

ФГБОУ ВО «БАШКИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

ФАКУЛЬТЕТ РОМАНО-ГЕРМАНСКОЙ ФИЛОЛОГИИ

Утверждено:

на заседании кафедры

протокол № 7 от «27» января 2021 г.

Зав. кафедрой М.А. /Морозкина Е.А.

Согласовано:

Председатель УМК факультета /института

Л.К. /Мазунова Л.К.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

дисциплина Программирование для лингвистов

(наименование дисциплины)

Часть формируемая участниками образовательных отношений

Дисциплина по выбору

(указать часть (обязательная часть или часть, формируемая участниками образовательных отношений, факультатив))

Программа бакалавриата

Направление подготовки

45.03.03 Фундаментальная и прикладная лингвистика

(указывается код и наименование направления подготовки (специальности))

Направленность (профиль) подготовки

Языковые технологии

(указывается наименование направленности (профиля) подготовки)

Квалификация

Бакалавр

(указывается квалификация)

Разработчик (составитель) <u>доц. к. филол.н.</u> (должность, ученая степень, ученое звание)	<u>Р.Г.</u> / Мифтахова Р.Г. (подпись, Фамилия И.О.)
--	---

Для приема: 2021 г.

Уфа 2021 г.

Составитель / составители: Мифтахова Р.Г.

Рабочая программа дисциплины *утверждена* на заседании кафедры протокол от «27»
января 2021 г. № 7

Заведующий кафедрой

_____ / Морозкина Е.А.

Список документов и материалов

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с установленными в образовательной программе индикаторами достижения компетенций
2. Цель и место дисциплины в структуре образовательной программы
3. Содержание рабочей программы (объем дисциплины, типы и виды учебных занятий, учебно-методическое обеспечение самостоятельной работы обучающихся)
4. Фонд оценочных средств по дисциплине
 - 4.1. Перечень компетенций и индикаторов достижения компетенций с указанием соотнесенных с ними запланированных результатов обучения по дисциплине. Описание критериев и шкал оценивания результатов обучения по дисциплине.
 - 4.2. Типовые контрольные задания или иные материалы, необходимые для оценивания результатов обучения по дисциплине, соотнесенных с установленными в образовательной программе индикаторами достижения компетенций. Методические материалы, определяющие процедуры оценивания результатов обучения по дисциплине.
5. Учебно-методическое и информационное обеспечение дисциплины
 - 5.1. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины
 - 5.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет» и программного обеспечения, необходимых для освоения дисциплины, включая профессиональные базы данных и информационные справочные системы
6. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с установленными в образовательной программе индикаторами достижения компетенций

По итогам освоения дисциплины обучающийся должен достичь следующих результатов обучения:

Категория (группа) компетенций¹ (при наличии ОПК)	Формируемая компетенция (с указанием кода)	Код и наименование индикатора достижения компетенции	Результаты обучения по дисциплине
	ПК-9 - Способен пользоваться лингвистически ориентированными программными продуктами	<i>ПК-9</i> Знает типы, характеристики и особенности основных доступных в Интернете лингвистических ресурсов.	<i>Знание</i> типов, характеристики и особенности основных доступных в Интернете лингвистических ресурсов.
		<i>ПК-9</i> Умеет сравнивать данные, полученные с использованием различных электронных лингвистических ресурсов и систем; применять методы математического анализа и моделирования в профессиональной деятельности	<i>Умение</i> сравнивать данные, полученные с использованием различных электронных лингвистических ресурсов и систем; применять методы математического анализа и моделирования в профессиональной деятельности
		<i>ПК-9</i> Владеет практическим опытом тестирования электронных лингвистических ресурсов, систем и компонентов	<i>Владение</i> практическим опытом тестирования электронных лингвистических ресурсов, систем и компонентов
	ПК-12 - Способен проводить квалифицированное тестирование	<i>ПК-12</i> Знает различные способы и инструментарий тестирования	<i>Знание</i> различных способов и инструментарий тестирования

¹ Указывается только для УК и ОПК (при наличии).

