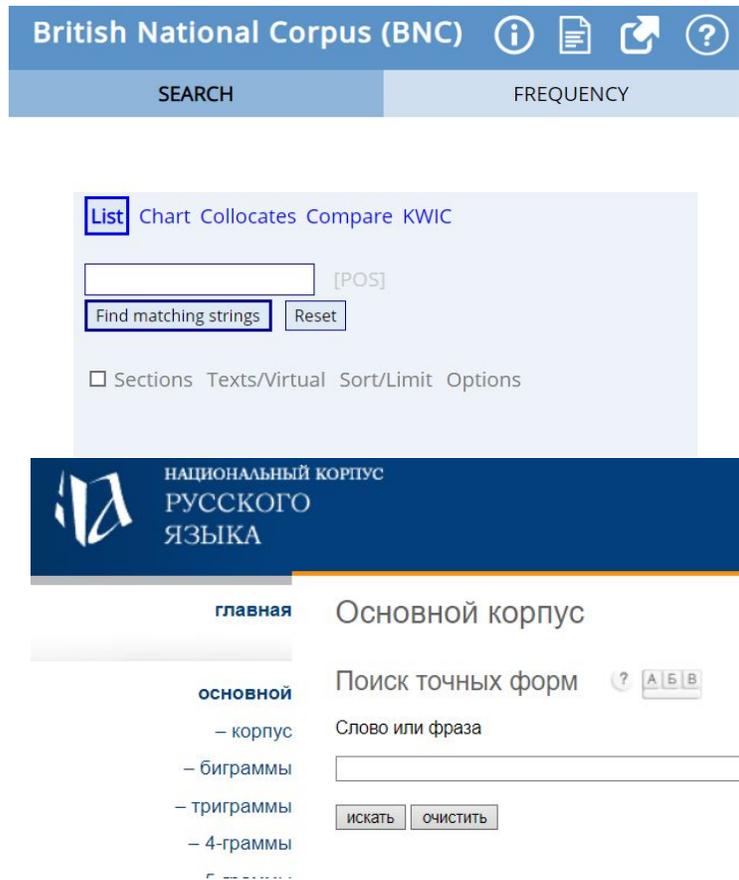


# Корпус текстов

В лингвистике

**Корпус** — подобранная и обработанная по определённым правилам **совокупность текстов**, используемых в качестве базы для **исследования языка**.



Они используются для статистического анализа и проверки статистических гипотез, подтверждения лингвистических правил в данном языке. Корпус текстов является предметом исследования корпусной лингвистики.

## Лексико-грамматический поиск ?

Слово ?

Доп. признаки ? [выбрать](#)

Расстояние: от  до  ?

Слово ?

Доп. признаки ? [выбрать](#)

Семант. признаки ? [выбрать](#)

1-е знач.  др. знач.  фильтр 1  фильтр 2 ?

Семант. признаки ? [выбрать](#)

1-е знач.  др. знач.  фильтр 1  фильтр 2 ?

Среди множества определений корпуса можно выделить его главные свойства:

- ▶ электронный — в современном понимании корпус должен быть в электронном виде;
- ▶ репрезентативный — должен хорошо «представлять» объект, который моделирует;
- ▶ размеченный — главное отличие корпуса от коллекции текстов;
- ▶ прагматически ориентированный — должен быть создан под определённую задачу.

## Многоязычные и двуязычные корпуса делятся на два типа:

параллельные —  
множество текстов и их  
переводов на один или  
несколько языков

сопоставимые  
(псевдопараллельные) —  
оригинальные тексты на  
двух или нескольких  
языках

# Параллельные корпуса текстов

являются своего рода наглядными пособиями для перевода

дерево

Найдено 162 документа, 575 вхождений.

Поискать в других корпусах: [основном](#), [акцентологическом](#), [газетном](#), [диалектном](#), [мул](#)

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

## 1. [Dan Brown. The Da Vinci Code \(2003\)](#) | Дэн Браун. Код Да Винчи (Н. Рейн, 2004)

en		Rosewood, Sophie realized. [Dan Brown. The Da Vinci Code (2003)   Дэн Браун. Код Да
ru		Розовое <b>дерево</b> , поняла Софи. [Дэн Браун. Код Да Винчи (Н. Рейн, 2004)] [ОМОНИ

en		There was a crash as the front bumper clipped a tree. [Dan Brown. The Da Vinci Co
ru		Раздался треск — это передний бампер зацепился за небольшое <b>дерево</b> .

В заключение можно отметить, что основной задачей создания корпусов текста в целом является обеспечение академических исследований и нужд преподавания.

Внедрение корпусных методов радикально изменило общий научный ландшафт в лингвистике. Теперь ограничений на объем анализируемого материала и скорость поиска информации в нем по существу нет, а это означает, что в распоряжении исследователя оказываются колоссальные массивы текстов самого разного типа.