

Рабочая программа дисциплины утверждена на заседании кафедры лингводидактики и переводоведения, протокол от «02» июня 2017 г. №10.

Заведующий кафедрой Мф / Морозкина Е.А.

Дополнения и изменения, внесенные в рабочую программу дисциплины, утверждены на заседании кафедры лингводидактики и переводоведения: обновлены ФОС, список литературы, программное обеспечение и базы данных, протокол № 18 от «23» мая 2018г.

Заведующий кафедрой Мф / Морозкина Е.А.

Дополнения и изменения, внесенные в рабочую программу дисциплины, утверждены на заседании кафедры лингводидактики и переводоведения, протокол № _____ от «____»

Заведующий кафедрой _____ / _____

Список документов и материалов

| | |
|--|----|
| 1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы | 4 |
| 2. Цель и место дисциплины в структуре образовательной программы | 5 |
| 3. Содержание рабочей программы (объем дисциплины, типы и виды учебных занятий, учебно-методическое обеспечение самостоятельной работы обучающихся) | 5 |
| 4. Фонд оценочных средств по дисциплине | 5 |
| 4.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания | 5 |
| 4.2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и опыта деятельности, характеризующих этапы формирования компетенций | 8 |
| 4.3. Рейтинг-план дисциплины | 9 |
| 5. Учебно-методическое и информационное обеспечение дисциплины | 15 |
| 5.1. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины | 15 |
| 5.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет» и программного обеспечения, необходимых для освоения дисциплины | 15 |
| 6. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине | 16 |

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

В результате освоения образовательной программы обучающийся должен овладеть следующими результатами обучения по дисциплине:

| Результаты обучения | | Формируемая компетенция (с указанием кода) | Примечание |
|--|---|---|------------|
| Знания | 1.знать основы математических дисциплин, которые используются при формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур: теории множеств, теории вероятностей и математической статистики, теории информации и кодирования, математической логики, математической теории грамматик; | ОПК-2 владение основами математических дисциплин, необходимых для формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур | |
| | 2. знать специфику применения ИТ при обработке текста и звучащей речи | ПК-2 владение основными методами инструментального анализа звучащей речи | |
| Умения | 1.уметь структурировать и моделировать базовые явления, относящиеся к сфере гуманитарных наук, с использованием математического аппарата; представлять в алгоритмическом виде процессы анализа и синтеза текста / дискурса; | ОПК-2 владение основами математических дисциплин, необходимых для формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур | |
| | 2.уметь автоматически обрабатывать тексты; | ПК-2 владение основными методами инструментального анализа звучащей речи | |
| Владения (навыки / опыт деятельности) | 1. владеть навыками применения основных лингвистических методик с целью решения исследовательских и прикладных задач современного языкознания – навыками анализа экспериментального материала и | ОПК-2 владение основами математических дисциплин, необходимых для формализации | |

| | | | |
|--|-------------------------------------|--|--|
| | обработки результатов эксперимента. | лингвистических знаний и процедур анализа и синтеза лингвистических структур | |
| | 2.владеть методикой тэггирования | ПК-2 владение основными методами инструментального анализа звучащей речи | |

2. Цель и место дисциплины в структуре образовательной программы

Дисциплина «Технологии обработки текста и звучащей речи» к относится к базовой части.

Дисциплина по выбору изучается на 2 курсе в 4 семестре.

Целью изучения дисциплины «Технологии обработки текста и звучащей речи» является формирование и развитие навыков математического и алгоритмического моделирования при изучении лингвистических феноменов и анализа лингвистических и алгоритмических возможностей таких моделей.

Освоение данной дисциплины может быть достигнуто при наличии у обучаемого уровня базовых знаний по информационным технологиям в лингвистике, сформированных навыков обработки естественного языка, навыков работы со средствами программирования.

3. Содержание рабочей программы (объем дисциплины, типы и виды учебных занятий, учебно-методическое обеспечение самостоятельной работы обучающихся)

Содержание рабочей программы представлено в Приложении № 1.