	<p>лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем</p>	<p>лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем</p>	<p>лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем</p>
		<p><i>ПК-12 Умеет</i> квалифицированно тестировать лингвистически ориентированные программные продукты, электронные ресурсы, лингвистически ориентированные системы и лингвистические компоненты интеллектуальных и информационных электронных систем.</p>	<p><i>Умение</i> квалифицированно тестировать лингвистически ориентированные программные продукты, электронные ресурсы, лингвистически ориентированные системы и лингвистические компоненты интеллектуальных и информационных электронных систем.</p>
		<p><i>ПК-12 Владеет</i> навыками, необходимыми для тестирования различных ресурсов и продуктов электронной среды, в том числе лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем</p>	<p><i>Владение</i> навыками, необходимыми для тестирования различных ресурсов и продуктов электронной среды, в том числе лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и</p>

			информационных электронных систем.
--	--	--	------------------------------------

2. Цель и место дисциплины в структуре образовательной программы

Дисциплина «Программирование для лингвистов» относится к обязательной части. Дисциплина изучается на 2 курсе(ах) в 3 семестре(ах).

Целью обучения программированию для лингвистов как виду профессиональной деятельности являются знакомство с инструментарием языка Python, базовых понятий, структур данных и алгоритмов, необходимых для изучения прикладных методов количественной лингвистики, а также для самостоятельного решения некоторых задач компьютерной лингвистики. Для освоения дисциплины необходимы компетенции, сформированные в рамках изучения следующих дисциплин «Компьютерные методы обработки языковой информации», «Иностранный язык», «Информационные технологии в лингвистике», «Лингвистические компоненты информационных электронных систем» и пр.

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 и индикаторов достижения компетенций с указанием соотнесенных с ними запланированных результатов обучения по дисциплине. Описание критериев и шкал оценивания результатов обучения по дисциплине.

Код и формулировка компетенции: ПК-9 Способен пользоваться лингвистически ориентированными программными продуктами;

Код и наименование индикатора достижения компетенции	Результаты обучения по дисциплине	Критерии оценивания результатов обучения	
		«Зачтено»	«Не зачтено»
ПК-9.1. Знает: типы, характеристики и особенности основных доступных в Интернете лингвистических ресурсов.	Знать: типы, характеристики и особенности основных доступных в Интернете лингвистических ресурсов	Обучающийся знает типы, характеристики и особенности основных доступных в Интернете лингвистических ресурсов;	Обучающийся не знает типы, характеристики и особенности основных доступных в Интернете лингвистических ресурсов;

ПК-9.2. Умеет: сравнивать данные, полученные с использованием различных электронных лингвистических ресурсов и систем; применять методы математического анализа и моделирования в профессиональной деятельности.	<i>Уметь:</i> сравнивать данные, полученные с использованием различных электронных лингвистических ресурсов и систем; применять методы математического анализа и моделирования в профессиональной деятельности	Обучающийся умеет сравнивать данные, полученные с использованием различных электронных лингвистических ресурсов и систем; применять методы математического анализа и моделирования в профессиональной деятельности	Обучающийся не умеет сравнивать данные, полученные с использованием различных электронных лингвистических ресурсов и систем; применять методы математического анализа и моделирования в профессиональной деятельности
ПК-9.3. Имеет практический опыт тестирования электронных лингвистических ресурсов, систем и компонентов	<i>Владеть:</i> практическим опытом тестирования электронных лингвистических ресурсов, систем и компонентов	Обучающийся владеет практическим опытом тестирования электронных лингвистических ресурсов, систем и компонентов	Обучающийся не владеет практическим опытом тестирования электронных лингвистических ресурсов, систем и компонентов

Код и формулировка компетенции: *ПК-12* Способен проводить квалифицированное тестирование лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем

Код и наименование индикатора достижения компетенции	Результаты обучения по дисциплине	Критерии оценивания результатов обучения	
		«Зачтено»	«Не зачтено»
ПК-12.1 Знает: различные способы и инструментарий тестирования лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем	Знать: различные способы и инструментарий тестирования лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем	Обучающийся знает различные способы и инструментарий тестирования лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем	Обучающийся не знает способы и инструментарий тестирования лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем

<p>ПК-12.2 Умеет: квалифицированно тестировать лингвистически ориентированные программные продукты, электронные ресурсы, лингвистически ориентированные системы и лингвистические компоненты интеллектуальных и информационных электронных систем</p>	<p><i>Уметь:</i> квалифицированно тестировать лингвистически ориентированные программные продукты, электронные ресурсы, лингвистически ориентированные системы и лингвистические компоненты интеллектуальных и информационных электронных систем</p>	<p>Обучающийся умеет квалифицированно тестировать лингвистически ориентированные программные продукты, электронные ресурсы, лингвистически ориентированные системы и лингвистические компоненты интеллектуальных и информационных электронных систем</p>	<p>Обучающийся не умеет квалифицированно тестировать лингвистически ориентированные программные продукты, электронные ресурсы, лингвистически ориентированные системы и лингвистические компоненты интеллектуальных и информационных электронных систем</p>
<p>ПК-12.3 Владеет: навыками, необходимыми для тестирования различных ресурсов и продуктов электронной среды, в том числе лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем</p>	<p><i>Владеть:</i> навыками, необходимыми для тестирования различных ресурсов и продуктов электронной среды, в том числе лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем</p>	<p>Обучающийся владеет навыками, необходимыми для тестирования различных ресурсов и продуктов электронной среды, в том числе лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем</p>	<p>Обучающийся не владеет навыками, необходимыми для тестирования различных ресурсов и продуктов электронной среды, в том числе лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем</p>

4.2. Типовые контрольные задания или иные материалы, необходимые для оценивания результатов обучения по дисциплине, соотнесенных с установленными в образовательной программе индикаторами достижения компетенций. Методические материалы, определяющие процедуры оценивания результатов обучения по дисциплине.

Код и наименование индикатора достижения компетенции	Результаты обучения по дисциплине	Оценочные средства
ПК-9.1. Знает: типы, характеристики и особенности основных доступных в Интернете лингвистических ресурсов.	<i>Знать</i> : типы, характеристики и особенности основных доступных в Интернете лингвистических ресурсов.	групповой опрос, индивидуальное задание, контрольная работа
ПК-9.2. Умеет: сравнивать данные, полученные с использованием различных электронных лингвистических ресурсов и систем; применять методы математического анализа и моделирования в профессиональной деятельности.	<i>Уметь</i> сравнивать данные, полученные с использованием различных электронных лингвистических ресурсов и систем; применять методы математического анализа и моделирования в профессиональной деятельности.	индивидуальное задание контрольная работа
ПК-9.3. Имеет практический опыт тестирования электронных лингвистических ресурсов, систем и компонентов	<i>Владеть</i> практическим опытом тестирования электронных лингвистических ресурсов, систем и компонентов.	индивидуальное задание контрольная работа
ПК-12.1 Знает: различные способы и инструментарий тестирования лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем	<i>Знать</i> различные способы и инструментарий тестирования лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем.	групповой опрос, индивидуальное задание.
ПК-12.2 Умеет: квалифицированно тестировать лингвистически	<i>Уметь</i> квалифицированно тестировать лингвистически ориентированные программные продукты,	Групповой опрос, индивидуальное задание, контрольная работа

<p>ориентированные программные продукты, электронные ресурсы, лингвистически ориентированные системы и лингвистические компоненты интеллектуальных и информационных электронных систем</p>	<p>электронные ресурсы, лингвистически ориентированные системы и лингвистические компоненты интеллектуальных и информационных электронных систем.</p>	
<p>ПК-12.3 Владеет: навыками, необходимыми для тестирования различных ресурсов и продуктов электронной среды, в том числе лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем</p>	<p><i>Владеть</i> навыками, необходимыми для тестирования различных ресурсов и продуктов электронной среды, в том числе лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем.</p>	<p>Групповой опрос, индивидуальное задание, контрольная работа</p>

Критериями оценивания при модульно-рейтинговой системе являются баллы, которые выставляются преподавателем за виды деятельности (оценочные средства) по итогам изучения модулей (разделов дисциплины), перечисленных в рейтинг-плане дисциплины (текущий контроль – максимум 50 баллов; рубежный контроль – максимум 50 баллов, поощрительные баллы – максимум 10). Шкалы оценивания: зачтено – от 60 до 110 рейтинговых баллов (включая 10 поощрительных баллов), не зачтено – от 0 до 59 рейтинговых баллов.

