Роды ошибок в корпусной лингвистике

Александр Пиперски

XVII Летняя лингвистическая школа

Дубна, 15.07.2015

И. В. фон Гёте

«Я жал посеянное другими, и дело моё есть дело коллективного существа, именуемого Гёте»

(Из разговора с Фредериком Соре, 17.02.1832)


ЯЗЫК ЧЕРЕЗ ПРИЗМУ
ПСИХОЛИНГВИСТИЧЕСК
ЭКСПЕРИМЕНТА (1)

Сюжет 1: «Посмотри в Корпусе»

Что такое «корпус»?

Что такое «корпус»?

- «У этих глаголов есть одна характеристика, которую можно узнать при помощи корпуса» (Н. С. Змановский, 14.07.15)

Что такое «корпус»?

- «У этих глаголов есть одна характеристика, которую можно узнать при помощи корпуса» (Н. С. Змановский, 14.07.15)
- «корпус» vs. «корпус ruTenTen» (Б. Л. Иомдин, 14.07.15)

Что такое «корпус»?

Что такое «корпус»?

- «[Т]еперь у любого языка, наряду с грамматикой и словарем, должен быть корпус» (В. А. Плунгян, postnauka.ru/talks/126)

Что такое «корпус»?

- «[Т]еперь у любого языка, наряду с грамматикой и словарем, должен быть корпус» (В. А. Плунгян, postnauka.ru/talks/126)

Что такое «корпус»?

- «[Т]еперь у любого языка, наряду с грамматикой и словарем, должен быть корпус» (В. А. Плунгян, postnauka.ru/talks/126)

Проблема

корпус = a corpus или the Corpus?

НКРЯ = the Corpus?

НКРЯ = the Corpus?

- НКРЯ (www.ruscorpora.ru) — самый известный корпус русского языка

НКРЯ = the Corpus?

- НКРЯ (www.ruscorpora.ru) — самый известный корпус русского языка
- Подходит для исследований языка художественной литературы и публицистики XIX–XX вв.

НКРЯ = the Corpus?

- НКРЯ (www.ruscorpora.ru) — самый известный корпус русского языка
- Подходит для исследований языка художественной литературы и публицистики XIX–XX вв.
- Не подходит для многого другого

Другие корпуса: ruTenTen

Другие корпуса: ruTenTen

- Система SketchEngine (the.sketchengine.co.uk)

Другие корпуса: ruTenTen

- Система SketchEngine (the.sketchengine.co.uk)
- ruTenTen — скачанные из Интернета текста разных эпох

Другие корпуса: ruTenTen

- Система SketchEngine (the.sketchengine.co.uk)
- ruTenTen — скачанные из Интернета текста разных эпох
- Значительная доля — новые тексты

Другие корпуса: ruTenTen

- Система SketchEngine (the.sketchengine.co.uk)
- ruTenTen — скачанные из Интернета текста разных эпох
- Значительная доля — новые тексты
- Почти нет метаразметки (информации об авторе и тексте)

Другие корпуса: ГИКРЯ

Другие корпуса: ГИКРЯ

- Under construction

Другие корпуса: ГИКРЯ

- Under construction
- ABBYY, РГГУ, МФТИ, Университет Лидса

Другие корпуса: ГИКРЯ

- Under construction
- ABBYY, РГГУ, МФТИ, Университет Лидса
- Скачанные из Интернета тексты с разных платформ (ЖЖ, ВКонтакте, Блоги@Mail.ru, ...)

Другие корпуса: ГИКРЯ

- Under construction
- ABBYY, РГГУ, МФТИ, Университет Лидса
- Скачанные из Интернета тексты с разных платформ (ЖЖ, ВКонтакте, Блоги@Mail.ru, ...)
- Обильная метаразметка

Другие корпуса: DIY

Другие корпуса: DIY

- DIY = Do it yourself

Другие корпуса: DIY

- DIY = Do it yourself
- См. лекцию Т. А. Архангельского

Другие корпуса: DIY

- DIY = Do it yourself
- См. лекцию Т. А. Архангельского
- SketchEngine позволяет загружать в систему свои тексты / коллекции текстов

Другие корпуса: DIY

- DIY = Do it yourself
- См. лекцию Т. А. Архангельского
- SketchEngine позволяет загружать в систему свои тексты / коллекции текстов
- Конкордансеры: AntConc, WordSmith, ...