4. Фонд оценочных средств по дисциплине

4.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

(ОПК-2)- Владение основами математических дисциплин, необходимых для формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур.

| Этап (уровень) освоения компетенции | Планируемые результаты обучения (показатели и достижения заданного уровня освоения компетенций) | Критерии оценивания результатов обучения | | | |
|-------------------------------------|---|--|-------------------------------|--------------------------|----------------------------------|
| | | «неудовлетворительно» | «удовлетворительно» | «хорошо» | «отлично» |
| Первый этап (уровень) | Знать: основы математически | Обучающийся не знает основы | Обучающийся удовлетворительно | Обучающийся хорошо знает | Обучающийся отлично знает основы |

| | | | | | |
|-----------------------|---|--|--|---|---|
| | х дисциплин, которые используются при формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур: теории множеств, теории вероятностей и математической статистики, теории информации и кодирования, математической логики, математической теории грамматик; | математических дисциплин, которые используются при формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур | тельно знает основы математических дисциплин, которые используются при формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур | основы математических дисциплин, которые используются при формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур | математических дисциплин, которые используются при формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур |
| Второй этап (уровень) | Уметь: уметь структурировать и моделировать базовые явления, относящиеся к сфере гуманитарных наук, с использованием математического аппарата; представлять в алгоритмическом виде процессы анализа и синтеза текста / дискурса; | Обучающийся не умеет структурировать и моделировать базовые явления, относящиеся к сфере гуманитарных наук, с использованием математического аппарата; представляет в алгоритмическом виде процессы анализа и синтеза текста / дискурса; | Обучающийся удовлетворительно умеет структурировать и моделировать базовые явления, относящиеся к сфере гуманитарных наук, с использованием математического аппарата; представляет в алгоритмическом виде процессы анализа и синтеза | Обучающийся хорошо умеет структурировать и моделировать базовые явления, относящиеся к сфере гуманитарных наук, с использованием математического аппарата; представляет в алгоритмическом виде процессы анализа и синтеза | Обучающийся отлично умеет структурировать и моделировать базовые явления, относящиеся к сфере гуманитарных наук, с использованием математического аппарата; представлять в алгоритмическом виде процессы анализа и синтеза текста / дискурса; |

| | | | | | |
|-----------------------|---|--|---|--|---|
| | | | текста / дискурса; | текста / дискурса; | |
| Третий этап (уровень) | Владеть: владеть навыками применения основных лингвистических методик с целью решения исследовательских и прикладных задач современного языкознания – навыками анализа экспериментального материала и обработки результатов эксперимента. | Обучающийся не владеет навыками применения основных лингвистических методик с целью решения исследовательских и прикладных задач современного языкознания – навыками анализа экспериментального материала и обработки результатов эксперимента | Обучающийся удовлетворительно владеет навыками применения основных лингвистических методик с целью решения исследовательских и прикладных задач современного языкознания – навыками анализа экспериментального материала и обработки результатов эксперимента | Обучающийся хорошо владеет навыками применения основных лингвистических методик с целью решения исследовательских и прикладных задач современного языкознания – навыками анализа экспериментального материала и обработки результатов эксперимента | Обучающийся отлично владеет навыками применения основных лингвистических методик с целью решения исследовательских и прикладных задач современного языкознания – навыками анализа экспериментального материала и обработки результатов эксперимента |

ПК-2 – владением основными методами инструментального анализа звучащей речи

| Этап (уровень) освоения компетенции | Планируемые результаты обучения (показатели и достижения заданного уровня освоения компетенций) | Критерии оценивания результатов обучения | | | |
|-------------------------------------|---|--|---------------------|----------|-----------|
| | | «неудовлетворительно» | «удовлетворительно» | «хорошо» | «отлично» |
| | | | | | |

| | | | | | |
|-----------------------|---|---|--|---|--|
| Первый этап (уровень) | Знать: знать специфику применения ИТ при обработке текста и звучащей речи | Обучающийся не знает специфику применения ИТ при обработке текста и звучащей речи | Обучающийся удовлетворительно знает специфику применения ИТ при обработке текста и звучащей речи | Обучающийся хорошо знает специфику применения ИТ при обработке текста и звучащей речи | Обучающийся отлично знает специфику применения ИТ при обработке текста и звучащей речи |
| Второй этап (уровень) | Уметь: уметь автоматически обрабатывать тексты; | Обучающийся не умеет автоматически обрабатывать тексты; | Обучающийся удовлетворительно умеет автоматически обрабатывать тексты; | Обучающийся хорошо умеет автоматически обрабатывать тексты; | Обучающийся отлично умеет автоматически обрабатывать тексты; |
| Третий этап (уровень) | Владеть: владеть методикой тэггирования | Обучающийся не владеет методикой тэггирования | Обучающийся удовлетворительно владеет методикой тэггирования | Обучающийся хорошо владеет методикой тэггирования | Обучающийся отлично владеет методикой тэггирования |

Показатели сформированности компетенций.

Критериями оценивания являются баллы, которые выставляются преподавателем за виды деятельности (оценочные средства) по итогам изучения модулей (разделов дисциплины), перечисленных в рейтинг-плане дисциплины для экзамена: текущий контроль – максимум 40 баллов; рубежный контроль – максимум 30 баллов, поощрительные баллы – максимум 10).
Шкалы оценивания:

для экзамена:

от 45 до 59 баллов – «удовлетворительно»;

от 60 до 79 баллов – «хорошо»;

от 80 баллов – «отлично».

