

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
"Амурский государственный университет"

УТВЕРЖДАЮ

Проректор по учебной и научной
работе

Лейфа А.В. Лейфа

26 июня 2024 г.

РАБОЧАЯ ПРОГРАММА
«ЛИНГВИСТИЧЕСКИЕ БАЗЫ ДАННЫХ»

Направление подготовки 45.04.03 Фундаментальная и прикладная лингвистика

Направленность (профиль) образовательной программы – Иностранные языки и речевые
технологии

Квалификация выпускника – Магистр

Год набора – 2024

Форма обучения – Очная

Курс 1 Семестр 2

Экзамен 2 сем

Общая трудоемкость дисциплины 216.0 (академ. час), 6.00 (з.е)

Составитель О.Н. Морозова, доцент, д-р. филол. наук

Филологический факультет

Кафедра иностранных языков

Рабочая программа составлена на основании Федерального государственного образовательного стандарта ВО для направления подготовки 45.04.03 Фундаментальная и прикладная лингвистика, утвержденного приказом Министерства науки и высшего образования Российской Федерации от 12.08.20 № 993

Рабочая программа обсуждена на заседании кафедры иностранных языков

_____ г. _____, протокол № _____

Заведующий кафедрой _____ Морозова О.Н. Морозова

СОГЛАСОВАНО

Учебно-методическое управление

_____ Чалкина Н.А. Чалкина

_____ 26 июня 2024 г.

СОГЛАСОВАНО

Выпускающая кафедра

_____ 26 июня 2024 г.

СОГЛАСОВАНО

Научная библиотека

_____ Петрович О.В. Петрович

_____ 26 июня 2024 г.

СОГЛАСОВАНО

Центр цифровой трансформации и
технического обеспечения

_____ Тодосейчук А.А. Тодосейчук

_____ 26 июня 2024 г.

		данные, дата/ время, денежные данные, счетчики, логические данные, поле «МЕМО», поле объекта «OLE», гиперссылки. Типы моделей: иерархические, сетевые, реляционные, объектно- ориентированные, объектно-реляционные.
2	Основные понятия, связанные с лингвистическими информационными ресурсами	1. Одноязычные и многоязычные ЛИР 2. Пассивные формы ЛИР. Активные формы ЛИР. 3. Проблемы разработки ЛИР: единство стандартов, экспертных требований, стратегий; защита ЛИР. Многофункциональные ЛИР большого объема.
3	Электронные словари и справочники	1. Одноязычный письменный лексикон. Словари. 2. Классификация словарей по: -лексикографической форме основной единицы -специфике отбираемой лексики -способу описания основных единиц -расположению материала в словарях -эпохе функционирования -назначению (адресату) 3. Примеры словарей (словоуказатель, конкорданс, контекстологический словарь) 4. Электронные энциклопедии и тезаурусы. 5. Онтологии (составляющие концептуальной схемы: релевантные классы объектов + теоремы и ограничения, отражающие связи объектов).
4	Корпусы данных различных языков	Банки данных устных и письменных текстов (текстовые массивы). Задачи построения корпусов в лексикографии и лексикологии, в грамматике, в лингвистике текста, при автоматическом переводе, в учебных целях. Исследовательские, иллюстративные, статические и динамические корпусы текстов; корпусы параллельных текстов. Особенности представления текстов в корпусе. Технология тэггирования. Проблемные вопросы создания корпусов данных.
5	Обработка массивов письменных текстов	Проблемы, связанные с созданием корпусов звучащих текстов. Лингвистические задачи, решаемые с помощью корпуса звучащих текстов. Основные принципы разметки звучащих текстов в программе PRAAT: создание и сохранение файла формата TextGrid, для составления аннотации, задание уровней, определение их оптимального количества для решения различных лингвистических задач, принципы сегментации звучащего текста, отделение лингвистических событий от паралингвистических, постановка меток и их удаление, заполнение уровней подписями, применение транскрипционных знаков, сохранение внесенных данных.

6	Обработка массивов звучащих текстов	Проблемы, связанные с созданием корпусов звучащих текстов. Лингвистические задачи, решаемые с помощью корпуса звучащих текстов. Основные принципы разметки звучащих текстов в программе PRAAT: создание и сохранение файла формата TextGrid, для составления аннотации, задание уровней, определение их оптимального количества для решения различных лингвистических задач, принципы сегментации звучащего текста, отделение лингвистических событий от паралингвистических, постановка меток и их удаление, заполнение уровней подписями, применение транскрипционных знаков, сохранение внесенных данных.
7	Алгоритм и способы его представления, языки программирования, универсальный язык программирования Free Pascal	Алгоритм, его понятие и свойства: задача, начальные условия, метод и способ ее решения; правило и алгоритм; свойства алгоритма (дискретность, результативность, массовость, детерминированность и формализованность). 4. Способы записи алгоритмов (словесный, графический, табличный, блок-схемой). 5. Средства решения лингвистических задач. 6. Алгоритмические языки и их элементы (Basic, Pascal, язык скриптов PRAAT и др): алфавит, типы исходных данных, операторы обработки исходных данных. 7. Основы языка скриптов к PRAAT.
8	Системы управления базами данных и способы доступа к информации в базах данных	1. Способы создания баз данных. 2. Поиск информации и модернизация данных. 3. Основные типы СУБД: настольные и серверные. 4. Способы доступа к информации в базах данных, основные режимы выдачи информации.