Промежуточные итоги

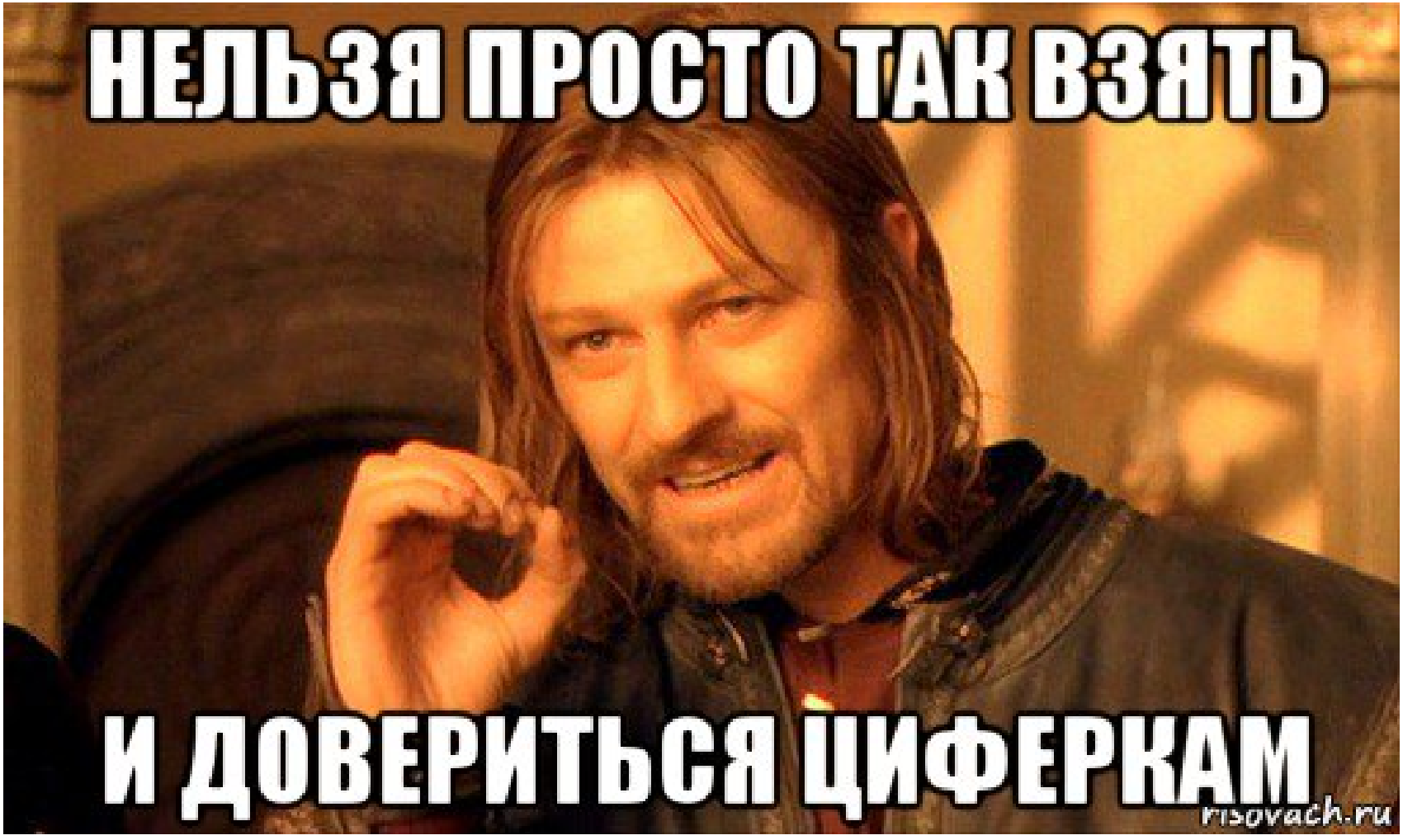
Промежуточные итоги

- Для каждой задачи необходимо подбирать адекватный корпус

Промежуточные итоги

- Для каждой задачи необходимо подбирать адекватный корпус
- По этому вопросу нужно принять экспертное решение в начале исследования

Сюжет 2: «Посмотри на цифры»


НЕЛЬЗЯ ПРОСТО ТАК ВЗЯТЬ
И ДОВЕРИТЬСЯ ЦИФЕРКАМ
risovach.ru

зачёркнуто

зачёркнуто

- Встречаются ли в устной речи аналоги письменного зачёркивания?

зачёркнуто

- Встречаются ли в устной речи аналоги письменного зачёркивания?
- Ищем слово *зачёркнуто* в устном подкорпусе НКРЯ

зачёркнуто

- Встречаются ли в устной речи аналоги письменного зачёркивания?
- Ищем слово *зачёркнуто* в устном подкорпусе НКРЯ
- 3 документа, 15 вхождений

зачёркнуто

- Встречаются ли в устной речи аналоги письменного зачёркивания?
- Ищем слово *зачёркнуто* в устном подкорпусе НКРЯ
- 3 документа, 15 вхождений
- Некоторые из вас уже знают, в чём проблема...

зачёркнуто

Найдено 3 документа, 15 вхождений.

Поискать в других корпусах: основном, акцентологическом, газетном, диалектном, мультимедийном, обучающем, параллельном, поэтическом, синтаксическом.

Страницы: **1**

1. **А. Сомин, А. Пиперски. Доклад на конференции «Диалог 2013» // Из коллекции НКРЯ, 2013** [омонимия не снята] **Все примеры (12)**

   [Сомин А., муж, 24, 1989] Но есть такая базовая классификация/ против которой не попрёшь/ это если замена/ то есть "по Ленинградке" **зачёркнуто** "пробки" / то

Промежуточные итоги

Промежуточные итоги

- Не заметить собственный текст было бы сложно

Промежуточные итоги

- Не заметить собственный текст было бы сложно
- Но могут быть и другие неожиданности в выдаче, которые не позволяют просто скопировать статистику

«Галла» (1921)

«Галла» (1921)

Там внизу, в отдаленной равнине, костры,
Точно красные звезды, повсюду.

«Галла» (1921)

Там внизу, в отдаленной равнине, костры,
Точно красные звезды, повсюду.

И помчались один за другими они,
Точно тучи в сияющей сини

«Галла» (1921)

Там внизу, в отдаленной равнине, костры,
Точно красные звезды, повсюду.

И помчались один за другими они,
Точно тучи в сияющей сини

Жирный негр восседал на персидских коврах
В полутемной неубранной зале,
Точно идол, в браслетах, серьгах и перстнях,
Лишь глаза его дивно сверкали.