4.2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и опыта деятельности, характеризующих этапы формирования компетенций

| Этапы освоения | Результаты обучения | Компетенция | Оценочные средства |
|----------------|---------------------|-------------|--------------------|
|----------------|---------------------|-------------|--------------------|

| | | | |
|-----------------------|--|-------|---|
| 1-й этап Знания | 1. Знание основ математических дисциплин, которые используются при формализации лингвистических знаний и процедур анализа и синтеза лингвистических структур: теории множеств, теории вероятностей и математической статистики, теории информации и кодирования, математической логики, математической теории грамматик; | ОПК-2 | групповой опрос, индивидуальное задание |
| | 2. Знание специфики применения ИТ при обработке текста и звучащей речи | ПК-2 | групповой опрос, индивидуальное задание |
| 2-й этап Умения | 1. Умение структурировать и моделировать базовые явления, относящиеся к сфере гуманитарных наук, с использованием математического аппарата; представлять в алгоритмическом виде процессы анализа и синтеза текста / дискурса; | ОПК-2 | групповой опрос, индивидуальное задание |
| | 2. Умение автоматически обрабатывать тексты; | ПК-2 | групповой опрос, индивидуальное задание |
| 3-ий этап Владения | 1. Владение навыками применения основных лингвистических методик с целью решения исследовательских и прикладных задач современного языкознания – навыками анализа экспериментального материала и обработки результатов эксперимента. | ОПК-2 | письменная контрольная работа |
| | 2. Владение методикой тэггирования | ПК-2 | групповой опрос, индивидуальное задание |

4.3 Рейтинг-план дисциплины

Рейтинг-план дисциплины представлен в приложении 2.

Экзаменационные билеты

Структура экзаменационного билета


Экзаменационный билет включает в себя один теоретический и один практический вопросы.

Перечень примерных вопросов для экзамена:

1. PyCharm Edu Editor. Guided Tour
2. Managing Editor Tabs
3. Opening and Reopening Files in the Editor
4. Closing Files in the Editor
5. Saving and Reverting Changes
6. 'Lens' Mode

7. Multicursor
8. Scratches
9. Zooming in the Editor
10. TODO Comments
11. Basic Editing Procedures
12. Advanced Editing Procedures
13. PyCharm Edu Usage Guidelines
14. Creating and Managing Projects
15. Creating Projects
16. Opening, Reopening, and Closing Projects
17. Opening Multiple Projects
18. Cleaning System Cache
19. Configuring Project and IDE Settings
20. Accessing Settings
21. Configuring Keyboard Shortcuts
22. Configuring Menus and Toolbars
23. Configuring Quick Lists
24. Configuring Scopes and File Colors
25. Configuring Colors and Fonts
26. Configuring Code Style
27. Copying Code Style Settings
28. Configuring Individual File Encoding
29. Configuring Line Separators
30. Creating and Registering File Types
31. Switching Between Schemes
32. Configuring Project Structure
33. Accessing Project Structure
34. Configuring Content Roots
35. Configuring Folders Within a Content Root

Образец экзаменационного билета:

| | |
|---|--|
| <p>МИНОБРНАУКИ РОССИИ ФГБОУ ВО «БАШКИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ» ФАКУЛЬТЕТ РОМАНО-ГЕРМАНСКОЙ ФИЛОЛОГИИ</p> | |
| <p>Дисциплина «Технологии обработки текста и звучащей речи»</p> | |
| <p>ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 1</p> | |
| <p>УТВЕРЖДАЮ Зав. кафедрой лингводидактики и переводоведения</p> | |
| <p>д.ф.н., проф. Морозкина Е.А. </p> | |
| <p>«_____» _____ 2018 г.</p> | |
| <p>1. Спектральный анализ гласных и согласных:</p> | |
| <p>2. Поределите вероятность $P(not\ do)$ на основе представленной базы данных</p> | |
| <p><input type="checkbox"/> <s> I am Sam </s></p> | |
| <p><input type="checkbox"/> <s> Sam I am </s></p> | |
| <p><input type="checkbox"/> <s> I do not like green eggs and ham </s></p> | |

Перевод оценки из 100-балльной в четырехбалльную производится следующим образом:

- отлично – от 80 до 110 баллов (включая 10 поощрительных баллов);
- хорошо – от 60 до 79 баллов;
- удовлетворительно – от 45 до 59 баллов;
- неудовлетворительно – менее 45 баллов.

Критерии итоговой оценки знаний, умений и навыков обучающегося по дисциплине «Технологии обработки текста и звучащей речи»:

При определении требований к экзаменационным оценкам по дисциплинам с преобладанием теоретического обучения предлагается руководствоваться следующим:

– **25-30 баллов** выставляются студенту, если студент, обнаруживает всестороннее, систематическое и глубокое знание программного материала, умение свободно выполнять задания, предусмотренные РПД, усвоивший основную и знакомый с дополнительной литературой, рекомендованной программой. Как правило, такие баллы выставляются студентам, усвоившим взаимосвязь основных понятий дисциплины в их значении для приобретаемой профессии, проявившим творческие способности в понимании, изложении и использовании учебного материала;

– **17-24 баллов** выставляется студенту, если студент, обнаруживает полное знание программного материала, успешно выполняющий предусмотренные в программе задания, усвоивший основную литературу, рекомендованную в программе. Как правило, такие баллы выставляются студентам, показавшим систематический характер знаний по дисциплине и способным к их самостоятельному пополнению и обновлению в ходе дальнейшей учебной работы и профессиональной деятельности;

– **10-16 баллов** выставляется студенту, если студент, обнаруживает знание основного программного материала в объёме, необходимом для дальнейшей учёбы и предстоящей работы по профессии, справляющийся с выполнением заданий, предусмотренных программой, знакомый с основной литературой, рекомендованной программой. Как правило, такие баллы выставляются студентам, допустившим погрешности не принципиального характера в ответе на экзамене и при выполнении экзаменационных заданий;

– **1-10 баллов** выставляется студенту, если студент, обнаруживает пробелы в знаниях основного программного материала, допустившему принципиальные ошибки в выполнении предусмотренных программой заданий. Как правило, такие баллы выставляются студентам, которые не могут продолжить обучение или приступить к профессиональной деятельности по окончании вуза без дополнительных занятий по соответствующей дисциплине.

Вопросы для группового опроса (коррелируют с темами практических занятий)

1.1. Write ten grapheme-to-phoneme rules in a language you know. Take a short newspaper article and apply the rules manually using the longest-match algorithm. Measure the error rate.

1.2. Write a Prolog accepting grapheme-to-phoneme rules and using the longestmatch algorithm. Apply the program on a short text with rules you have written.

1.6 Further Reading 39 1.3. Take five sentences in a language that you know and, using four levels of pitch – 1, 2, 3, 4 – from low to high, try to annotate syllables of the sentences.

1.4. Write a search algorithm in Prolog decoding words from a sequence of phonemes. Use a brute-force search strategy. 1.5. Write a search algorithm decoding words from a sequence of phonemes. Incorporate a trigram language model and use a beam search strategy.

1.6. Using a speech API, program an interactive dialogue system based upon the kiosk grammar.

Для успешного выполнения тестовых заданий рекомендуется выполнение практических

заданий из Л.Р. Зиндер. Введение в языкознание: Сб. задач.М, 1998 (1987). раздел «Фонетика», «Морфонология» Калабина Н.С. Практикум по введению в языкознание. Разделы «Фонетика», «Письмо».

Критерии оценки (в баллах): 5 баллов за задание

- 4-5 баллов выставляется студенту, если студент продемонстрировал полные знания, умения, навыки;

- 1-3 баллов выставляется студенту, если студент в основном справился с поставленными целями и задачами и продемонстрировал неполные знания, умения, навыки;

-0- баллов выставляется студенту, если студент не продемонстрировал требуемые от него знания, умения, навыки;

Индивидуальные задания:

(а) На основе изучения нескольких Интернет-обзоров, содержащих сравнительный анализ возможностей различных поисковых систем (обзоры найти самим, в работе указать какие, т.е. название и адрес) написать обоснование Вашего выбора поисковика, который Вы используете постоянно (либо регулярно, либо чаще остальных). Для этого необходимо описать функциональные возможности данной поисковой системы и сравнить его с другими поисковыми системами по параметрам, представляющим для Вас особую значимость.

(б) На основе изучения информации о назначении и функциональных возможностях Британского национального корпуса (British National Corpus (BNC) [<http://www.natcorp.ox.ac.uk/>]) либо Национального корпуса русского языка [<http://www.ruscorpora.ru/>] подробно описать возможные варианты использования данного корпуса филологом (преподавателем иностранного языка (ИЯ), переводчиком) в процессе изучения/преподавания ИЯ, перевода или проведения исследований в области языка. Оба задания №1 и №2 выполняются письменно (набираются на компьютере в текстовом редакторе) и представляются преподавателю. Имя файла должно содержать следующие сведения, разделённые нижним пробелом (_): 1) информацию о модуле (Mod1); 2) информацию о группе (например, gr401); 3) фамилию студента, выполнившего, задание латинскими буквами. Пример имени файла с выполненным заданием: Mod1_gr401_Petrov.

