**Правительство Российской Федерации**

**Федеральное государственное автономное образовательное учреждение высшего профессионального образования**

**"Национальный исследовательский университет   
"Высшая школа экономики"**

**(Нижний Новгород)**

**Программа дисциплины**

**«Корпусные методы в политической лингвистике»**

для направления 45.04.03 «Фундаментальная и прикладная лингвистика»

подготовки магистра

магистерская программа «Прикладная лингвистика»

Автор программы:

Хоменко А.Ю., преподаватель кафедры прикладной лингвистики и межкультурной коммуникации, [lili-th89@narod.ru](mailto:lili-th89@narod.ru)

Одобрена на заседании кафедры прикладной лингвистики и межкультурной коммуникации «28»августа 2014 г.

Зав. кафедрой В.Г. Зусман

Рекомендована секцией УМС «Гуманитарные науки» «28»августа 2014 г.

Председатель В.Г. Зусман

Утверждена УМС НИУ ВШЭ – Нижний Новгород «22»сентября 2014 г.

Председатель В.М. Бухаров

Нижний Новгород, 2014

*Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения кафедры-разработчика программы*

# Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих данную дисциплину, учебных ассистентов и студентов-магистрантов направления, изучающих дисциплину «Корпусные методы в политической лингвистике».

Программа разработана в соответствии с:

* Образовательным стандартом НИУ ВШЭ для направления 45.04.03 «Фундаментальная и прикладная лингвистика»;
* Образовательной программой направления 45.04.03 «Фундаментальная и прикладная лингвистика»;
* Рабочим учебным планом университета по направлению подготовки 45.04.03 «Фундаментальная и прикладная лингвистика», утвержденным в 2014г.

# Цели освоения дисциплины

Цель курса:

- ознакомление с теоретическими основами корпусного подхода изучения речевой и языковой информации;

- ознакомление с терминологическим аппаратом дисциплины;

- практическое применение студентами корпусных методов для анализа текстов политической тематики.

Задачи изучения курса: научить анализировать языковые средства выражения политического дискурса; продемонстрировать применение специальных методик и приемов, основанных на цифровых (корпусных) технологиях, использующихся в современной политической лингвистике, и выполнять собственные экспериментальные исследования в данных областях науки.

# Компетенции обучающегося, формируемые в результате освоения дисциплины

Обучающиеся приобретают знания основ корпусного подхода изучения языка, знакомятся с основными корпусами русского, английского, немецкого, испанского, французского и пр. языков, знакомятся с классификацией корпусов по различным основаниям. Получают знания теоретических и практических основ создания корпусов, знания о корпусных менеджерах и функциях интерфейсов различных корпусов.

Студенты учатся извлекать из корпусов различного типа грамматическую, синтаксическую, семантическую и пр. информацию.

Студенты также учатся на практике применять полученные знания для своих индивидуальных квалификационных исследований.

В результате освоения дисциплины студент осваивает следующие компетенции:

| Компетенция | Код по НИУ | Дескрипторы – основные признаки освоения (показатели достижения результата) | Формы и методы обучения, способствующие формированию и развитию компетенции |
| --- | --- | --- | --- |
| способен рефлексировать, оценивать и перерабатывать освоенные научные методы и способы деятельности | СК-М1 | - тоды обучения:остранном языке подразумевает умению контактировать с иностранными коллегамиеской экспертизыописывать её в соотспособен к обобщению, анализу и восприятию информации;  - способен свободно пользоваться корпусами как инструментами лингвистического анализа;  - работа с корпусами на иностранном языке подразумевает умению контактировать с иностранными коллегами  -способен производить анализ проблем, осуществлять постановку задач, выбирать теории и методы для их решения  -способен анализировать и воспроизводить смысл междисциплинарных текстов с использованием языка и аппарата прикладной лингвистики  -распознаёт различные типы дискурсов /текстов | Формы обучения:  – интерактивные лекции,  – семинарские занятия,  – самостоятельная работа  Методы обучения:  - анализ научного исследования на предмет вычленения проблемы, исследовательского вопроса, целей, задач и рабочей гипотезы автора;  - обсуждение соответствия выбранных методов исследования поставленным целям и задачам;  - мониторинг и анализ массивов политической информации с использованием современных компьютерных инструментов лингвистического исследования;  - обсуждение проблематики и языка политических текстов;  - формулирование научно-исследовательской задачи на основе анализа массива политических текстов;  - обсуждение научно-исследовательской задачи и методов ее решения; |
| способен предлагать концепции, модели, изобретать и  апробировать способы и инструменты профессиональной  деятельности | СК-М2 |
| способен вести профессиональную, в том числе научно-  исследовательскую деятельность в международной среде | СК-М8 |
| cпособен планировать научно-исследовательскую  деятельность, проводить самостоятельные  исследования и получать новые научные  результаты в области профессиональной  деятельности | ИК-М1.2н\_1.2и\_1.2п\_1.2пр |
| способен произвести углубленный анализ  проблем, постановку и обоснование проектных  задач, связанных с лингвистической  проблематикой, моделированием и разработкой  лингвистических компонент в  междисциплинарных проектах | ИК–М.1.1-2\_М5.1-6 |
| способен самостоятельно разработать  методический инструментарий для осуществления  исследовательской и проектной деятельности в  области фундаментальной и прикладной  лингвистики | ИК-М.5.1-6 |
| способен проводить квалифицированный анализ  различных типов устного и письменного дискурса  в производственно-практических целях | ИК- М1.2н\_1.2и\_1.2п\_1.2пр |

