

Правительство Российской Федерации

**Федеральное государственное автономное образовательное учреждение
высшего профессионального образования
«Национальный исследовательский университет
«Высшая школа экономики»**

Факультет Компьютерных наук
Департамент больших данных и информационного поиска
Базовая кафедра Яндекс

УТВЕРЖДАЮ
Академический руководитель
образовательной программы
«Науки о данных»
по направлению 01.04.02
«Прикладная математика и информатика»
С.О. Кузнецов

«__» _____ 2015 г.

Программа дисциплины «Алгоритмы обработки больших данных»
для направления 01.04.02 "Прикладная математика и информатика" подготовки
магистра
для магистерской программы "Науки о данных"

Автор программы:
Пузыревский И.В. (ipuzyrevskiy@hse.ru)

Одобрена на заседании базовой кафедры Яндекс «__» _____ 2015 г.

Заведующий кафедрой _____ М.А. Бабенко

Рекомендована Академическим советом образовательной программы
«Науки о данных» «__» _____ 2015 г.

Менеджер базовой кафедры Яндекс _____ И.И. Алескерова

Москва, 2015

Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения подразделения разработчика программы.



Пояснительная записка

Автор программы

Пузыревский И.В.

Аннотация

Дисциплина «Введение в лингвистику» предназначена для подготовки магистров 01.04.02 – Прикладная математика и информатика.

Данный курс рассчитан на тех, кто не учился лингвистике, но хочет узнать о разнообразии языков мира, о том, как устроен человеческий язык, и о том, что в нем доступно для автоматической обработки, а что пока недоступно. В рамках курса предусматривается решение лингвистических задач разных типов (самодостаточных и не совсем).

Программа курса предусматривает лекции (26 часов) и практические занятия (38 часов).

Учебные задачи курса

Целью данного курса является ознакомление слушателей с естественными языками, их разнообразием, структурой, письменностью. Все темы курса снабжены практическими лингвистическими заданиями, призванными продемонстрировать изучаемую теорию на примере различных языков.

Тематический план дисциплины «Введение в лингвистику»

№	Название темы	Всего часов по дисциплине	Аудиторные часы		Самостоятельная работа
			Лекции	Сем. и практика	
1	Вводная часть: язык как система	10	2	2	6
2	Происхождение, функционирование и развитие естественных языков	14	2	4	8
3	Языки мира	12	2	4	6
4	Грамматика и лексика	14	4	4	8
5	Фонетика	18	2	6	10
6	Морфология	22	4	6	10
7	Синтаксис	22	4	8	10
8	Лексическая семантика	14	2	2	10
9	Лексикография	14	2	2	10
10	Корпусная лингвистика	14	2	2	10



Итого	152	26	38	88
-------	-----	----	----	----

I. Источники информации

Список литературы

Основная литература

1. Ю. Д. Апресян. Лингвистическая терминология словаря // В. Ю. Апресян, Ю. Д. Апресян, Е. Э. Бабаева, О. Ю. Богуславская, И. В. Галактионова, М. Я. Гловинская, С. А. Григорьева, Б. Л. Иомдин, Т. В. Крылова, И. Б. Левонтина, А. В. Птенцова, А. В. Санников, Е. В. Урысон. Новый объяснительный словарь синонимов русского языка. Москва – Вена, 2004. VIII–XI.
2. Ю. Д. Апресян. Основные принципы и понятия системной лексикографии // В. Ю. Апресян, Ю. Д. Апресян, Е. Э. Бабаева, О. Ю. Богуславская, Б. Л. Иомдин, Т. В. Крылова, И. Б. Левонтина, А. В. Санников, Е. В. Урысон. Языковая картина мира и системная лексикография. М., 2006, 33–74.
3. Большой энциклопедический словарь. Языкознание. М., 2003.
4. С. А. Бурлак, С. А. Старостин. Сравнительно-историческое языкознание. М., 2005.
5. Задачи лингвистических олимпиад. 1965–1975. Ред.-сост. В. И. Беликов, Е. В. Муравенко, М. Е. Алексеев. М., 2007.
6. А. Зализняк. Из заметок о любительской лингвистике. М., 2009.
7. Л. Л. Иомдин. О модели русского синтаксиса // Ю. Д. Апресян, И. М. Богуславский, Л. Л. Иомдин, В. З. Санников. Теоретические проблемы русского синтаксиса. Взаимодействие грамматики и словаря. М., 2010, с. 21–43.
8. С. Пинкер. Язык как инстинкт. М., 2009.
9. В. А. Плунгян. Почему языки такие разные. М., 2010.