Критерии оценки (в баллах): 5 баллов за задание

- 4-5 баллов выставляется студенту, если студент продемонстрировал полные знания, умения, навыки;

- 1-3 баллов выставляется студенту, если студент в основном справился с поставленными целями и задачами и продемонстрировал неполные знания, умения, навыки;

-0- баллов выставляется студенту, если студент не продемонстрировал требуемые от него знания, умения, навыки;

Пример рубежной письменной контрольной работы

1. . Suppose "the" is the most frequent word in our corpus and occurs 200,000 times. The second most frequent word is "of", which occurs 190,000 times. Given just what we have explained so far, what would be the Good-Turing estimated count of "the"?

2. Suppose we train a language model with add-one smoothing on a given corpus C . How does the perplexity of the smoothed language model on that corpus, $PP_{smoothed}(C)$, compare with the perplexity of the maximum-likelihood language model, $PPML(C)$?

1. $PP_{smoothed}(C) = PPML(C)$
2. $PP_{smoothed}(C) \leq PPML(C)$
3. $PPML(C) \leq PP_{smoothed}(C)$
4. The answer depends on the corpus.

3. A traffic signal has three colors: green, yellow, and red, which appear the following probabilities. Using a unigram model, what is the perplexity of the sequence (green, yellow, red)?

- $P(\text{green}) = 2/5$
- $P(\text{yellow}) = 1/5$
- $P(\text{red}) = 2/5$
- $(2/5 \cdot 2/5 \cdot 1/5)^{1/3}$
- $(2/5 \cdot 2/5 \cdot 1/5) - 1/3$
- $(5/2 \cdot 5/2 \cdot 5/1) - 1/3$
- 3

4. If we estimate a bigram language model from the following corpus, what is $P(\text{not}|\text{do})$?

- $\langle s \rangle$ I am Sam $\langle /s \rangle$
- $\langle s \rangle$ Sam I am $\langle /s \rangle$
- $\langle s \rangle$ I do not like green eggs and ham $\langle /s \rangle$
- $2/3$
- $1/3$
- 1
- $1/2$

5. Which is assigned higher probability by a unigram language model for English?

- $P(\text{I like ice cream})$
- $P(\text{the thethe the})$
- $P(\text{Go to class daily})$
- $P(\text{class daily go to})$

6. Compute min edit distance.

7. What is the value of the cell labeled X in the following partially-completed Levenshtein distance table (substitutions cost 2):

| | | | | | |
|--|--|--|--|--|--|
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

- 5
- 4
- 3
- 2

8. A period (".") occurs at the end of four words that each have one of the following four wordshapes. Which of the four periods (all else being equal) is more likely than the other to be a sentence boundary?

- Upper
- Lower
- Cap
- Number

9. The Porter stemmer would stem the word 'aching' as

- aching
- ach
- ache
- aches

10. How many wordform types and tokens are in this sentence:

I want to cook a dish that I never cooked before

- 10 types 11 tokens
- 11 types 10 tokens
- 11 types 11 tokens
- 10 types 10 tokens

11. We saw that the words 'd' and 's' appear frequently in the word tokenization results. Which of following reasons might explain this?

- The words "s" and "d" are archaic words that were more popular during the time that Shakespeare was writing.
- The words "s" and "d" were not very common during the time that Shakespeare was writing, but they are among the many words which Shakespeare appears to have made up himself.
- The tokenizer splits on all non-alphabetical characters, which includes the apostrophe. This mean that words like "Ophelia's" and "'dimm'd" will be tokenized to "Ophelia s" and "dimm d".
- The Tokenizer splits words into smaller subparts named morphemes. "s" and "d" are two very common such morphemes because they tend to occur at the end of words.

Критерии оценки (в баллах): 15 баллов за тест

-11- 15 баллов выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы модуля и умение применять их на практике при решении конкретных задач, свободное и правильное обоснование проблемных ситуаций;

- 6-10 баллов выставляется студенту, показавшему не достаточно глубокое знание теории межкультурной коммуникации, не умеющему в полной мере аргументировано обосновать решение конкретных задач;

-0-5 баллов выставляется студенту, который не знает большей части основного содержания учебной программы модуля, допускает грубые ошибки в формулировках основных понятий дисциплины и не умеет использовать полученные знания при решении типовых практических задач.

