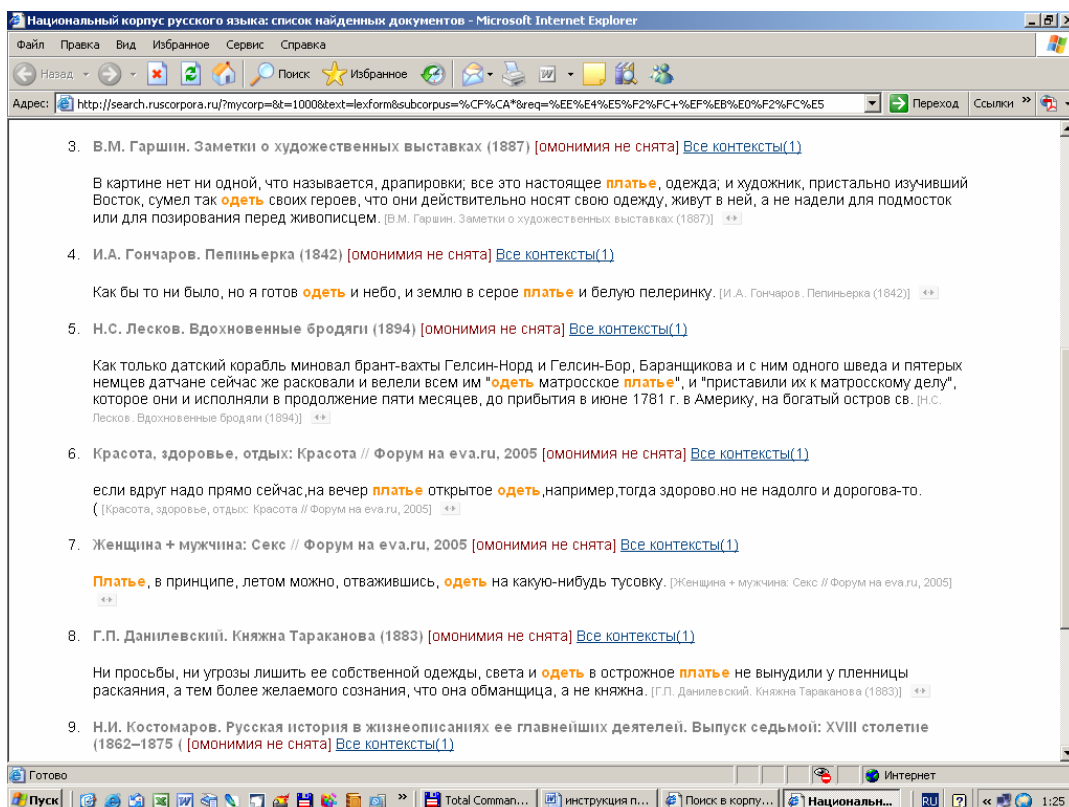


39 шаг

Тридцать девятый шаг. Как найти в Корпусе словосочетание? Если Вам нужно словосочетание в какой-то конкретной грамматической форме, то Вы просто набираете его в строке «Поиск точных форм» и щелкаете мышкой на кнопке «Искать». Например, Вам нужно найти все случаи употребления неправильного словосочетания «одеть платье» именно в этой грамматической форме:

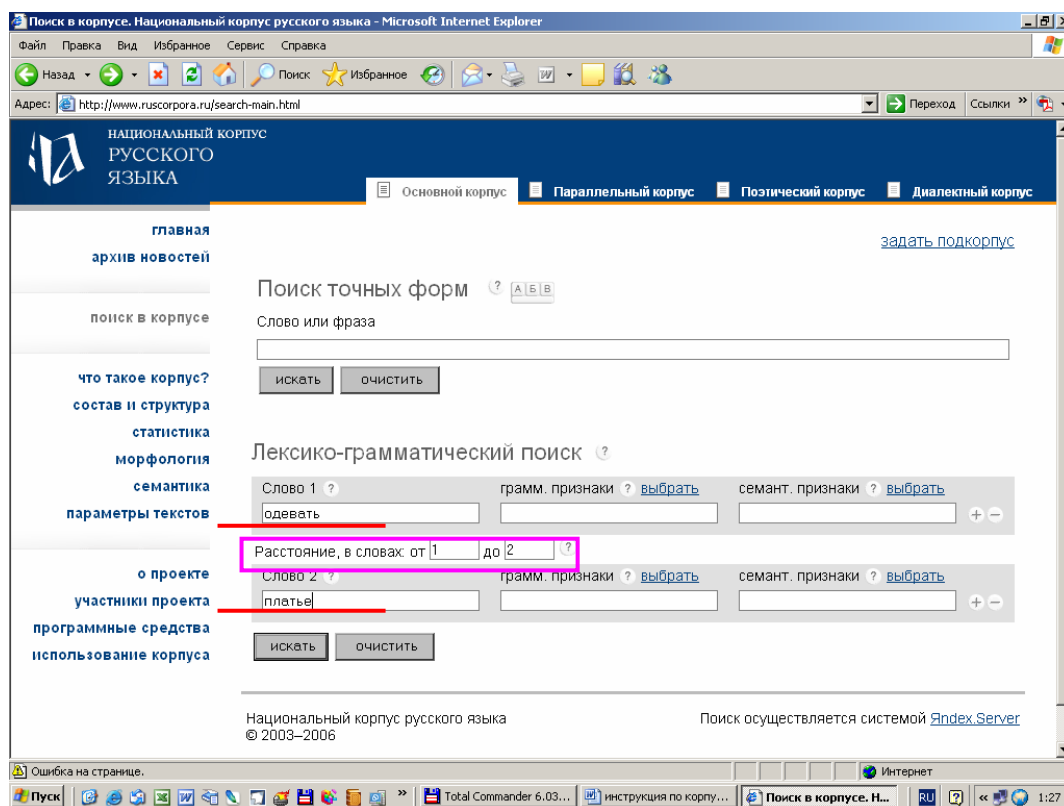


Вы щелкаете на кнопке «Искать» и получаете ответ:

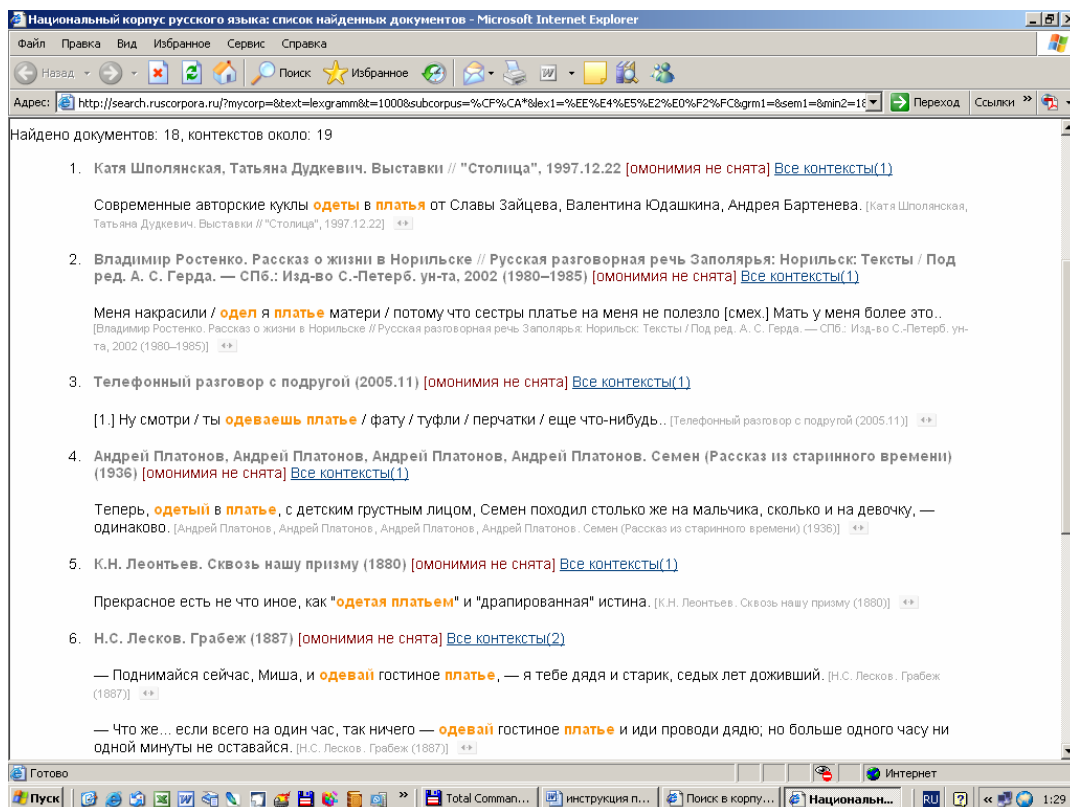


39 шаг

Если же Вам нужно получить из корпуса это словосочетание во всех возможных грамматических формах, то Вы в поле «Лексико-грамматический поиск» 1) в графе «Слово 1» набираете слово «одевать» (обязательно в несовершенном виде!!!), 2) в графе «Расстояние в словах» устанавливаете нужное расстояние (например, расстояние 1–1 означает, что слова будут примыкать друг к другу непосредственно, расстояние 1–2 – между ними будет одно слово, 1–3 – два слова и т.д.), 3) в графе «Слово 2» набираете слово «платье» и щелкаете на кнопке «Искать»:

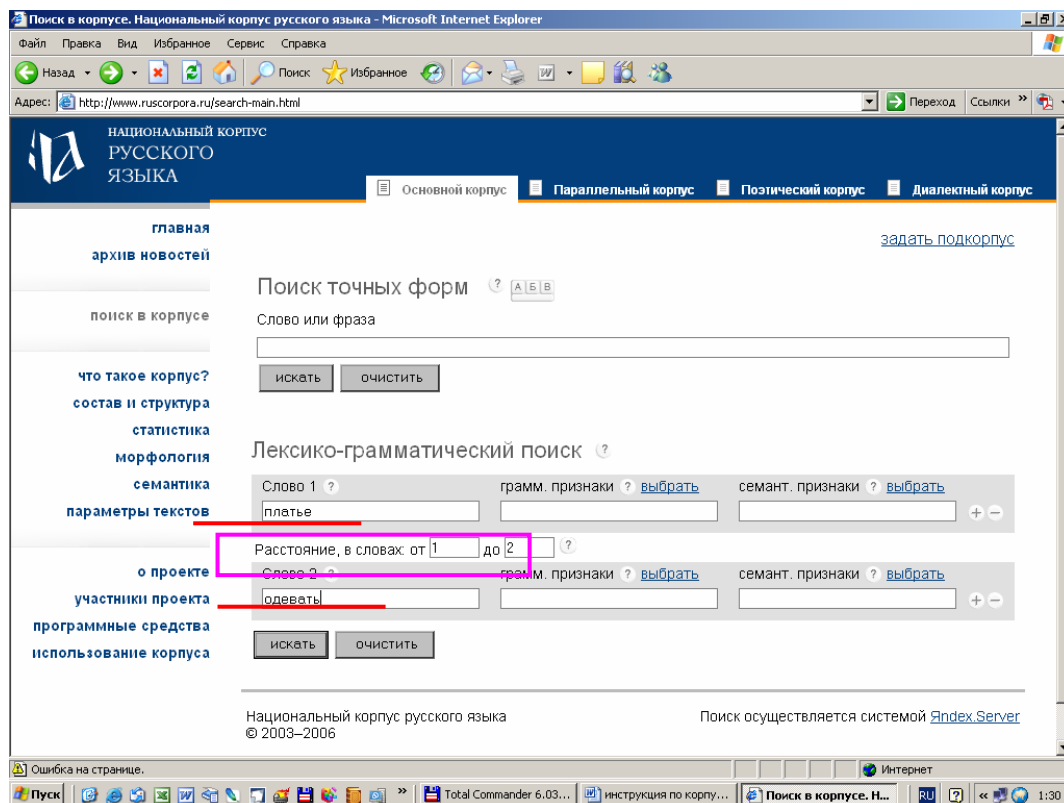


Вот результат поиска:

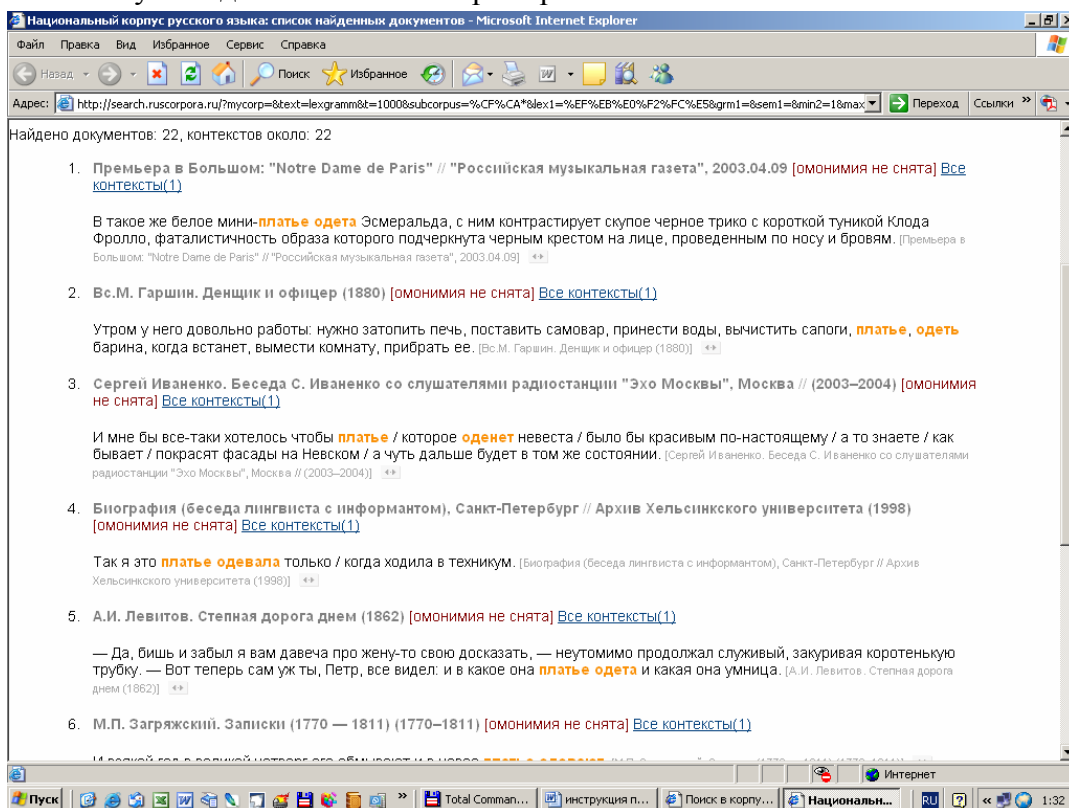


39 шаг

Следует, однако, помнить, что в случае поиска всех возможных словоформ порядок «Слово 1» и «Слово 2» имеет значение. Поэтому для полноты картины поиск следует повторить в следующем виде:

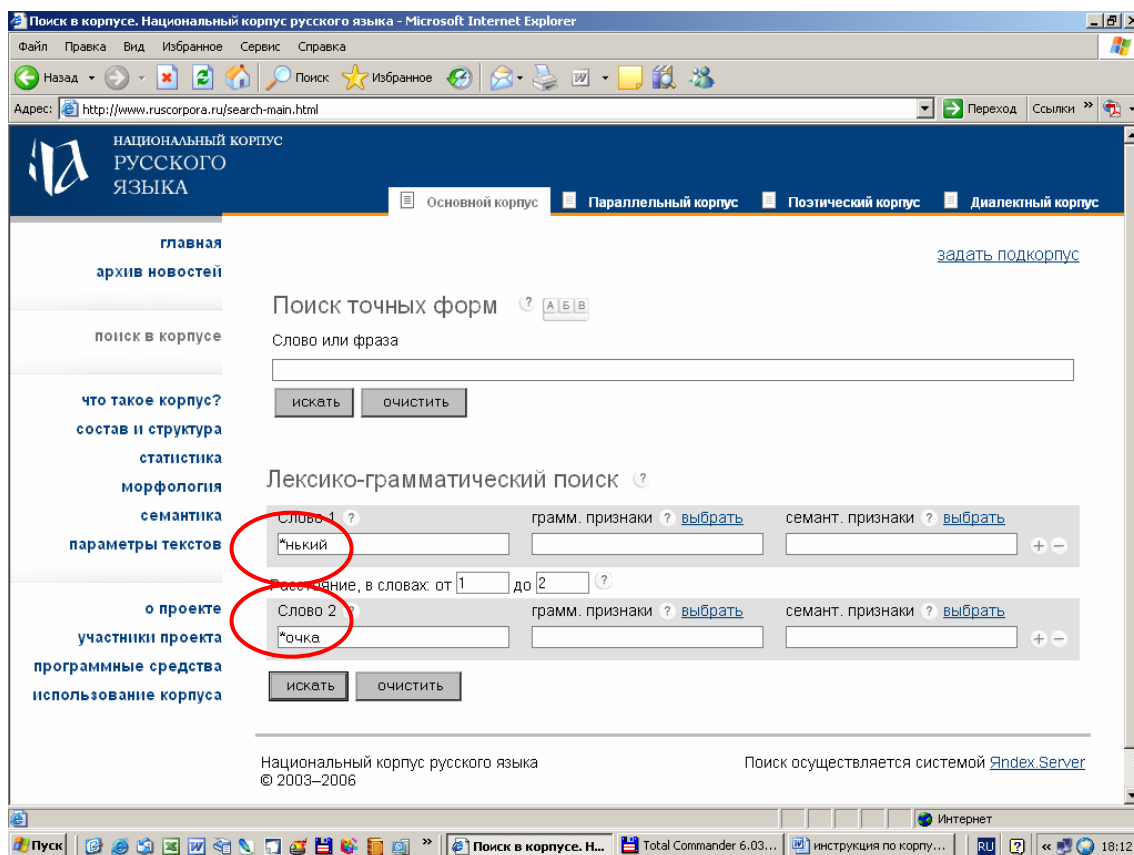


После этого Вы получите дополнительные примеры:

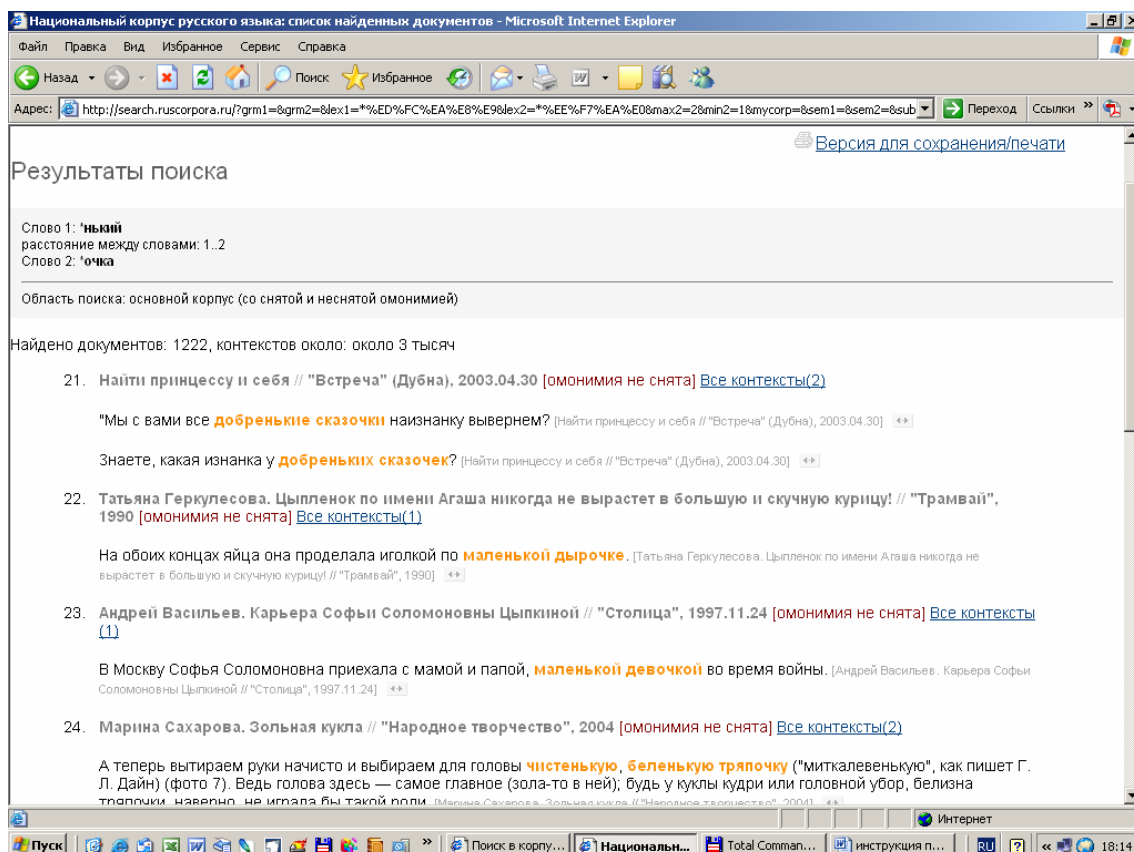


39 шаг

Разумеется, Вы можете искать сочетания слов, которые имеют одинаковую конечную или начальную часть. Например, можно найти примеры двусловных словосочетаний, в которых первое слово заканчивается на *-нький*, а второе – на *-очка* (с расстоянием, например, в ноль и одно слово):

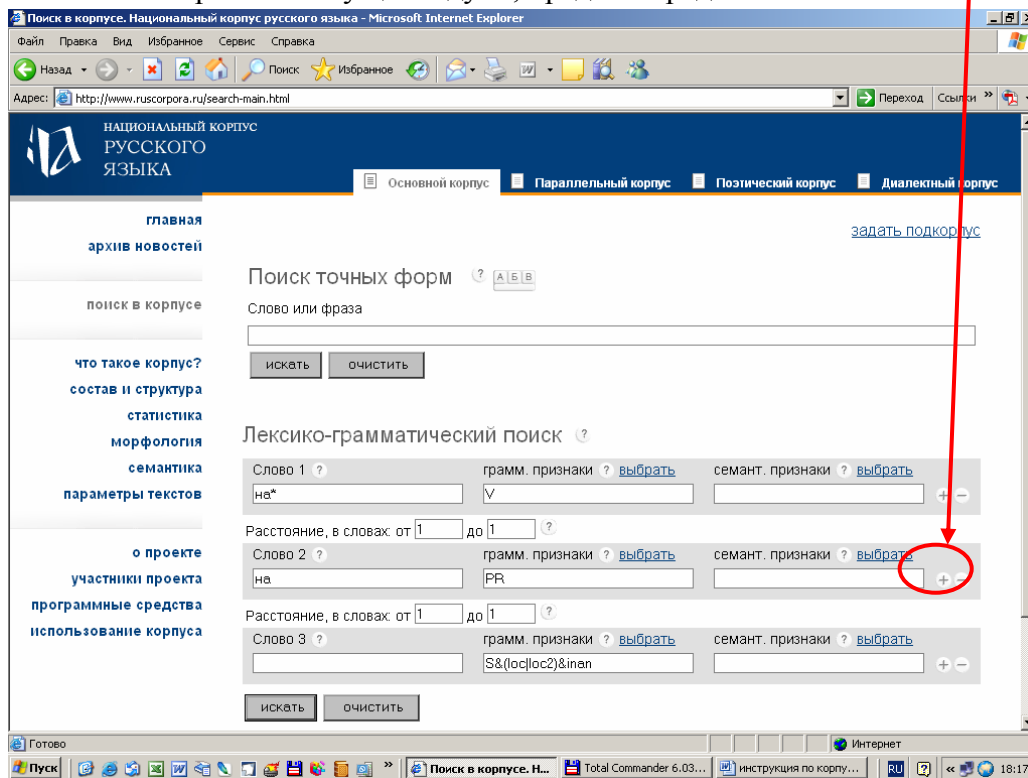


Вот часть результатов, которые Вы получите на этот запрос:

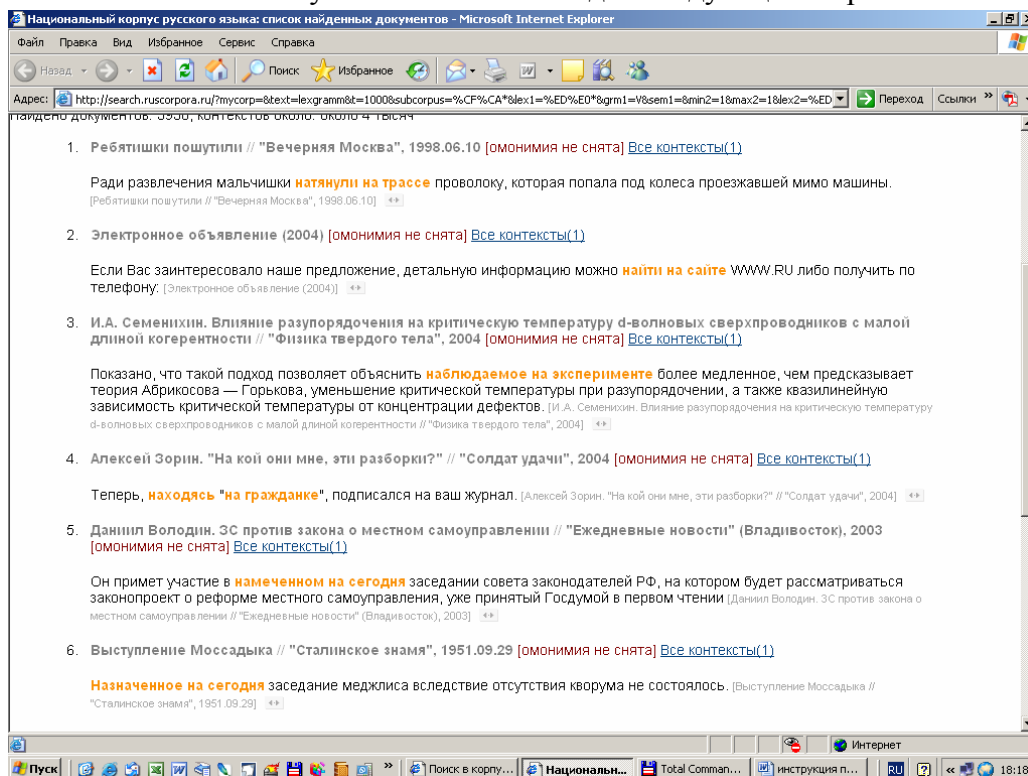


39 шаг

Возможны случаи, когда Вам нужно найти словосочетание из более чем двух слов. Например, как построить запрос, чтобы получить словосочетание, состоящее из трех элементов: 1) глагол с приставкой *на-* + 2) предлог *на* + 3) неодушевленное существительное в предложном падеже (в т.ч. и во втором предложном)? В поле «Слово 1» набираем *на**, в грамматических признаках выбираем «глагол», устанавливаем расстояние между словами 1–1, в поле «Слово 2» набираем *на* и в поле «грамм. признаки» выбираем «предлог», после этого щелкаем мышкой на знаке «+» рядом с полем «Слово 2» и в появившемся поле «Слово 3» устанавливаем расстояние между словами 1–1 и выбираем грамматические признаки «сущ. неодуш., предл. и предл.2»:

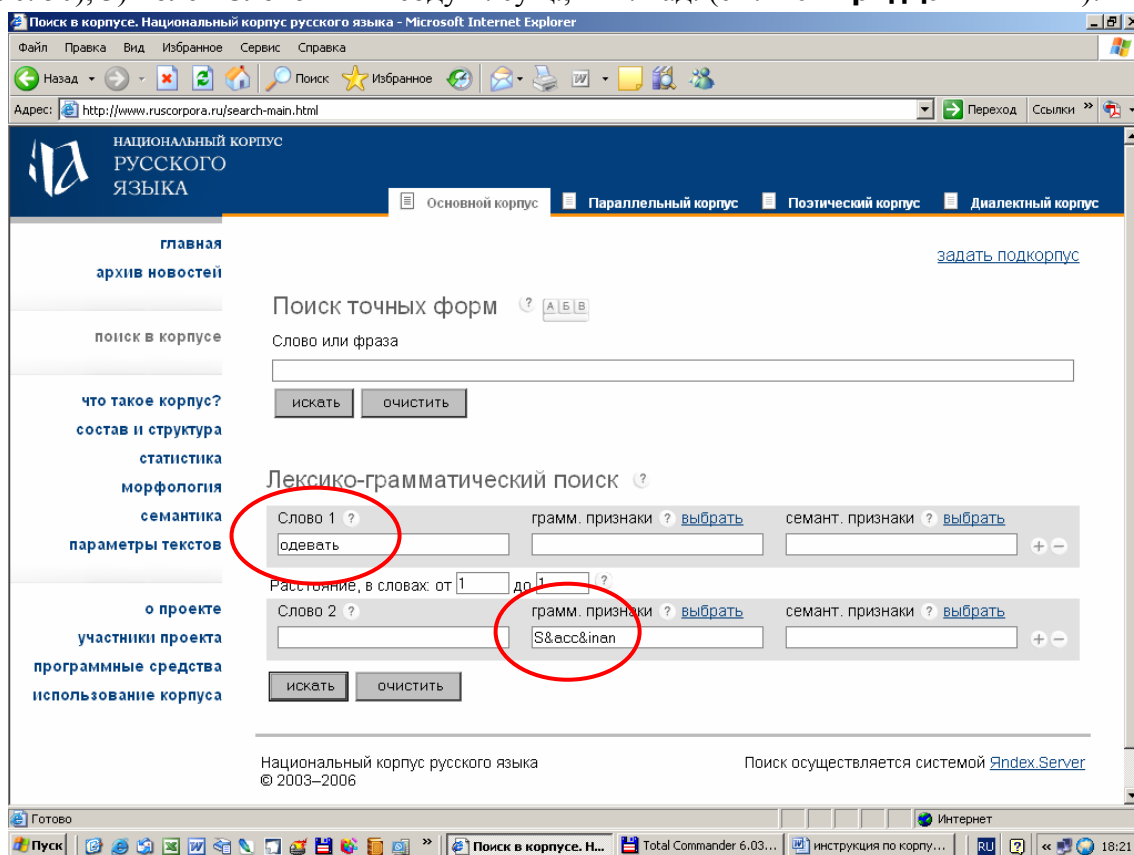


Щелкаем на кнопке «Искать». Результат поиска выглядит следующим образом:

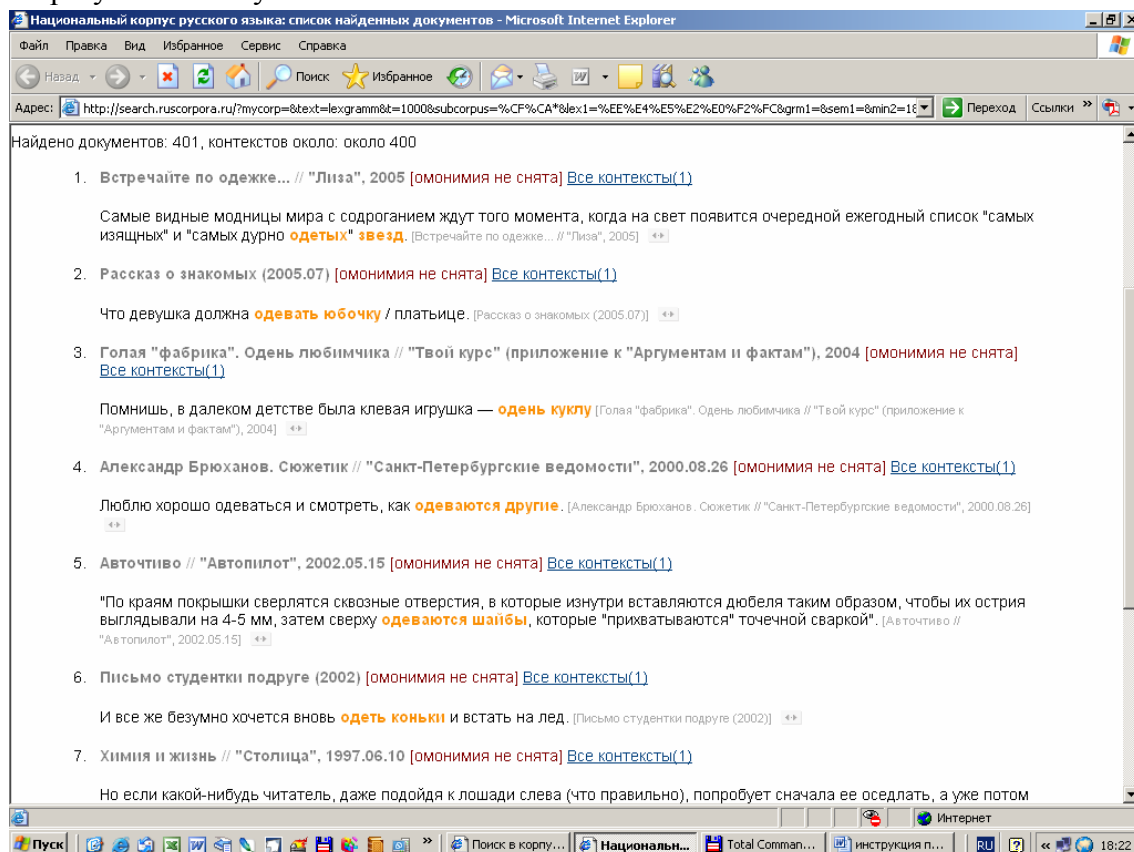


40 шаг

Сороковой шаг. Как найти словосочетание, в котором задано первое слово, а второе характеризуется какими-либо грамматическими признаками? Например, как найти сочетание глагола *одевать* и любого неодушевленного существительного в винительном падеже? На странице поиска набираем: 1) поле «Слово 1» – *одевать* (см. шаг тридцать третий), 2) расстояние между словами 1–1 (см. с. 50), 3) поле «Слово 2» – неодуш. сущ., вин. пад. (см. шаг тридцать пятый):

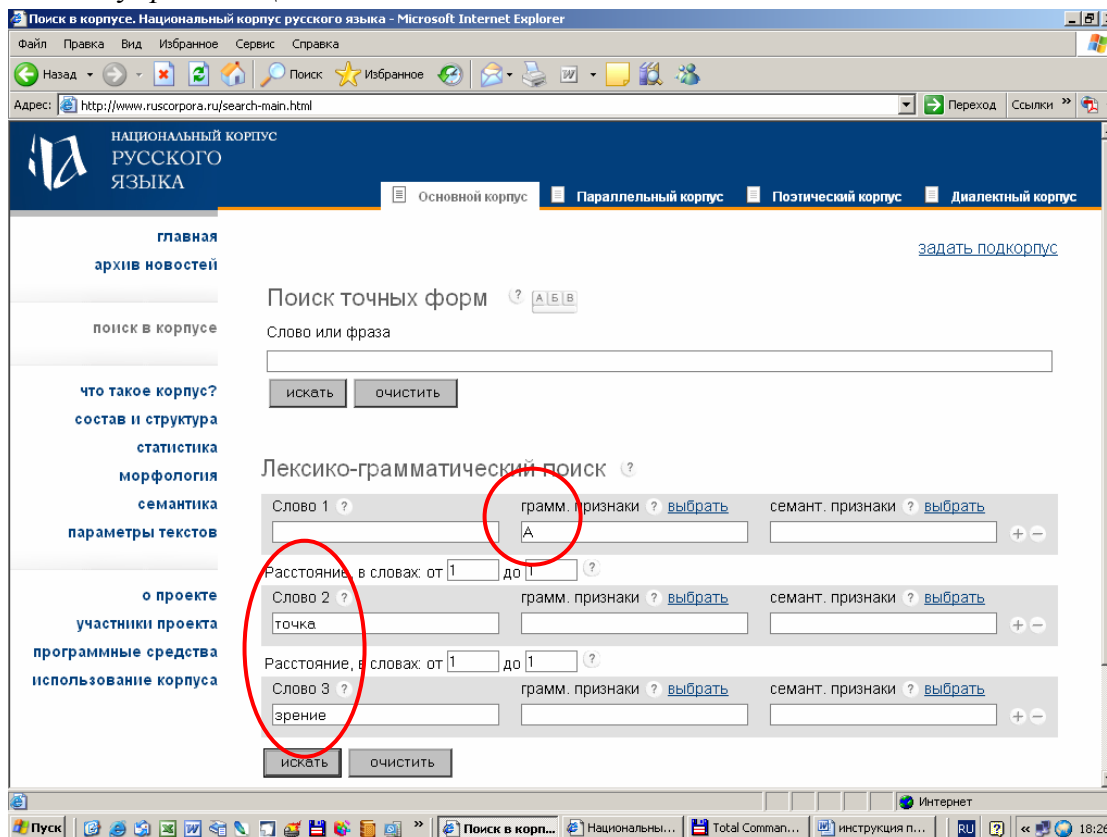


В качестве результата получаем:

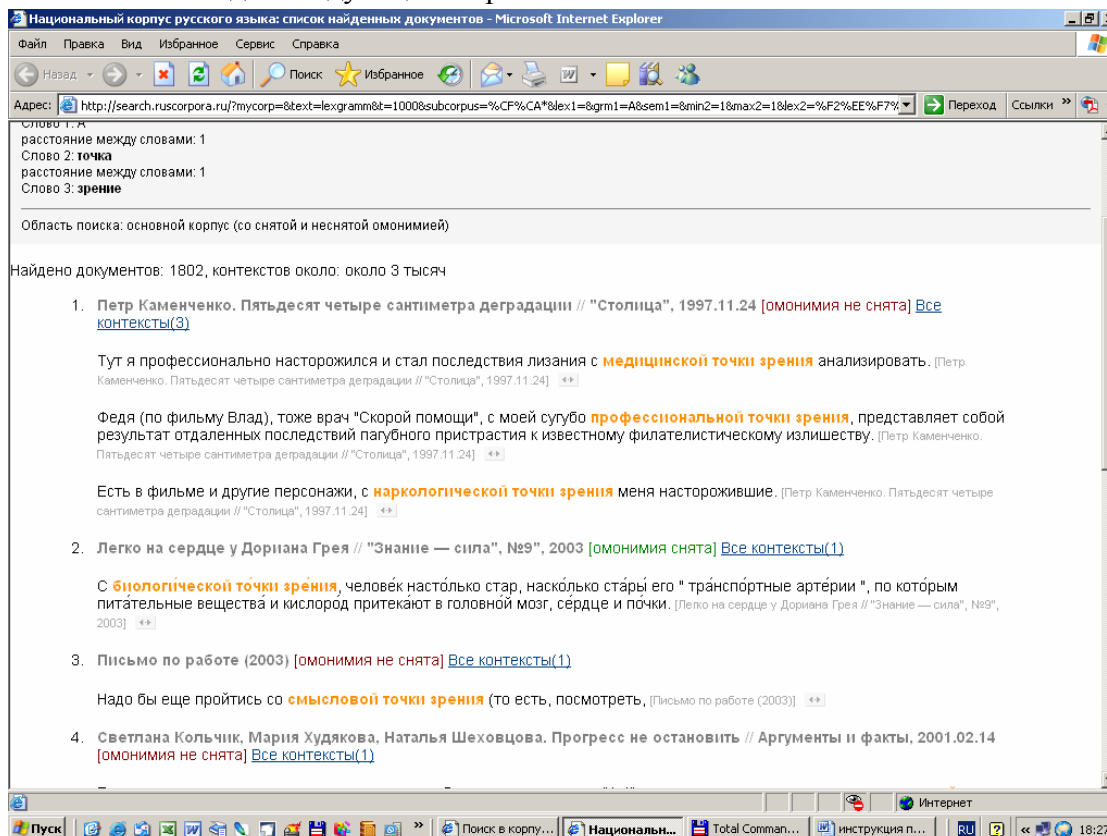


40 шаг

А как найти в корпусе словосочетание, в котором заданы грамматические характеристики первого слова и точно определено второе слово? Например, как найти все определения, которые имеются в Корпусе для словосочетания *точка зрения*? На странице поиска в поле «Слово 1» выбираем грамматическую характеристику «прилагательное» (см. шаг тридцать пятый); в поле «Слово 2» набираем лексему *точка*; в поле «Слово 3» (см. шаг тридцать девятый, о выборе поля «Слово 3») набираем лексему *зрение*. Щелкаем мышкой на кнопке «Искать»:

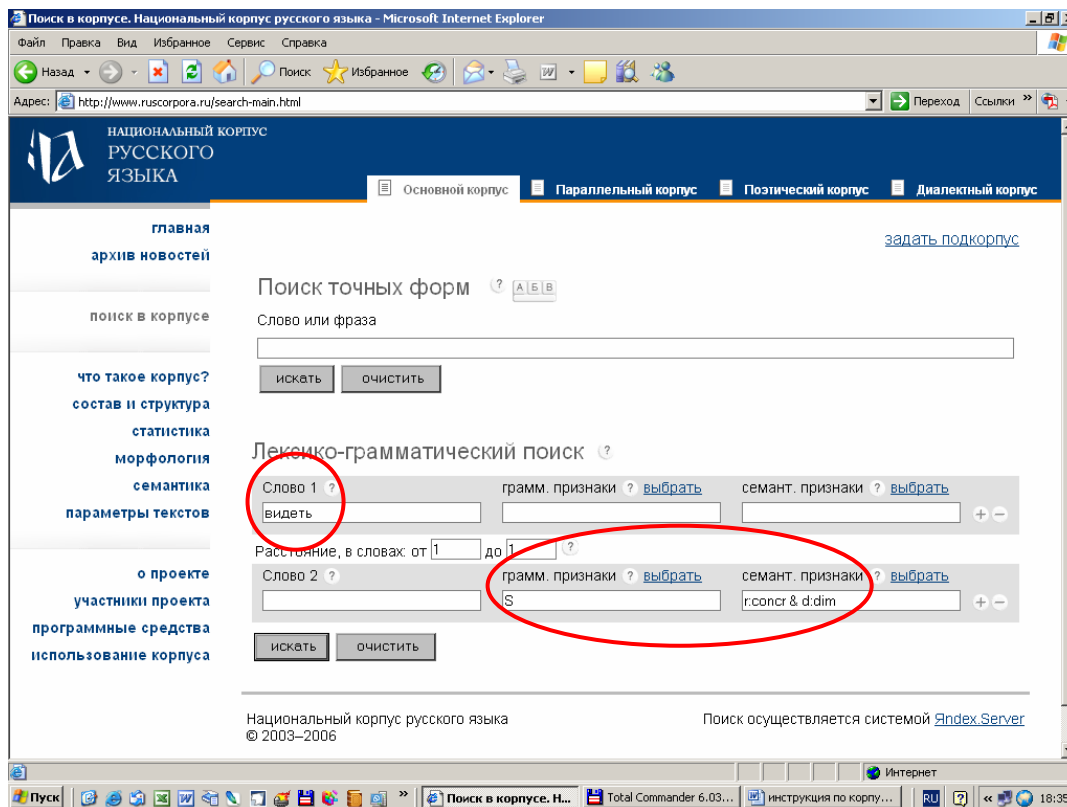


Результаты поиска выглядят следующим образом:

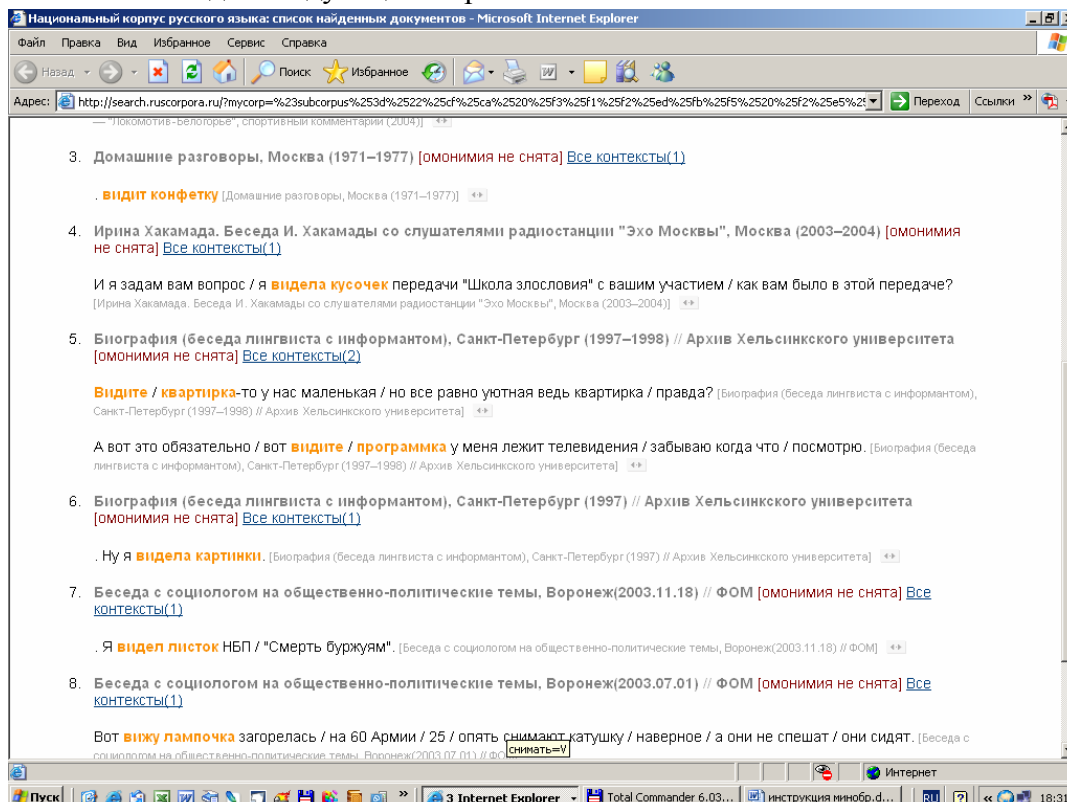


41 шаг

Сорок первый шаг. Как найти словосочетания, в которых точно определено первое слово и заданы семантические признаки второго слова? Например, как Вам найти сочетание глагола *видеть* с **дими́нutiveми**, т.е. с существительными, содержащими уменьшительный суффикс? В «Лексико-грамматическом поиске» в поле «Слово 1» наберите глагол *видеть* (см. **шаг тридцать третий**), в поле «Слово 2» в графе «семант. признаки» выберите пункт «Дими́нutiveы» в разделе «Предметные имена» (см. **шаг тридцать шестой**). Окно поиска после этого имеет следующий вид:

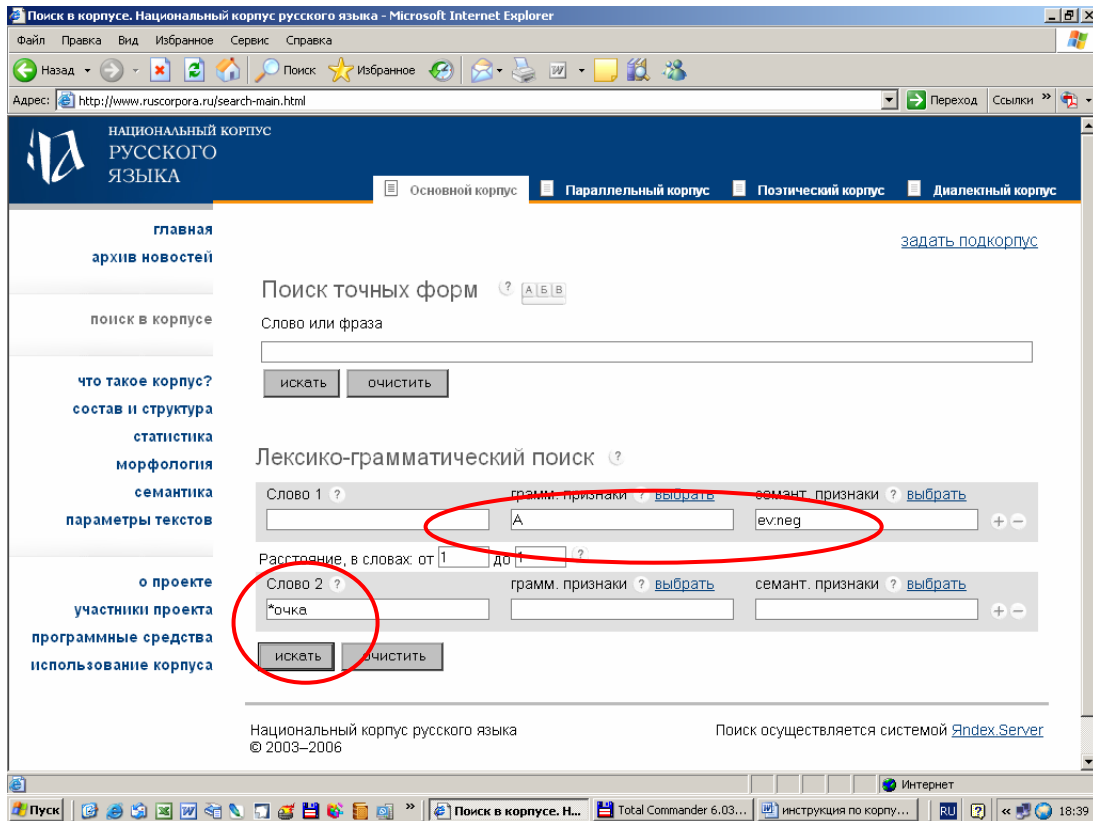


Результаты поиска выглядят следующим образом:

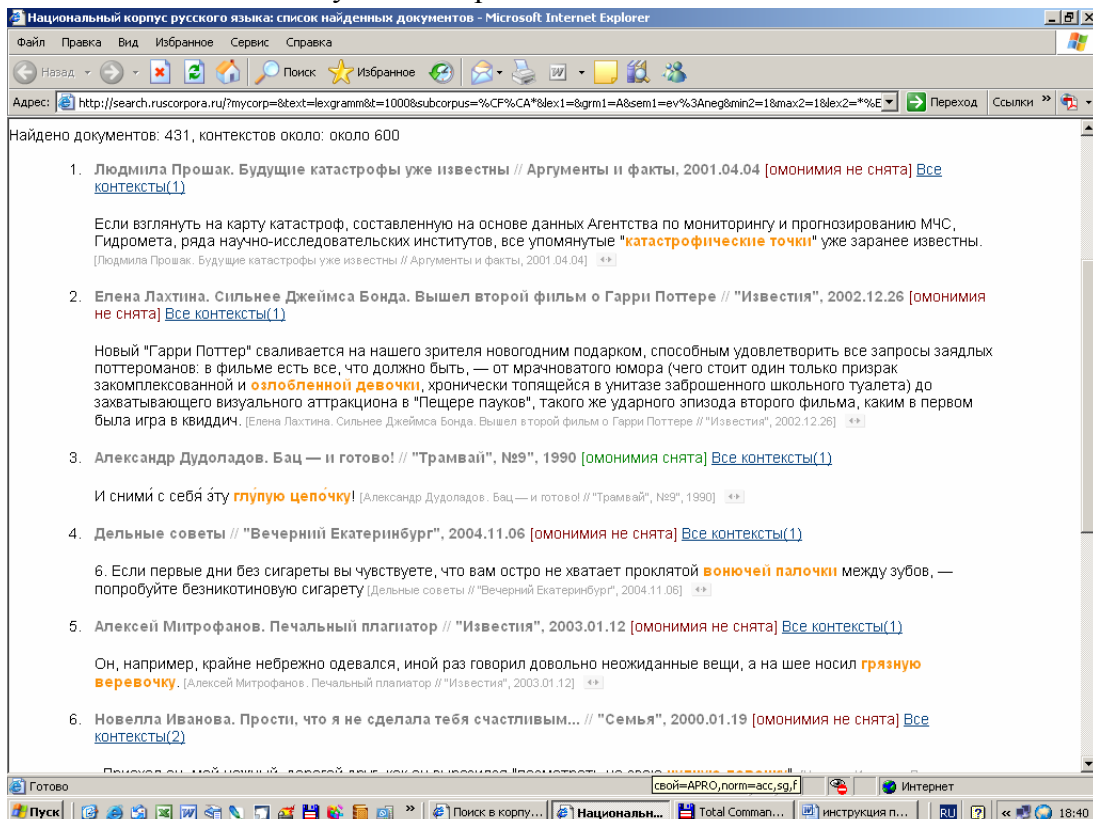


41 шаг

А как найти в корпусе словосочетание, в котором первое слово характеризуется какими-либо семантическими признаками, а второе слово точно определено? Например, как построить запрос, если Вам нужно найти все прилагательные отрицательной оценки в сочетании с существительными, заканчивающимися на *-очка*? В поле «Слово 1» в графе «семант. признаки» в разделе «Прилагательные» выбираем «Отрицательная оценка» (см. шаг тридцать шестой), а в поле «Слово 2» набираем **очка*:



Результаты поиска выглядят следующим образом:



42 шаг

Сорок второй шаг. Как найти в корпусе словосочетание, в котором заданы грамматические характеристики первого и второго слова? Например, как найти сочетания двух глаголов в 1-2 л.? На странице поиска в полях «Слово 1» и «Слово 2» в графе «грамм. признаки» выбираем глагол 1-2 л. (см. шаг тридцать пятый). Страница поиска имеет следующий вид:

Поиск в корпусе. Национальный корпус русского языка – Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

Назад Поиск Избранное

Адрес: <http://www.ruscorpora.ru/search-main.html> Переход Ссылки

НАЦИОНАЛЬНЫЙ КОРПУС РУССКОГО ЯЗЫКА

Основной корпус Параллельный корпус Поэтический корпус Диалектный корпус

главная архив новостей

поиск в корпусе

что такое корпус? состав и структура статистика морфология семантика параметры текстов

о проекте участники проекта программные средства использование корпуса

задать подкорпус

Поиск точных форм

Слово или фраза

искать очистить

Лексико-грамматический поиск

Слово 1 ? грамм. признаки ? [выбрать](#) семант. признаки ? [выбрать](#)

Расстояние, в словах: от до

Слово 2 ? грамм. признаки ? [выбрать](#) семант. признаки ? [выбрать](#)

искать очистить

Национальный корпус русского языка © 2003–2006

Поиск осуществляется системой [Index Server](#)

Результаты поиска выглядят следующим образом:

Национальный корпус русского языка: список найденных документов – Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

Назад Поиск Избранное

Адрес: <http://search.ruscorpora.ru/grm1=V%26%281p%7C2p%29&grm2=V%26%281p%7C2p%29&lex1=8&lex2=8&max2=18&min2=18&mscorp=%23&subcorpus%23> Переход Ссылки

Ну хотя / я же не **весь день буду** учиться / вечером можем к брату на концерт сходить. [Телефонный разговор двух женщин (2005.04.07)] [**](#)

. Ладно / пошла дальше зубрить / хотя в голове ничего не остаётся / я тут 2 часа **сижу учу** 9 билетов / даже не учу / а повторяю / а потом ловлю себя на том / что разглядываю этикетку от кока-колы. [Телефонный разговор двух женщин (2005.04.07)] [**](#)

36. Телефонный разговор, Московская область (2005) [омонимия не снята] [Все контексты\(2\)](#)

Дом сижу телевизор смотрю / сегодня может с Леркой **пойду погуляю** / она сегодня вот сдала последний экзамен / так что теперь все / гуляем на всю! [Телефонный разговор, Московская область (2005)] [**](#)

. Ну ладно / Лен / я еще вечером **позвоню / узнаю** как там это. [Телефонный разговор, Московская область (2005)] [**](#)

37. Домашние разговоры (2000–2003) // М. В. Китайгородская, Н. Н. Розанова. Речь москвичей: Коммуникативно-культурологический аспект. М.: ИРЯ РАН, 1999 [омонимия не снята] [Все контексты\(2\)](#)

Заходите / заходите. @h2* [Домашние разговоры (2000–2003) // М. В. Китайгородская, Н. Н. Розанова. Речь москвичей: Коммуникативно-культурологический аспект. М.: ИРЯ РАН, 1999] [**](#)

Дорогие **гости / прошу** к столу. @h2* ** — Ну как будем садиться? — Тут наверно Лена сядет. [Домашние разговоры (2000–2003) // М. В. Китайгородская, Н. Н. Розанова. Речь москвичей: Коммуникативно-культурологический аспект. М.: ИРЯ РАН, 1999] [**](#)

38. Из беседы трех женщин (1991) // М. В. Китайгородская, Н. Н. Розанова. Речь москвичей: Коммуникативно-культурологический аспект. М.: ИРЯ РАН, 1999 [омонимия не снята] [Все контексты\(1\)](#)

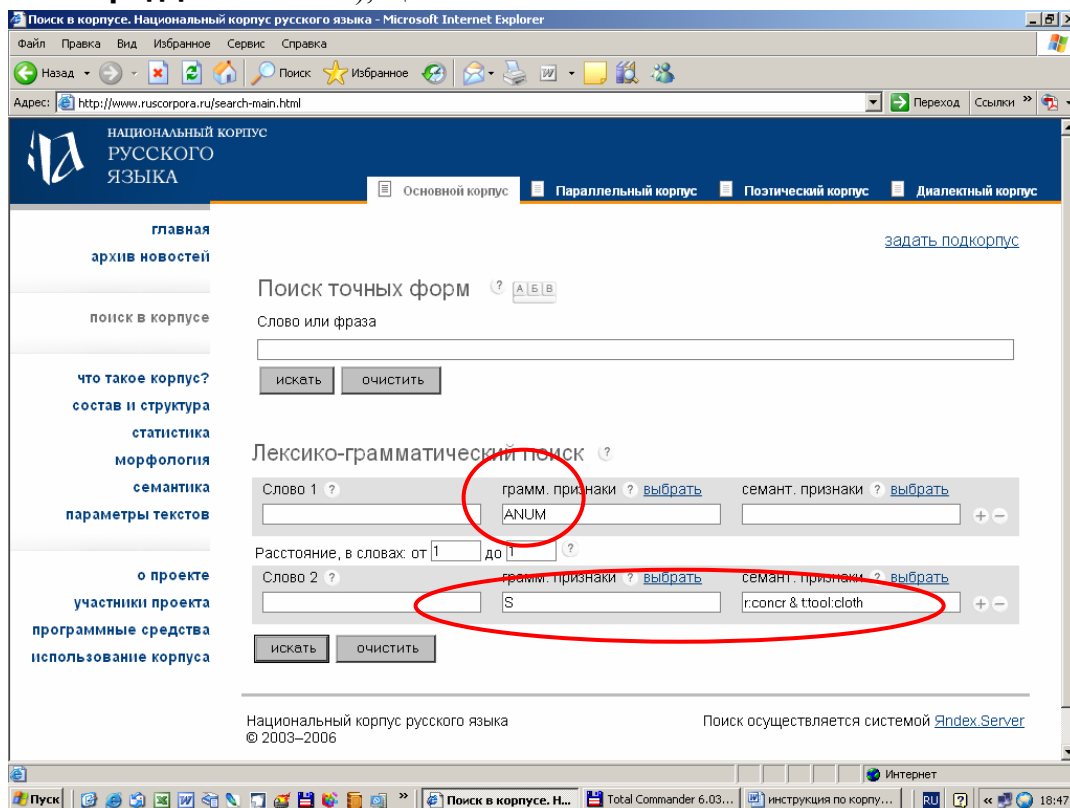
Ну вот сейчас даже **живём / скажем** / очень / очень уединённо / [Из беседы трех женщин (1991) // М. В. Китайгородская, Н. Н. Розанова. Речь москвичей: Коммуникативно-культурологический аспект. М.: ИРЯ РАН, 1999] [**](#)

39. Евгений Лебедев. Рассказ о жизни в Норильске (1980–1985) // Русская разговорная речь Заполярья: Норильск: Тексты / Под ред. А. С. Герда. — СПб.: Изд-во С—Петербург, ун-та, 2002 [омонимия не снята] [Все контексты\(2\)](#)

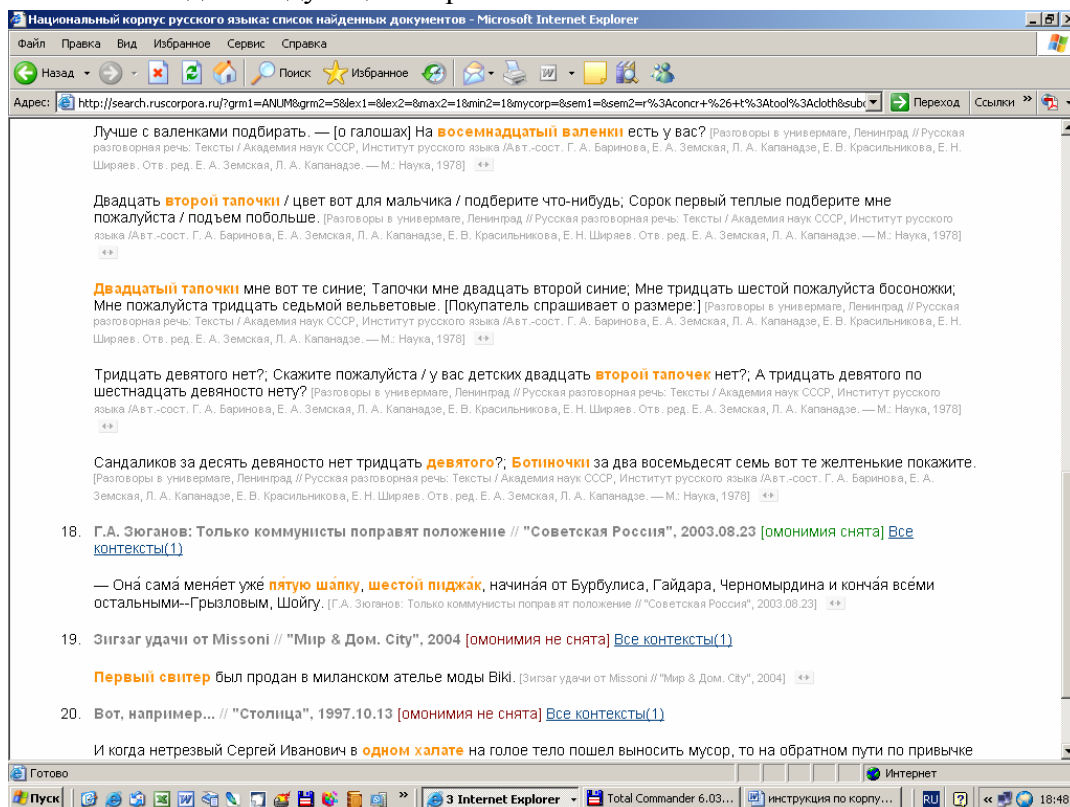
Работаешь / работаешь / а они говорят / что ты спишь / вообще / нехорошие они люди / они мне не нравятся / так и сказать им не бойтесь / они поймут / [Евгений Лебедев. Рассказ о жизни в Норильске (1980–1985) // Русская разговорная речь Заполярья: Норильск: Тексты / Под ред. А. С. Герда. — СПб.: Изд-во С—Петербург, ун-та, 2002] [**](#)

43 шаг

Сорок третий шаг. Как найти словосочетание, в котором заданы грамматические признаки первого слова и семантические признаки второго слова? Например, как построить запрос, если Вы хотите найти сочетания «порядковое числительное + названия одежды и обуви»? В поле «грамм. признаки первого слова» выбираем «числительное-прилагательное» (см. шаг тридцать пятый), в поле «семант. признаки» второго слова в разделе «Предметные имена» выбираем «Одежда и обувь» (см. шаг тридцать шестой), щелкаем на кнопке «Искать»:

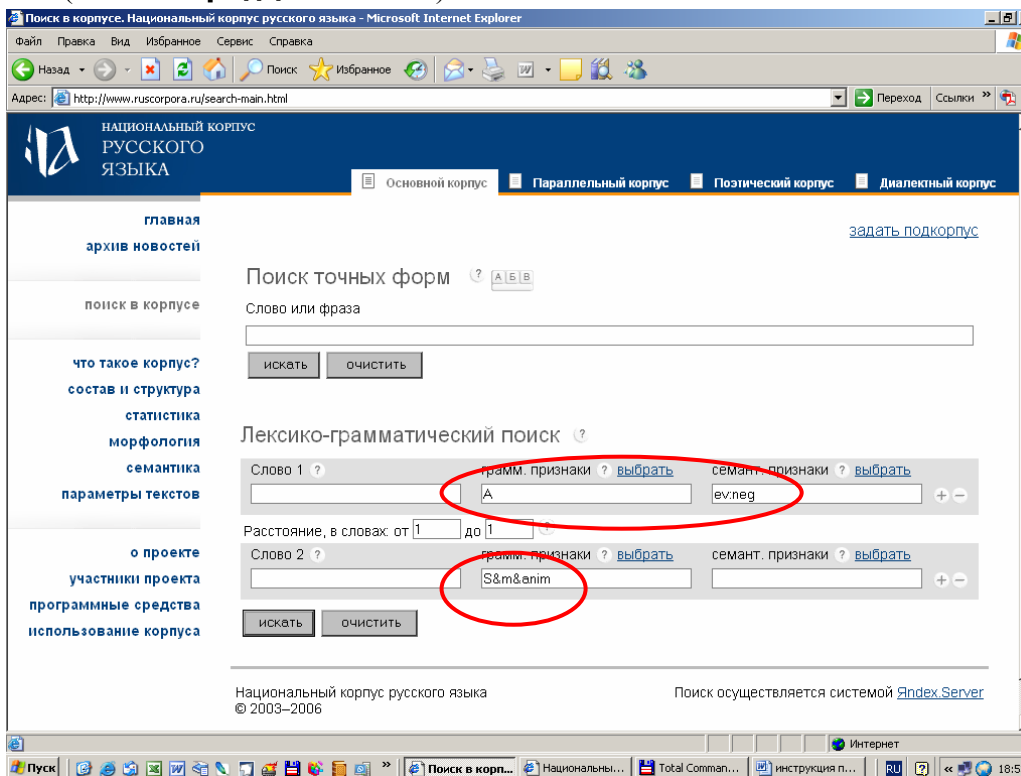


Результат поиска выглядит следующим образом:

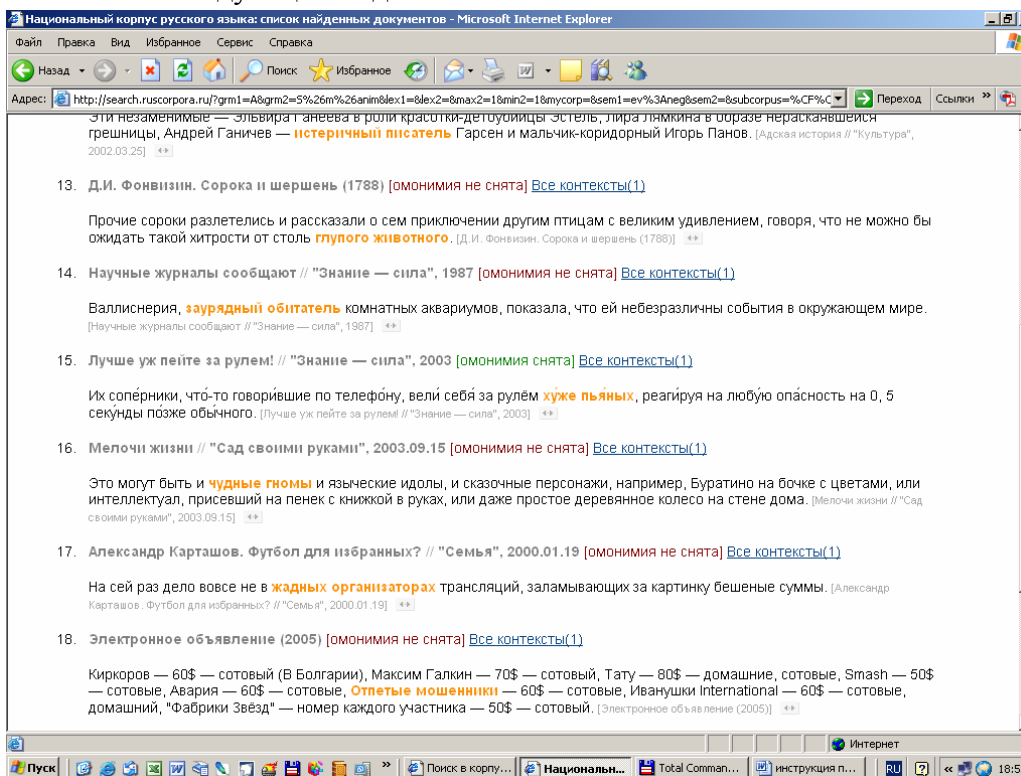


43 шаг

А как найти в корпусе словосочетания, в которых первое слово имеет какие-либо семантические, а второе – какие-либо грамматические характеристики? Например, как найти словосочетания, в которых одушевленные существительные мужского рода имели бы в качестве определения прилагательное отрицательной оценки? В окне поиска в поле «Слово 1» в графе «семант. признаки» выбираем значение «Прилагательное отрицательной оценки» (см. **шаг тридцать шестой**), а в поле «Слово 2» в графе «грамм. признаки» выбираем значение «существительное + мужской род + одушевленное» (см. **шаг тридцать пятый**):