5.2. Практические занятия

Наименование темы	Содержание темы
Базы данных и их классификация.	1. Основные понятия: данное, объект, свойство объекта, атрибут. 2. Организация данных: кортеж, n-мерный кортеж, идентификатор объекта. 3. Типы данных: текстовые данные, числовые данные, дата/ время, денежные данные, счетчики, логические данные, поле «МЕМО», поле объекта «OLE», гиперссылки. 4. Типы моделей: иерархические, сетевые, реляционные, объектно- ориентированные, объектно-реляционные.
Основные понятия, связанные с лингвистическими информационными ресурсами (ЛИР).	1. Одноязычные и многоязычные ЛИР. 2. Пассивные формы ЛИР. Активные формы ЛИР. 3. Проблемы разработки ЛИР: единство стандартов, экспертных требований, стратегий; защита ЛИР. Многофункциональные ЛИР большого объема.
Электронные словари, справочники и онтологии.	1. Одноязычный письменный лексикон. Словари. 2. Классификация словарей по: лексикографической

	<p>форме основной единицы, специфике отбираемой лексики, способу описания основных единиц, расположению материала в словарях, эпохе функционирования, назначению (адресату). 3. Примеры словарей (словоуказатель, конкорданс, контекстологический словарь) 4. Электронные энциклопедии и тезаурусы. 5. Онтологии (составляющие концептуальной схемы: релевантные классы объектов + теоремы и ограничения, отражающие связи объектов).</p>
<p>Корпусы данных различных языков.</p>	<p>1. Банки данных устных и письменных текстов (текстовые массивы). 2. Задачи построения корпусов в лексикографии и лексикологии, в грамматике, в лингвистике текста, при автоматическом переводе, в учебных целях. 3. Исследовательские, иллюстративные, статические и динамические корпусы текстов; корпусы параллельных текстов. 4. Особенности представления текстов в корпусе. 5. Технология тэггирования. 6. Проблемные вопросы создания корпусов данных.</p>
<p>Обработка массивов письменных текстов.</p>	<p>1. Понятие корпуса текста. 2. Лингвистические задачи, для решения которых используются корпусы письменных текстов. 3. Классификация корпусов текстов. 4. Особенности представления текстов в корпусе. 5. Проблемы, связанные с созданием корпусов письменных текстов. 6. Задачи, решаемые при конкретизации качественного состава корпуса текстов. 7. Разметка письменного текста (таггирование). 8. Проблематика создания корпусов параллельных текстов и пути решения.</p>
<p>Обработка массивов звучащих текстов.</p>	<p>1. Проблемы, связанные с созданием корпусов звучащих текстов. 2. Лингвистические задачи, решаемые с помощью корпуса звучащих текстов. 3. Основные принципы разметки звучащих текстов в программе PRAAT. 4. Программа обработки речевого сигнала PRAAT. 5. Создание TextGrid, задание количества уровней и их названий, переименование уровней. 6. Постановка, перемещение и удаление меток. 7. Правила заполнения уровней. 8. Знаки для фонемной транскрипции, правила их набора и корректировки. 9. Знаки для аллофонной транскрипции, правила их набора и корректировки.</p>
<p>Алгоритм и способы его представления.</p>	<p>1. Алгоритм, его понятие и свойства: задача, начальные условия, метод и способ ее решения; правило и алгоритм; свойства алгоритма (дискретность, результативность, массовость, детерминированность и формализованность). 2. Способы записи алгоритмов (словесный, графический, табличный, блок-схемой). 3. Средства решения лингвистических задач. 4. Алгоритмические языки и их элементы. 5. Основы языка скриптов к PRAAT.</p>

Система управления базами данных.	1. Способы создания баз данных. 2. Поиск информации и модернизация данных. 3. Основные типы СУБД: настольные и серверные. 4. Способы доступа к информации в базах данных, основные режимы выдачи информации. 5. Основные понятия о данных. Организация данных. 6. Типы данных: история возникновения и развитие. 7. Типы моделей баз данных: основные черты, достоинства и недостатки. 8. Электронные лингвистические ресурсы: словари, тезаурусы, онтологии, энциклопедии.
-----------------------------------	---

5.3. Лабораторные занятия

Наименование темы	Содержание темы
Основы программирования на языке Pascal на примере программ чтения английских буквосочетаний.	1. Структурные особенности языка Pascal. 2. Программа для чтения английского диграфа -ea-. 3. Программа для чтения английского диграфов -ou/ow-. 4. Программа для чтения английского диграфа -th-.