5. Учебно-методическое и информационное обеспечение дисциплины

5.1. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Основная литература:

1. Захарова, Т.В. Практические основы компьютерных технологий в переводе : учебное пособие / Т.В. Захарова, Е.В. Турлова ; Министерство образования и науки Российской Федерации, Оренбургский Государственный Университет. - Оренбург : Оренбургский государственный университет, 2017. - 109 с. : табл., граф., ил. - Библиогр.: с. 104. - ISBN 978-5-7410-1736-4 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=481823> (19.11.2018).
2. Михеева, Е.В. Информационные технологии в профессиональной деятельности : учебное пособие / Е.В. Михеева. - Москва : Проспект, 2014. - 448 с. - Библиогр. в кн. - ISBN 978-5-392-12318-6 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=251602> (19.11.2018).

Дополнительная литература

3. Бовтенко, М.А. Язык пользователя персонального компьютера: учебное пособие / М.А. Бовтенко, Е.В. Кугаевская. – Новосибирск: НГТУ, 2011. – Ч. 2. – 75 с. – ISBN 978-5-7782-1873-4 – ЭВИ в ЭБС «Универс. библио. онлайн» – URL: <http://biblioclub.ru/index.php?page=book&id=228749>
4. Бойченко, Г.Н. Информационные сервисы Интернет в профессиональной деятельности педагога: учебное пособие / Г.Н. Бойченко. – Новокузнецк: Кузбасская государственная педагогическая академия, 2008. – 106 с. – ISBN 978-5-85117-320-2 – ЭВИ в ЭБС «Универс. библио. онлайн» – URL: <http://biblioclub.ru/index.php?page=book&id=88672>
5. Гафурова, Н.В. Методика обучения информационным технологиям. Теоретические основы: учебное пособие / Н.В. Гафурова, Е.Ю. Чурилова. – Красноярск: СибФУ, 2012. – 111 с. – ISBN 978-5-7638-2234-2 – ЭВИ в ЭБС «Универс. библио. онлайн» – URL: <http://biblioclub.ru/index.php?page=book&id=229302> .
6. Гунина, Н.А. Компьютер для работы и досуга = Computer for Work and Leisure: учебное пособие / Н.А. Гунина, И.В. Шеленкова, А.А. Шиповская. – Тамбов: Издательство ФГБОУ ВПО «ТГТУ», 2012. – 137 с. – ЭВИ в ЭБС «Универс. библио. онлайн» – URL: <http://biblioclub.ru/index.php?page=book&id=277672>

5.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет» и программного обеспечения, необходимых для освоения дисциплины

7. Научно-образовательный портал «Лингвистика в России: ресурсы для исследователей»: http://uisrussia.msu.ru/linguist/_B7_komp_tehn_v_prepodavanii_jazykov.jsp
8. Универсальная научно-популярная онлайн-энциклопедия. http://www.krugosvet.ru/enc/gumanitarnye_nauki/lingvistika/GERMANSKIE_YAZIKI.html
9. Свободная энциклопедия Википедия.: http://ru.wikipedia.org/wiki/Германские_языки
10. Библиотека Русского гуманитарного интернет-университета: <http://www.i-u.ru/biblio>
11. Лингвоинфо: интернет-журнал: <http://www.lingvoinfo.com>
12. Онлайн-энциклопедия <http://www.krugosvet.ru>
13. Русский филологический портал Philology.ru.: <http://philology.ru/linguistics1.htm1>
14. Центр лингвистической документации: <http://www.mccme.ru/ling/index.htm>

15. http://www.nyu.edu/projects/politicsdatalab/localdata/workshops/NLTK_Presentation.pdf
 16. Windows 8 Russian. Windows Professional 8 Russian Upgrade.
 17. Microsoft Office Standard 2013 Russian.

6. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

Компьютерный класс, мультимедийный проектор, экран, доска, оргтехника, аудиоаппаратура (в стандартной комплектации для лекционных занятий и самостоятельной работы студентов); доступ к сети Интернет (во время самостоятельной подготовки и на лекционных занятиях).