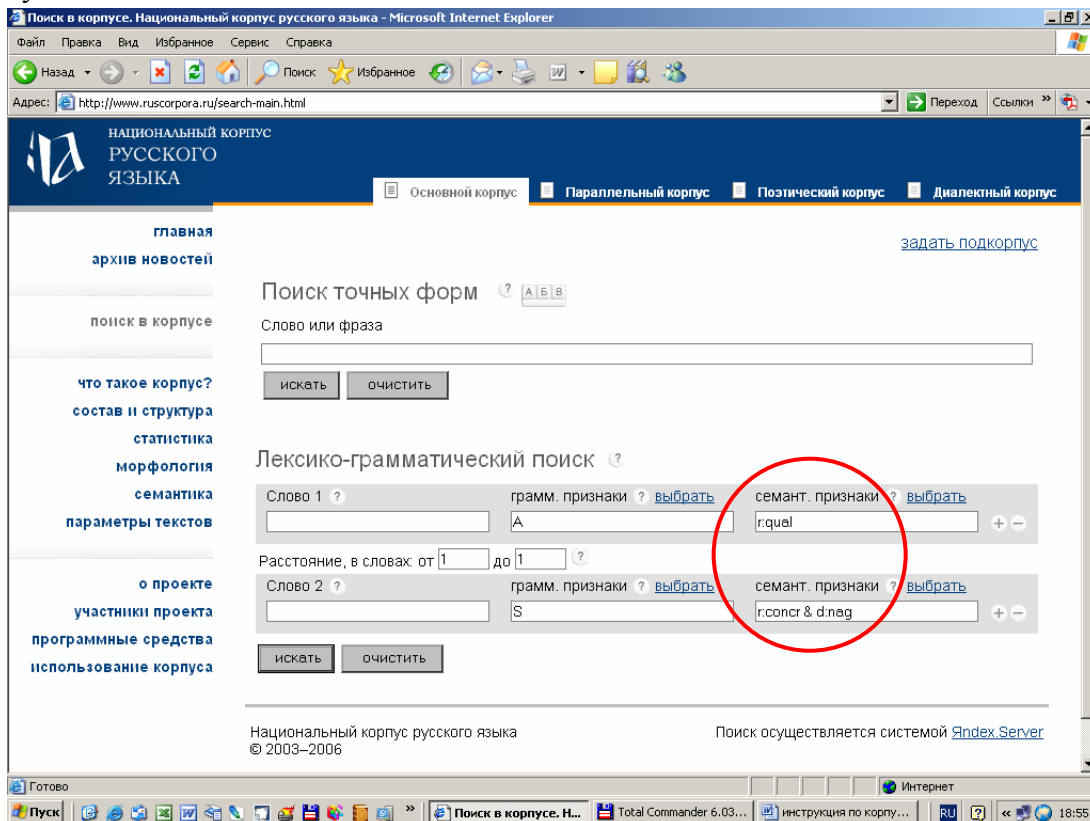


Результат поиска имеет следующий вид:

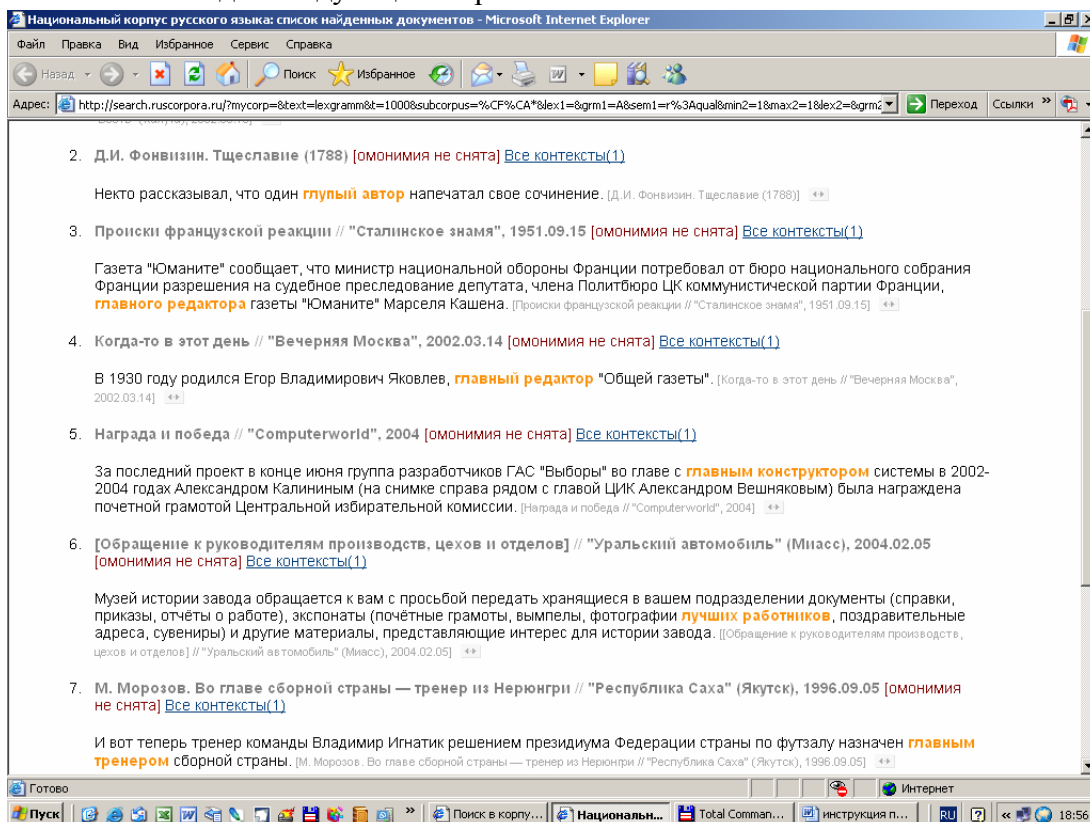


44 шаг

Сорок четвертый шаг. Как найти в Корпусе словосочетания, первое и второе слово в которых характеризуются какими-либо семантическими признаками? Например, как найти сочетания качественных прилагательных с именами деятеля? На странице поиска в поле «Слово 1» в графе «семант. признаки» в разделе «Прилагательные» выбираем значение «Качественные прилагательные», а в поле «Слово 2» в графе «семант. признаки» в разделе «Предметные имена» в графе «Словообразование» выбираем *nomina agentis* (см. шаг тридцать шестой). Страница поиска имеет следующий вид:



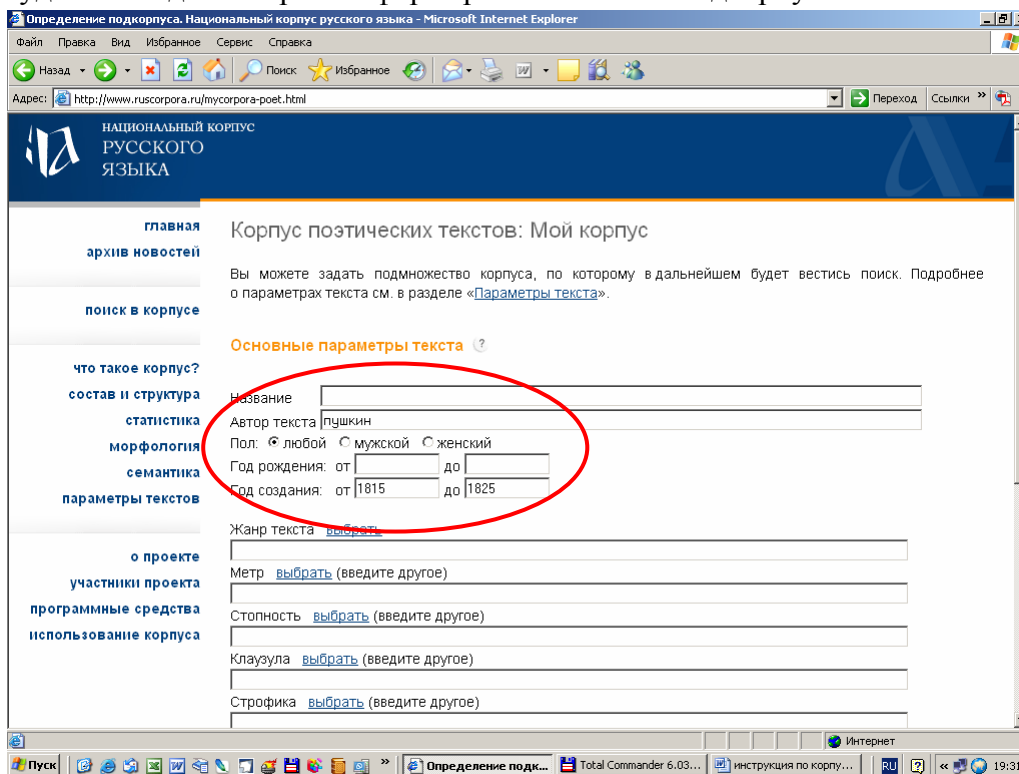
Результаты поиска выглядят следующим образом:



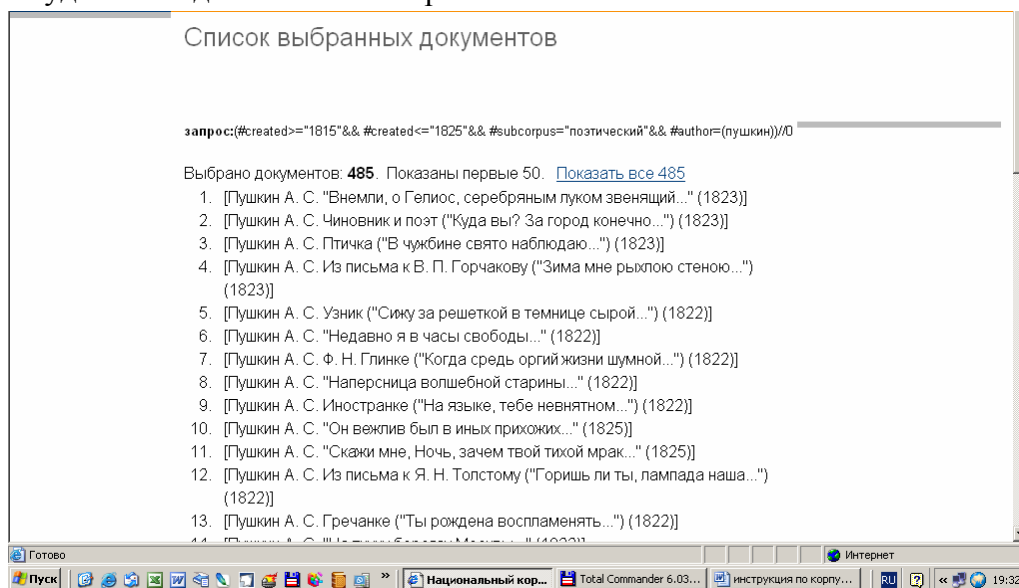
Сорок пятый шаг. Стандартные типы поиска в Поэтическом корпусе.

Поэтическим корпусом можно пользоваться так же, как и Основным. После того, как Вы осуществили **шаг четвертый**, Вы можете сделать **шаги с девятого по тринадцатый**, т.е. задать свой поэтический подкорпус – точно так же, как мы задавали свой прозаический подкорпус, и в дальнейшем осуществлять поиск уже в этом подкорпусе. Так, например, осуществив шаг **десятый** (найти тексты одного автора), **тринадцатый** (отобрать тексты по году их создания) и шаг **тридцать третий** (найти в корпусе определенную лексему), мы можем поставить перед собой задачу найти все случаи использования лексемы *красный* в произведениях Пушкина, написанных между 1815 и 1825 годами.

Вот так будет выглядеть запрос на формирование такого подкорпуса:

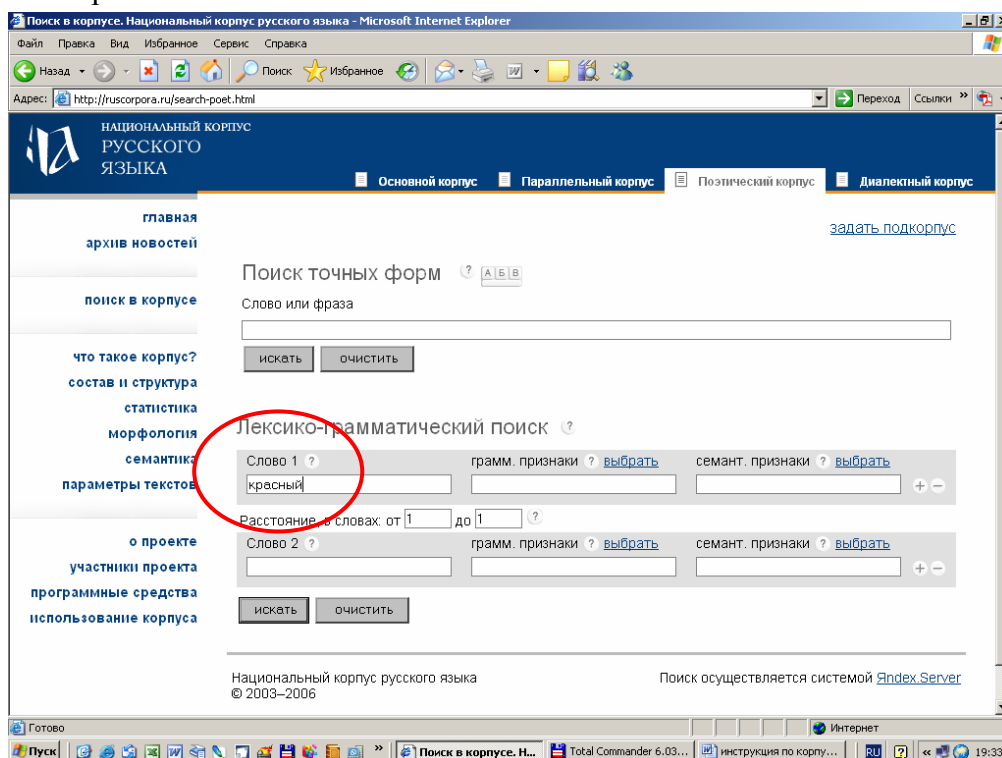


А вот так будет выглядеть список отображенных текстов:

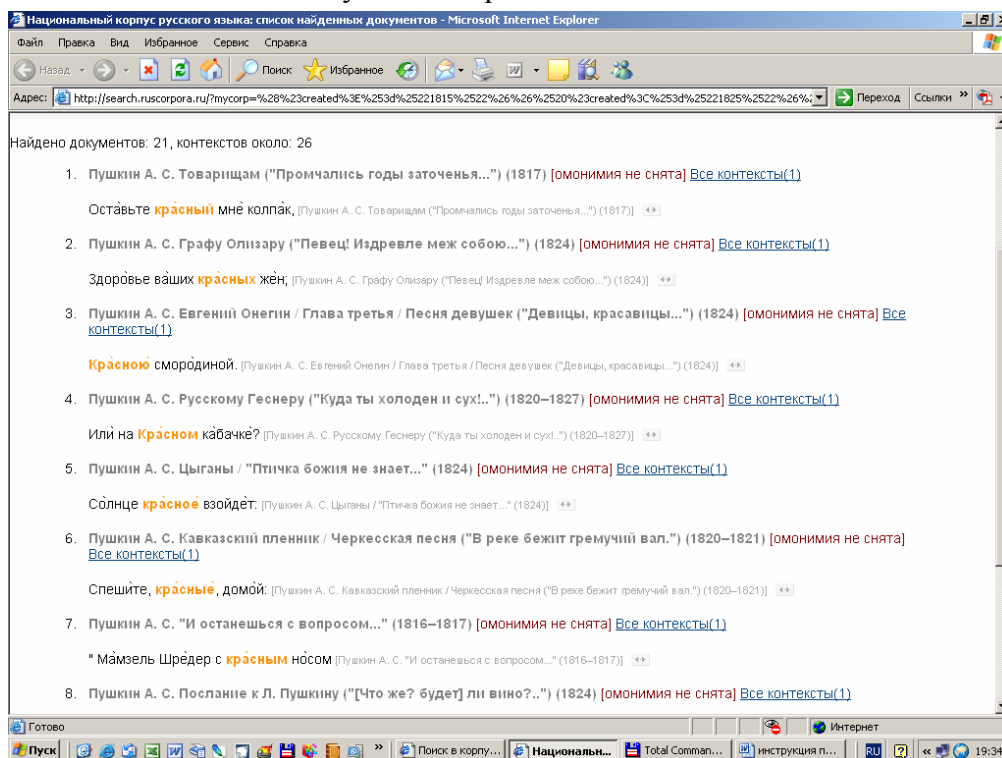


45 шаг

После этого, в соответствии с **шагом тридцатым**, мы возвращаемся на страницу поиска и формируем запрос:



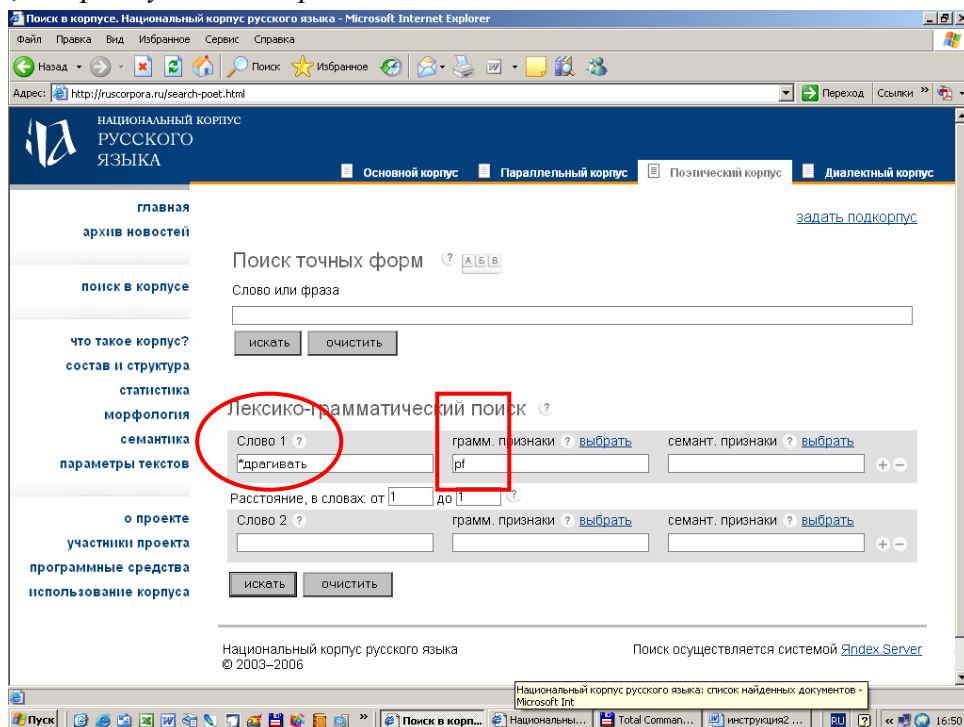
Результаты поиска выглядят следующим образом:



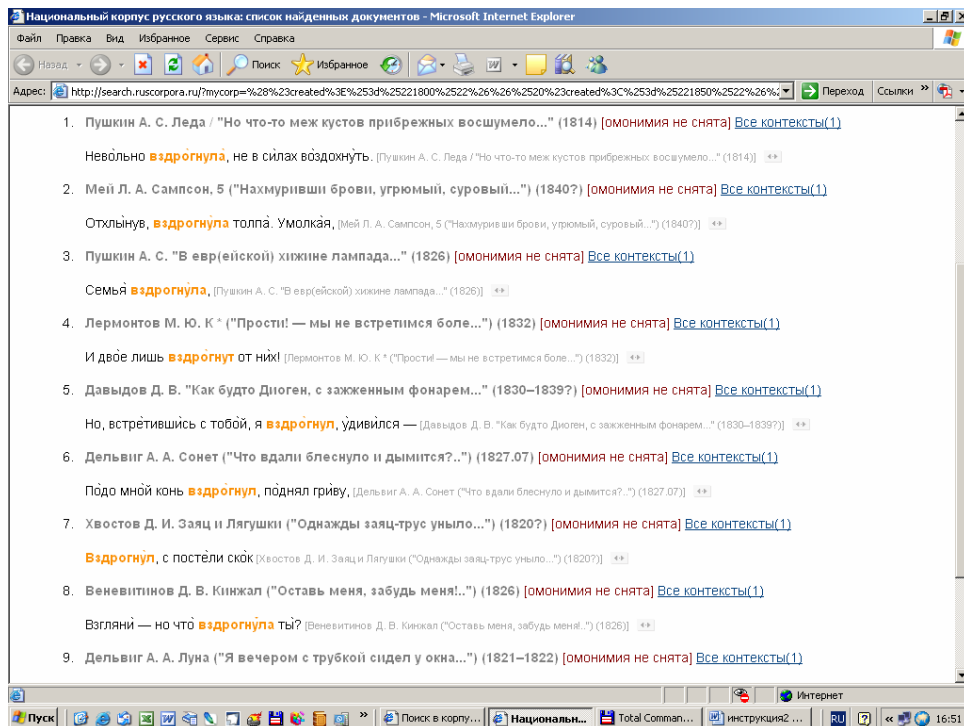
Как видим, основное отличие от результатов аналогичных запросов в Основном корпусе заключается в том, что в поэтическом тексте «обратным» знаком ударения отмечены **сильные доли**. Сильные доли проставлены в соответствии с метром стихотворения. Обращаем Ваше внимание на то, что «обратные» знаки ударения на сильных долях не являются знаками ударения в обычном смысле! Но поскольку в нормальном случае словесное ударение в силлабо-тоническом стихе падает именно на сильную, а не на слабую до-

45 шаг

лю, то отмеченные сильные доли в стихотворном тексте являются косвенными свидетельствами современных автору нормативных, «правильных» ударений. Так, например, сформировав подкорпус русской поэзии первой половины 19 века (1800-1850) – так, как указано выше, – и запросив в поисковой форме лексему «*драгивать» и совершенный вид, мы сможем определить, какое ударение в большей степени характерно для русского языка этой эпохи, – *дрОгнуть или *дроГнУть:



Вот результат поиска:



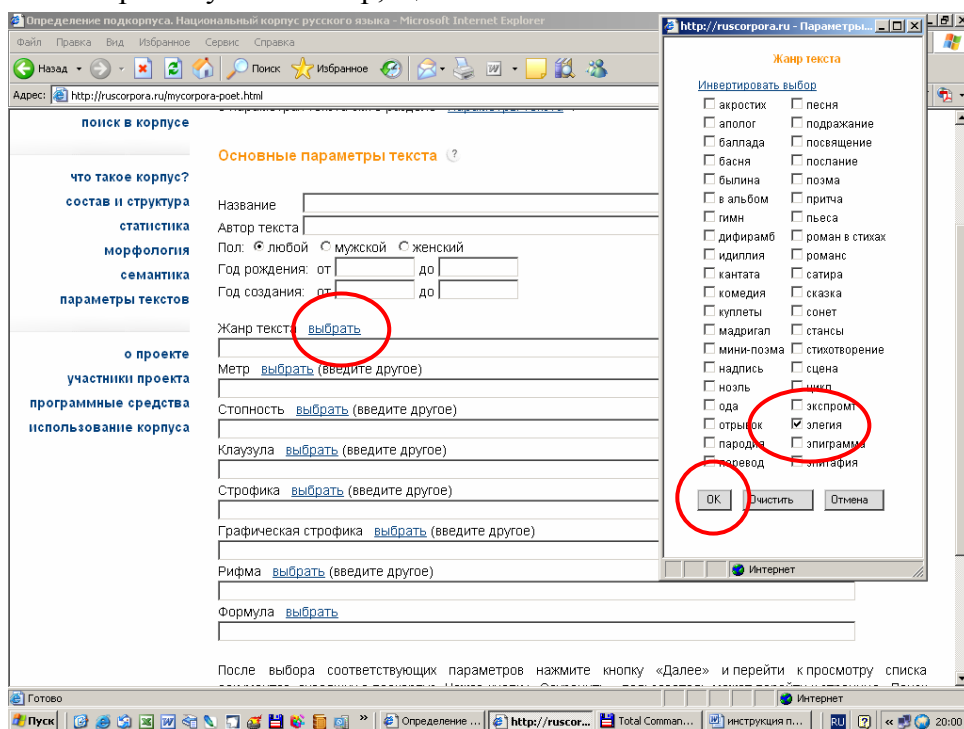
Можно видеть, что ударение в эту эпоху реально колебалось – из 54 полученных контекстов (примеров), в 29 случаях употребляется современное ударение *дрОгнуть, а в 25 случаях – старое *дроГнУть, причем колебания происходят у одного автора и даже в одном и том же тексте.