# Место дисциплины в структуре образовательной программы

Настоящая дисциплина относится к вариативной части цикла дисциплин программы.

Для освоения учебной дисциплины, студенты должны освоить следующие дисциплины:

- Морфология;

- Синтаксис.

Основные положения дисциплины должны быть использованы в дальнейшем при изучении следующих дисциплин:

* функциональные и когнитивные модели в лингвистике;
* лингвистическая конфликтология и др.

# Тематический план учебной дисциплины

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № | Название раздела | Всего часов | Аудиторные часы | | Самостоя­тельная работа |
| Лекции | Семина-ры |
| 1 | Основная информация о корпусах | 8 | 2 |  | 6 |
| 2 | Основные постулаты корпусного подхода и терминологический аппарат корпусной лингвистики | 8 | 2 |  | 6 |
| 3 | Зарубежные корпусы | 8 | 2 | 2 | 6 |
| 4 | Классификация корпусов по разным основаниям | 8 | 2 |  | 6 |
| 5 | Инструменты корпусной лингвистики | 8 | 2 | 2 | 6 |
| 6 | Диахронический аспект политического дискурса и его современное состояние | 8 | 2 | 2 | 6 |
| 7 | Выбор темы, направленность речи, оценка состава слушателей и речевой ситуации | 10 | 2 | 2 | 4 |
| 8 | Отбор информации: как работать с письменными источниками; работа над фактами на примере интернета как лингвистического корпуса | 12 | 2 | 2 | 6 |
| 9 | Анализ лексического состава политических выступлений | 12 | 2 | 2 | 4 |
| 10 | Репрезентативность корпусов при анализе | 8 | 2 |  | 4 |
| 11 | Анализ типичных для политического дискурса синтаксических конструкций | 12 | 2 | 2 | 6 |
| 12 | Построение синтаксических моделей политических выступлений | 6 | 2 | 2 | 6 |
| 13 | Анализ предвыборной речи | 6 |  | 2 | 4 |
| 14 | Анализ тематических текстов с помощью корпусных методов: анализ рекламных текстов | 6 |  | 2 | 6 |
| 15 | Анализ тематических текстов с помощью корпусных методов: тексты СМИ | 6 |  | 2 | 4 |
| 16 | Анализ тематических текстов с помощью корпусных методов: тексты, связанные многонациональностью русскоязычного лингвокультурного сообщества | 6 |  | 2 | 4 |
| 17 | Анализ тематических текстов с помощью корпусных методов: тексты, связанные с информационной безопасностью | 6 |  | 2 | 4 |
| 18 | Анализ тематических текстов с помощью корпусных методов: анализ биополитики с помощью текстовых ресурсов | 6 |  | 2 | 4 |
|  |  |  |  |  |  |
|  | **Всего** | **144** | **24** | **28** | **92** |

# Формы контроля знаний студентов

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Тип  контроля | Форма  контроля | Модуль | | | | Параметры |
|  | 2 | 3 | 4 |
| Текущий  (неделя) | Домашнее задание |  |  |  | \* | письменная работа |
| Итоговый | Экзамен |  |  |  | \* | устный, защита письменной работы, презентация |

## Критерии оценки знаний, навыков

* Активность на лекциях и семинарских занятиях;
* Домашнее задание;
* Успешная подготовка к экзаменационному заданию;
* Успешная сдача экзамена.

Оценки по всем формам текущего контроля выставляются по 10-ти балльной шкале.

Для получения положительной оценки за курс обучающиеся должны посетить лекционные и семинарские аудиторные занятия, выполнить как в составе рабочей группы, так и самостоятельно небольшие исследования по тематике одного из семинаров или одной из лекций и выступить на семинаре с презентацией результатов исследования. В презентации обучающиеся показывают знание методов корпусной лингвистики и умение их использования на практике.