| Наименование специализированных аудиторий, кабинетов, лабораторий | Вид занятий | Наименование оборудования, программного обеспечения |
|---|--|---|
| 1 | 2 | 3 |
| <p>1. Учебная аудитория для проведения занятий семинарского типа: аудитория №10 (Лаборатория синхронного перевода, компьютерный класс), аудитория №15 (Лаборатория моделирования естественных языков), аудитория №04 (Лингафонный кабинет) (ул. Коммунистическая, д. 19, лит. А, А1)</p> | <p>Семинарские занятия</p> | <p>Аудитория № 15 Учебная мебель, учебно-наглядные пособия, доска, учебно-методическая литература, многофункциональное устройство Canon – 1 шт.; персональный компьютер в комплекте: системный блок HP Pavilion Slimline S3500F AMD Athlon 64X2 5400+/2.8 GHz, 4Gb, монитор LG 19 L1942S SF 1280*1024, 5ms, 8000:1, black, клавиатура, мышь; МФУ, ноутбуки, 4 компьютера.</p> <p>Аудитория №10 <i>Оборудование для синхронного перевода:</i> Процессор управления конференций до 120 пультов Televic Confidea; многофункциональные микрофонные пульты переводчика Televic ID 2500D – 5 шт; Стандартный (базовый) пульт делегата Televic L-DD – 1 шт; Блок расширения на 8 аналоговых выходов Televic AOP2500;. Инфракрасный 8-канальный делегатский приемник Televic R8; Инфракрасный 8-канальный передатчик Televic T8; Компьютер в составе: Корпус Cooler Master, Процессор Intel Core i5/2*2Гб/ GigaByte/Накопитель Seagate/ Видеоадаптер Zotac / Оптический привод Optiarc AD, Клавиатура Genius, Оптический манипулятор Genius 130X, Монитор Viewsonic VG2239M 2.</p> <p><i>Проекционное (демонстрационное) оборудование:</i> Проектор NEC M402W (M402WG). Full3D.DLP.4000 ANSI</p> |
| <p>2. Учебная аудитория для проведения групповых и индивидуальных консультаций, учебная аудитория для текущего контроля и промежуточной аттестации: аудитория №15 (Лаборатория моделирования естественных языков), аудитория №10 (Лаборатория синхронного перевода, компьютерный класс), аудитория №04 (Лингафонный кабинет) (ул. Коммунистическая, д.</p> | <p>Групповые и индивидуальные консультации, промежуточная аттестация</p> | <p>120 пультов Televic Confidea; многофункциональные микрофонные пульты переводчика Televic ID 2500D – 5 шт; Стандартный (базовый) пульт делегата Televic L-DD – 1 шт; Блок расширения на 8 аналоговых выходов Televic AOP2500;. Инфракрасный 8-канальный делегатский приемник Televic R8; Инфракрасный 8-канальный передатчик Televic T8; Компьютер в составе: Корпус Cooler Master, Процессор Intel Core i5/2*2Гб/ GigaByte/Накопитель Seagate/ Видеоадаптер Zotac / Оптический привод Optiarc AD, Клавиатура Genius, Оптический манипулятор Genius 130X, Монитор Viewsonic VG2239M 2.</p> <p><i>Проекционное (демонстрационное) оборудование:</i> Проектор NEC M402W (M402WG). Full3D.DLP.4000 ANSI</p> |

| | | |
|--|------------------------|---|
| 19, лит. А, А1) | | Lm.WXGA; Экран с электроприводом Lumien Master Control 129x200см (88"); Интерактивная доска SmartBoard 680, 77"/105,6 |
| 3. Помещения для самостоятельной работы: аудитория №13 (читальный зал) (ул. Коммунистическая, д. 19, лит. А, А1) | Самостоятельная работа | <p>Персональные компьютеры с выходом в сеть Интернет (10 шт.) и обеспечивающие доступ к электронной информационно-образовательной среде (ЭИОС) вуза.</p> <p>Аудитория №04</p> <p>Учебная мебель, учебно-наглядные пособия, доска, моноблоки – 12 шт. с выходом в Интернет, обеспечивающие доступ к электронной информационно-образовательной среде (ЭИОС) вуза</p> <p>Аудитория №13</p> <p>Учебная мебель, учебно-наглядные пособия, доска, учебно-методическая литература, многофункциональное устройство – 1 шт., моноблоки – 2 шт. с выходом в Интернет, обеспечивающие доступ к электронной информационно-образовательной среде (ЭИОС) вуза, книжный фонд читального зала ФРГФ</p> <p>1. Windows 8 Russian. Windows Professional 8 Russian Upgrade. Договор № 104 от 17.06.2013 г. Лицензии бессрочные</p> <p>2. Microsoft Office Standard 2013 Russian. Договор № 114 от 12.11.2014 г. Лицензии бессрочные</p> <p>3. Adobe Audition CS6 5 AcademicEdition License International English MultiplePlatforms № 228 от 06.11.2012г</p> |

Приложение №1

МИНОБРНАУКИ РОССИИ
 ФГБОУ ВО «БАШКИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
 ФАКУЛЬТЕТ РОМАНО-ГЕРМАНСКОЙ ФИЛОЛОГИИ