Сравнительные союзы

Сравнительные союзы

- Кажется, что Николай Гумилёв предпочитает *точно* другим двусложным сравнительным союзам (*будто* и *словно*)

Сравнительные союзы

- Кажется, что Николай Гумилёв предпочитает *точно* другим двусложным сравнительным союзам (*будто* и *словно*)
- Как исследовать его предпочтения на фоне других поэтов?

Сравнительные союзы

- Кажется, что Николай Гумилёв предпочитает *точно* другим двусложным сравнительным союзам (*будто* и *словно*)
- Как исследовать его предпочтения на фоне других поэтов?
- $\longrightarrow$ Поэтический подкорпус НКРЯ

будто, словно, точно: Гумилёв

будто, словно, точно: Гумилёв

- *будто* — 31 / 67 705 = 458 ipm

будто, словно, точно: Гумилёв

- *будто* — 31 / 67 705 = 458 ipm
- *словно* — 89 / 67 705 = 1315 ipm

будто, словно, точно: Гумилёв

- *будто* — 31 / 67 705 = 458 ipm
- *словно* — 89 / 67 705 = 1315 ipm
- *точно* — 48 / 67 705 = 709 ipm

будто, словно, точно: Гумилёв

- *будто* — 31 / 67 705 = 458 ipm
- *словно* — 89 / 67 705 = 1315 ipm
- *точно* — 48 / 67 705 = 709 ipm
- В чём проблемы?

будто, словно, точно: проблемы

будто, словно, точно: проблемы

- *будто* часто встречается в сочетании *как будто*, где оно не взаимозаменимо с другими сравнительными союзами

будто, словно, точно: проблемы

- *будто* часто встречается в сочетании *как будто*, где оно не взаимозаменимо с другими сравнительными союзами
- *точно* может быть не только сравнительным союзом