Дополнительная литература

1. С. В. Князев, С. К. Пожарицкая. Современный русский литературный язык: Фонетика, графика, орфография, орфоэпия. М., 2012.
2. М. А. Кронгауз. Семантика. М., 2005.
3. М. Копотев. Введение в корпусную лингвистику. [Электронноеиздание](#). Прага, 2014.
4. Дж. Лайонз. Язык и лингвистика: Вводный курс. М., 2004.
5. В. А. Плунгян. Общая морфология: Введение в проблематику. М., 2012.
6. Я. Г. Тестелец. Введение в общий синтаксис. М., 2001.
7. А. Я. Шайкевич. Введение в лингвистику. М., 2010.
8. A. Akmajian, R. A. Demers, A. K. Farmer, R. M. Harnish. Linguistics. An Introduction to Language and Communication. MIT Press, 2010.



9. Ю. Д. Апресян. Исследования по семантике и лексикографии. Т.1. Парадигматика. М., 2009.
10. Ю. Д. Апресян, И. М. Богуславский, Л. Л. Иомдин, А. В. Лазурский, Н. В. Перцов, В.З. Санников, Л. Л. Цинман. Лингвистическое обеспечение системы ЭТАП-2. М., 1989.
11. В. Ю. Апресян, Ю. Д. Апресян, Е. Э. Бабаева, О. Ю. Богуславская, И. В. Галактионова, М. Я. Гловинская, Б. Л. Иомдин, Т. В. Крылова, И. Б. Левонтина, А. В. Птенцова, А. В. Санников, Е. В. Урысон. Активный словарь русского языка. Тт. 1–2: А–Г. Отв. ред. акад. Ю. Д. Апресян. М., 2014.
12. Т. В. Булыгина, А. Д. Шмелев. Языковая концептуализация мира (на материале русской грамматики). М., 1997.
13. С. А. Бурлак. Происхождение языка. Факты, исследования, гипотезы. М., 2011.
14. А. Вежбицкая. Семантические универсалии и описание языков. М., 1999.
15. А. В. Гладкий. Синтаксические структуры естественного языка. 2-е изд. М., 2007.
16. А. А. Зализняк. Русское именное словоизменение. С приложением избранных работ. М., 2002.
17. З. Зорина, А. Смирнова. О чем рассказали «говорящие» обезьяны. М., 2006.
18. Б. Л. Иомдин. Как составляются словари. Конспект семинаров // Лингвистика для всех. Летние лингвистические школы 2005–2006. М., 2008, с. 75–87.
19. Б. Л. Иомдин. Многозначные слова в контексте и вне контекста // Вопросы языкознания, 2014, №4. С. 87–103.
20. Б. Л. Иомдин, А. А. Лопухина, Г. В. Носырев. К созданию частотного словаря значений слов // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2014» (Бекасово, 4–8 июня 2014 г.). Вып. 13 (20). М.: РГГУ, 2014. С. 199–212.
21. А. Леонтьев. Путешествие по карте языков мира. М., 2008.
22. И. А. Мельчук. Опыт теории лингвистических моделей «Смысл \Leftrightarrow Текст». М., 1974.
23. А. Пиперски. Артланги: искусственные языки в литературе и в кино
24. И. Фридрих. История письма. М., 1979.
25. Н. Хомский. Аспекты теории синтаксиса. Пер. В. А. Звегинцева. М., 1972.
26. Recreational Linguistics. Puzzles in Logic, Languages and Computation. Ed. By Dragomir Radev. Springer, 2013.
27. G. Yule. The Study of Language. Cambridge University Press, 2010.