Рейтинг – план дисциплины (при необходимости)

Программирование для лингвистов

(название дисциплины согласно рабочему учебному плану)

направление/специальность _____ 45.03.03 Фундаментальная и прикладная лингвистика ____
курс _____ 2 _____, семестр _____ 3 _____

Виды учебной деятельности студентов	Балл за конкретное задание	Число заданий за семестр	Баллы	
			Минимальный	Максимальный
Модуль 1 Основы обработки естественного языка на языке программирования Питон				
Текущий контроль				
1. Аудиторная работа (групповой опрос)	10	1	0	10
2. Индивидуальное задание	10	1	0	10
Рубежный контроль				
1. Контрольная работа	10	1	0	10
Итого				30
Модуль 2 Загрузка библиотек, анализ корпусных данных.				
Текущий контроль				
1. Аудиторная работа (групповой опрос)	10	1	0	10
2. Индивидуальное задание	10	1	0	10
Рубежный контроль				
1. Индивидуальное задание	20	2	0	20
				40
Модуль 3 Работа с модулем условного частотного распределения.				
Текущий контроль				
1. Аудиторная работа (групповой опрос)	10	1	0	10
2. Индивидуальное задание	10	1	0	10
1. Аудиторная работа (групповой опрос)				
Рубежный контроль				
1. Контрольная работа	10	2	0	10
				30
Поощрительные баллы				
1. Участие в научных конференциях и фестивалях, культурных мероприятиях, публикация статей				3
2. Публикация статей				5
3. Работа со школьниками (кружок, конкурсы, олимпиады)				2
4 ...				
Посещаемость (баллы вычитаются из общей суммы набранных баллов)				

1. Посещение практических занятий			0	-6
Итоговый контроль				
Всего				110
1. Зачет				-

1. Задания для группового опроса.

По модулю 1 Основы обработки естественного языка на языке программирования Питон

1. Create a variable phrase containing a list of words. Experiment with the operations described in this chapter, including addition, multiplication, indexing, slicing, and sorting.
2. Use the corpus module to explore austen-persuasion.txt. How many word tokens does this book have? How many word types?
3. Use the Brown Corpus reader `nlk.corpus.brown.words()` or the Web Text Corpus reader `nlk.corpus.webtext.words()` to access some sample text in two different genres.
4. Read in the texts of the State of the Union addresses, using the `state_union` corpus reader. Count occurrences of men, women, and people in each document. What has happened to the usage of these words over time?
- 5 The CMU Pronouncing Dictionary contains multiple pronunciations for certain words. How many distinct words does it contain? What fraction of words in this dictionary have more than one possible pronunciation?
6. What percentage of noun synsets have no hyponyms? You can get all noun synsets using `wn.all_synsets('n')`.
7. Define a function `supergloss(s)` that takes a synset `s` as its argument and returns a string consisting of the concatenation of the definition of `s`, and the definitions of all the hypernyms and hyponyms of `s`.
8. Write a program to find all words that occur at least three times in the Brown Corpus.
9. Write a function that finds the 50 most frequently occurring words of a text that are not stopwords.
10. Write a program to print the 50 most frequent bigrams (pairs of adjacent words) of a text, omitting bigrams that contain stopwords.
11. Write a program to create a table of word frequencies by genre, like the one given in Section 2.1 for modals. Choose your own words and try to find words whose presence (or absence) is typical of a genre. Discuss your findings.
12. Write a function `word_freq()` that takes a word and the name of a section of the Brown Corpus as arguments, and computes the frequency of the word in that section of the corpus.
13. Write a program to guess the number of syllables contained in a text, making use of the CMU Pronouncing Dictionary.
14. Define a function `hedge(text)` that processes a text and produces a new version with the word 'like' between every third word

По модулю 2 Загрузка библиотек, анализ корпусных данных.

Some strings that will be discussed:

A text corpus is a large, structured collection of texts. NLTK comes with many corpora, e.g., the Brown Corpus, `nlk.corpus.brown`.

- Some text corpora are categorized, e.g., by genre or topic; sometimes the categories of a corpus overlap each other.
- A conditional frequency distribution is a collection of frequency distributions, each one for a different condition. They can be used for counting word frequencies, given a context or a genre.

- Python programs more than a few lines long should be entered using a text editor, saved to a file with a .py extension, and accessed using an import statement.
- Python functions permit you to associate a name with a particular block of code, and reuse that code as often as necessary.
- Some functions, known as “methods,” are associated with an object, and we give the object name followed by a period followed by the method name, like this: x.func(y), e.g., word.isalpha().
- To find out about some variable v, type help(v) in the Python interactive interpreter to read the help entry for this kind of object.
- WordNet is a semantically oriented dictionary of English, consisting of synonym sets—or synsets—and organized into a network.
- Some functions are not available by default, but must be accessed using Python’s import statement.