точно: Пушкин

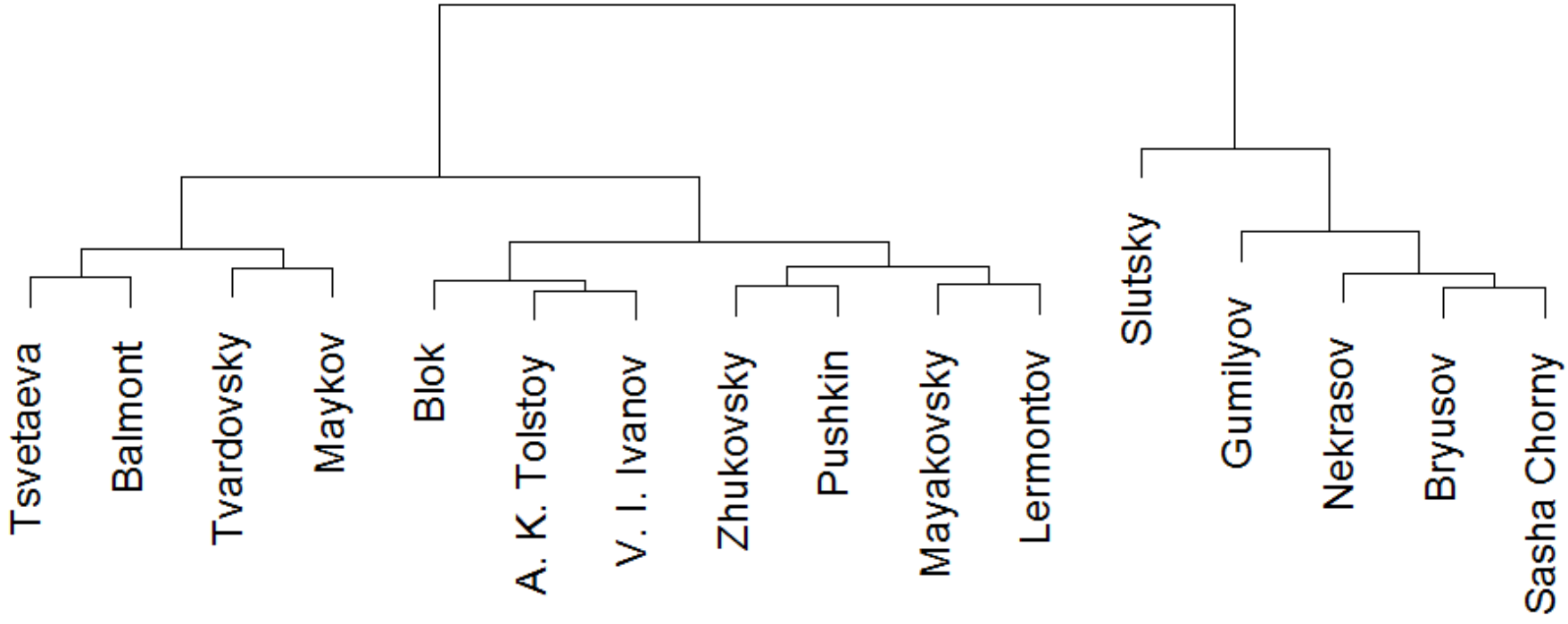
точно: Пушкин

- 32 вхождения на 198 163 слова

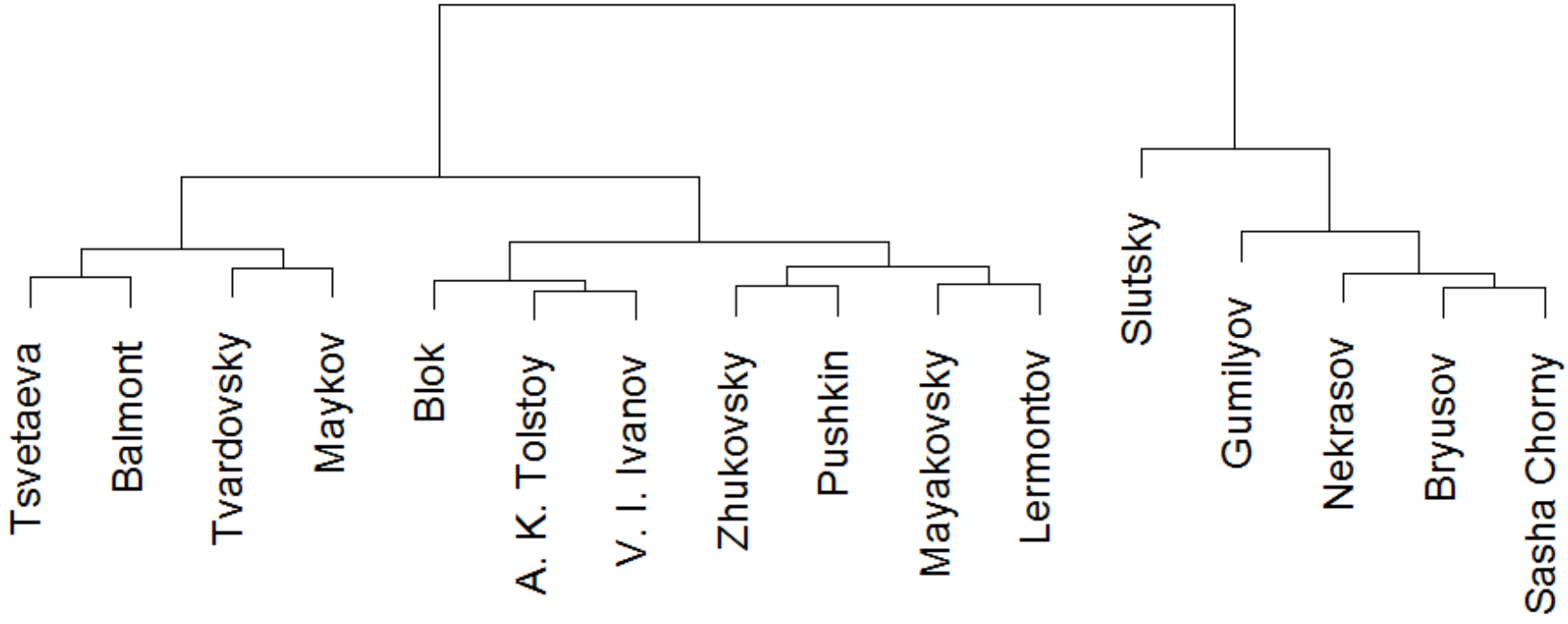
точно: Пушкин

- 32 вхождения на 198 163 слова
- Из них: ~ 2 сравнительных союза

Дендрограмма



Дендрограмма



Подробнее: в моем докладе на конференции Corpora–2015

Промежуточные итоги

Промежуточные итоги

- Нельзя смотреть только на статистику, не просматривая контексты

Промежуточные итоги

- Нельзя смотреть только на статистику, не просматривая контексты
- Возможны ошибки:

Промежуточные итоги

- Нельзя смотреть только на статистику, не просматривая контексты
- Возможны ошибки:
  - в запросе

Промежуточные итоги

- Нельзя смотреть только на статистику, не просматривая контексты
- Возможны ошибки:
  - в запросе
  - в разметке

Промежуточные итоги

- Нельзя смотреть только на статистику, не просматривая контексты
- Возможны ошибки:
  - в запросе
  - в разметке
- Они не всегда исправимы

ТОЧНО

ТОЧНО

- В НКРЯ есть грамматическая разметка

точно

- В НКРЯ есть грамматическая разметка
- Почему бы не указать при слове точно грамматический признак "союз"?

Экскурс: разметка в корпусах

Метаразметка

Метаразметка

- Обычно содержит доступную и правильную информацию о текстах

Метаразметка

- Обычно содержит доступную и правильную информацию о текстах
- Тоже бывают ошибки

Найдено 3 документа, 15 вхождений.

Поискать в других корпусах: основном, акцентологическом, газетном, диалектном, мультимедийном, обучающем, параллельном, поэтическом, синтаксическом.

Страницы: **1**

1. **А. Сомин, А. Пиперски. Доклад на конференции «Диалог 2013» // Из коллекции НКРЯ, 2013** [омонимия не снята] **Все примеры (12)**

   [Сомин А., муж, 24, 1989] Но есть такая базовая классификация/ против которой не попрёшь/ это если замена/ то есть "по Ленинградке" **зачёркнуто** "пробки" / то

Найдено 3 документа, 15 вхождений.

Поискать в других корпусах: основном, акцентологическом, газетном, диалектном, мультимедийном, обучающем, параллельном, поэтическом, синтаксическом.