## Порядок формирования оценок по дисциплине

Итоговая оценка складывается из:

а) активности на семинарах (активное участие на протяжении всего семинара 10, средняя активность 5, полное или практически полное отсутствие активности 0), усредненной по числу семинаров;

б) самостоятельной работы (домашнего задания);

г) устного экзамена;

Накопленная оценка по 10-ти балльной шкале за самостоятельную работу определяется перед промежуточным или итоговым контролем – *Осам. работа*.

Накопленная оценка за текущий контроль учитывает результаты студента по текущему контролю следующим образом:

*Онакопл. = 0,3\*Оауд. + 0,3\*Осам.раб. + 0,4\*О д/з.*

Способ округления накопленной оценки текущего контроля: арифметический.

Результирующая оценка за итоговый контроль в форме экзамена выставляется по следующей формуле, где *Оэкз.* – оценка за работу непосредственно на экзамене:

*Оитоговый = 0,3·Оэкз. + 0,7·Онакопл.*

Округление оценки производится в пользу студента.

# Содержание дисциплины

**1. Основная информация о корпусах.**

Знакомство с основными понятиями. Практическое применение корпусов. Обзорное знакомство с историей корпусной лингвистики.

**2. Основные постулаты корпусного подхода и терминологический аппарат корпусной лингвистики**.

Основные понятия, составляющие терминологический аппарат корпусной лингвистики. Отличия корпусной лингвистики от лингвистики Н. Хомского.

**3. Зарубежные корпусы.**

Зарубежные корпусы в своём историческом развитии. Исследованиях Библии (Cruden и другие). Лексикографии (Oxford English Dictionary, Webster Dictionary). Преподавании языков (Thorndike, 1920). Дескриптивной грамматике (Fries, 1940, Quirk, 1968).

**4. Классификация корпусов по разным основаниям**.

Виды, типы корпусов. Задачи компьютерной лингвистики в связи с видами корпусов. Как можно использовать корпус для анализа политического дискурса.

**5. Инструменты корпусной лингвистики**.

Конкордансы. N-граммы. Задание подкорпусов. Частотность. Корреляция.

**6. Диахронический аспект политического дискурса и его современное состояние.**

Тексты отечественного политического дискурса с 30-х гг XX в. по 10-е гг XXI в. Тексты зарубежного политического дискурса. Открытый корпус Европарламента, <http://www.statmt.org/europarl/>.

**7. Выбор темы, направленность речи, оценка состава слушателей и речевой ситуации.**

Анализ метаразметки корпуса как способ установить коммуникативную ситуацию анализ коммуникативной ситуации как первый этап исследования текста политического дискурса.

**8. Отбор информации: как работать с письменными источниками; работа над фактами на примере интернета как лингвистического корпуса**.

Алгоритм отбора информации. Поисковые запросы в системе интернет.

**9. Анализ лексического состава политических выступлений.**

Анализ семантики текстов с помощью корпусных методов. Составление тематических сеток текстов. Анализ стилистики текстов с помощью корпуса.

**10. Репрезентативность корпусов при анализе.**

Понятие репрезентативности. Относительность репрезентативности. Опыт составления малых корпусов и сбора малых коллекций текстов.

**11. Анализ типичных для политического дискурса синтаксических конструкций.**

Анализ ресурсов типа tree-bank. Извлечение синтаксической информации из национальных корпусов. Извлечение синтаксической информации из корпусов без синтаксической разметки.

**12. Построение синтаксических моделей политических выступлений.**

Анализ текстов политических выступлений с точки зрения актуального членения конструкций. Анализ синтаксических стилистических фигур. Сценарии, скрипты, фреймы политических выступлений.

***Семинар 1.*** Зарубежные корпусы.

***Семинар 2.*** Классификация корпусов по разным основаниям.

***Семинар 3.*** Инструменты корпусной лингвистики.

***Семинар 4.*** Диахронический аспект политического дискурса и его современное состояние.

***Семинар 5.*** Выбор темы, направленность речи, оценка состава слушателей и речевой ситуации.

***Семинар 6.*** Отбор информации: как работать с письменными источниками; работа над фактами на примере интернета как лингвистического корпуса.

***Семинар 7.*** Анализ лексического состава политических выступлений.

***Семинар 8.*** Анализ типичных для политического дискурса синтаксических конструкций.

***Семинар 9.*** Построение синтаксических моделей политических выступлений.

***Семинар 10.*** Анализ предвыборной речи

***Семинар 11.*** Анализ тематических текстов с помощью корпусных методов: анализ рекламных текстов

***Семинар 11.*** Анализ тематических текстов с помощью корпусных методов: тексты СМИ.

***Семинар 12.*** Анализ тематических текстов с помощью корпусных методов: тексты, связанные многонациональностью русскоязычного лингвокультурного сообщества.