46 шаг

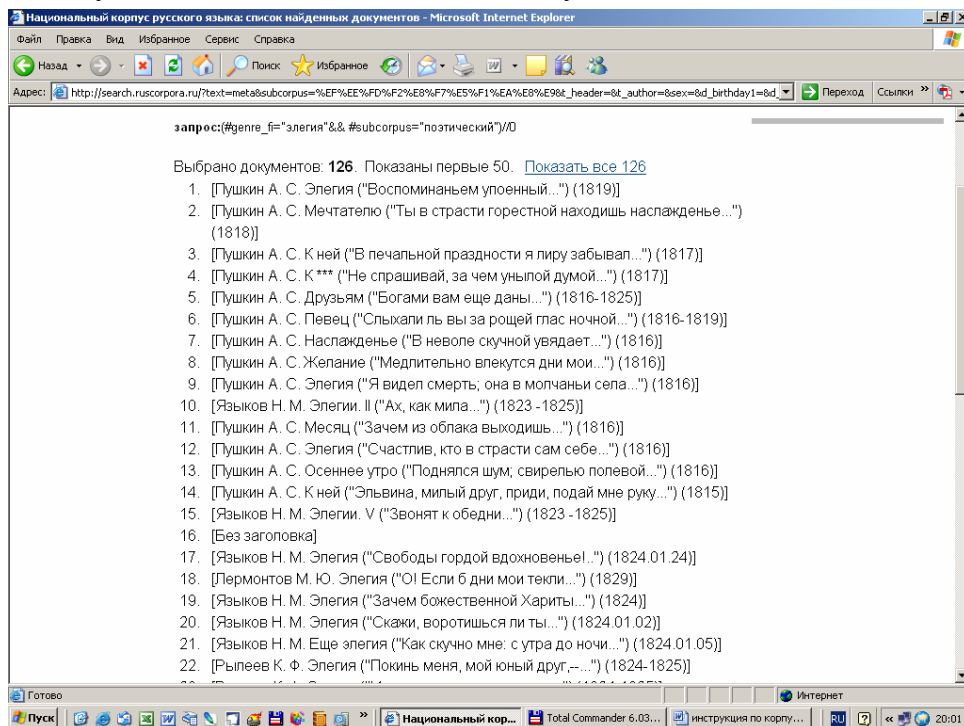
Сорок шестой шаг. Нестандартные типы поиска в Поэтическом корпусе.

Поэтический корпус позволяет Вам находить и подбирать тексты по специальным поэтическим характеристикам. В данный момент Корпус выдает только списки (названия) отобранных текстов, но в ближайшем будущем у Вас появится возможность переходить от списка напрямую к соответствующему тексту. Ниже мы проиллюстрируем возможности отбора, которые есть в Поэтическом корпусе и, естественно, либо отсутствуют в Основном, прозаическом корпусе, либо чем-то отличаются от аналогичных позиций в Основном корпусе.

1. Отбор текстов по жанру. Осуществите **седьмой шаг, первую часть**, в Поэтическом корпусе. В строке «Жанр текста» щелкните по кнопке «Выбрать». В открывшемся окне выберите нужный жанр, щелкните по кнопке ОК.

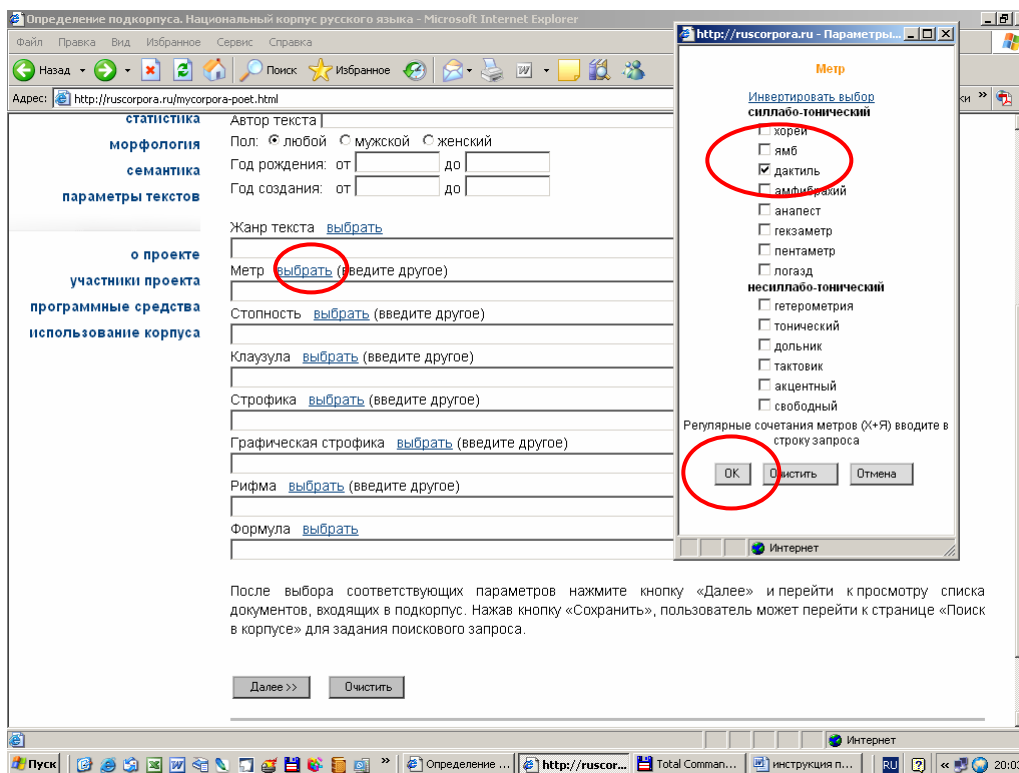


После этого осуществите **шаг тридцатый** и получите список элегий:

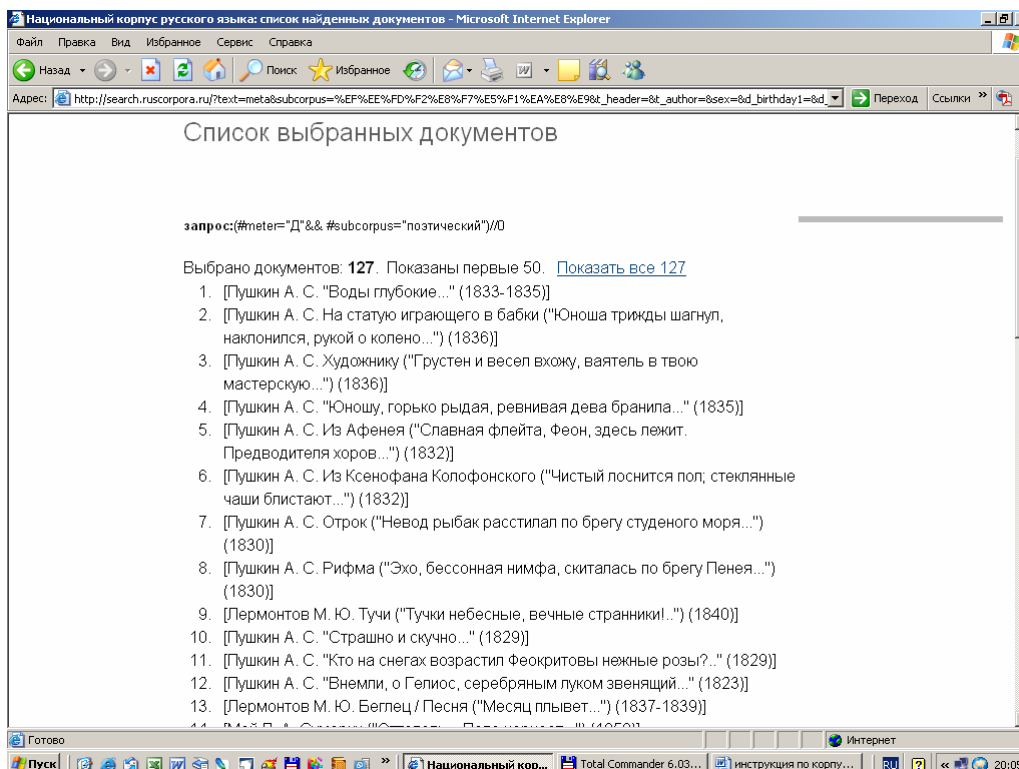


46 шаг

2. Отбор текстов по метру. Осуществите **седьмой шаг, первую часть**, в Поэтическом корпусе. В строке «Метр» щелкните по кнопке «Выбрать». В открывшемся окне выберите нужный метр, щелкните по кнопке ОК.



После этого осуществите **шаг тридцатый** и получите список стихотворений, написанных дактилем:



46 шаг

3. Отбор текстов по стопности. Осуществите **седьмой шаг, первую часть**, в Поэтическом корпусе. В строке «Стопность» щелкните по кнопке «Выбрать». В открывшемся окне выберите нужное число стоп в строке, щелкните по кнопке ОК.

После выбора соответствующих параметров нажмите кнопку «Далее» и перейти к просмотру списка документов, входящих в подкорпус. Нажав кнопку «Сохранить», пользователь может перейти к странице «Поиск в корпусе» для задания поискового запроса.

После этого осуществите **шаг тридцатый** и получите список восьмистопных стихотворений:

НАЦИОНАЛЬНЫЙ КОРПУС
РУССКОГО
ЯЗЫКА

Список выбранных документов

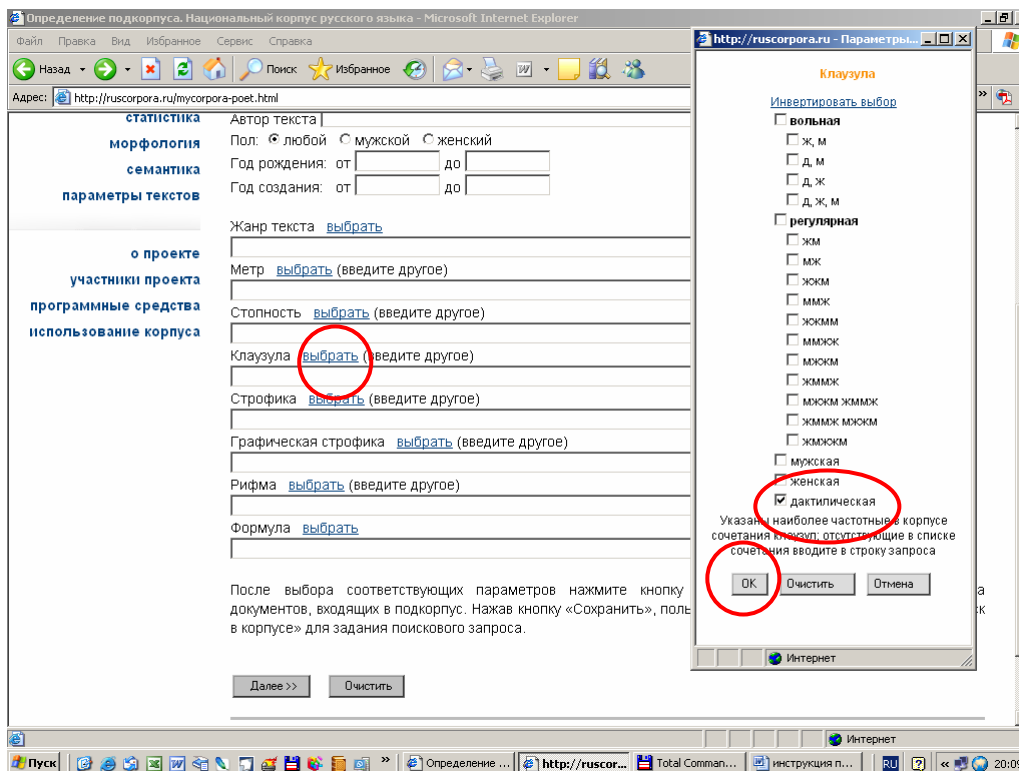
запрос:(#clausula="8"&& #subcorpus="поэтический")//D

Выбрано документов: 1.

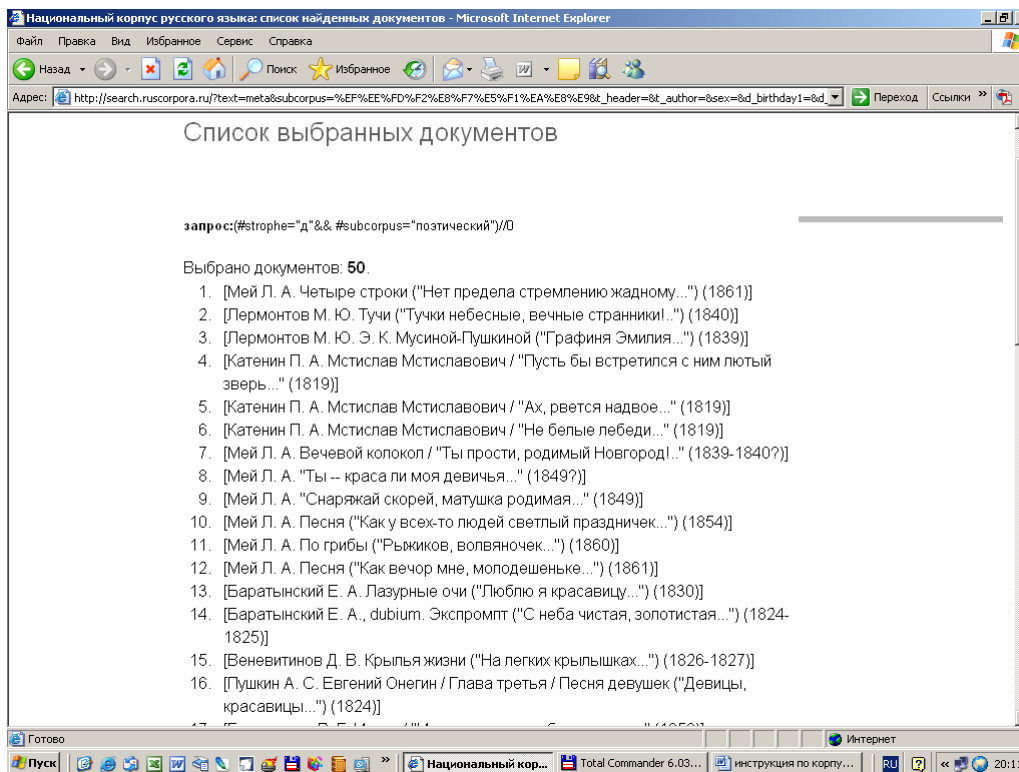
1. [Мей Л. А. Вечевой колокол / "Ты прости, мой брат названный, буйный Волхов мой, прости..." (1839-1840?)]

46 шаг

4. Отбор текстов по клаузуле (т.е. по количеству слогов после последнего ударного слога в строке). Осуществите **седьмой шаг, первую часть**, в Поэтическом корпусе. В строке «Клаузула» щелкните по кнопке «Выбрать». В открывшемся окне выберите нужный тип клаузулы, щелкните по кнопке ОК.

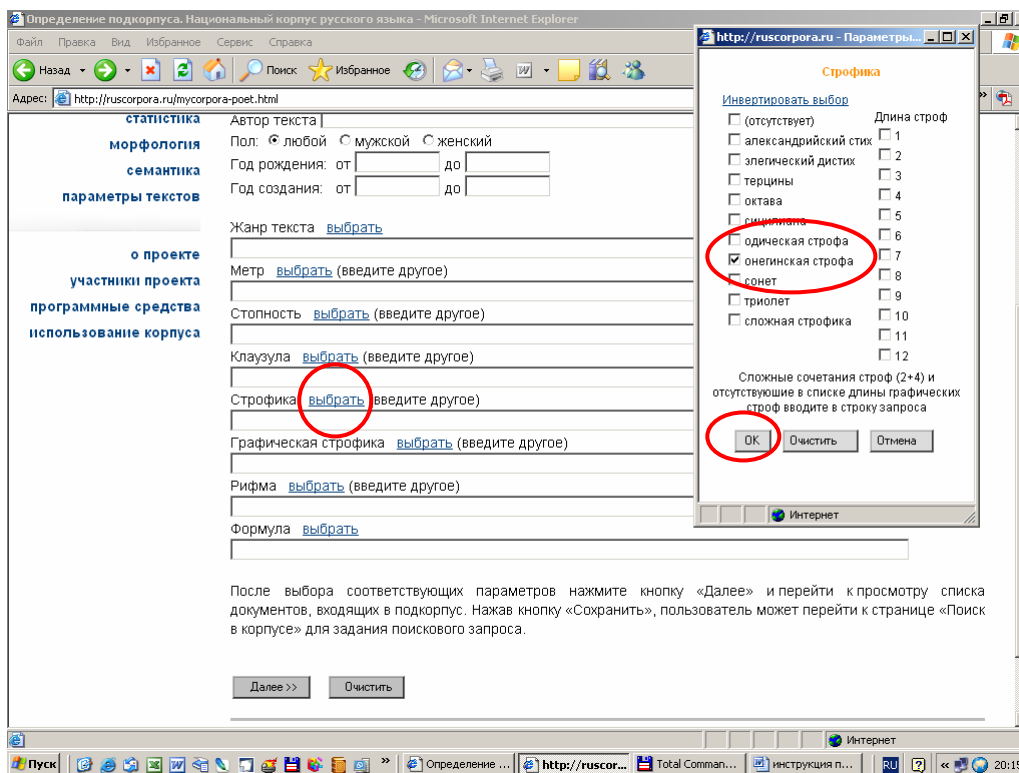


После этого осуществите **шаг тридцатый** и получите список стихотворений, имеющих дактилическую клаузулу

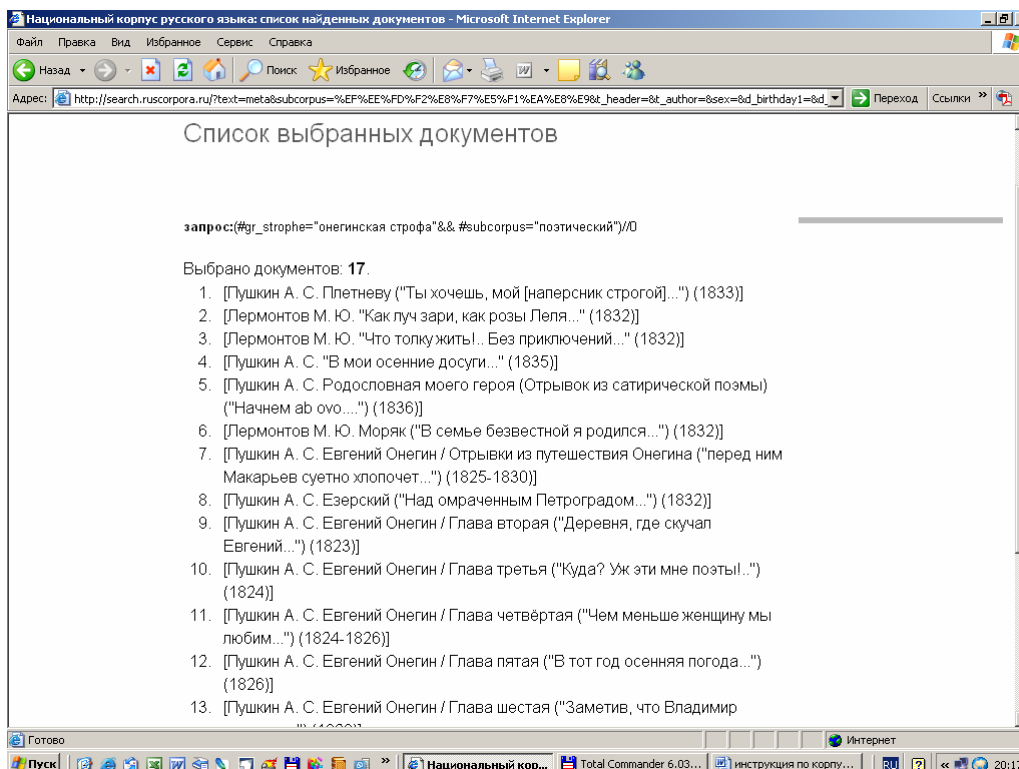


46 шаг

5. Отбор текстов по типу строфы. Осуществите **седьмой шаг, первую часть**, в Поэтическом корпусе. В строке «Строфика» щелкните по кнопке «Выбрать». В открывшемся окне выберите нужный тип строфы, щелкните по кнопке ОК.