По модулю 3 Работа с модулем условного частотного распределения.

17. Write a function that finds the 50 most frequently occurring words of a text that are not stopwords.
18. Write a program to print the 50 most frequent bigrams (pairs of adjacent words) of a text, omitting bigrams that contain stopwords.
19. Write a program to create a table of word frequencies by genre, like the one given in Section 2.1 for modals. Choose your own words and try to find words whose presence (or absence) is typical of a genre. Discuss your findings.
20. Write a function word_freq() that takes a word and the name of a section of the Brown Corpus as arguments, and computes the frequency of the word in that section of the corpus.
21. Write a program to guess the number of syllables contained in a text, making use of the CMU Pronouncing Dictionary.
22. Define a function hedge(text) that processes a text and produces a new version with the word 'like' between every third word.
23. Zipf’s Law: Let $f(w)$ be the frequency of a word w in free text. Suppose that all the words of a text are ranked according to their frequency, with the most frequent word first. Zipf’s Law states that the frequency of a word type is inversely proportional to its rank (i.e., $f \times r = k$, for some constant k). For example, the 50th most common word type should occur three times as frequently as the 150th most common word type.
 - a. Write a function to process a large text and plot word frequency against word rank using pylab.plot. Do you confirm Zipf’s law? (Hint: it helps to use a logarithmic scale.) What is going on at the extreme ends of the plotted line?
 - b. Generate random text, e.g., using random.choice("abcdefg "), taking care to include the space character. You will need to import random first. Use the string concatenation operator to accumulate characters into a (very) long string. Then tokenize this string, generate the Zipf plot as before, and compare the two plots. What do you make of Zipf’s Law in the light of this?
24. Modify the text generation program in Example 2-1 further, to do the following tasks:

Критерии оценки (в баллах) результатов для группового опроса:

5 балла выставляется студенту, показавшему умение применять знания теории межкультурной коммуникации на практике, свободно и аргументировано обосновывать решение конкретных задач;

3 балл выставляется студенту, показавшему не достаточно глубокое знание теории межкультурной коммуникации, не умеющему в полной мере свободно и аргументировано обосновать решение конкретных задач;

0 баллов выставляется студенту, который не понимает поставленной задачи и не

способнее верно решить.

Индивидуальные задания

Примеры индивидуальных заданий студентов:

1. Вычислите коэффициент лексического разнообразия для жанра *News* в корпусе *Brown*.
2. Какая из записей ошибочна. Исправьте. Какие данные будут получены на выходе?

```
1. <<< cfd = nltk.ConditionalFreqDist
( ... (target, file[:4])
... for fileid in inaugural.fileids() ... for w in inaugural.words(fileid)
... for target in ['america', 'citizen']
... if w.lower().startswith(target))
```

```
2. . <<< cfd = nltk.ConditionalFreqDist
( ... (target, fileid[:4])
... for fileid in inaugural.fileids() ... for w in inaugural.words(fileid)
... for target in ['america', 'citizen']
... if w.lower().startswith(target))
2.
```

Индивидуальные задания выполняются в письменной форме и являются частью текущего контроля. Оцениваются в соответствии с рейтинг-планом дисциплины.

Контрольная работа по Модулю 1.

1. What is the difference between the following two lines? Which one will give a larger value? Will this be the case for other texts?

```
>>> sorted(set([w.lower() for w in text1]))
>>> sorted([w.lower() for w in set(text1)])
```

2. What is the difference between the following two tests: `w.isupper()` and `not w.islower()`?
3. Write the slice expression that extracts the last two words of `text2`.
4. Find all the four-letter words in the Chat Corpus (`text5`). With the help of a frequency distribution (`FreqDist`), show these words in decreasing order of frequency.
5. Review the discussion of looping with conditions in Section 1.4. Use a combination of `for` and `if` statements to loop over the words of the movie script for *Monty Python and the Holy Grail* (`text6`) and print all the uppercase words, one per line.
6. Write expressions for finding all words in `text6` that meet the following conditions. The result should be in the form of a list of words: `['word1', 'word2', ...]`.
 - a. Ending in `ize`
 - b. Containing the letter `z`
 - c. Containing the sequence of letters `pt`
 - d. All lowercase letters except for an initial capital (i.e., titlecase)

Контрольная работа по Модулю 3.