СОДЕРЖАНИЕ РАБОЧЕЙ ПРОГРАММЫ

дисциплины «Технологии обработки текста и звучащей речи»
 на 4 семестр
 форма обучения очная

| Вид работы | Объем дисциплины |
|--|------------------|
| Общая трудоемкость дисциплины (ЗЕТ / часов) | 3/108 |
| Учебных часов на контактную работу с преподавателем: | 33,2 |

| | |
|---|------|
| лекций | |
| практических/ семинарских | 32 |
| лабораторных | |
| других (групповая, индивидуальная консультация и иные виды учебной деятельности, предусматривающие работу обучающихся с преподавателем) (ФКР) | 1,2 |
| Учебных часов на самостоятельную работу обучающихся (СР) | 11,8 |
| Учебных часов на подготовку к экзамену/зачету/дифференцированному зачету (Контроль) | 63 |

Форма контроля: экзамен: 4 семестр

| № п/п | Тема и содержание | Форма изучения материалов: лекции, практические занятия, семинарские занятия, лабораторные работы, самостоятельная работа и трудоемкость (в часах) | | | | Основная и дополнительная литература, рекомендуемая студентам (номера из списка) | Задания по самостоятельной работе студентов | Форма текущего контроля успеваемости (коллоквиумы, контрольные работы, компьютерные тесты и т.п.) |
|----------|--|--|------------|----|-----|--|--|---|
| | | ЛК | ПР/ СЕМ | ЛР | СР | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1. | <p>Модуль 1:</p> <p>Тема 1: Объект, цель и методы исследования дисциплины. Уровни восприятия звука. Отбор и оцифровка сигналов.</p> <p>Тема 2: Фонемы, фоны, аллофоны, форманты, фреймы. Спектральный анализ гласных и согласных.</p> <p>Тема 3: Правила преобразования графем в фонемы на основе алгоритма самого длинного совпадения. Измерение коэффициента ошибок. Абсолютные и относительные значения интенсивности</p> | | 16 | | 10 | <p>Основная литература: 1,2</p> <p>Дополнительная литература: 1,3</p> | <p>Подготовка к компьютерному тестированию, реферат</p> <p>а) 1,2,3 б) 1,2,6</p> | <p>групповой опрос, индивидуальное задание, письменная контрольная работа</p> |
| 2. | <p>Модуль 2:</p> <p>Тема 4: Быстрое преобразование Фурье. Частотный интервал. Преобразования Фурье для некоторых функций.</p> <p>Тема 5: Обработка текста для синтеза речи. Токенизация. Правила преобразования графем</p> | | 16 | | 1,8 | <p>Основная литература: 1,2</p> <p>Дополнительная литература: 1, 3</p> | <p>Подготовка к компьютерному тестированию, реферат, доклад</p> <p>а) 1,2,3 б) 1,2,5,6</p> | <p>групповой опрос, индивидуальное задание, письменная контрольная работа</p> |

| | | | | | | | | |
|--|--|--|----|--|------|--|--|--|
| | <p>в фонемы Тема 6 Частотный диапазон формантов гласных. Восприятие просодических характеристик речи Тема 7 Программные интерфейсы для обработки звука Тема 8. Восприятие просодических характеристик речи</p> | | | | | | | |
| | Всего часов: | | 32 | | 11,8 | | | |

Рейтинг-план дисциплины

| Виды учебной деятельности студентов | Балл за конкретное задание | Число заданий за семестр | Баллы | |
|---|----------------------------|--------------------------|-------------|--------------|
| | | | Минимальный | Максимальный |
| Модуль 1 | | | | |
| Моделирование фонов по скрытой модели Маркова | | | | |
| Текущий контроль | | | | |
| 1. Аудиторная работа (групповой опрос) | 5 | 2 | 0 | 10 |
| 2. Выполнение индивидуальных заданий | 5 | 2 | 0 | 10 |
| Рубежный контроль | | | | |
| 1. письменная контрольная работа | 15 | | 0 | 15 |
| Итоговый балл за модуль 1 | | | 0 | 35 |
| Модуль 2 | | | | |
| Реферативная работа с различными информационными структурами. Презентации рефератов. | | | | |
| Текущий контроль | | | | |
| 1. Аудиторная работа (групповой опрос) | 5 | 2 | 0 | 10 |
| 2. Выполнение индивидуальных заданий | 5 | 2 | 0 | 10 |
| Рубежный контроль | | | | |
| 1. письменная контрольная работа | 15 | | 0 | 15 |
| Итоговый балл за модуль 2 | | | 0 | 35 |
| Поощрительные баллы | | | | |
| Участие в научных, культурных и спортивных мероприятиях факультета, написание научных статей, призовые места на языковых конкурсах и т.д. | | | 0 | 10 |
| Посещаемость (баллы вычитаются из общей суммы набранных баллов) | | | | |
| Посещение лекционных занятий | | | | -6 |
| Посещение лабораторных занятий | | | | -10 |
| Итоговый контроль | | | | |
| Экзамен | | | | 30 |
| Всего | | | | 110 |