***Семинар 13.*** Анализ тематических текстов с помощью корпусных методов: тексты, связанные с информационной безопасностью.

***Семинар 14.*** Анализ тематических текстов с помощью корпусных методов: анализ биополитики с помощью текстовых ресурсов.

# Образовательные технологии

Данная дисциплина преподаётся в форме лекционных и семинарских занятий, а также с помощью освоения студентами информации посредством самостоятельной работы.

# Оценочные средства для текущего контроля и аттестации студента

В основу текущего контроля положено посещение учебных занятий и участие в их работе. При посещении от 39 до 46 часов занятий обучающиеся получают 10 баллов в накопленную оценку, что составляет 0,7 итоговой оценки. При посещении менее 39 часов – 0 баллов, что не является, однако, блокирующей оценкой.

Экзаменационная оценка – оценка за презентацию самостоятельного исследования на семинаре выставляется каждому участнику рабочей группы с учетом степени его личного участия в исследовании.

Критерии оценки презентации исследования на семинаре:

1. Полнота изложения теоретического материала – до 3 баллов
2. Объем использованных источников – до 1 балла
3. Участие в дискуссии – до 1 балла
4. Качество выполнения и интерпретации экспериментального материала (индивидуально для каждого участника рабочей группы) – до 4 баллов
5. Качество ppt-презентации – до 1 балла

# Учебно-методическое и информационное обеспечение дисциплины

## Основная литература

Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. – Иркутск: ИГЛУ, 2011.

## Дополнительная литература

1.  Анищенко О. А. Эволюция обозначения молодежной речи: от технического языка до жаргона. Вопросы языкознания, № 2, 2009.

2.  Венцов А. В., Грудева Е. В. О корпусе русского литературного языка (www. narusco. ru) // Russian Linguistics, Vol. 33, No 2, 2009.

3.  Добросклонская Т. Г. Лингвистические способы выражения идеологической модальности в медиатекстах // Вестник МГУ, Серия 19, Лингвистика и межкультурная коммуникация, 2/2009. С.

4.  Копотев М. В., Мустайоки А. Современная корпусная русистика. // Инструментарий русистики: корпусные подходы. Slavica Helsingiensia, 34. Helsinki University Press, 2010.

5.  Кронгауз М. А. Номинация по созвучию как способ освоения слов // Russian Linguistics, Vol. 33, No 1, 2009.

6.  Крысин Л. П. Некоторые принципы словарного описания русской разговорной речи (постановка задачи) // Русский язык в научном освещении, № 2(16), 2009.

7.  Никипорец-Такигава Г. Неологизмы: метод поиска при помощи маркеров новизны («как сейчас говорят») и пределы компьютерных возможностей. //Инструментарий русистики: корпусные подходы. Slavica Helsingiensia, 34. Helsinki University Press, 2009.

8.  Плунгян В. А. Корпус как инструмент и как идеология: о некоторых уроках современной корпусной лингвистики // Русский язык в научном освещении, № 2(16), 2009.

9.  Шейгал Е. И. Семиотка политического дискурса. – М.: ИТДГК «Гнозис». – 2011. – 326с.

**10.3. Электронные источники и корпусы:**

1. Национальный корпус русского языка, НКРЯ, URL: [http://www.ruscorpora.ru](http://www.ruscorpora.ru/).
2. Открытый корпус, http://opencorpora.org/.
3. Корпус устной речи, <http://spokencorpora.ru/>
4. Брауновский корпус, Brown Corpus ,  
   <http://corpus.leeds.ac.uk/protected/>.
5. The Lancaster-Oslo/Bergen Corpus   
   (often abbreviated as LOB Corpus),   
    <http://www.hit.uib.no/icame/lob/lob-dir.htm>.
6. Машинный фонд русского языка.
7. The Uppsala Corpus (Upsal'skij korpus russkix tekstov), <http://www.moderna.uu.se/slaviska/ryska/corpus>.
8. Британский национальный корпус   
   (BNC – British National Corpus),  
    <http://corpus.byu.edu/bnc/>.
9. Bank of English,  
    <http://www2.lingsoft.fi/doc/engcg/Bank-of-English.html>.
10. Corpus of Contemporary American English,  
     <http://corpus.byu.edu/coca/>.
11. American national corpus, http://www.americannationalcorpus.org/oanc/index.html,  
     <http://www.americannationalcorpus.org/>.
12. [Синтаксически размеченный корпус – банк деревьев (Penn TreeBank, http://www.cis.upenn.edu/~treebank/](http://www.cis.upenn.edu/~treebank).
13. Параллельные корпуса:

* EUROPARL - 20.000.000 словоупотреблений, открытый корпус Европарламента на 11 языках (автор Philipp Koehn) <http://www.statmt.org/europarl/> - можно скачать;
* CHEMNITZ GERMAN-ENGLISH TRANSLATION CORPUS - 1.000.000 словоупотреблений (тексты - академические, политика, туризм) <http://ell.phil.tu-chemnitz.de/search/> ;
* KACENKA (Korpus anglicko-cesky; Czech) - 3.000.000 словоупотребленийй <http://www.phil.muni.cz/angl/kacenka/kachna.html> - платный;
* English-French Canadian Hansard (Brown et al, 1991; Gale and Church, 1991);
* Lancaster's UCREL, <http://ucrel.lancs.ac.uk/corpora.html> ;
* OPUS parallel corpus, <http://opus.lingfil.uu.se/>
* [Параллельный корпус переводов «Слова о полку Игореве»](http://www.nevmenandr.net/slovo/) (доступен для онлайн-поиска) http://www.nevmenandr.net/slovo/.

1. Мультисловари:

- <http://www.dwds.de/>

- <http://www.lingvo-online.ru/>

1. Многоязычные корпусы:

- European Parliament Proceedings Parallel Corpus 1996-2011, [Корпус слушаний Европарламента](http://www.statmt.org/europarl/) (доступен для скачивания), <http://www.statmt.org/europarl/>;

- [Корпус документов Евросоюза (более 20 языков)](http://ipsc.jrc.ec.europa.eu/index.php?id=198) (доступен для скачивания), <http://ipsc.jrc.ec.europa.eu/index.php?id=198>;

- [ParaSol](http://www-korpus.uni-r.de/ParaSol/): параллельный корпус славянских и других языков Бернского университета (бывший Регенсбургский) (доступен для онлайн-поиска, требуется регистрация по e-mail) [http://www.parasolcorpus.org](http://www.parasolcorpus.org/) ;

- [InterCorp](http://www.korpus.cz/intercorp/): параллельные корпуса Пражского университета (доступен для онлайн-поиска, требуется регистрация, дающая доступ также к Чешскому национальному корпусу) <http://www.nevmenandr.net/slovo/>

- [Параллельный корпус переводов «Слова о полку Игореве»](http://www.nevmenandr.net/slovo/) (доступен для онлайн-поиска) http://www.nevmenandr.net/slovo/

1. Германоязычные корпусы:

- <http://www.aac.ac.at/lab.html>;

- <http://www.dwds.de/>;

* <http://www1.ids-mannheim.de/kl/projekte/korpora/> ;
* <http://dgd.ids-mannheim.de:8080/dgd/pragdb.dgd_extern.welcome> ;
* <http://corp.hum.sdu.dk/cqp.de.html> ;

1. Франкоязычные корпусы:

* <http://corp.hum.sdu.dk/cqp.fr.html>
* [http://wortschatz.uni-leipzig.de/ws\_fra/](https://vk.com/away.php?utf=1&to=http://wortschatz.uni-leipzig.de/ws_fra/)
* [http://www.cnrtl.fr/corpus/](https://vk.com/away.php?utf=1&to=http://www.cnrtl.fr/corpus/)
* [http://lextutor.ca/conc/fr/](https://vk.com/away.php?utf=1&to=http://lextutor.ca/conc/fr/)
* [http://www.lextutor.ca/](https://vk.com/away.php?utf=1&to=http://www.lextutor.ca/)
* [http://cfpp2000.univ-paris3.fr/](https://vk.com/away.php?utf=1&to=http://cfpp2000.univ-paris3.fr/) -

1. Корпусы СМИ:

* <http://www.integrumworld.com/rus/subscribe.php> - Интегрум
* БОКР: <http://www.comp.leeds.ac.uk/ssharoff/bokrcorpora/index-en.html>
* [http://www.philol.msu.ru/~lex/corpus/corp\_descr.html#5](http://www.philol.msu.ru/~lex/corpus/corp_descr.html)
* Информация: <http://corpora.phil.spbu.ru/Works2011/%D0%A1%D0%B0%D0%B2%D1%87%D1%83%D0%BA_309.pdf>

# Материально-техническое обеспечение дисциплины и условие участия

Используется проектор, ноутбук. Условие участия – умение обучающихся работать с основными офисными программами, включая Excel.

Автор программы А.Ю. Хоменко