1. Define `sent` to be the list of words `['she', 'sells', 'sea', 'shells', 'by', 'the', 'sea', 'shore']`. Now write code to perform the following tasks:
 - a. Print all words beginning with `sh`.
 - b. Print all words longer than four characters
2. What does the following Python code do? `sum([len(w) for w in text1])` Can you use it to work out the average word length of a text?

3. Define a function called `vocab_size(text)` that has a single parameter for the text, and which returns the vocabulary size of the text.
4. Define a function `percent(word, text)` that calculates how often a given word occurs in a text and expresses the result as a percentage.
5. We have been using sets to store vocabularies. Try the following Python expression: `set(sent3) < set(text1)`. Experiment with this using different arguments to `set()`. What does it do? Can you think of a practical application for this?

Критерии оценки (в баллах) письменной контрольной работы:

- 10 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;
- 5 баллов выставляется студенту, показавшему не достаточно глубокое знание теоретического материала, не умеющему в полной мере аргументировано обосновать решение конкретных задач;
- 0 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.

Учебно-методическое и информационное обеспечение дисциплины

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

Основная литература:

1. Елагина, Ю.С. Практикум по устному переводу: учебное пособие / Ю.С. Елагина; Министерство образования и науки Российской Федерации, Оренбургский Государственный Университет. - Оренбург: ОГУ, 2017. - 107 с. - Библиогр.: с. 95-98 - ISBN 978-5-7410-1648-0; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=481754> .
2. Терехова, Е.В. Двусторонний перевод общественно-политических текстов (с элементами скорописи в английском языке): учебное пособие / Е.В. Терехова. - 3-е изд., стер. - Москва: Издательство «Флинта», 2017. - 320 с. - Библиогр. в кн. - ISBN 978-5-89349-955-1; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=115136>

Дополнительная 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>).

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. библиотеч. 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. библиотеч. 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. библиотеч. онлайн» – URL: <http://biblioclub.ru/index.php?page=book&id=229302> .
6. Гунина, Н.А. Компьютер для работы и досуга = ComputerforWorkandLeisure: учебное пособие / Н.А. Гунина, И.В. Шеленкова, А.А. Шиповская. – Тамбов: Издательство ФГБОУ ВПО «ТГТУ», 2012. – 137 с. – ЭВИ в ЭБС «Универс. библиотеч. 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

9. Научно-образовательный портал «Лингвистика в России: ресурсы для исследователей»: http://uisrussia.msu.ru/linguist/_B7_komp_tehn_v_prepodavanii_jazykov.jsp
10. Универсальная научно-популярная онлайн-энциклопедия. http://www.krugosvet.ru/enc/gumanitarnye_nauki/lingvistika/GERMANSKIE_YAZIKI.html
11. Свободная энциклопедия Википедия: http://ru.wikipedia.org/wiki/Германские_языки
12. Библиотека Русского гуманитарного интернет-университета: <http://www.i-u.ru/biblio>
13. Лингвоинфо: интернет-журнал: <http://www.lingvoinfo.com>
14. Онлайн-энциклопедия <http://www.krugosvet.ru>
15. Русский филологический портал Philology.ru.: <http://philology.ru/linguistics1.html>
16. Центр лингвистической документации: <http://www.mccme.ru/ling/index.htm>
17. Windows 8 Russian. Windows Professional 8 Russian Upgrade. Договор № 104 от 17.06.2013 г. Лицензии бессрочные
18. Microsoft Office Standard 2013 Russian. Договор № 114 от 12.11.2014 г. Лицензии бессрочные