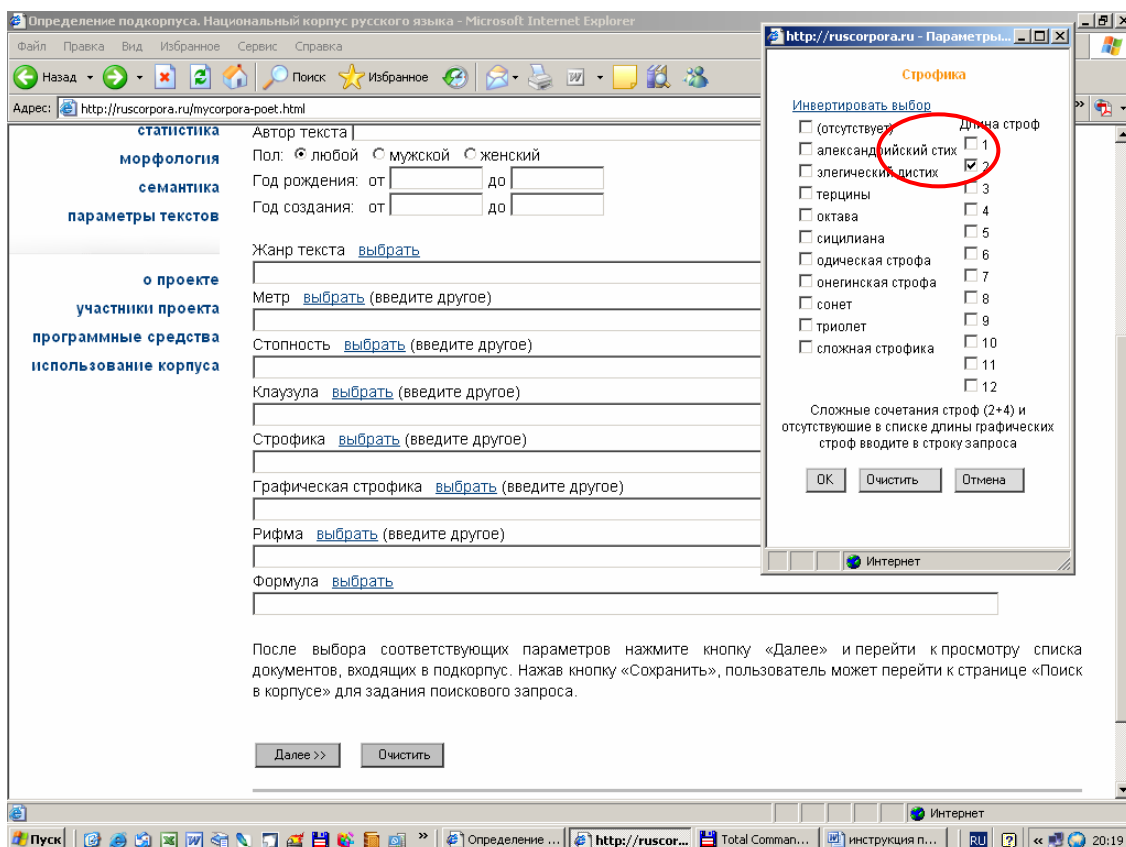
После этого осуществите **шаг тридцатый** и получите список произведений, написанных онегинской строфой:



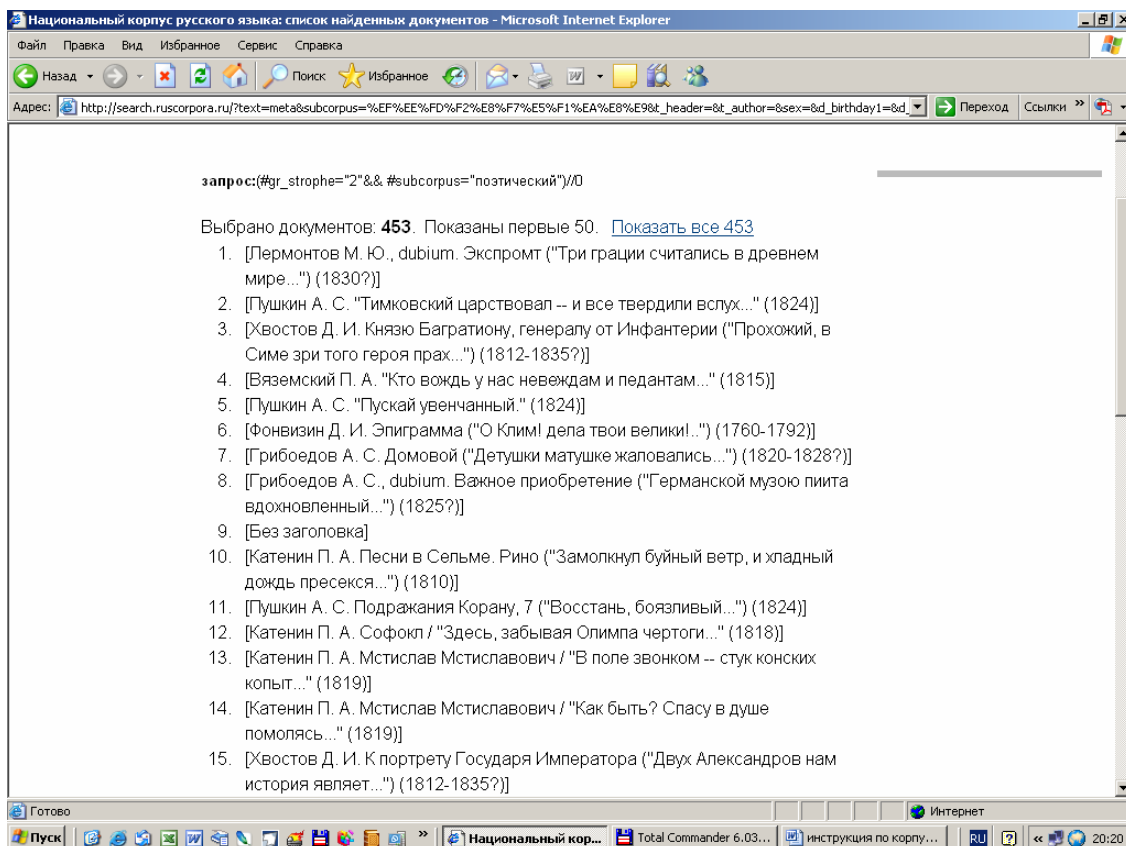
Как видите, онегинской строфой был написан не только «Евгений Онегин», и использовал этот тип строфы не только Пушкин.

46 шаг

Обратим внимание, что в качестве типа строфы может быть задан ее размер, т.е. количество строк в строфе:

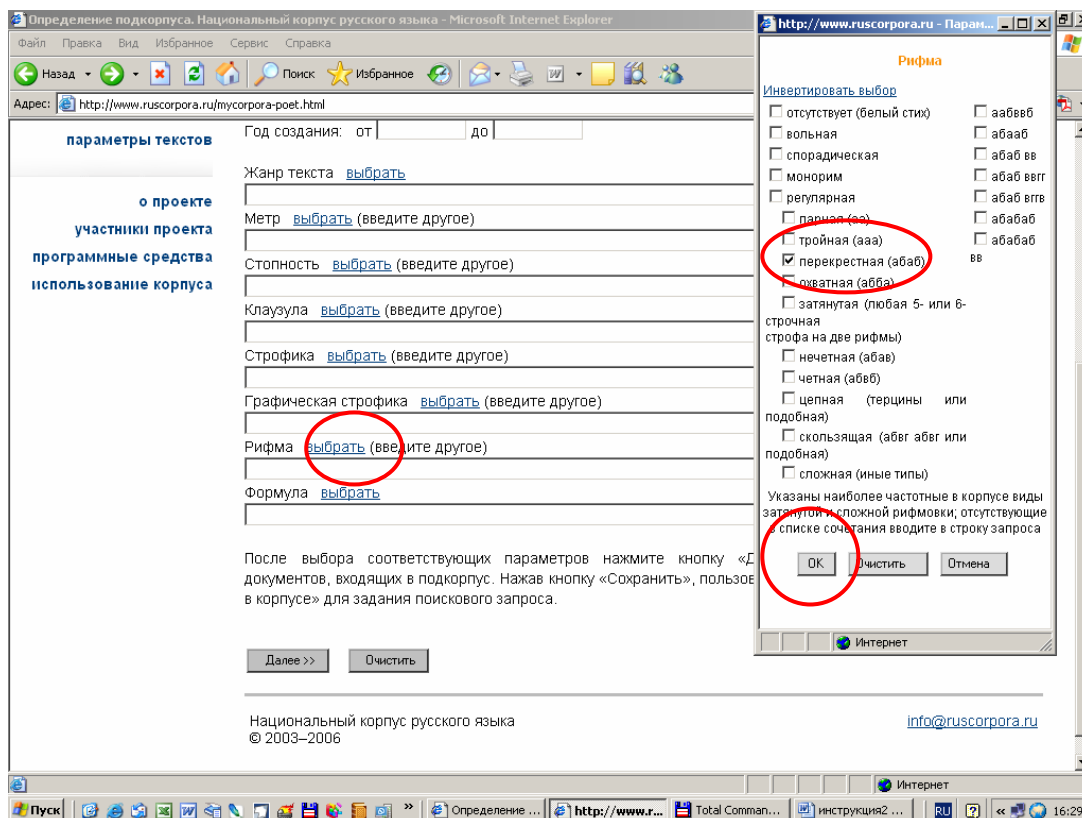


При таком запросе мы можем получить все стихотворения, написанные двустишиями:

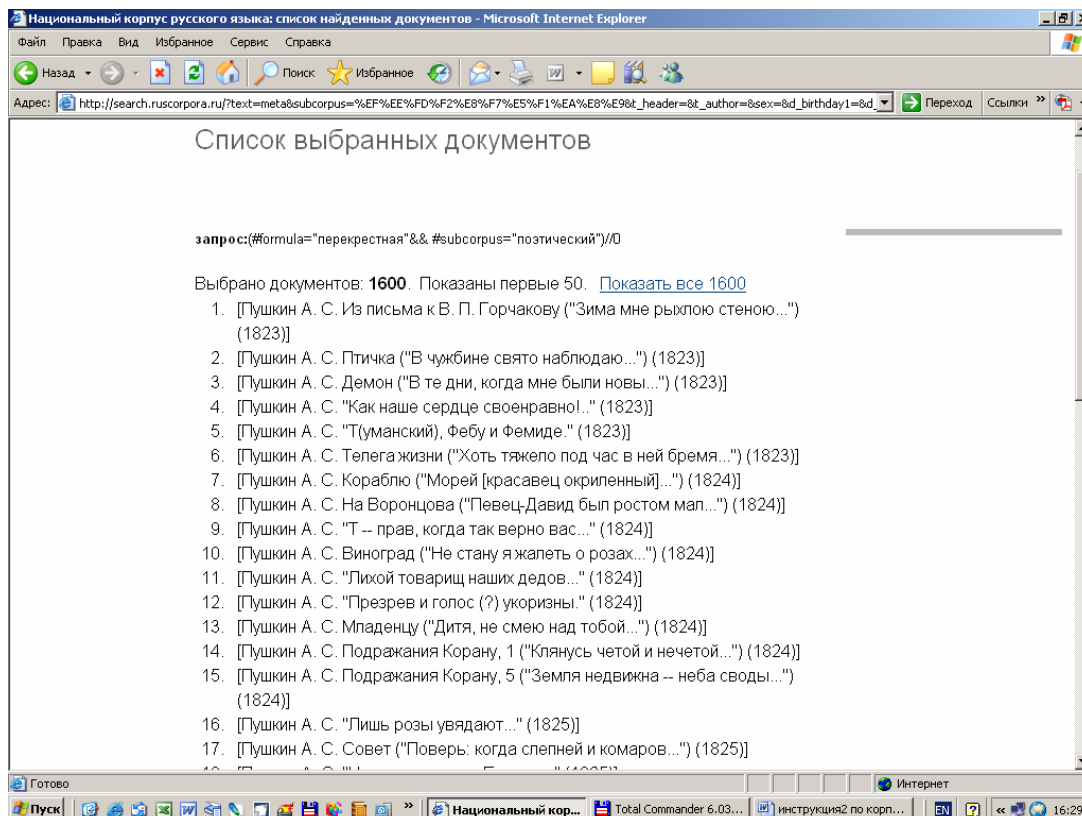


46 шаг

6. Отбор текстов по типу рифмы. Осуществите **седьмой шаг, первую часть**, в Поэтическом подкорпусе. В строке «Рифма» щелкните по кнопке «Выбрать». В открывшемся окне выберите нужный тип рифмы, щелкните по кнопке ОК.



После этого осуществите **шаг тридцатый** и получите список стихотворений, имеющих перекрестную рифму:



46 шаг

7. Поиск по всем параметрам. Разумеется, все перечисленные выше параметры поиска можно использовать одновременно. Поставим, например, задачу найти 1) романы, 2) сочиненные поэтами-мужчинами, 3) написанные четырехстопным ямбом 4) с перекрестной рифмой и 5) регулярным чередованием женской и мужской клаузулы (жм); 6) строфа при этом должна содержать 4 строки. Соответствующий запрос будет выглядеть следующим образом:

Определение подкорпуса. Национальный корпус русского языка - Microsoft Internet Explorer

Адрес: <http://www.ruscorpora.ru/mycorporapoeet.html>

статистика морфология семантика параметры текстов

Автор текста |

Пол: любой мужской женский

Год рождения: от до

Год создания: от до

Жанр текста [выбрать](#)

романс

Метр [выбрать](#) (введите другое)

Я

Стопность [выбрать](#) (введите другое)

4

Клаузула [выбрать](#) (введите другое)

жм

Строфика [выбрать](#) (введите другое)

4

Графическая строфика [выбрать](#) (введите другое)

Рифма [выбрать](#) (введите другое)

перекрестная

Формула [выбрать](#)

После выбора соответствующих параметров нажмите кнопку «Далее» и перейти к просмотру списка документов, входящих в подкорпус. Нажав кнопку «Сохранить», пользователь может перейти к странице «Поиск в корпусе» для задания поискового запроса.

А вот список текстов, найденных по этому запросу:

Адрес: http://search.ruscorpora.ru/?text=meta&subcorpus=%EF%EE%FD%F2%E8%F7%E5%F1%EA%E8%E9&t_header=&t_author=&sex=%EC%F3%E6&d...

НАЦИОНАЛЬНЫЙ КОРПУС РУССКОГО ЯЗЫКА

Список выбранных документов

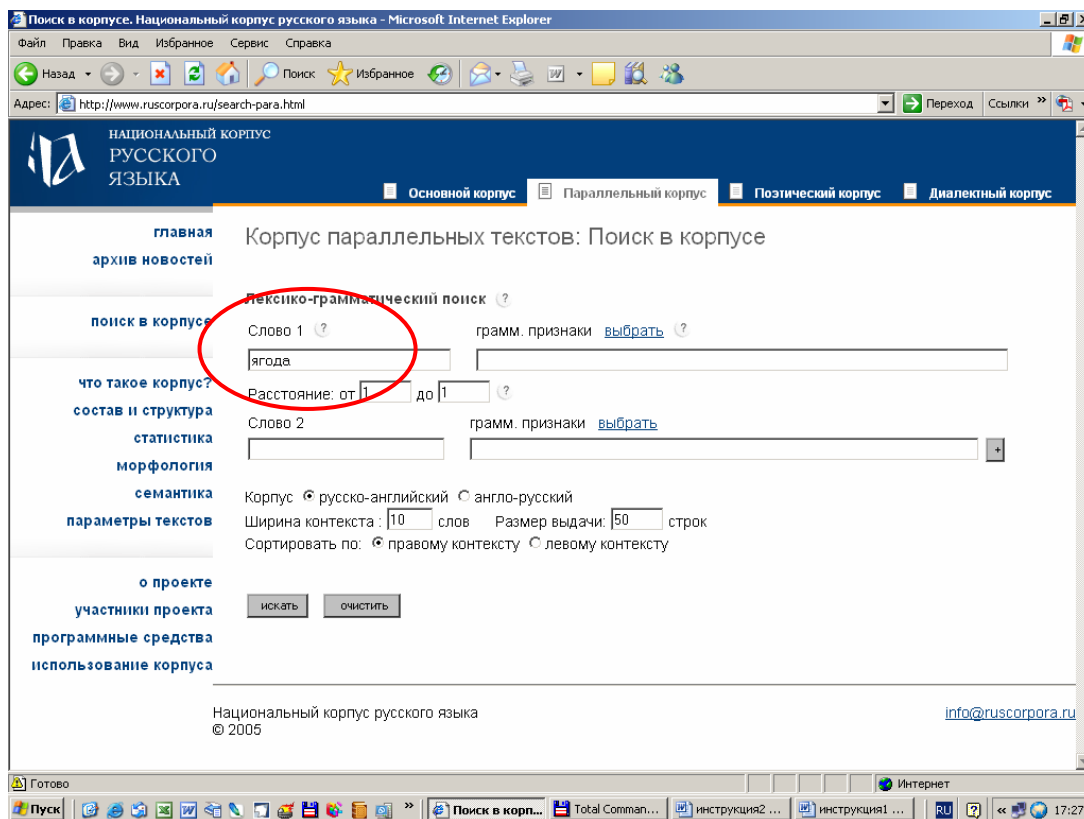
запрос:({#clausula="4"&& #formula="перекрестная"&& #genre_fi="романс"&& #gr_strophe="4"&& #meter="Я"&& #sex="муж"&& #strophe="жм"&& #subcorpus="поэтический"/})

Выбрано документов: 3.

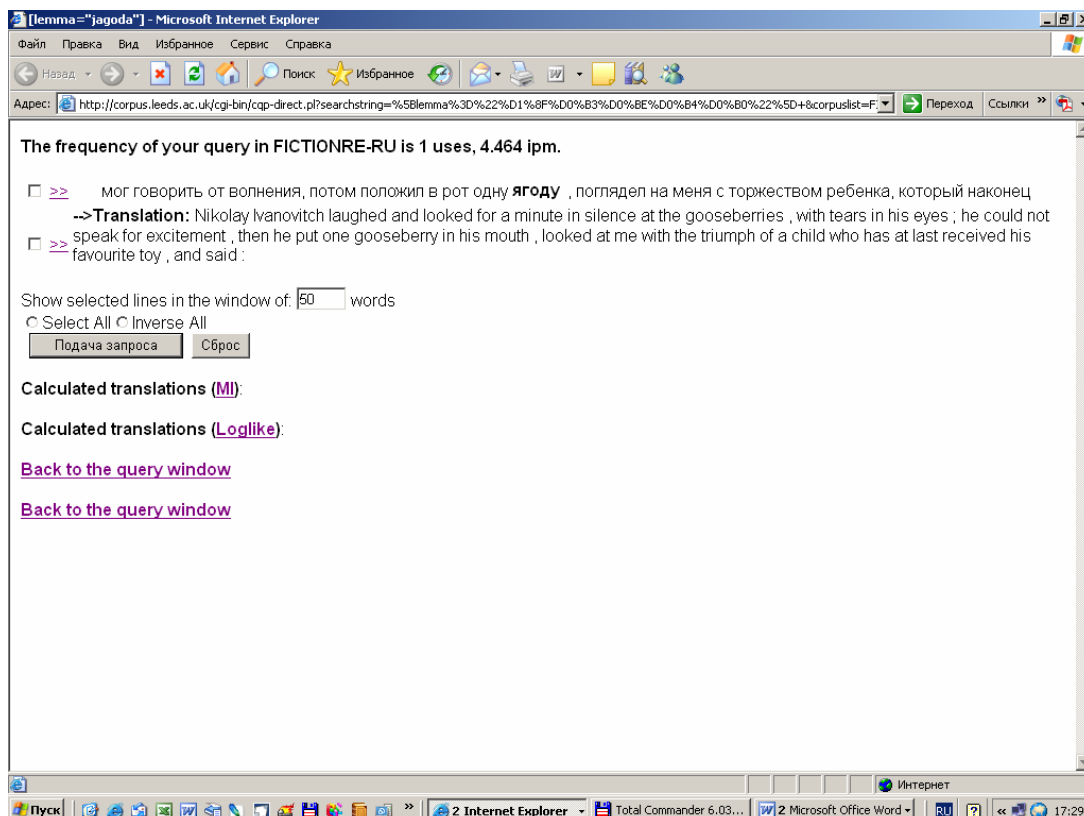
1. [Языков Н. М. Романс ("Красой небесною прекрасна...") (1824)]
2. [Пермонтов М. Ю. Романс ("Коварной жизнью недовольный...") (1829)]
3. [Пушкин А. С. Романс ("Под вечер, осенью ненастной...") (1814)]

47 шаг

Сорок седьмой шаг. Как найти слово в русско-английском параллельном корпусе? После того, как Вы сделаете **шаг пятый**, Вы окажетесь на странице поиска Параллельного корпуса. Предположим, Вы хотите найти в нем слово *ягода* и примеры его перевода на английский язык, зафиксированные в Параллельном корпусе. Ваш запрос будет выглядеть следующим образом:

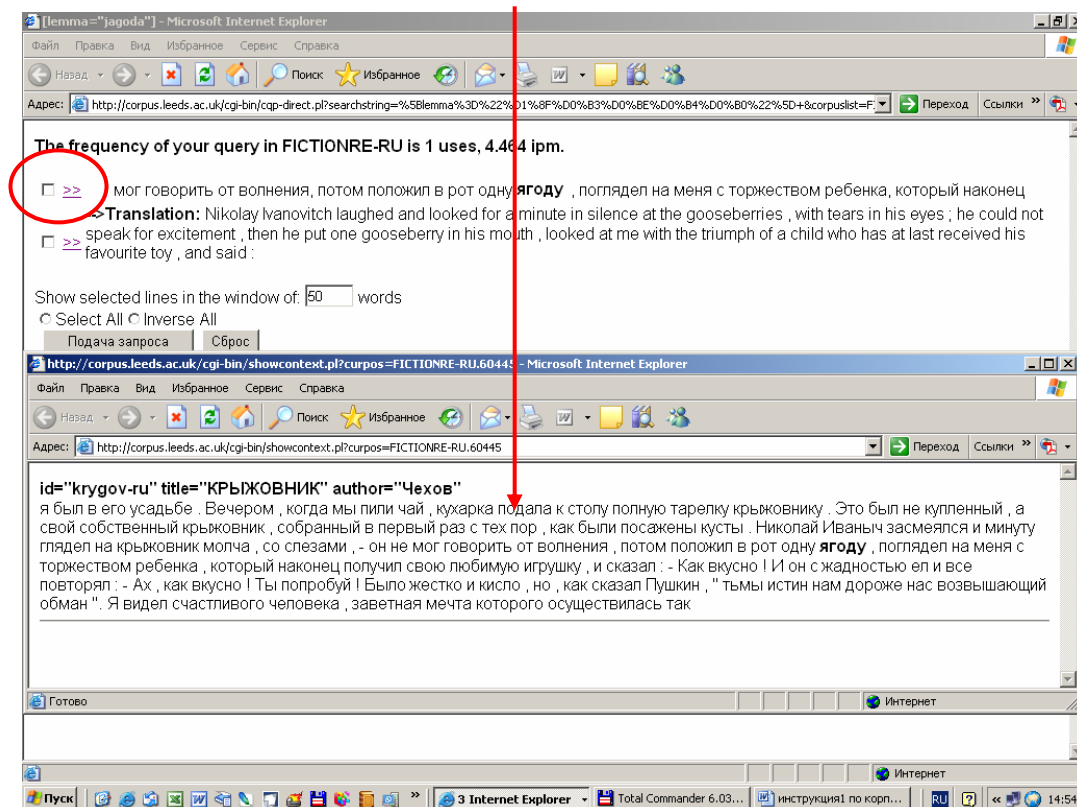


Щелкните мышкой на кнопке «Искать» и получите ответ на этот запрос:

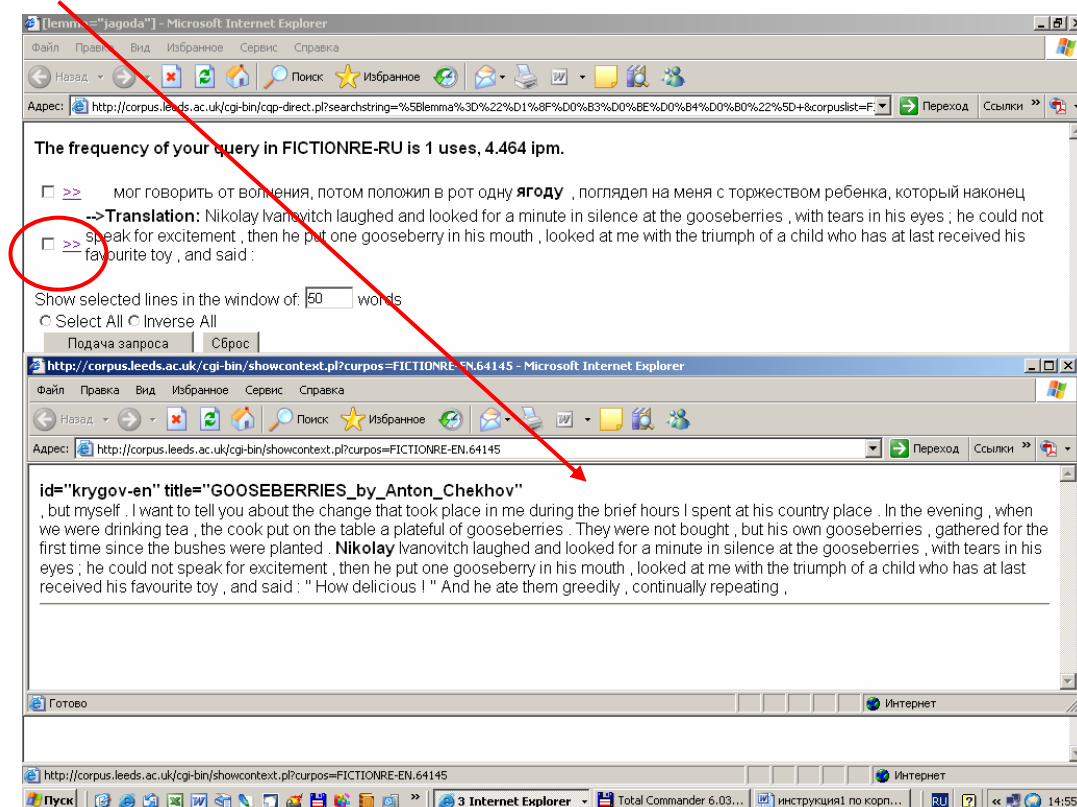


47 шаг

Если Вы хотите расширить контекст и узнать источник русского текста, щелкните мышкой на стрелке напротив русского текста. В отдельном окне Вы увидите исходный русский текст (из рассказа А.П.Чехова «Крыжовник»):

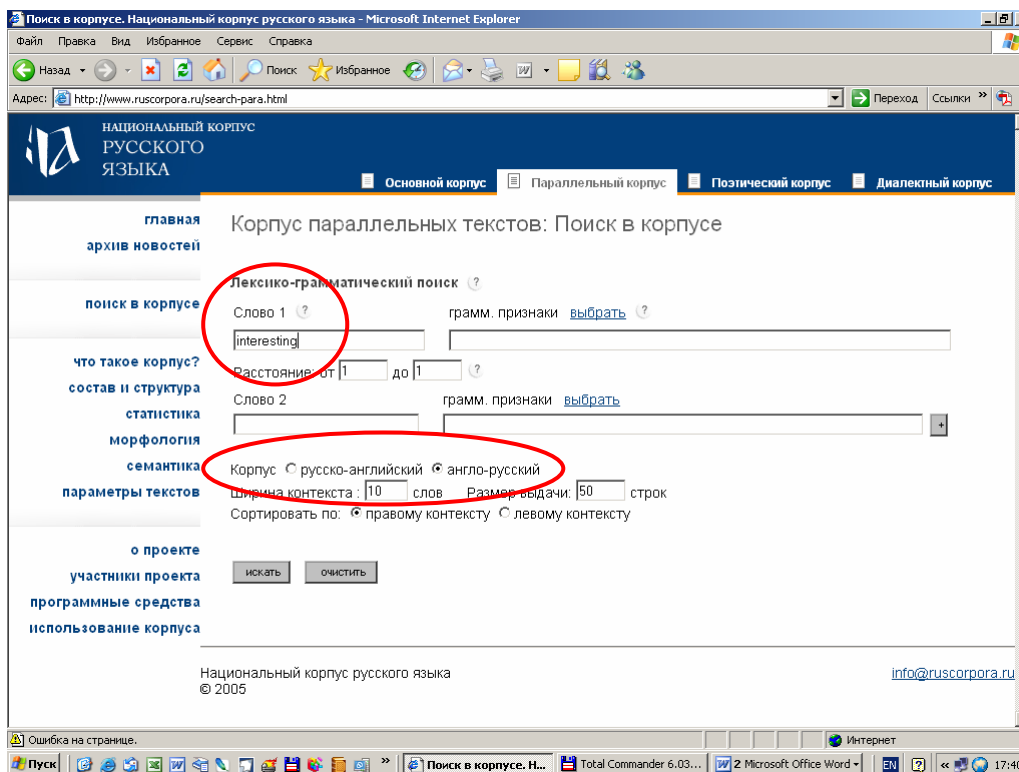


Если Вы захотите расширить контекст английского перевода, щелкните на соответствующей стрелочке напротив английского текста и получите расширенный английский контекст:

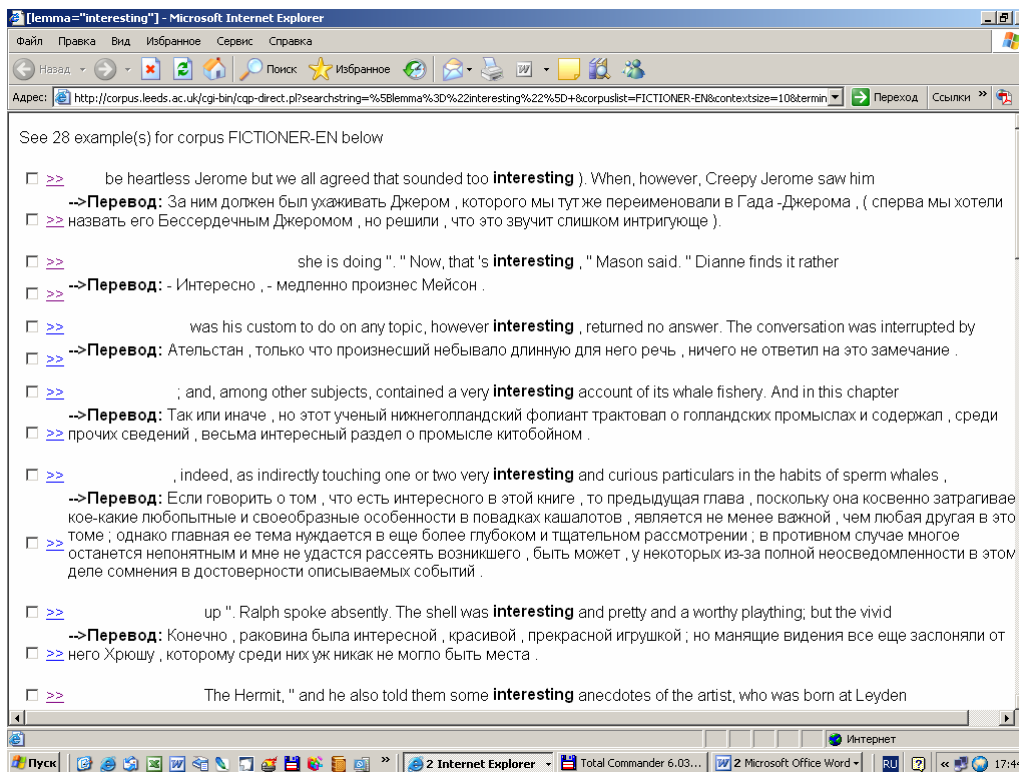


48 шаг

Сорок восьмой шаг. Как найти слово в англо-русском параллельном корпусе? После **шага пятого** Вы, находясь на странице поиска Параллельного корпуса, должны переключиться из русско-английского корпуса в англо-русский, затем в строке Слово 1 набрать интересующее Вас слово, например, interesting, и щелкнуть мышкой на кнопке «Искать»:



Результат поиска выглядит следующим образом:



48 шаг

В окне результата Вы можете использовать стрелочки для расширения английского и русского контекстов (см. шаг сорок седьмой):

The screenshot shows a Microsoft Internet Explorer browser window displaying search results from the Corpus Leeds website. The address bar shows the URL: <http://corpus.leeds.ac.uk/cgi-bin/showcontext.pl?curpos=FICTIONER-EN.652116>. The search results are for the author "Walter Scott" and the title "IVANHOE".

The English text displayed is:

id="walter_scott_new-en" author="Walter_Scott" title="IVANHOE"
carelessness; `` for he whose hands are tied should have his tongue at freedom . But thy glibness of reply , comrade , `` rejoined he , speaking to Athelstane , `` will not win the freedom of the Lady Rowena `` . To this Athelstane , who had already made a longer speech than was his custom to do on any topic , however **interesting** , returned no answer . The conversation was interrupted by the arrival of a menial , who announced that a monk demanded admittance at the postern gate . `` In the name of Saint Bennet , the prince of these bull-beggars , `` said Front-de-Buf , `` have we a real monk this time , or another impostor ? Search him , slaves for an ye suffer

The Russian translation displayed is:

id="walter_scott_new-ru" author="Вальтер Скотт" title="АЙВЕНГО"
, полученной его приятелем , - сакс отдал тебе довольно метко " . " Да , насколько пленник способен попадать в цель , - отвечал де Браси с притворной беззаботностью , - у кого руки связаны , тот и дает волю своему языку " . Но хотя ты и боек на язык , дружок , - продолжал он , обращаясь к Ательстану , - леди Ровена от этого не станет свободнее . **Ательстан** , только что произнесший небывало длинную для него речь , ничего не ответил на это замечание . Но тут разговор был прерван приходом слуги , который доложил , что у ворот стоит монах и умоляет впустить его . " Ради святого Беннета , покровителя этой нищей братии , - сказал Фрон де Беф , - желал бы я знать , настоящий ли это монах или опять переодетый обманщик " ! Общайте

Below the Russian text, there is a section for expanding the translation:

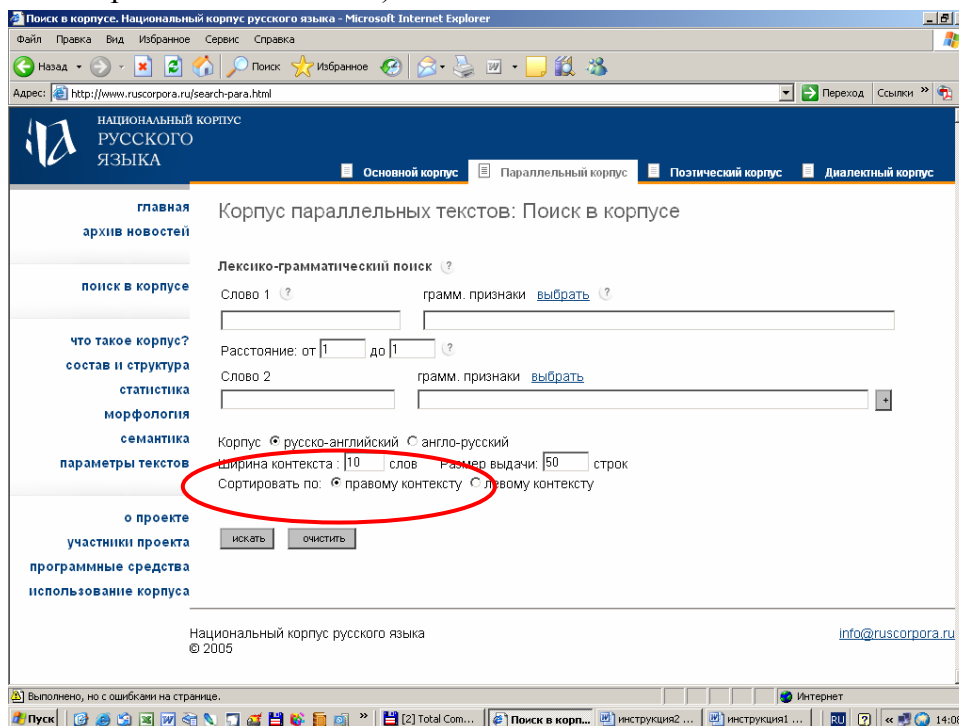
>> up " . Ralph spoke absently. The shell was **interesting** and pretty and a worthy plaything; but the vivid
-->Перевод: Конечно , раковина была интересной , красивой , прекрасной игрушкой ; но манящие видения все еще заслоняли от
 >> него Хрюшу , которому среди них уж никак не могло быть места .
 >> The Hermit , " and he also told them some **interesting** anecdotes of the artist , who was born at Leyden

The taskbar at the bottom shows the system clock at 18:00 and several open applications including Internet Explorer, Total Commander, and Microsoft Office Word.

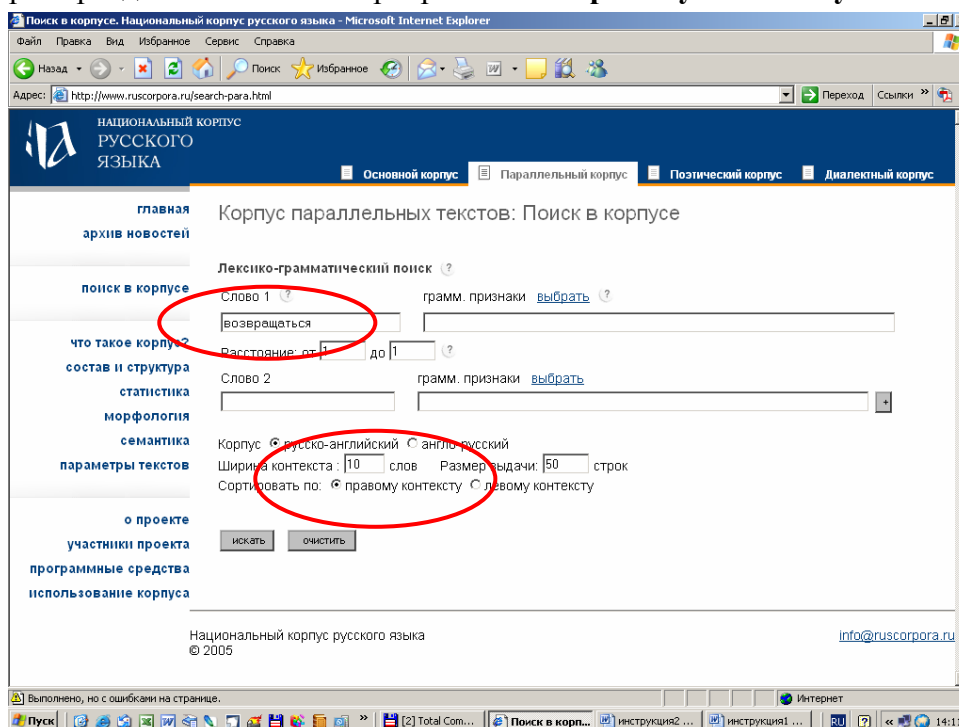
49 шаг

Сорок девятый шаг. Как сортировать примеры в Параллельном корпусе?

После того, как Вы осуществили **шаг сорок седьмой** или **сорок восьмой**, Вы получили некоторый список примеров, с которыми можете работать дальше. Однако для дальнейшей работы бывает удобно, чтобы примеры были определенным образом отсортированы, т.е. расположены в том или ином порядке, особенно когда примеров много. Параллельный корпус предлагает Вам выбрать между сортировкой по правому контексту (эта возможность задается автоматически) или по левому контексту (сортировку по левому контексту нужно выбирать специально перед началом поиска):

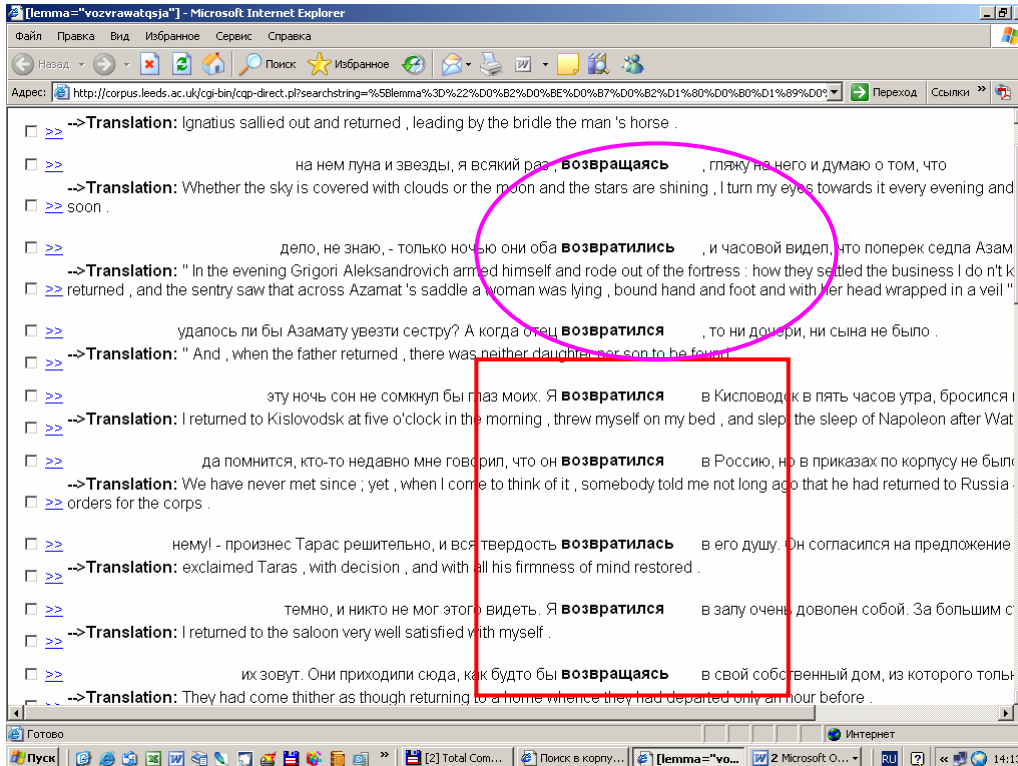


Если Вы выберете **сортировку по правому контексту**, то полученные Вами в результате поиска примеры будут расположены в соответствии с алфавитным порядком слов, следующих непосредственно за словом, которое Вы запросили. Например, Вы ищете глагол *возвращаться*, а примеры должны быть отсортированы **по правому контексту**:

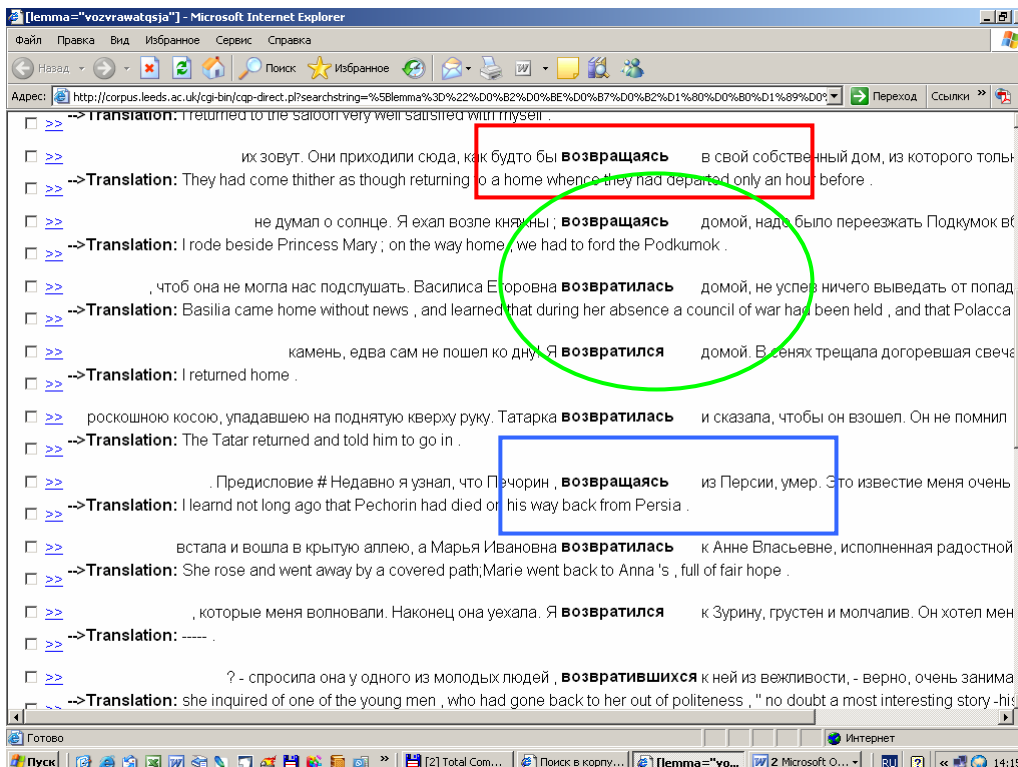


49 шаг

Вот результат, который Вы получаете:

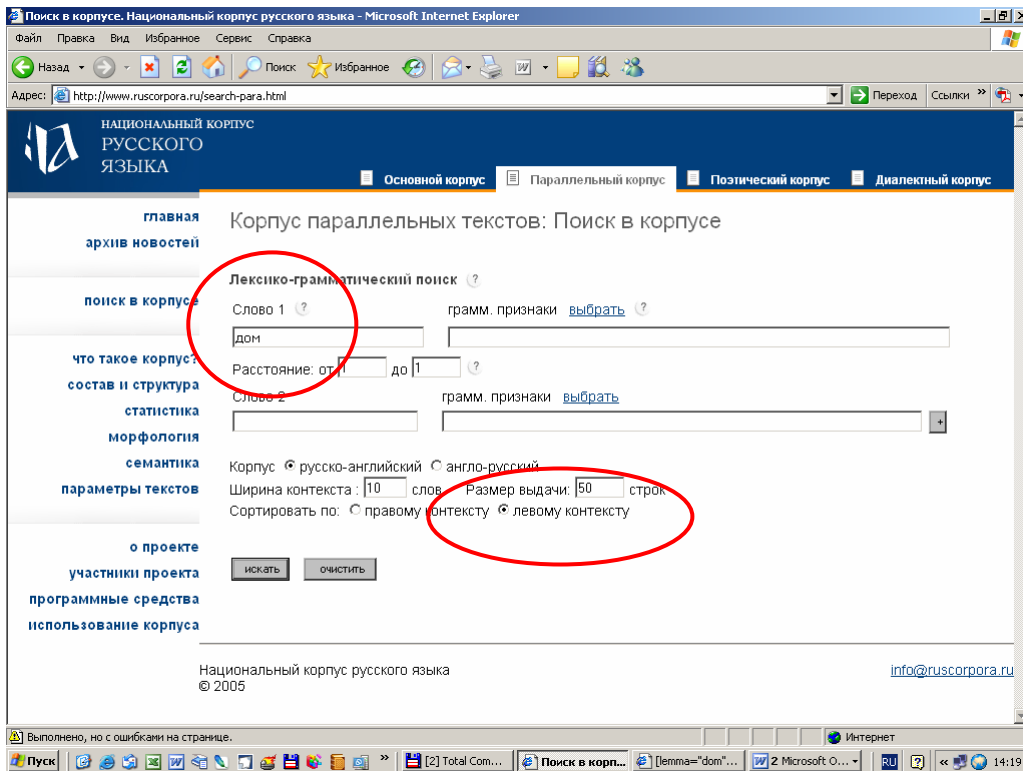


При выдаче контекстов сначала следуют примеры употребления глагола **возвращаться** без зависимых слов справа, а затем – все случаи использования предлога **в** после глагола **возвращаться**. После примеров с предлогом **в** следуют примеры с наречием **домой**, затем с предлогом **из**, **к** и прочее – в алфавитном порядке:

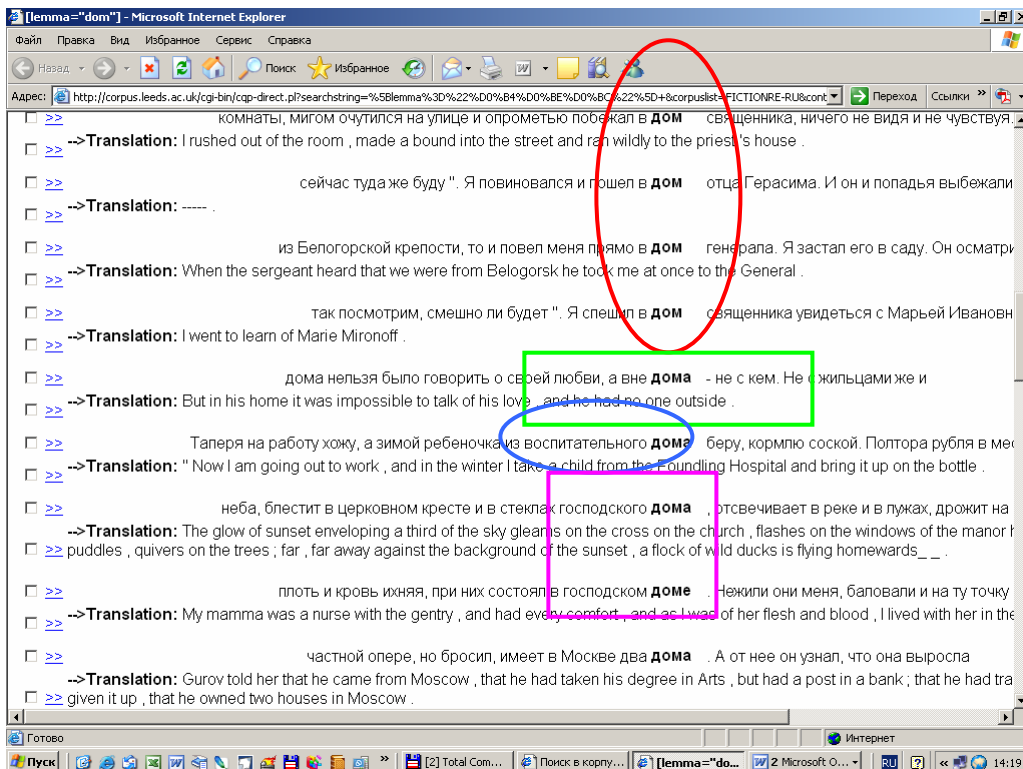


49 шаг

Если Вы выберете **сортировку по левому контексту**, то полученные Вами в результате поиска примеры будут расположены в соответствии с алфавитным порядком слов, непосредственно предшествующих слову, которое Вы запросили. Например, Вы ищете существительное *дом*, а примеры должны быть отсортированы по левому контексту:



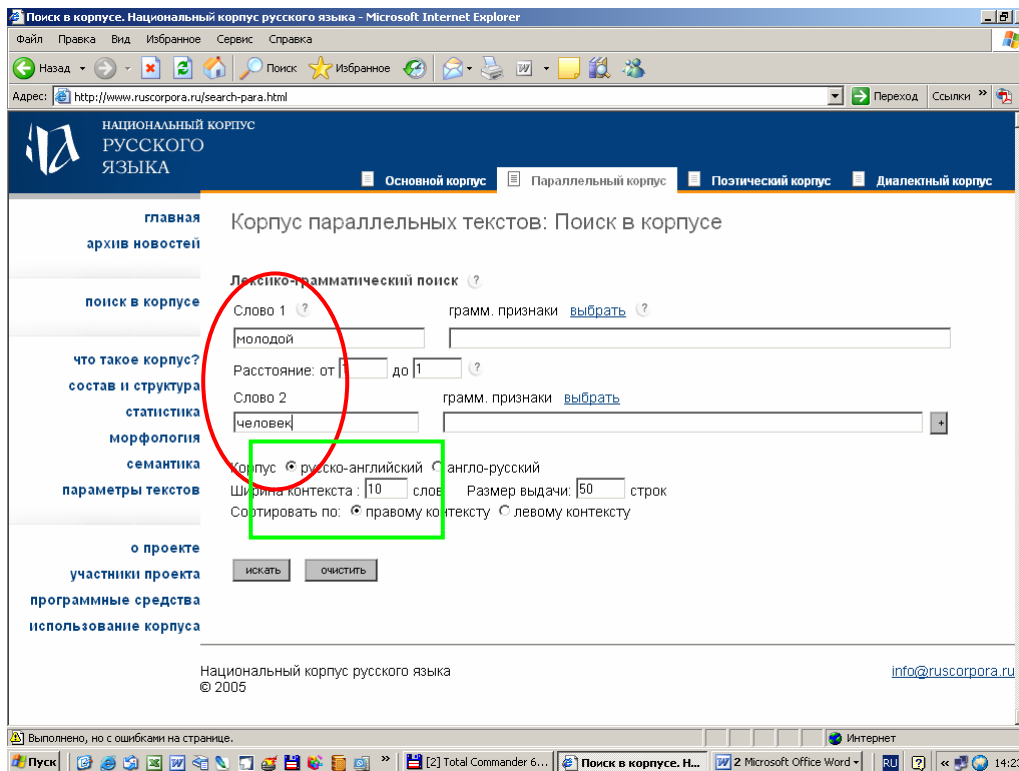
А качестве результата Вы получите следующий набор примеров:



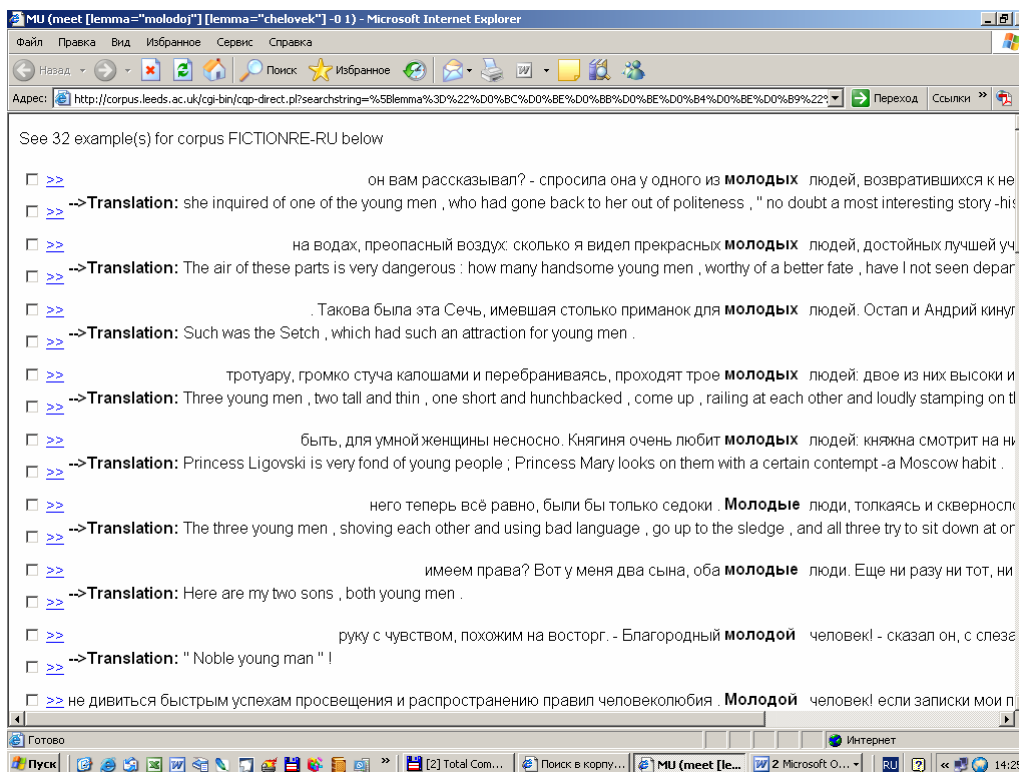
50 шаг

Пятидесятый шаг. Как найти словосочетание в русско-английском и в англо-русском параллельном корпусе?

Словосочетания в Параллельном корпусе ищутся так же, как и в Основном корпусе (см. **шаги с тридцать девятый, сороковой и сорок второй**). Предположим, Вы хотите узнать, как переводится на английский язык русское словосочетание *молодой человек*. Ваш запрос будет выглядеть следующим образом:

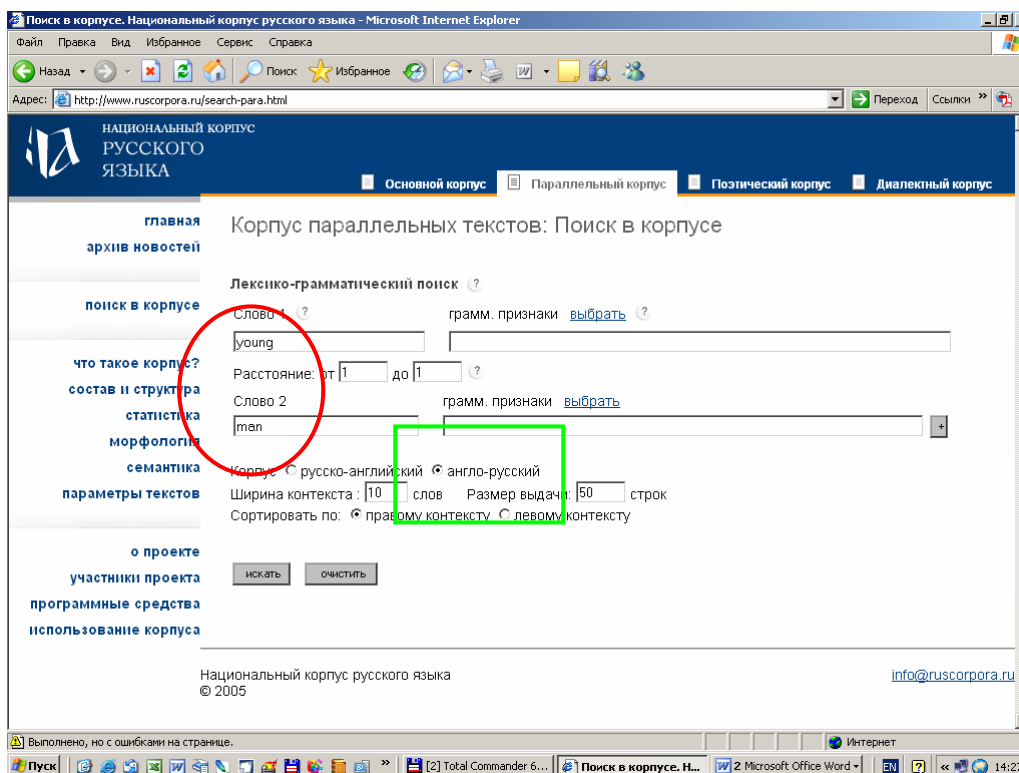


Вот результат Вашего запроса:

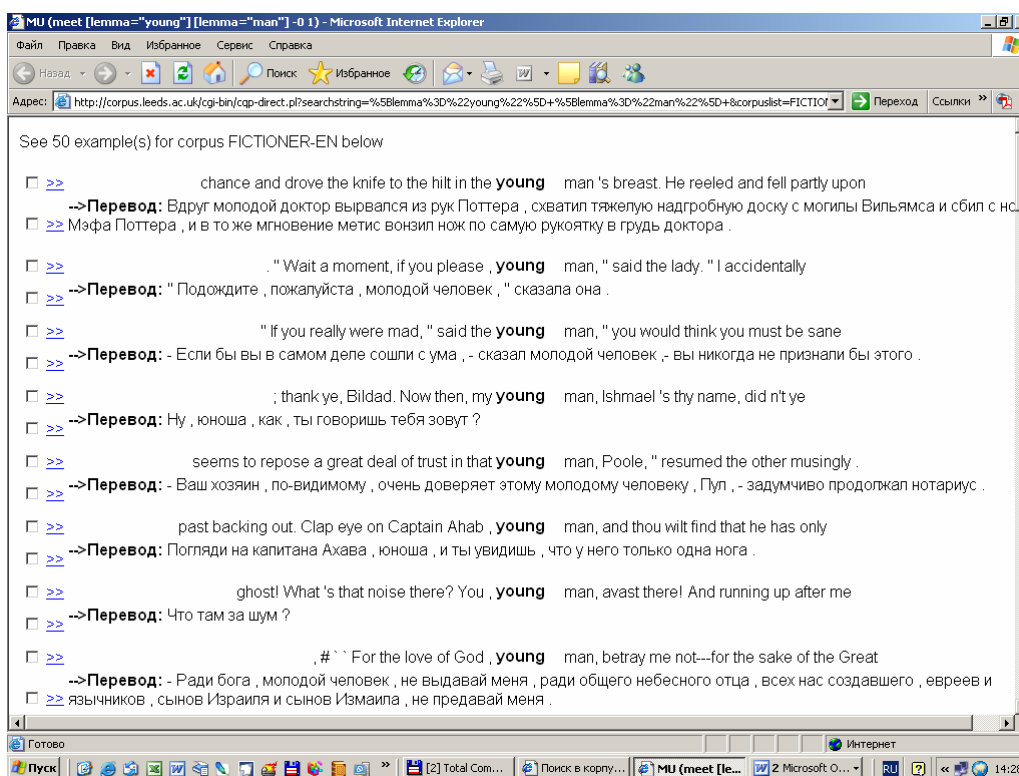


50 шаг

Аналогичным образом, Вы можете узнать, как переводится на русский язык английское словосочетание *young man*. Запрос будет выглядеть следующим образом:



А ответ на него будет таким:



Некоторые необходимые термины¹

- акростих – стихотворение, в котором начальные буквы строк составляют слово или фразу
- акцентный стих – основная форма чисто тонического стихосложения — стих с приблизительно урегулированным количеством ударений в стихе и произвольным количеством безударных слогов между ними
- александрийский стих – двестише шестистопного ямба с парной рифмовкой
- амфибрахий – стихотворный метр, образуемый трехсложными стопами с сильной долей на втором слоге (напр., *Не ветер бушует над бором*)
- анапест – стихотворный метр, образуемый трехсложными стопами с сильной долей на третьем слоге (напр., *Укажи мне такую обитель*)
- аннотация, 1
- аполог – разновидность нравоучительной поэзии в России в XVIII — начале XIX века, короткая басня с резюме в последней строке
- баллада – сюжетное стихотворение, построенное на фантастическом, фольклорном, легендарно-историческом, бытовом материале, часто с мрачным таинственным колоритом
- басня – небольшое произведение нравоучительного характера.
- былина – песня-сказание о богатырях, народных героях и исторических событиях.
- виртуальная клавиатура, 38
- вольная рифмовка – рифмовка, не подчиненная никакой регулярной схеме
- второй винительный падеж – винительный падеж одушевленных существительных во множественном числе с предлогом *в*, совпадающий с именительным множественного (*пойти в солдаты*)
- второй предложный падеж – предложный падеж на -у с предлогами *в*, *на* (*в отпуску*)
- второй родительный падеж, 48
- гекзаметр – стихотворный размер со строкой, состоящей из шести стоп, в которой первые четыре стопы являются преимущественно дактилическими, но могут заменяться хорейми, пятая стопа дактилическая, а последняя – хорейческая
- гетерометрия – употребление в поэтическом произведении нескольких стихотворных размеров, зачастую в соседних строках
- гимн – одическое (см. ода) стихотворение на возвышенную тему
- гипердактилическая клаузула – клаузула с ударением на четвертом от конца слоге в строке
- грамматическая омонимия, 13
- дактилическая клаузула – клаузула с ударением на третьем от конца слоге в строке
- дактиль – стихотворный метр, образуемый трехсложными стопами с сильной долей на первом слоге (напр., *Тучки небесные, вечные странники*)
- Диалектный корпус, 10
- диминутивы, 48, 56
- дифирамб – жанр торжественной лирики, близкий к оде или гимну
- дольник – промежуточная форма между силлабо-тоническим и чисто-тоническим стихосложением: стих, в котором длина слабых (неударных) промежутков между сильными (преимущественно ударными) местами колеблется в пределах 1-2 слогов (напр., *Чёрный ворон в сумраке снежном*)
- женская клаузула – клаузула с ударением на предпоследнем от конца слоге в строке
- идиллия – жанр лироэпической поэзии с описанием мирных бытовых картин и пейзажей, безмятежной жизни земледельцев, пастухов и рыбаков с их простыми наивными характеристиками
- кантата – стихотворение, написанное на определенный торжественный случай
- клаузула, 68
- корпус, 1
- куплеты – строфическое стихотворение с рефреном, предназначенное для пения в водевилях

¹ Цифры по слова обозначают номер страницы в Инструкции, где объяснен данный термин. В словарики частично использованы материалы, подготовленные Н.В. Перцовым.

Некоторые термины

- логаэд – стихотворный размер, образуемый сочетанием неодинаковых стоп (например, дактилей и хореев), последовательность которых правильно повторяется из стиха в стих или из строфы в строфу
- мадригал – небольшое лирическое стихотворение, содержащее комплимент или лестную характеристику лица, к которому оно обращено (обычно к женщине)
- метр – упорядоченное чередование в стихе сильных долей (иктов) и слабых долей
- мини-поэма – поэма небольшого объема
- монорим – стихотворение или часть его с однозвучной рифмовкой
- морфологическая омонимия, 13
- мужская клаузула – клаузула с единственным последним в строке ударным слогом в строке
- надпись – краткая надпись к статуе, портрету, на книге, в альбом и пр.
- национальный корпус, 1
- нежанровая проза, 32
- нечетная рифма – строфа со схемой рифмовки *abac*
- ноэль – сатирическая песенка французского происхождения со строфой из восьми стихов с перекрестной рифмовкой в первом четверостишии и охватной во втором
- ода – торжественное лирическое стихотворение, написанное в приподнятом тоне
- одическая строфа – строфа, написанная четырехстопным ямбом, состоящая из десяти стихов со схемой рифмовки *AbAbCCdEEd*
- октава – строфа из восьми стихов со схемой рифмовки *abababcc*
- онегинская строфа – строфа, написанная четырехстопным ямбом из четырнадцати стихов со схемой рифмовки *AbAbCCddEffEgg*
- Основной корпус Национального корпуса русского языка, 7
- охватная рифма – рифма по схеме *abba*
- Параллельный корпус, 9
- парная рифма – рифма по схеме *aabbvv...*
- пентаметр – стихотворный размер, представляющий собой двукратное соединение двух дактилических стоп и стопы из одного ударного слога (вторая строка элегического дистиха)
- перекрестная рифма – рифма по схеме *abab*
- подкорпус текстов со снятой омонимией, 13
- послание – стихотворение, обращенное к определенному лицу, в котором автор, как бы беседуя с адресатом, высказывает свои суждения по какому-либо важному вопросу
- поэма – большое стихотворное произведение эпического или лирического характера
- представительность, 1
- притча – небольшое стихотворение, сходное с басней, поучительного или сатирического характера
- разметка, 1
- расстояние между словами, 50
- расширение контекста, 40
- романс – небольшое лирическое стихотворение напевного типа, преимущественно на любовную тему
- сбалансированность, 1
- свободный стих – нерифмованные стихи, расчлененные на стихотворные строки, но не обладающие постоянными признаками их соизмеримости
- силлабо-тонический стих – стих, подчиняющийся упорядоченной схеме расположения слогов, в соответствии с которой ударные слоги помещаются (исключительно или преимущественно) на сильных долях метра, а безударные — на слабых
- сильные доли, 63
- сицилиана – строфа из восьми стихов со схемой рифмовки *abababab* или *abbabaab*
- сонет – стихотворение из четырнадцати строк, образующих два четверостишия-катрена (на две рифмы) и два трехстишия-терцета (на две или три рифмы), чаще всего со схемами рифмовки *abab abab cdc dcd* (итальянская) или *abba abba ccd eed* (французская)
- сортировка по левому контексту, 79
- сортировка по правому контексту, 77
- сортировка, 77

Некоторые термины

- стансы – небольшое лирическое стихотворение, написанное четверостишиями, четырехстопным ямбом с преимущественно перекрестными рифмами и с обязательной строфической замкнутостью
- стопа – повторяющееся сочетание сильной и слабой доли в стихотворном метре, служащее единицей длины стиха
- тактовик – промежуточная форма между силлабо-тоническим и тоническим стихосложением: стих, в котором длина слабых (неударных) промежутков между сильными (преимущественно ударными) долями колеблется в пределах 1-3 слогов
- терцины – строфическая форма трехстиший, в которых рифмы идут в порядке *aba bcb cdc ded* и т. д.
- тонический стих – стих, организованный определенным количеством ударных слов в строке, с произвольным количеством слогов между смежными ударениями
- транскрипт, 34
- триолет – строфа из восьми строк, в которой совпадают, во-первых, первое и последнее двустишие и, во-вторых, первая и четвертая строка
- хорей – стихотворный метр с сильными долями на нечетных слогах
- хронотоп, 22
- элегический дистих – двустишие, состоящее из гекзаметра и пентаметра
- элегия – лирическое стихотворение, проникнутое смешанным чувством радости и печали или только грустью, раздумьем, размышлением, с оттенком поэтической интимности
- эпиграмма – короткое сатирическое стихотворение
- эпитафия – надгробная надпись
- ямб – стихотворный метр с сильными долями на четных слогах