Страницы: **1**

1. **А. Сомин, А. Пиперски. Доклад на конференции «Диалог 2013» // Из коллекции НКРЯ, 2013** [омонимия не снята] **Все примеры (12)**

   [Сомин А., муж, 24, 1989] Но есть такая базовая классификация/ против которой не попрёшь/ это если замена/ то есть "по Ленинградке" **зачёркнуто** "пробки" / то

Мелочь, но:
Антон Сомин — 1988 года рождения

Лингвистическая разметка

Лингвистическая разметка

- Морфологическая и синтаксическая разметка обычно содержит гораздо больше ошибок, чем метаразметка

Лингвистическая разметка

- Морфологическая и синтаксическая разметка обычно содержит гораздо больше ошибок, чем метаразметка
- Проблема: омонимия

Морфология в НКРЯ

Морфология в НКРЯ

- Генерируются все возможные разборы без учёта синтаксиса (программа Mystem)

Морфология в НКРЯ

- Генерируются все возможные разборы без учёта синтаксиса (программа Mystem)
- В небольшой части корпуса (6 млн словоформ из 230 млн основного подкорпуса) омонимия снята вручную

Предложение из НКРЯ

Предложение из НКРЯ

- *В кодексе мой грех стоит три года общего режима* [Андрей Рубанов. Сажайте, и вырастет (2005)]

Предложение из НКРЯ

- *В кодексе мой грех стоит три года общего режима* [Андрей Рубанов. Сажайте, и вырастет (2005)]
- Сколько слов здесь имеет неоднозначный разбор?

Предложение из НКРЯ

Предложение из НКРЯ

- *мой*, *стоит*, *три*: непонятна начальная форма

Предложение из НКРЯ

- *мой*, *стоит*, *три*: непонятна начальная форма
- *грех*, *года*, *общего*: непонятны грамматические признаки

Предложение из НКРЯ

- *мой*, *стоит*, *три*: непонятна начальная форма
- *грех*, *года*, *общего*: непонятны грамматические признаки
- *В*, *кодексе*, *режима*: однозначно

Вывод Mystem (1)

Вывод Mystem (1)

В{в=PR=|в=S,сокр=пр,мн|=S,сокр=пр,ед|
=S,сокр=вин,мн|=S,сокр=вин,ед|=S,сокр=
дат,мн|=S,сокр=дат,ед|=S,сокр=род,мн|=S,
сокр=род,ед|=S,сокр=твор,мн|=S,сокр=твор,
ед|=S,сокр=им,мн|=S,сокр=им,ед}
кодексе{кодекс=S,муж,неод=пр,ед}
мой{мой=APRO=вин,ед,муж,неод|=APRO
=им,ед,муж|мыть=V,несов,пе=ед,пов,2-л}

Вывод Mystem (2)

Вывод Mystem (2)

грех{грех=S,муж,неод=вин,ед|=S,муж,неод=им,ед|грех=ADV,прдк=}
стоит{стоять=V,несов,нп=непрош,ед,изъяв,3-л|стоить=V,несов=непрош,ед,изъяв,3-л}
три{три=NUM=им|=NUM=вин,неод|тереть=V,несов,пе=ед,пов,2-л}
года{год=S,муж,неод=вин,мн|=S,муж,неод=род,ед|=S,муж,неод=им,мн}

Вывод Mystem (3)

Вывод Mystem (3)

общего{общий=A=вин,ед,полн,муж,од|=A=род,ед,полн,муж|=A=род,ед,полн,сред}
режима{режим=S,муж,неод=род,ед}

Вывод Mystem (3)

общего{общий=A=вин,ед,полн,муж,од|=A=род,ед,полн,муж|=A=род,ед,полн,сред}
режима{режим=S,муж,неод=род,ед}

NB: чуть больше неоднозначностей в определении начальной формы, чем мы ожидали

Ручное снятие омонимии

Ручное снятие омонимии

- Кто снимает омонимию?

Ручное снятие омонимии

- Кто снимает омонимию?
- Обычно — слабо мотивированные и неквалифицированные разметчики

Ручное снятие омонимии

- Кто снимает омонимию?
- Обычно — слабо мотивированные и неквалифицированные разметчики
- Краудсорсинг — попытка привлечь заинтересованных людей

OpenCorpora

OpenCorpora

- OpenCorpora (Открытый корпус) — Санкт-Петербург

OpenCorpora

- OpenCorpora (Открытый корпус) — Санкт-Петербург
- Более 2 млн ответов

OpenCorpora

- OpenCorpora (Открытый корпус) — Санкт-Петербург
- Более 2 млн ответов
- Образец — задание по разметке им./вин. мн.

Ручное снятие омонимии

Ручное снятие омонимии

- Ручное снятие омонимии не достигает 100%-ной точности

Ручное снятие омонимии

- Ручное снятие омонимии не достигает 100%-ной точности
- Исследователи работают с огромными корпусами, которые невозможно разметить вручную (ruTenTen: 14 млрд словоупотреблений)

Ручное снятие омонимии

- Ручное снятие омонимии не достигает 100%-ной точности
- Исследователи работают с огромными корпусами, которые невозможно разметить вручную (ruTenTen: 14 млрд словоупотреблений)
- Нужна автоматическая разметка!

Автоматическое снятие омонимии

Автоматическое снятие омонимии

- *о случайной отставке*

Автоматическое снятие омонимии

- *о случайной отставке*
- *случайной* — 4 возможных разбора

Автоматическое снятие омонимии

- *о случайной отставке*
- *случайной* — 4 возможных разбора
- Как выбрать верный разбор?