II. Формы контроля и структура итоговой оценки

Текущий контроль - домашняя работа в первом модуле, контрольная работа в первом модуле.

Итоговый контроль – письменный экзамен (120 мин.)

Итоговая оценка вычисляется следующим образом:



0,1*оценка за домашнюю + 0,2*оценка за контрольную + 0,7*оценка за экзамен.

Таблица соответствия оценок по десятибалльной и системе зачет/незачет

Оценка по 10-балльной шкале	Оценка по 5-балльной шкале
1	Незачет
2	
3	
4	Зачет
5	
6	
7	
8	
9	
10	

Таблица соответствия оценок по десятибалльной и пятибалльной системе

По десятибалльной шкале	По пятибалльной системе
1 – неудовлетворительно 2 – очень плохо 3 – плохо	неудовлетворительно – 2
4 – удовлетворительно 5 – весьма удовлетворительно	удовлетворительно – 3
6 – хорошо 7 – очень хорошо	хорошо – 4
8 – почти отлично 9 – отлично 10 - блестяще	отлично – 5

III. Программа дисциплины «Введение в лингвистику»

Тема 1. Вводная часть: язык как система

Когда появилась лингвистика
Что изучает лингвистика
Синхрония и диахрония
Семиотика
Свойства языковых знаков
Омонимия и многозначность
Многозначность на разных уровнях языка

Тема 2. Происхождение, функционирование и развитие естественных языков

Происхождение языка



От языка животных к языку людей
Избыточность языкового знака
Свойства естественного языка
Функции языка
Социальные формы существования языка
Диалектный континуум
Основные понятия лингвистической компаративистики
Регулярные фонетические соответствия
Глоттохронология
Языковая систематика
Генеалогическая классификация языков

Тема 3. Языки мира

Обзор языков мира
Основные графические системы

Тема 4. Грамматика и лексика

Лексическое и грамматическое значение
Свойства грамматических значений
Грамматикализация и деграмматикализация
Грамматические категории
Интегральное описание языка

Тема 5. Фонетика

Фонетика: основные понятия
Артикуляционная классификация звуков
Акустическая классификация звуков
Фонология

Тема 6. Морфология

Морфема и морфемика
Основные типы аффиксов
Словообразование и словоизменение
Основа и флексия
Грамматические средства языка
Типологическая классификация языков
Историческая морфология

Тема 7. Синтаксис

Предложение и высказывание
Синтаксические структуры: непосредственно составляющие, деревья
зависимостей, гибридные структуры
Типы синтаксических отношений в модели «Смысл↔Текст»
Синтаксически размеченный корпус SynTagRus
Свойства синтаксических отношений
Свойства синтаксических структур



Синтаксические признаки
Гипотеза глубины

Тема 8. Лексическая семантика

История семантики
Идеи современной семантики
Омонимия и синонимия
Омонимия и полисемия
Типы полисемии
Мышление и язык
Языковая картина мира

Тема 9. Лексикография

История лексикографии
Толковые словари
Принципы системной лексикографии
Естественные метаязыки: Natural Semantic Metalanguage, Московская семантическая школа
Особая область: бытовая лексика

Тема 10. Корпусная лингвистика

Лингвистика в докомпьютерную эпоху
«Революция в лингвистике»
Корпуса текстов и корпусная лингвистика

IV. Методические указания студентам

Самостоятельная работа студента предусматривает решение лингвистических задач.

Автор программы: _____ / <Пузыревский И.В.> /