



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

МАЙНОР «ЛИНГВИСТИКА»

2015 – 2017

Part 2

Майнор совмещает теоретическую

Языки мира

Русский язык и история языка

и практическую составляющие науки о языке

Корпусная лингвистика

Цифровые методы в гуманитарных науках



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Корпусная лингвистика

Преподаватели



Ольга Николаевна
Ляшевская

Светлана Юрьевна
Толдова



Как изучать язык?

Язык — это набор слов и система правил.

Нет такого места, где «живёт» язык, нельзя пойти туда и посмотреть на язык непосредственно.

Поэтому мы не можем изучать язык напрямую, а можем только исследовать тексты, которые на нём созданы.

Что такое корпус?

Исследование тем точнее, чем больше текстов оно охватывает.

А если текстов слишком много, чтобы один человек их внимательно прочёл?

Тогда нужен корпус — большая электронная поисковая система, в которой лежат специальным образом подготовленные тексты.

Этих текстов может быть больше, чем один человек способен прочесть и запомнить.



Что такое корпус?

Вот так, например, выглядит корпус русского языка

search.ruscorpora.ru/search.xml?env=alpha&mycorp=&mysent=&mysize=&mysentsize=&mydocsize=&spd=&text=lexgramm&mode=main&sort=gr_tagging&lang=ru&nodia=1&parent1=0&level1=0&lex1=%E4%EE%E1%F0%EE%EF%EE%F0

Результаты поиска в основном корпусе

[перейти на страницу поиска](#) [выбрать подкорпус](#) [версия с ударениями](#) [настройки](#) [формат KWIC](#) [English](#)

Объем всего корпуса: 85 996 документов, 19 362 746 предложений, 229 968 798 слов.

добропорядочный

Найдено 370 документов, 443 вхождения.

[Распределение по годам](#) [Статистика](#)

Поискать в других корпусах: [акцентологическом](#), [газетном](#), [диалектном](#), [мультимедийном](#), [обучающем](#), [параллельном](#), [поэтическом](#), [синтаксическом](#), [устном](#).

Страницы: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [следующая страница](#)

1. [Георгий Летов. Героинтовая шкатулка \(2003\) // «Еженедельный журнал», 2003.04.08](#) [\[омонимия снята\]](#) [Все примеры \(1\)](#)

Город без наркотиков" умело канализировал беспокойство **добропорядочных** граждан за собственных детей-подростков. [Георгий Летов. Героинтовая шкатулка (2003) // «Еженедельный журнал», 2003.04.08] [\[омонимия снята\]](#) [←...→](#)

2. [Александр Кривцов. БАБская защита демократии \(2002\) // «Финансовая Россия», 2002.09.19](#) [\[омонимия снята\]](#) [Все примеры \(1\)](#)

Вот, например, перманентный автор газетных публикаций — **добропорядочный** эсквайр и почётный лондонец Борис Абрамович Березовский, он же миллиардер и бывший депутат Государственной Думы. [Александр Кривцов. БАБская защита демократии (2002) // «Финансовая Россия», 2002.09.19] [\[омонимия снята\]](#) [←...→](#)

3. [И. А. Архипова. Музыка жизни \(1996\)](#) [\[омонимия снята\]](#) [Все примеры \(2\)](#)

Она сама иронизировала на этот счёт: "Такое долголетие не подобает **добропорядочной** жене". [И. А. Архипова. Музыка жизни (1996)] [\[омонимия снята\]](#) [←...→](#)
Когда на доме, где в последние годы (с 1964 по 1969) жил академик В. В. Виноградов (это в Калашном переулке), была установлена мемориальная доска, его "**добропорядочная** жена" заметила с горькой иронией: "Не похож. [И. А. Архипова. Музыка жизни (1996)] [\[омонимия снята\]](#) [←...→](#)

4. [Юрий Давыдов. Синие тюльпаны \(1988-1989\)](#) [\[омонимия снята\]](#) [Все примеры \(1\)](#)

Годы спустя, после лагеря Башуцкий с неделю работал в Тартуском архиве, что на улице Лливий; ночевал неподалёку, в уютном, **добропорядочном**, тихо пьющем отельчике "Тооме"; дни падали сумрачные, зимние, с каплей; домой возвращался автобусом, ехал долго, сугробы кружили, поля, перелески... [Юрий Давыдов. Синие тюльпаны (1988-1989)] [\[омонимия снята\]](#) [←...→](#)

5. [Ю. О. Домбровский. Факультет ненужных вещей, часть 4 \(1978\)](#) [\[омонимия снята\]](#) [Все примеры \(1\)](#)

Все, с кем он тут встречался — от наркома до такого же маленького администратора, как он, — все друг с другом были вежливы, **добропорядочны**, честны. [Ю. О. Домбровский. Факультет ненужных вещей, часть 4 (1978)] [\[омонимия снята\]](#) [←...→](#)

Что такое корпус?

Корпус отражает, как устроен язык.

Если говорят *«нет в языке такого слова»*, корпус даст это проверить: действительно ли нет?

Все ли знают слово «онучи»?

Или «сегун»?

А корпус знает!

Корпус может быть специальным. Скажем, сделанным только для того, чтобы выяснить, как обычно люди делают ударения в словах.

Мы все учимся тому, как говорить правильно.

Это — «норма».

Но лингвист... ещё и должен выяснять, как принято говорить на самом деле.

Человек «звонит» или «звони'т»?

Как правильно? «Звони'т».

Как на самом деле? И так и так.

Какой вариант более распространён? Ответ в корпусе!

Может ли корпус помочь не лингвисту, а историку, культурологу, социологу?

Да! Чтобы выяснить,

- как обычно говорят или не говорят по-русски, чтобы понять, поймут ли тебя;
- когда появилось какое-то слово или понятие (корпус знает про все свои тексты, когда они возникли);
- с чем некоторое слово обычно сочетается в языке, а с чем нет.



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Цифровые методы в гуманитарных науках



Анастасия Александровна
Бонч-Осмоловская

Борис Валерьевич
Орехов



«Идеалом» науки является физика.

- В ней многое (скорость, масса, энергия...) может быть переведено в числа, посчитано и передано математической формулой.

Гуманитарные науки устроены не так.

- Считать что-то можно, но не обязательно эти подсчёты будут полезны гуманитариям.

Компьютер лучше всего умеет именно считать.

Неужели это значит, что он не нужен гуманитариям?

Чем компьютер может помочь гуманитарию?

Систематизировать!

Только нужно при этом найти и выделить главное.

Визуализировать!

Построить графики, диаграммы, нарисовать географические карты и графы.

Всё это возможно только в сотрудничестве между гуманитарием, знающим специфику своего предметного поля, и специалистом по цифровым методам.

Что вы изучаете?

- **Фольклор.** Компьютер поможет вам понять, чем варианты текстов отличаются друг от друга, построить карту распространения сюжетов и мотивов.
- **История.** Компьютер поможет в систематизации данных и моделировании процессов.
- **Культура.** Компьютер даст возможность анализировать качественные данные любой структуры.

Что такое «хорошо» и что такое «плохо»?

Какой инструмент для гуманитария лучше?

С точки зрения современной ситуации в IT лучше то, что находится в интернете и доступно через браузер или мобильное приложение.

Лучше такие инструменты, где нет заранее заданных ответов, и они получаются только когда пользователь изобретает вопрос.

И так далее, и так далее...





НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Спасибо за внимание!

101000, Россия, Москва, Мясницкая ул., д. 20

Тел.: (495) 621-7983, факс: (495) 628-7931

www.hse.ru