Автоматическое снятие омонимии

- *о случайной отставке*
- *случайной* — 4 возможных разбора
- Как выбрать верный разбор?
- **По контексту**

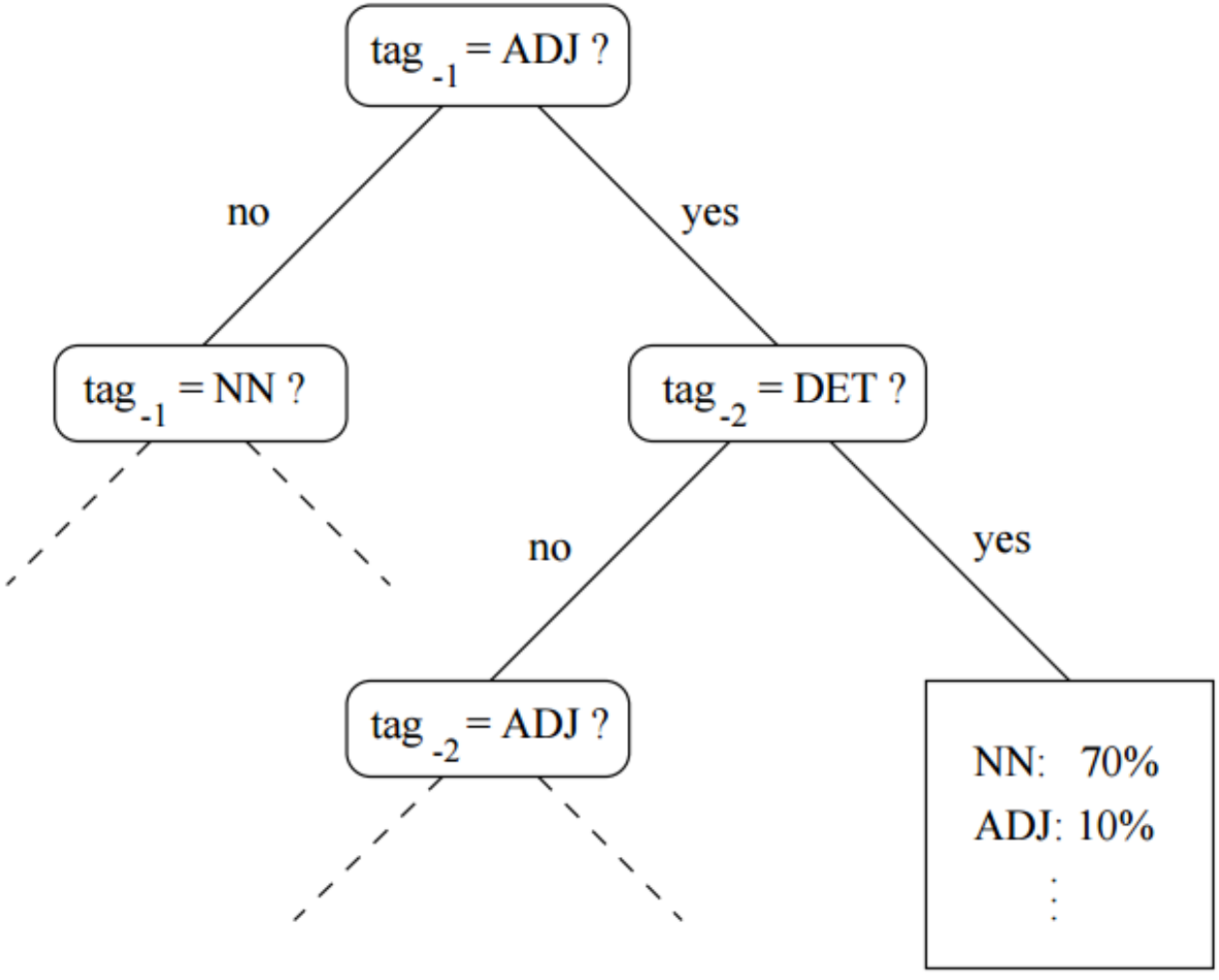
Автоматическое снятие омонимии

- *о случайной отставке*
- *случайной* — 4 возможных разбора
- Как выбрать верный разбор?
- **По контексту**

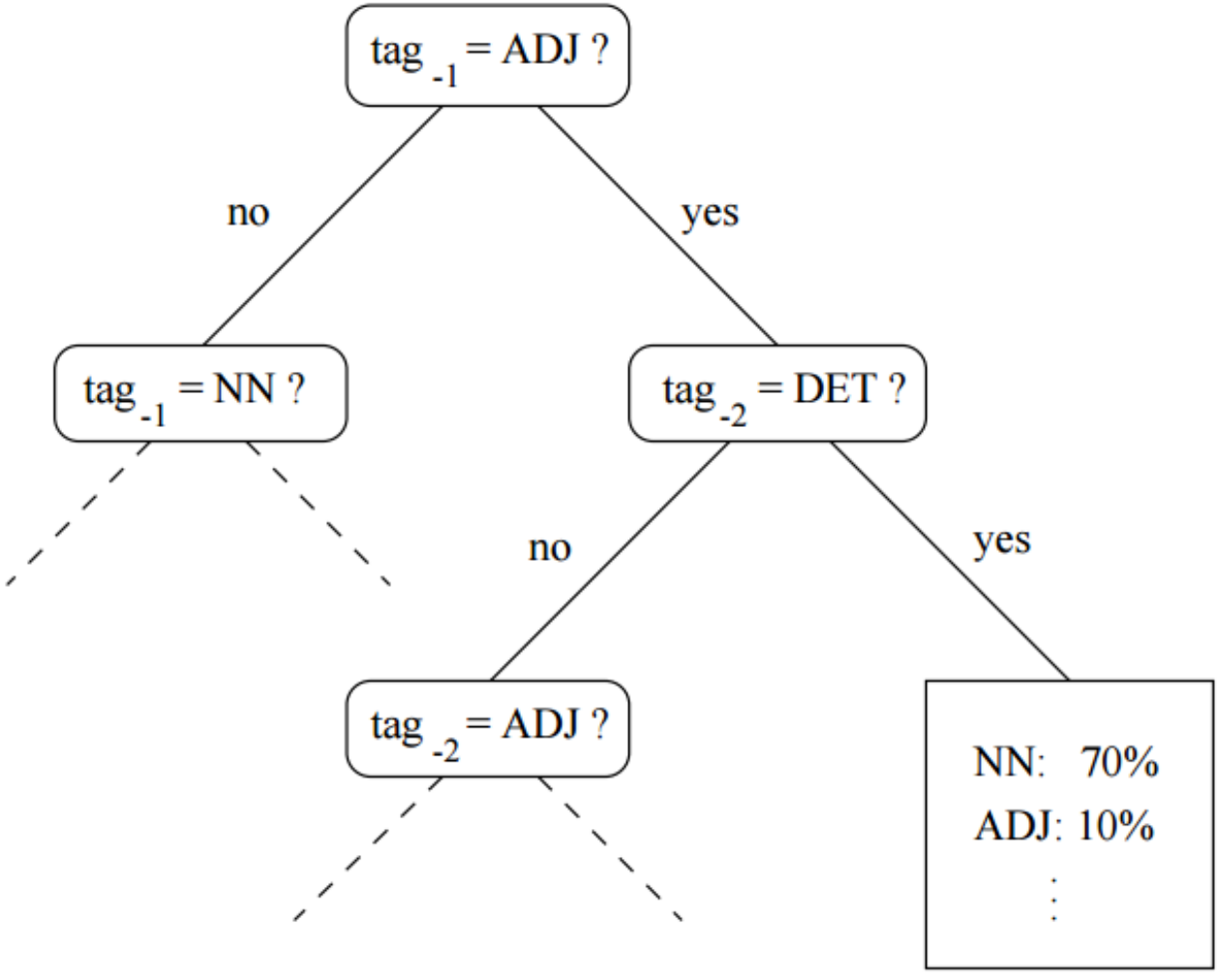
Автоматическое снятие омонимии

- *о случайной отставке*
- *случайной* — 4 возможных разбора
- Как выбрать верный разбор?
- **По контексту** = по атрибутам соседних слов

TreeTagger (Schmidt 1994)



TreeTagger (Schmidt 1994)



Автоматическое снятие омонимии: проблемы

Автоматическое снятие омонимии: проблемы

- Непредсказуемость результатов
  $\Longrightarrow$ нельзя подкрутить алгоритм

Сюжет 3: ошибки в разметке

Вопрос с ЧГК

- ЭТО есть у карасей и щук, у скатов и акул и еще у множества других рыб. ЭТОГО нет у плотвы и трески, хотя то, что вы слышали, можно принять за ЭТО. Назовите ЭТО.

Вопрос с ЧГК

- ЭТО есть у карасей и щук, у скатов и акул и еще у множества других рыб. ЭТОГО нет у плотвы и трески, хотя то, что вы слышали, можно принять за ЭТО. Назовите ЭТО.
- **Ответ:** множественное число

Вопрос с ЧГК

- ЭТО есть у карасей и щук, у скатов и акул и еще у множества других рыб. ЭТОГО нет у плотвы и трески, хотя то, что вы слышали, можно принять за ЭТО. Назовите ЭТО.
- **Ответ:** множественное число
- Проверим по ruTenTen?

Плотвы и трески

Запросы вида
[lemma="плотва" & tag="Nc.s.*"]

	ед. ч.	мн. ч.	
карась	35 971 (74%)	12 653 (26%)	48 624
щука	109 337 (91%)	11 150 (9%)	120 487
скат	33 594 (57%)	24 852 (43%)	120 487
акула	59 166 (44%)	74 324 (56%)	120 487
плотва	23 320 (96%)	1069 (4%)	24 389
треска	22 401 (96%)	981 (4%)	23 382

Плотвы и трески

Плотвы и трески

- *плотвы* и *трески* не так уж далеки от щук

Плотвы и трески

- *плотвы* и *трески* не так уж далеки от щук
- Но все *плотвы* и *трески* — неправильные разборы!

Плотвы и трески

- *плотвы* и *трески* не так уж далеки от щук
- Но все *плотвы* и *трески* — неправильные разборы!
  - *приготовление тушки трески*

Плотвы и трески

- *плотвы* и *трески* не так уж далеки от щук
- Но все *плотвы* и *трески* — неправильные разборы!
  - *приготовление тушки трески*
  - *сквозь помехи и ночные трески*

мультипликационн.*

мультипликационн.*

ruTenTen

*мультипликационн.**

ruTenTen

- [word="мультипликационн.*" & word!="мультипликационност.*"]: 25996

мультипликационн.*

ruTenTen

- [word="мультипликационн.*" & word!="мультипликационност.*"]: 25996
- [word="мультипликационн.*" & lemma!="мультипликационный" & word!="мультипликационност.*"]: 1897

мультипликационн.*

*мультипликационн.**

- Около 7% форм слова *мультипликационный* не приводятся к правильной начальной форме

*мультипликационн.**

- Около 7% форм слова *мультипликационный* не приводятся к правильной начальной форме
- Почему, необъяснимо

*мультипликационн.**

- Около 7% форм слова *мультипликационный* не приводятся к правильной начальной форме
- Почему, необъяснимо
- Будет ли этот процент таким же у других слов или нет?..

белки, жиры, углеводы

белки, жиры, углеводы

- Что чаще упоминается в языке Интернета — *белки*, *жиры* или *углеводы* и насколько?

белки, жиры, углеводы

- Что чаще упоминается в языке Интернета — *белки*, *жиры* или *углеводы* и насколько?
- Ищем в корпусе ruTenTen

белки, жиры, углеводы

белки, жиры, углеводы

- [lemma="белок" & tag="Ncmp.*"]: 299 420

белки, жиры, углеводы

- [lemma="белок" & tag="Ncmp.*"]: 299 420
- [lemma="жир" & tag="Ncmp.*"]: 258 635

белки, жиры, углеводы

- [lemma="белок" & tag="Ncmp.*"]: 299 420
- [lemma="жир" & tag="Ncmp.*"]: 258 635
- [lemma="углевод" & tag="Ncmp.*"]: 209 712

белки, жиры, углеводы

белки, жиры, углеводы

- [word="[бБ]елк(и|ов|ам|ами|ах)"]: 344 205

белки, жиры, углеводы

- [word="[бБ]елк(и|ов|ам|ами|ах)"]: 344 205
- [word="[жЖ]ир(ы|ов|ам|ами|ах)"]: 261 953

белки, жиры, углеводы

- [word="[бБ]елк(и|ов|ам|ами|ах)"]: 344 205
- [word="[жЖ]ир(ы|ов|ам|ами|ах)"]: 261 953
- [word="[уУ]глевод(ы|ов|ам|ами|ах)"]: 210 678

белки, жиры, углеводы

белки, жиры, углеводы

- Для *жиров* и *углеводов* результаты почти не изменились

белки, жиры, углеводы

- Для *жиров* и *углеводов* результаты почти не изменились
- *белков* оказалось примерно на 20% больше

белки, жиры, углеводы

- Для *жиров* и *углеводов* результаты почти не изменились
- *белков* оказалось примерно на 20% больше
- Добавились слова, разобранные неправильно — как формы слова *белка* или *белок* женского рода (sic!)

Промежуточные итоги

Промежуточные итоги

- Надо не просто просматривать выдачу, но и понимать, чего в ней могло не оказаться

Аналогия: анализы

Аналогия: анализы

- Для здорового пациента анализ в 1% случаев ошибочно говорит, что он болен

Аналогия: анализы

- Для здорового пациента анализ в 1% случаев ошибочно говорит, что он болен
- Для больного пациента анализ в 1% случаев ошибочно говорит, что он здоров

Аналогия: анализы

Аналогия: анализы

Задача 1

Результат анализа — болен. С какой вероятностью пациент на самом деле болен?

Аналогия: анализы

Задача 1

Результат анализа — болен. С какой вероятностью пациент на самом деле болен?

Задача 2

Результат анализа — здоров. С какой вероятностью пациент на самом деле здоров?

Аналогия: анализы

- Надо знать долю больных и здоровых!

Аналогия: анализы

- Надо знать долю больных и здоровых!
- NB: при корпусном исследовании у нас нет такого априорного знания

Аналогия: анализы

Аналогия: анализы

- Болен 1% населения

Аналогия: анализы

- Болен 1% населения

Аналогия: анализы

- Болен 1% населения

	болен	здоров	
болен	99	1	100
здоров	99	9801	990
	198	9802	10000

Аналогия: анализы

- Болен 1% населения

	болен	здоров	
болен	99	1	100
здоров	99	9801	990
	198	9802	10000

- Из 198 пациентов с положительным анализом болен только каждый второй ($p = 50\%$)!

Аналогия: анализы

- Болен 1% населения

	болен	здоров	
болен	99	1	100
здоров	99	9801	990
	198	9802	10000

- Из 9802 пациентов с отрицательным анализом здоровы почти все ($p = 99{,}99\%$)!

Роды ошибок

Роды ошибок

- Ложноположительные результаты (ошибки 1-го рода): пациент здоров, но анализ говорит, что болен

Роды ошибок

- Ложноположительные результаты (ошибки 1-го рода): пациент здоров, но анализ говорит, что болен
- Ложноотрицательные результаты (ошибки 2-го рода): пациент болен, но анализ говорит, что здоров

Роды ошибок

Роды ошибок

Вопрос для размышления

Какого рода ошибки хуже для корпусного лингвиста?

Роды ошибок

Вопрос для размышления

Какого рода ошибки хуже для корпусного лингвиста?

- Ложноположительные результаты легко увидеть глазами

Роды ошибок

Вопрос для размышления

Какого рода ошибки хуже для корпусного лингвиста?

- Ложноположительные результаты легко увидеть глазами
- Ложноотрицательные результаты хуже тем, что незаметны!

Выводы

Выводы

Выводы

- Корпуса — очень мощный инструмент лингвистического исследования

Выводы

- Корпуса — очень мощный инструмент лингвистического исследования
- Пользуясь этим инструментом, надо хорошо осознавать его ограничения и недостатки, чтобы избегать ошибок


CAUTION
CAUTION
CAUTION
CAUTION
CAUTION
CAUTION