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

<i>Наименование специализированных аудиторий, кабинетов, лабораторий</i>	<i>Вид занятий</i>	<i>Наименование оборудования, программного обеспечения</i>
1	2	3
Учебная аудитория для проведения занятий семинарского типа: аудитория №31(мультимедийный класс), (ул. Коммунистическая, д. 19, лит. А, А1)	Практические занятия	<p align="center">Аудитория №31</p> Учебная мебель, учебно-наглядные пособия, доска, моноблоки – 13 шт. с выходом в Интернет, обеспечивающие доступ к электронной информационно-образовательной среде (ЭИОС) вуза, мультимедиа-проектор MitsubishiEX320U XGA, экран настенный Classic Norma 244*183, ноутбук ASUSX51RL (место хранения деканат ФРГФ, ауд.№ 6а)
Аудитория для текущего контроля и промежуточной аттестации: Аудитория №4, Аудитория № 11 (ул. Коммунистическая, д. 19, лит. А, А1)	Групповые и индивидуальные консультации, текущий контроль и промежуточная аттестация	<p align="center">Аудитория №4</p> Учебная мебель, учебно-наглядные пособия, доска, моноблоки – 12 шт. с выходом в Интернет, обеспечивающие доступ к электронной информационно-образовательной среде (ЭИОС) вуза <p align="center">Аудитория №11</p> Учебная мебель, учебно-наглядные пособия, доска
Помещения для самостоятельной работы: аудитория №13 (читальный зал) (ул. Коммунистическая, д. 19, лит. А, А1)	Самостоятельная работа	<p align="center">Аудитория №13</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 Windows 8 Russian. Windows Professional 8 Russian Upgrade. Договор № 104 от 17.06.2013 г. Лицензии бессрочные Microsoft Office Standard 2013 Russian. Договор № 114 от 12.11.2014 г. Лицензии бессрочные.

ФГБОУ ВО «БАШКИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

ФАКУЛЬТЕТ РОМАНО-ГЕРМАНСКОЙ ФИЛОЛОГИИ

СОДЕРЖАНИЕ РАБОЧЕЙ ПРОГРАММЫ²дисциплины _____ Программирование для лингвистов ___ на ___ 3 ___ семестр
(наименование дисциплины)_____ очная _____

форма обучения

Вид работы	Объем дисциплины
Общая трудоемкость дисциплины (з.е. / часов)	2/72
Учебных часов на контактную работу с преподавателем:	18,2
лекций	-
практических/ семинарских	18
лабораторных	-
других (групповая, индивидуальная консультация и иные виды учебной деятельности, предусматривающие работу обучающихся с преподавателем) (ФКР)	0,2
из них, предусмотренные на выполнение курсовой работы / курсового проекта ³	
Учебных часов на самостоятельную работу обучающихся (СР)	53,8
из них, предусмотренные на выполнение курсовой работы / курсового проекта ⁴	-
Учебных часов на подготовку к экзамену/зачету/дифференцированному зачету (Контроль)	-

Форма(ы) контроля:

экзамен _____ - _____ семестр

зачет _____ 1 _____ семестр

курсовая работа / курсовой проект ___ - ___ семестр

² Количество часов/з.е. указывается в соответствии с учебным планом, таблицы заполняются отдельно по каждой форме обучения (очной, очно-заочной, заочной).

³ Контактных часов – 2

⁴ Количество часов на самостоятельную работу указывается на усмотрение разработчика, но не более 20 часов

№ п/п	Тема и содержание	Форма изучения материалов: лекции, практические занятия, семинарские занятия, лабораторные работы, самостоятельная работа и трудоемкость (в часах)				Задания по самостоятельной работе студентов ⁵	Форма текущего контроля успеваемости (коллоквиумы, контрольные работы, компьютерные тесты и т.п.)
		ЛК	ПР/СЕМ	ЛР	СР		
1	2	3	4	5	6	8	9
1.	Основы обработки естественного языка на языке программирования Питон. Токенизация. Лемматизация. Стемминг. Обработка корпусных данных.	-	6	-	18	Основная литература 1	Групповой опрос, индивидуальные задания, контрольная работа
2.	Загрузка библиотек, построение графов. Вычисление коэффициента лексического разнообразия текста. Определение модальности текста путем переформатирования корпуса.	-	6	-	18	Основная литература 1	Групповой опрос, индивидуальные задания
3.	Работа с модулем условного частотного	-	6	-	18.8	Основная литература 1, 2	Групповой опрос, индивидуальные

⁵ К заданиям для самостоятельной работы можно отнести, например: подготовку к индивидуальному или групповому опросу; выполнение домашних заданий; подготовку к лабораторным работам, контрольным работам, собеседованиям, коллоквиумам; изучение теоретического материала; подготовку докладов и сообщений; написание эссе, рефератов и статей; подготовку проектов и творческих заданий (выступлений, презентаций, кроссвордов и пр.) и т.д.

	распределения. Построение кортежей Работа с фонетическими корпусами						задания, контрольная работа
	Всего часов:		18		53.8		