6. САМОСТОЯТЕЛЬНАЯ РАБОТА

№ п/п	Наименование темы (раздела)	Содержание темы (раздела)	Трудоемкость в академических часах
1	Базы данных и их классификация	Подготовка к терминологическим диктантам и тестам, выполнение практических заданий.	14,5
2	Основные понятия, связанные с лингвистическими информационными ресурсами	Подготовка к терминологическим диктантам и тестам, выполнение практических заданий.	14,5
3	Электронные словари и справочники	Подготовка к терминологическим диктантам и тестам, выполнение практических заданий.	14,5
4	Корпусы данных различных языков	Подготовка к терминологическим диктантам и тестам, выполнение практических заданий.	14,5
5	Обработка массивов письменных текстов	Подготовка к терминологическим диктантам и тестам, выполнение практических заданий.	14,5
6	Обработка массивов звучащих текстов	Подготовка к терминологическим диктантам и тестам, выполнение практических заданий.	14,5
7	Алгоритм и способы его представления, языки программирования, универсальный язык	Подготовка к терминологическим диктантам и тестам, выполнение практических заданий.	14,5

	программирования Free Pascal		
8	Системы управления базами данных и способы доступа к информации в базах данных	Подготовка к терминологическим диктантам и тестам, выполнение практических заданий.	14,5

7. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

В ходе преподавания дисциплины лекции с традиционным и проблемным изложением теоретического материала, использование интерактивных обучающих мультимедиа средств; практические занятия с текущим устным опросом, интерактивные методы решения задач, использование наглядных средств, выполнение тестовых заданий, подготовка и презентация проекта.

8. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания, типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков, характеризующих этапы формирования компетенций, а так же методические материалы, определяющие процедуры оценивания знаний, умений и навыков отражены в оценочных средствах по дисциплине.

В процессе изучения дисциплины осуществляется текущий, промежуточный и итоговый контроль знаний.

Текущий контроль знаний проводится в виде опроса на занятиях. Одной из форм контроля является использование большого количества практических заданий на сегментацию и аннотирование звучащей и письменной речи, составление алгоритмов. Данные задания призваны способствовать формированию навыков сбора, обработки, использования и хранения лингвистической информации.

Промежуточный контроль осуществляется в виде терминологических диктантов и многовыборных тестов по пройденному материалу и практических заданий.

Итоговый контроль проводится в виде экзамена по билетам с одним вопросом и одним практическим заданием в билете.

Примерные вопросы к экзамену:

1. Базы данных и их классификация.
2. Основные понятия, связанные с лингвистическими информационными ресурсами (ЛИР).
3. Электронные словари, справочники и онтологии.
4. Корпусы данных различных языков.
5. Обработка массивов письменных текстов.
6. Обработка массивов звучащих текстов.
7. Алгоритм и способы его представления.
8. Система управления базами данных.
9. Основы программирования на языке Pascal на примере программ чтения английских буквосочетаний.

9. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

а) литература

1. Королева, О. Н. Базы данных : курс лекций / О. Н. Королева, А. В. Мажукин, Т. В. Королева ; под редакцией В. И. Мажукин. — Москва : Московский гуманитарный университет, 2012. — 66 с. — ISBN 978-5-98079-838-3. — Текст : электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprbookshop.ru/14515.html> (дата обращения: 25.06.2024). — Режим доступа: для авторизир. пользователей
2. Федорова, О. В. Экспериментальный анализ дискурса / О. В. Федорова. — Москва :

Языки славянской культуры, 2014. — 511 с. — ISBN 978-5-9905856-0-7. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/35742.html> (дата обращения: 25.06.2024). — Режим доступа: для авторизир. пользователей

3.Алексеев, В. А. Основы проектирования и реализации баз данных : методические указания к проведению лабораторных работ по курсу «Базы данных» / В. А. Алексеев. — Липецк : Липецкий государственный технический университет, ЭБС АСВ, 2014. — 26 с. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/55122.html> (дата обращения: 25.06.2024). — Режим доступа: для авторизир. пользователей

б) программное обеспечение и Интернет-ресурсы

в) профессиональные базы данных и информационные справочные системы

№	Наименование	Описание
1	https://elibrary.ru/	Научная электронная библиотека eLIBRARY.RU - российский информационно- аналитический портал в области науки, технологии, медицины и образования
2	http://www.ruscorpora.ru	Национальный корпус русского языка. Информационно- справочная система, основанная на собрании русских текстов в электронной форме
3	http://diss.rsl.ru/	Электронная библиотека диссертаций
4	http://www.gumer.info/bibliotek_Buks/Linguist/Index_Ling.php	Библиотека Гумер - Языкознание. Лингвистика. Филология. Языкознание.
5	http://superlinguist.ru/	Superlinguist – электронная научная библиотека, посвященная теоретическим и прикладным вопросам лингвистики, а также изучению различным языков.

10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

В учебном процессе используются учебные аудитории, лаборатория экспериментально- фонетических исследований, компьютерный класс, лингафонный кабинет при кафедре ИЯ ФФ АмГУ для проведения лекционных и практических занятий с возможностью использования мультимедийных средств. Самостоятельная работа обучающихся осуществляется в помещениях, оснащенных компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета.